

ABSTRACT

LI, WUYUAN. Fundamental Limits of Estimation Using Arbitrary Compact Arrays. (Under the direction of Dr. Brian L. Hughes.)

Antenna arrays play a central role in a wide variety of important estimation problems. Most existing literatures on array signal processing focus on conventional arrays, where antennas are separated by a relatively large distance so that they are uncoupled. While such arrays usually have good estimation accuracy and are easy to design, they may be too large for platforms with size limitations, such as cellular handsets and wireless sensor nodes. Instead, deploying multiple antennas on such platforms requires placing antennas close together, which can cause interactions among the elements. Such interactions can profoundly impact received power and estimation error. Moreover, estimation performance depends not only on the properties of the array, but also on aspects of the receiver front-end, such as antenna impedance matching, amplifier properties, and the dominant sources of noise. Although many prior works have studied the performance of compact arrays from different perspectives, to the best of our knowledge, no one has yet considered the impact of impedance matching on the performance of estimators using physical models of observation noise. Besides studying the performance of any given arrays, perhaps a more interesting yet challenging question to ask is how the properties of the array itself changes the estimation accuracy, which very few papers have looked into.

In this dissertation, we investigate three aspects of multiple antennas receiver design. Firstly, we consider a general class of Bayesian and non-Bayesian estimation problems, in which the signal of interest is observed through a sensor front-end consisting of coupled antennas, an impedance matching network, amplifiers, and physical noise sources. We derive the Fisher information associated with each problem and prove that one kind of

matching network is universally optimal for all estimation problems in the class. Secondly, we consider the general problem of estimating the parameters of an incident Gaussian electromagnetic field using a sensor array that observes the field through the currents in an arbitrary conductor. We characterize the maximum Fisher information that can be achieved with this array and derive conditions under which the upper bound can be attained. Lastly, we study properties of the antenna conductor itself and show how the size and shape of the conductor impact the estimation accuracy. We apply our results to square and spherical conductors to demonstrate how the Fisher information increases with the antenna size. We further study the performance of several widely-studied compact arrays and evaluate their efficiency in observing the information contained in the spaces they occupy.

© Copyright 2018 by Wuyuan Li

All Rights Reserved

Fundamental Limits of Estimation Using Arbitrary Compact Arrays

by
Wuyuan Li

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Electrical Engineering

Raleigh, North Carolina

2018

APPROVED BY:

Dr. Jacob J. Adams

Dr. Brian A. FLoyd

Dr. J. Keith Townsend

Dr. Muhammad Shahzad

Dr. Brian L. Hughes
Chair of Advisory Committee

DEDICATION

To my parents and Yi.

BIOGRAPHY

Wuyuan Li received his bachelor degree in Information Engineering from Beijing Institute of Technology, China in June 2011. In August 2011, Wuyuan began working toward his PhD in Electrical Engineering at North Carolina State University under the advisement of Prof. Brian L. Hughes in Wireless Systems Engineering Lab. During the course of his PhD, he worked at Nokia Technologies as a research intern from Jun 2016 to August 2016, where he worked on radar signal processing.

His research interests include Wireless Communication, Estimation Theory, MIMO System, and Digital Signal Processing.

ACKNOWLEDGEMENTS

I have been very fortunate to have had Prof. Brian L. Hughes as my advisor, without whom the work in this dissertation would not have been possible. The continuous guidance and support from him made my PhD experience truly rewarding. I have benefited tremendously from his technical expertise and analytical skills.

I am very grateful to the professors serving in my committee board, Dr. Jacob J. Adams, Dr. Brian A. Floyd, Dr. J. Keith Townsend, and Dr. Muhammad Shahzad, for their helpful comments and inspiring questions. I would like to especially thank my co-authors Dr. Jacob J. Adams and Binbin Yang for their help in antenna simulation. The numerous discussions I had with them greatly enriched my understanding of electromagnetic fields. I also want to thank National Science Foundation for supporting the research projects I worked on.

I want to sincerely thank Dr. Prasad Shamain and Dr. Klaus Doppler of Nokia Bell Labs for offering me the research intern position in June 2016.

I am thankful to all my fellow research group members, Lopamudra Kundu, Shaohan Wu, and Divyakumar Mahiman Badheka, for the collaboration and camaraderie. I would also like to thank all my friends for all the memorable moments we had together.

I am indebted to my parents and grandparents. Their unconditional love has laid the foundation for any success I have ever achieved. I also want to express my gratitude to my wife, Yi. Her enduring support helped me so much in going through all the highs and lows of graduate school.

TABLE OF CONTENTS

List of Figures	vi
Chapter 1 INTRODUCTION	1
1.1 Antenna Coupling	2
1.2 Matching Network	5
1.3 Compact Antenna Arrays	8
1.4 Outline	10
Chapter 2 Matching Network for Compact Sensor Arrays	12
2.1 System Model	14
2.1.1 Sensor Array	15
2.1.2 Front-End	16
2.1.3 Matching Networks	17
2.2 Fundamental Bounds on Estimation	19
2.2.1 Cramér-Rao Bound for Real Parameters	19
2.2.2 Optimal Matching	20
2.2.3 Bayesian Cramér-Rao Bound	23
2.2.4 Complex Parameters	24
2.3 Two Spatial Estimation Problems	25
2.3.1 MIMO Channel Estimation	25
2.3.2 Direction-of-Arrival Estimation	28
2.4 Numerical Results	30
2.4.1 MIMO Channel Estimation	31
2.4.2 Direction-of-Arrival Estimation	34
2.5 Conclusions	39
Chapter 3 Performance Limits of Arbitrary Compact Arrays	40
3.1 System Model	42
3.1.1 Sensor Array	42
3.1.2 Front-End	45
3.1.3 Cramér-Rao Bound	46
3.1.4 CRB Optimal Matching	47
3.2 Arrays from Arbitrary Conductors	49
3.2.1 Fields and Currents on Lossy Conductors	50
3.2.2 Embedded Ports	51
3.3 Aperture Fisher Information	56
3.3.1 Bounds on Arbitrary Arrays	57
3.3.2 Optimal Array Design: $\Sigma_{\mathbf{y}_V}$ Independent of Parameters	61

3.3.3	Optimal Array Design: μ_{y_V} Independent of Parameters	63
3.4	Numerical Examples	69
3.4.1	Estimating the Temperature of Black-Body Radiation	69
3.4.2	Estimating Direction-of-arrival of a Known Signal	74
3.4.3	Two-dimensional DoA Estimation using Tripoles	78
3.5	Conclusion	80
Chapter 4 Impact of Antenna Aperture on Fisher Information		82
4.1	Antenna Shape	83
4.1.1	Black-Body Radiation Temperature Estimation	86
4.1.2	One-dimensional DoA Estimation	90
4.1.3	Two-dimensional DoA Estimation	92
4.2	Conventional Sensor Arrays	95
4.2.1	Dipole ULA	95
4.2.2	Tripole	97
4.2.3	Vector Sensor	98
4.3	Conclusion	102
Chapter 5 Conclusion and Future Work		104
5.1	Summary of Dissertation	104
5.2	Future Work	107
References		109
APPENDICES		117
Appendix A	Proof of theorem 2.1	118

LIST OF FIGURES

Figure 2.1	Circuit model of a sensor array front-end	15
Figure 2.2	BCRB on mean-squared channel estimation error for a 4×4 MIMO channel	32
Figure 2.3	Ratio BCRB-SM/BCRB-MNF versus F_{min} for SNR=10 dB	33
Figure 2.4	CRB on RMS angular error for 4-element array with one target at $\phi = 90^\circ$ and $T = 200$	35
Figure 2.5	CRB on RMS angular error versus target angle ϕ for $d = 0.1\lambda$	36
Figure 2.6	CRB on RMS angular error for targets at 90° and 120° with $\rho = 0$, SNR=10 dB.	37
Figure 2.7	CRB on RMSE of target at 90° with second target at $90^\circ + \Delta\phi$ and SNR=10 dB	38
Figure 3.1	Circuit model of a sensor array front-end	43
Figure 3.2	An array formed from an arbitrary conductor V	52
Figure 3.3	Three arrays formed from the wire V	70
Figure 3.4	Eigencurrent efficiencies and CRBs for temperature estimation example.	73
Figure 3.5	Spherical coordinates	75
Figure 3.6	CRB on RMSE versus target angle θ using half-wavelength wire.	77
Figure 3.7	CRB on RMSE versus arrival angle ϕ for signal arriving in $\theta = 90^\circ$ plane.	78
Figure 3.8	CRB on RMSE versus arrival angle ϕ for signal arriving in $\theta = 60^\circ$ plane.	80
Figure 4.1	Eigencurrent efficiencies of a Square	86
Figure 4.2	$\mathcal{I}_V(1)$ versus side length l for a square aperture	87
Figure 4.3	Aperture CRB on Temperature Estimation using a Square	87
Figure 4.4	Eigencurrent efficiencies of a sphere	88
Figure 4.5	$\mathcal{I}_V(1)$ versus radius r for a spherical aperture	89
Figure 4.6	Aperture CRB on Temperature Estimation using a sphere	89
Figure 4.7	A uniform linear array of 2 to 5 dipoles	91
Figure 4.8	Aperture CRB on ϕ estimation using a ULA	91
Figure 4.9	Aperture CRB on DoA estimation	94
Figure 4.10	CRBs for 1 port per wire	96
Figure 4.11	Array CRB versus Aperture CRB for ϕ estimation	97
Figure 4.12	Array CRB for a tripole, $r = 0.1\lambda$	99
Figure 4.13	Array CRB for a tripole, $r = 0.25\lambda$	99
Figure 4.14	Antenna ports on vector sensor	101
Figure 4.15	Array CRB for a vector sensor, $r = 0.1\lambda$	101

Figure 4.16 Array CRB for a vector sensor, $r = 0.25\lambda$ 102

Chapter 1

INTRODUCTION

For decades, antenna arrays have been used in a series of real-world estimation problems like radar, astronomy, seismology, and medical diagnosis and treatment [1, 2]. The radio signals collected by array elements are usually jointly considered to enable accurate temporal or spatial parameter estimation. One typical example of early algorithms used in array signal processing is beamforming, which performs spatial filtering to separate signals coming from different directions [3]. Later, more complex subspace-based approaches such as MUSIC and its variations have been proposed to obtain information about multiple targets or multipath channels [4, 5].

Besides estimation applications, researches on wireless communications show that arrays can also increase achievable data rates under multipath channels, especially when both the transmitter and receiver are equipped with multiple antennas [6, 7]. This is usually referred as Multiple-Input, Multiple-Output (MIMO) techniques, and has become part of many modern communication standards (e.g. LTE, 802.11n, WiMax). Most studies on MIMO assume perfect channel state information (CSI) is available at least to the receiver. However, such information has to be estimated in practice, usually by sending

a sequence of known training symbols. Therefore, parameter estimation is also playing an important role in multiple-antenna wireless communications.

In this dissertation, we study the performance of multiple-antenna receivers in general parameter estimation applications. We firstly introduce circuit models for different RF components in a receiver that incorporates potential antenna mutual coupling, then analyze the impact of matching network for any given receiver. After deriving the optimal matching network, we look at the properties of antenna array itself and propose a new method to model antenna ports by the currents they excite on the antenna aperture. Based on this model, we introduce a new quantity called Aperture Fisher information, which is the upper bound on the Fisher information obtainable using any antenna arrays built from an arbitrary conducting aperture. Finally, we check the antenna aperture itself, explore how its shape and conductivity affect the estimation accuracy, then use this result as a benchmark to evaluate the performance of several widely-studied arrays.

1.1 Antenna Coupling

Conventionally, antenna elements in an array are separated by a relatively large distance so that they do not interact with each other. While this kind of electrically large arrays have advantages like easy to design and utilize, they can not be applied to many applications. For example, in low frequency bands like HF and UHF, a several wavelengths long array can be too big for vehicle-mounted or person-born systems; even for applications at a relatively high frequency, size-limited platforms such as cell-phones or wireless access points may not support such arrays. Therefore, it is important to study the performance of compact arrays, where antennas are placed close to each other.

Antennas in close proximity cannot be modeled as independent elements anymore.

Instead, they interact with each other in several aspects. Current flowing in one antenna can induce voltages in other antennas, changing the array impedance and increasing the difficulty of impedance matching, this is commonly known as the mutual coupling effect. Radiation patterns of elements are distorted by the scattering effect. Furthermore, the signal and noise may become spatially correlated. Without proper handling, the system performance of a compact array may be much worse than arrays composed of well-separated elements.

Early works on compact sensor arrays mostly consider mutual coupling effect only, which is modeled to change the output voltage of the antenna array. Based on this model, researchers proposed a simple “decoupling” method, which is to multiply the output signal by the inverse of mutual coupling matrix [8, 9]. A large number of works have been conducted on Direction-of-Arrival (DoA) estimation following this method. Yeh *et al* [10] considered how antenna patterns are changed by known mutual coupling and showed that DoA estimation accuracy can be improved by incorporating mutual coupling into the corresponding search vector. Friedlander and Weiss [11] considered the problem of jointly estimating the DoA, path gain, and mutual coupling, and proposed an algorithm to estimate these parameters iteratively. Instead of evaluating the performance of specific algorithms, Svantesson [12, 13] calculated the Cramér-Rao Bound (CRB), the lower bound on error covariance for any unbiased estimators, for DoA estimation in the presence of mutual coupling. He reached the conclusion that mutual coupling does not significantly impact the estimation accuracy. However, he mostly focused on Uniform Linear Array with large inter-element spacing (half-wavelength), where the mutual coupling effect is relatively weak. Lau and Bach-Anderson [14] explored similar problem with various array size and showed that CRB increases significantly when the coupling is strong, but this effect can be greatly canceled by using a well-designed multi-port impedance match-

ing network, which will be explained in further details in the next section. There are many other publications on array pattern calibration or mutual coupling compensation techniques for better DoA estimation accuracy, such as [15–18].

Unlike DoA estimation, relatively few studies have considered channel estimation using compact antenna arrays directly. However, there are a significant amount of publications on a related problem, correlated MIMO channel estimation. Channel estimation methods can be divided into three categories based on whether a training sequence (pilot symbols) is used: training sequence based estimation, blind estimation, and semi-blind estimation. The CRB for each case has been derived in [19]. For training sequence based correlated channel estimation, Biguesh and Gershman [20] studied the performance of four estimators including the popular least square estimator and minimum-mean-square-error estimator, and derived the optimal training sequences for each estimator under different channel correlation conditions. Blind estimation has been widely explored in OFDM-related applications to enable the coherent signal detection, methods to perform blind estimation for MIMO-OFDM channels include subspace-based estimator, precoding method, noise subspace method, second-order statistics, and so on (e.g. [21] and references therein). For compact arrays with coupling effect, training sequence based channel estimation has been considered in [22,23] by Liu *et al*, where they investigated the performance of several estimators. Their results show that coupling can significantly increase the channel estimation error.

Although effects of mutual coupling on estimation problems are modeled in these papers, people mostly consider spatially white noise with a fixed distribution, which is independent of the antennas and matching network. A few papers on DoA estimation use a physical model of antenna and amplifier noise instead and study how the noise varies with coupling effect, but only for antennas with a fixed 50Ω load [24,25]. Existing works

on MIMO communication systems have shown that impedance matching network can significantly cancel the coupling effect and increase the capacity of a compact array [27], which makes it reasonable to expect the matching network would change the estimation accuracy as well. It is therefore of interest to study how the mutual coupling effect and matching network change the system performance of a compact array using practical noise model. In the next section, we briefly introduce the matching network concept and discuss the performance of different matching networks.

1.2 Matching Network

In circuit theory, it is well known that maximum power can be transferred from the source to the load when the load impedance is matched to the source impedance. Antennas can be viewed as the power source in the sense that electromagnetic waves are received by antennas then transferred into low-noise amplifiers (LNA) and following circuits. For Single-Input, Single-Output (SISO) systems, various matching methods have been introduced in [35, Chap. 5]. For an uncoupled array, each antenna is independent and can be matched individually using similar methods. However, with mutual coupling, the entire antenna array should be viewed as one multiport system instead. Multiport matching technique for maximum power transfer is discussed in [36], examples include *multiport conjugate match* [37].

Instead of focusing on the signal power only, perhaps a better goal for matching network is to consider the signal and noise jointly and try to maximize the Signal-to-Noise Ratio (SNR). RF front-end usually induces a certain amount of noise, which decreases the SNR by a ratio referred as the noise factor (or *noise figure* in dB). Studies have shown that matching network can decrease the noise factor in a SISO system [35, Chap. 10.2].

We call any matching networks that minimize the noise factor *minimum-noise-factor (MNF) match*. The multiport version of MNF match is derived in [27] and [38] under independent noisy amplifiers assumption, the basic idea is to decouple the antenna array impedance and match each output port to the optimal input impedance for the amplifier.

Although multiport matching networks like the multiport conjugate match and MNF match have many advantages in theory, the actual implementation can be difficult. Therefore, people also proposed a series of suboptimal yet easier to implement single port matching methods, where the signal received by one antenna will only be transferred to one amplifier. Single port matching can be seen as a special case of multiport matching, where the cross-connection between antennas and amplifiers are set to be open. Examples of single port matching are self-match, input-impedance match, 50Ω match, etc [40–42]. These matching networks have limited abilities in eliminating the mutual coupling effect, therefore their performance is usually worse than multiport matching, especially when the sensor array is compact.

The impact of different matching networks towards system performance has been studied from various aspects. Based on network analysis, receiver noise in compact antenna arrays is analyzed in [27, 39, 43, 44], where the noise sources considered are exterior environment (antenna noise) and LNAs. Gans studied both the antenna noise dominant and amplifier noise dominant scenarios and showed that in the former case matching network does not change the system performance [43], yet in the latter case the receiver noise can be significantly reduced by matching, resulting in a much better system performance [44]. A more general noise model is considered in [27], where formulas for noise covariance and system capacity are derived. One important conclusion from [27] is that assuming amplifiers are independent, the system capacity under various fading and CSI conditions are all maximized by using the MNF match. It is also found that MNF match

can eliminate the mutual coupling to such a degree that the capacity of a compact MIMO system is largely comparable to the ideal uncoupled system. While multi-port matching like MNF match provides a theoretical upper bound for system performance, people also studied the performance of single port matching such as self-match [27]. It is shown that when the amplifier noise is non-negligible, there may be a significant performance gap between self-match and MNF match under strong coupling conditions. The optimal single port matching network has been derived for a coupled 2×2 MIMO system in high SNR scenarios in [42], which is referred as input-impedance match. For circularly symmetric arrays, Taluja and Hughes [45] considered a different approach, where they decouple the antenna impedance first then apply single port match to the decoupled ports.

On estimation problems, although there are some efforts in designing “coupling compensation” algorithm from signal processing perspective (e.g. [16]), not many works are conducted based on circuit analysis under a physical noise model. As mentioned before, Lau and Anderson [14] considered DoA estimation problem using a compact Uniform Linear Array with different inter-element spacing, then compared the performance of 50Ω match, self-match, and Multiport-conjugate match. Under the fixed spatially white noise model, they concluded that matching network can significantly reduce estimation error under strong coupling conditions.

While so many research results have shown that matching network can improve the system performance of a compact array, so far we have not discussed how properties of the array itself affect the estimation accuracy. In the next section, we summarize prior works on this topic and discuss different antenna array structures.

1.3 Compact Antenna Arrays

As discussed before, many researches on compact arrays are focused on uniform arrays consist of simple antenna elements like omni-directional dipoles [14–18]. Meanwhile, people have also considered a series of more complicated structures, among which the vector sensor and its variations (tripole, triloop, etc) are exemplary. A vector sensor is a six-antenna array composed of three co-located orthogonal dipoles and three co-located orthogonal loops. An ideal vector sensor is considered to be capable of observing all the electromagnetic field components at one point while exhibiting almost no mutual coupling between antennas. Many studies have been performed on DoA and polarization estimation using ideal vector sensor(s) [46, 47, 51, 53] and its variations [48–50, 52, 54]. A structure similar to the triloop is Giselle or ground-symmetric loop (GSL) array, which consists of three co-located orthogonal cross-connected twin-loops, built for DoA estimation purpose on the 3 to 30 MHz HF band [57]. The reason for using cross-connected Twin-loops is that the null in the radiation pattern is constant over the frequency of interest, which simplifies the estimation algorithm in broadband application. Recently, a few two-element arrays, inspired by the biological acoustic orienting mechanism, have been proposed for DoA estimation purpose. Xin and Ding [55] put a lossy scatter between two antennas to emulate the low-pass filtering function of the human head, and Behdad et al. [56] use a coupling network between two monopole antennas to mimic the hearing mechanism of *Ormia ochracea*. Other examples of electrically small antenna array structures can be found in [58, 59] and the references therein.

Although so many compact array structures have been explored already, it is still difficult to draw a conclusion on which one is the optimal for parameter estimation applications. Perhaps a more systematic way to study a compact array is to use *Theory of*

Characteristic Modes [60–63]. Antennas can generally be divided into two parts - a conducting structure whose (surface) current is excited by incoming electromagnetic waves, and antenna ports through which such current can be indirectly observed. Instead of analyzing the radiation and impedance of any particular array, the theory of characteristic modes analyzes the entire underlying antenna structure. Based on this theory, the current and radiation pattern of any structure can be represented as a combination of a series of orthogonal “characteristic modes”. There are several examples of antenna array design and analysis based on this theory, either on signal processing applications or communication applications. Akers and Corbin [68] calculated the radiation pattern of 16 characteristic modes of a cross-shaped patch, then studied its DoA estimation accuracy when different combinations of these modes are used. Li et al. [69] designed an array whose size fits a cellphone, where they use two ports to excite two resonant characteristic modes of the antenna structure. Gustafsson et al. performed a series of studies on communication and estimation using small antenna arrays based on spherical wave expansion, which is equivalent to the characteristic modes of a sphere aperture [64–67]. Although the theory of characteristic modes gives us some insight into the underlying antenna aperture, it is not clear which modes are preferable to perform estimation. Moreover, exciting certain modes in particular may require external matching network [70], which can be difficult to implement.

Very recently, a new method of systematic antenna analysis for communication purpose was proposed by Kundu and Hughes [79], where the current on the antenna is divided into a set of eigen-modes according to how much ergodic capacity they contain. The interaction between ports and the underlying structure are modeled by a *coupling network* which works similarly to the impedance matching network. Their results show that any lossy antenna array built on certain aperture has a fundamental limit on the

antenna ergodic capacity, which is determined by the shape and conductivity of the aperture. For any given amount of antenna ports, there exists a coupling network which can maximize the system capacity. It is possible that similar method can be developed for estimation applications.

1.4 Outline

In this dissertation, we study general parameter estimation problems based on antenna theory and circuit theory, look for performance bound for antenna arrays with size limitation, and re-evaluate the performance of existing arrays. This dissertation is organized as follows.

In chapter 2, we consider a general class of Bayesian and non-Bayesian estimation problems, in which the signal of interest is observed through a sensor front-end consisting of coupled antennas, an impedance matching network, amplifiers, and physical noise sources. We derive the Cramér-Rao Bound (CRB) associated with each problem and explore how it varies with impedance matching and the physical characteristics of the signals and noise. We show that minimum-noise-factor matching is universally optimal in the sense of minimizing the CRB for all estimation problems in the class. We then apply these results to multi-antenna channel estimation and direction-of-arrival estimation.

In chapter 3, we consider the general problem of estimating the parameters of an incident Gaussian electromagnetic field using a sensor array that observes the field through the currents in an arbitrary conductor V . We characterize the maximum Fisher information that can be achieved with this array using a front-end consisting of a matching network followed by a bank of noisy amplifiers. This Fisher information is always bounded above by a quantity we call *aperture Fisher information* (AFI), which intuitively repre-

sents all of the useful information contained in the currents of V . For a broad class of estimation problems, we show that AFI can be associated with a system of unit-power eigencurrents on V , which decompose the volume currents into a sequence of independent scalar signal observations. We show the Fisher information of any M -port array is bounded above by the information in the best M eigencurrents, and derive conditions under which equality can be attained. Finally, we apply our results to estimation examples involving wire arrays.

In chapter 4, using the framework we developed in previous chapters, we analyze how the estimation accuracy changes with the aperture properties. We prove that under the same radiation environment, the AFI of a structure is always greater than or equal to the AFI of its substructures. For radiation intensity and DoA estimation applications, we use square and spherical apertures as examples to study how the AFI increases with the aperture size. We then consider several widely-studied antenna arrays, which are substructures of squares and spheres, and evaluate their ability to observe the Fisher information contained in the spaces they occupy.

Finally, we present conclusions and discuss potential future works in Chapter 5.

Chapter 2

Matching Network for Compact Sensor Arrays

We begin our study by building a framework to analyze the performance of arbitrary MIMO system in estimation applications. In addition to introducing the system model and performance metric, this chapter also studies what kind of matching network maximizes the estimation accuracy. Results in this chapter lay the foundation to explore other interesting topics in following chapters.

Most literatures on array signal processing focus on well-separated antenna arrays, where antenna elements are assumed to be independent. However, deploying arrays on size-limited platforms like cellular handsets requires placing antennas close together. As discussed in the previous chapter, this can cause interactions among the elements: antennas exhibit mutual coupling, radiation patterns are distorted by scattering, and the signal and noise fields detected by different elements may be spatially correlated. These interactions can profoundly impact received power and estimation error. Moreover, estimation performance depends not only on the properties of the array, but also on aspects

of the receiver front-end, such as antenna impedance matching, amplifier properties, and the dominant sources of noise (both internal and external).

Several authors have considered the impact of antenna coupling on important estimation problems. Some have investigated the effect of coupling on receiver antenna patterns [10, 11]. Svantesson [12] derived a Cramér-Rao Bound (CRB) for direction-of-arrival (DoA) estimation in the presence of coupling. Using a similar approach, Lau and Bach-Anderson [14] showed that strong coupling can significantly degrade this CRB, but losses can be greatly reduced by using multi-port impedance matching between the array and front end. In [22, 23], Liu *et al* investigated the performance of certain MIMO channel estimators in the presence of coupling and showed that strong coupling can significantly increase estimation error. All of these papers model the effects of coupling on the received signal, but none consider the impact of coupling and matching on physical noise sources. Instead, observation noise is modeled as spatially white, with a fixed distribution that is independent of the antennas and matching. Conversely, other studies of DoA estimation with coupling have considered physical models of antenna and amplifier noise [24, 25], but only for antennas with a fixed 50Ω termination. To the best of our knowledge, no one has yet considered the impact of impedance matching on the performance of estimators using physical models of observation noise.

In this chapter, we consider a general class of Bayesian and non-Bayesian estimation problems, in which the parameters of interest are observed through a front-end consisting of coupled antennas, an impedance matching network, amplifiers, and physical sources of antenna and amplifier noise. To gain insight into the performance of the best estimators, we derive the CRB associated with each problem and explore how it varies with impedance matching and the physical characteristics of the signals and noise. We prove that one type of matching network, minimum-noise factor (MNF) matching, is universally

optimal in the sense of minimizing the CRB for all estimation problems in the class. MNF matching was previously shown to maximize the capacity of MIMO channels [26,27]. We then apply these results to MIMO channel estimation and DoA estimation for targets in the plane. Numerical results suggest the CRB is generally sensitive to the physical composition of observation noise. However, optimal matching can eliminate this sensitivity and significantly reduce estimation error, particularly when amplifier noise dominates.

The rest of the chapter is organized as follows. In Sec. 2.1, we present our system model and the class of estimation problems considered. In Sec. 2.2, we derive the associated Cramér-Rao Bound and prove it is universally minimized by MNF matching. We then apply these results to MIMO channel estimation and DoA estimation in Sec. 2.3, and consider specific numerical examples in Sec. 2.4. Finally, we summarize our conclusions in Sec. 2.5.

2.1 System Model

Consider a compact sensor array that estimates a collection of parameters $\mathbf{u} \in \mathbb{C}^p$ using observations from M antennas. Since the antennas are closely spaced, they may experience mutual coupling and the signals detected by each may be correlated. Before the signals are processed by the estimation algorithms, they traverse an analog front-end that consists of an impedance matching network in series with a bank of noisy amplifiers. A model for this system is illustrated in Fig. 2.1. Similar models have been used to study capacity in MIMO communication systems [27]; however, here we consider more general kinds of signals and noise. Each element of the model is described in detail below.

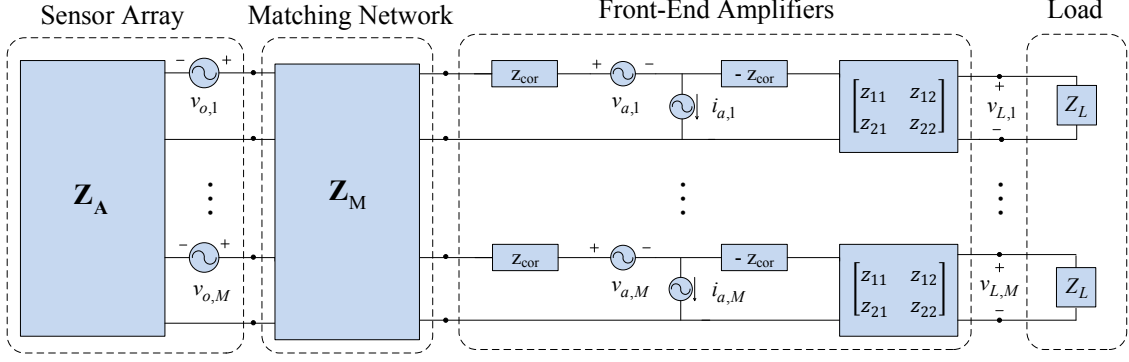


Figure 2.1: Circuit model of a sensor array front-end

2.1.1 Sensor Array

Let $\mathbf{v} \in \mathbb{C}^M$ denote the complex-baseband voltage across the terminals of the array and let $\mathbf{i} \in \mathbb{C}^M$ be the corresponding current flowing into the array. The relationship between the two can be described by

$$\mathbf{v} = \mathbf{Z}_A \mathbf{i} + \mathbf{v}_o, \quad (2.1)$$

where $\mathbf{Z}_A \in \mathbb{C}^{M \times M}$ is the array impedance matrix and \mathbf{v}_o is the open-circuit voltage induced by the incident field. Here, diagonal elements of \mathbf{Z}_A represent antenna self-impedances, while off-diagonal elements represent mutual coupling between antennas. We can always write

$$\mathbf{Z}_A = \mathbf{R}_A + j\mathbf{X}_A, \quad (2.2)$$

where $\mathbf{R}_A \triangleq \frac{1}{2}(\mathbf{Z}_A + \mathbf{Z}_A^H)$, $\mathbf{X}_A \triangleq \frac{1}{2j}(\mathbf{Z}_A - \mathbf{Z}_A^H)$, and $(\cdot)^H$ denotes the conjugate-transpose. We assume the antennas are reciprocal $\mathbf{Z}_A = \mathbf{Z}_A^T$, so $\mathbf{R}_A = \text{Re}[\mathbf{Z}_A]$ and $\mathbf{X}_A = \text{Im}[\mathbf{Z}_A]$ are both real, symmetric matrices.

Consider the problem of estimating \mathbf{u} based on indirect observations of the open-

circuit voltage \mathbf{v}_o in (2.1). In general, we assume \mathbf{v}_o can be expressed as

$$\mathbf{v}_o = \mathbf{s} + \mathbf{n}_o , \quad (2.3)$$

where $\mathbf{s} \in \mathbb{C}^M$ is the target signal and \mathbf{n}_o represents noise. Here we take \mathbf{s} to be a circularly-symmetric Gaussian random vector [74, Sec. 15.7] with mean $\boldsymbol{\mu}_s(\mathbf{u})$ and covariance $\boldsymbol{\Sigma}_s(\mathbf{u})$, which we denote by $\mathbf{s} \sim \mathcal{CN}(\boldsymbol{\mu}_s(\mathbf{u}), \boldsymbol{\Sigma}_s(\mathbf{u}))$. Two specific estimation problems, with specific choices of $\boldsymbol{\mu}_s(\mathbf{u})$ and $\boldsymbol{\Sigma}_s(\mathbf{u})$, will be considered in Sec. 2.3.

The term \mathbf{n}_o in (2.3) represents antenna noise induced in the sensor array by the surrounding environment. Here we model the noise as $\mathbf{n}_o \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Sigma}_o)$, where $\mathbf{0} \in \mathbb{C}^M$ is the all-zero vector and $\boldsymbol{\Sigma}_o$ is independent of \mathbf{u} . For example, if the sensor array is lossless and surrounded by an isotropic black-body enclosure of temperature T_0 , Twiss [75] has shown that the antenna noise covariance is given by

$$\boldsymbol{\Sigma}_o = 4k_b T_0 B \mathbf{R}_A , \quad (2.4)$$

where $k_b = 1.38 \times 10^{-23}$ J/K is Boltzmann's constant, B is the system bandwidth in Hz, and \mathbf{R}_A is the array resistance matrix in (2.2).

2.1.2 Front-End

In Fig. 2.1, the signals observed by the estimation algorithms are the load voltages $\mathbf{v}_L = (v_{L,1}, \dots, v_{L,M})^T$. Prior to reaching the load, the signals pass through a front-end consisting of a matching network followed by a bank of identical amplifiers. As in [27], we represent the amplifiers using the standard Rothe-Dahlke equivalent network, where $\mathbf{v}_a \sim \mathcal{CN}(\mathbf{0}, 4k_b T_0 B r_a \mathbf{I})$ and $\mathbf{i}_a \sim \mathcal{CN}(\mathbf{0}, 4k_b T_0 B g_a \mathbf{I})$ are independent random vectors that

represent internal amplifier noise sources, and \mathbf{I} is the identity matrix. For simplicity, we call these circuits amplifiers, but any sequence of linear, noisy circuits can be represented this way [27].

In this section, we first consider the case of no matching network in Fig. 2.1. Using basic circuit analysis, we can show the load voltages in Fig. 2.1 are given by (e.g. [39, eq. 14])

$$\mathbf{v}_L = \mathbf{K} \left[\mathbf{v}_o - \mathbf{v}_a - (\mathbf{Z}_A + z_{\text{cor}} \mathbf{I}) \mathbf{i}_a \right], \quad (2.5)$$

where $\mathbf{K} = \frac{z_{21} z_L}{z_L + z_{22}} \left(\mathbf{Z}_A + \frac{z_{11}(z_L + z_{22}) - z_{12} z_{21}}{z_L + z_{22}} \mathbf{I} \right)^{-1}$. Since optimal estimation performance is not affected by an invertible transformation of the observation, we can define the observation to be

$$\begin{aligned} \mathbf{y} &= \mathbf{K}^{-1} \mathbf{v}_L = \mathbf{v}_o - \mathbf{v}_a - (\mathbf{Z}_A + z_{\text{cor}} \mathbf{I}) \mathbf{i}_a \\ &= \mathbf{s} + \mathbf{n}_A, \end{aligned} \quad (2.6)$$

where $\mathbf{n}_A \triangleq \mathbf{v}_o - \mathbf{v}_a - (\mathbf{Z}_A + z_{\text{cor}} \mathbf{I}) \mathbf{i}_a$ is the overall system noise. Intuitively, (2.6) expresses all sources in (2.5) in terms of equivalent sources at the open-circuit array terminals.

2.1.3 Matching Networks

We have so far assumed no matching between the sensors and front-end in Fig. 2.1. However, several studies have shown that impedance matching can significantly enhance capacity in MIMO communication systems (e.g. [27] and references therein). In particular, for receivers that observe the antennas through a bank of noisy amplifiers, capacity is maximized by a multi-port (coupled) matching network designed to minimize the ampli-

fier noise factor, called the *minimum-noise-factor match* [26, 27, 44]. It seems reasonable to expect that matching can also improve the performance of estimation algorithms.

For $M = 1$, the noise contributed by the amplifier in (2.6) is often measured by the *noise factor* $F \triangleq \sigma_{\mathbf{n}_A}^2 / \sigma_{\mathbf{n}_o}^2$, which depends on antenna impedance. The noise factor takes its minimum value (e.g. [27, Appendix])

$$F_{min} = 1 + 2 \left(g_a r_{cor} + \sqrt{g_a r_a + (g_a r_{cor})^2} \right), \quad (2.7)$$

when the antenna impedance equals $z_{opt} = \sqrt{r_a/g_a + r_{cor}^2} - jx_{cor}$, where g_a , r_a and $z_{cor} = r_{cor} + jx_{cor}$ are internal amplifier noise parameters.

For general M , the matching network in Fig. 2.1 is described by a block impedance matrix

$$\mathbf{Z}_M \triangleq \begin{bmatrix} \mathbf{Z}_{M11} & \mathbf{Z}_{M12} \\ \mathbf{Z}_{M21} & \mathbf{Z}_{M22} \end{bmatrix}, \quad (2.8)$$

where each submatrix lies in $\mathbb{C}^{M \times M}$. Ideally the network is designed with passive, reactive elements so it is lossless, reciprocal and noiseless. The network is lossless and noiseless if [39, eq. 30] $\mathbf{Z}_M = -\mathbf{Z}_M^H$ and reciprocal if $\mathbf{Z}_M = \mathbf{Z}_M^T$ where $[\cdot]^T$ denotes the transpose.

When matching is applied, the noise in the observation model (2.6) is changed to

$$\begin{aligned} \mathbf{y} &= \mathbf{s} + \mathbf{n}, \\ \mathbf{n} &\triangleq \mathbf{n}_o - \mathbf{M}^{-1} [\mathbf{v}_a + (\bar{\mathbf{Z}}_A + z_{cor}\mathbf{I}) \mathbf{i}_a], \end{aligned} \quad (2.9)$$

where $\mathbf{M} \triangleq \mathbf{Z}_{M21} (\mathbf{Z}_A + \mathbf{Z}_{M11})^{-1}$ and $\bar{\mathbf{Z}}_A \triangleq -\mathbf{M}\mathbf{Z}_{M12} + \mathbf{Z}_{M22}$. Since \mathbf{n}_o , \mathbf{v}_a and \mathbf{i}_a are Gaussian and independent, the total observation noise is distributed as $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \Sigma_n)$

where

$$\boldsymbol{\Sigma}_{\mathbf{n}} = \boldsymbol{\Sigma}_{\mathbf{o}} + 4k_b T_0 \mathbf{M}^{-1} \left[r_a \mathbf{I} + g_a (\bar{\mathbf{Z}}_{\mathbf{A}} + z_{\text{cor}} \mathbf{I}) (\bar{\mathbf{Z}}_{\mathbf{A}} + z_{\text{cor}} \mathbf{I})^H \right] \mathbf{M}^{-H}. \quad (2.10)$$

For a given matching network, we conclude the relationship between the observation \mathbf{y} and parameter vector \mathbf{u} is described by the Gaussian probability density function (pdf) [74, pg. 524]

$$p(\mathbf{y}|\mathbf{u}) \triangleq \frac{1}{\det[\pi \boldsymbol{\Sigma}_{\mathbf{y}}(\mathbf{u})]} \exp \left[-(\mathbf{y} - \boldsymbol{\mu}_{\mathbf{s}}(\mathbf{u}))^H \boldsymbol{\Sigma}_{\mathbf{y}}^{-1}(\mathbf{u}) (\mathbf{y} - \boldsymbol{\mu}_{\mathbf{s}}(\mathbf{u})) \right], \quad (2.11)$$

where $\boldsymbol{\Sigma}_{\mathbf{y}}(\mathbf{u}) \triangleq \boldsymbol{\Sigma}_{\mathbf{s}}(\mathbf{u}) + \boldsymbol{\Sigma}_{\mathbf{n}}$.

2.2 Fundamental Bounds on Estimation

The aim of this chapter is to explore the impact of matching and physical noise sources on our ability to estimate the parameters \mathbf{u} in the observation model (2.11). We are interested in both Bayesian and non-Bayesian estimation, and both real and complex \mathbf{u} . To gain insight into the performance of the best estimators, we now derive Cramér-Rao Lower Bounds for each scenario. We further show, for any given noise model, that all of these bounds are minimized by one matching network: minimum-noise-factor matching.

2.2.1 Cramér-Rao Bound for Real Parameters

Consider first the non-Bayesian scenario, where \mathbf{u} is fixed and real. For a given matching network, let $\hat{\mathbf{u}} = \hat{\mathbf{u}}(\mathbf{y})$ be any unbiased estimator of $\mathbf{u} \in \mathbb{R}^p$ based on the observation of \mathbf{y} , so $E_{\mathbf{y}|\mathbf{u}}[\hat{\mathbf{u}}] = \mathbf{u}$, where $E_{\mathbf{y}|\mathbf{u}}$ denotes expectation with respect to the pdf (2.11). The

Cramér-Rao Bound (CRB) provides a fundamental lower bound on error covariance of $\hat{\mathbf{u}}$: [74, Sec. 15.7]

$$\boldsymbol{\Sigma}_{\hat{\mathbf{u}}}(\mathbf{u}) \triangleq E_{\mathbf{y}|\mathbf{u}}[(\hat{\mathbf{u}} - \mathbf{u})(\hat{\mathbf{u}} - \mathbf{u})^T] \geq \boldsymbol{\mathcal{I}}^{-1}(\mathbf{u}) , \quad (2.12)$$

where $\boldsymbol{\mathcal{I}}(\mathbf{u})$ is the Fisher information matrix. Here, we use the notation $\mathbf{A} \geq \mathbf{B}$ for Hermitian matrices \mathbf{A} and \mathbf{B} , to indicate $\mathbf{A} - \mathbf{B}$ is positive semi-definite.

For complex Gaussian observations, the Fisher information is given by [74, Sec. 15.7]

$$[\boldsymbol{\mathcal{I}}(\mathbf{u})]_{ij} = 2\text{Re} \left[\frac{\partial \boldsymbol{\mu}_s^H}{\partial u_i} \boldsymbol{\Sigma}_{\mathbf{y}}^{-1}(\mathbf{u}) \frac{\partial \boldsymbol{\mu}_s}{\partial u_j} \right] + \text{tr} \left[\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}(\mathbf{u}) \frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}}(\mathbf{u})}{\partial u_i} \boldsymbol{\Sigma}_{\mathbf{y}}^{-1}(\mathbf{u}) \frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}}(\mathbf{u})}{\partial u_j} \right] , \quad (2.13)$$

where tr , Re and $(\cdot)^H$ denote the trace, real part, and conjugate transpose, respectively.

The lower bound is achieved by any estimator $\hat{\mathbf{u}}$ that satisfies

$$\frac{\partial p(\mathbf{y}|\mathbf{u})}{\partial \mathbf{u}} = \boldsymbol{\mathcal{I}}(\mathbf{u}) [\hat{\mathbf{u}}(\mathbf{y}) - \mathbf{u}] , \quad (2.14)$$

in which case the estimator is called *efficient*.

2.2.2 Optimal Matching

We now consider impedance matching networks that maximize the Fisher information (and hence minimize the CRB) for all systems of the form described in Sec. 2.1. As noted earlier, a matching network that maximizes the ergodic capacity of MIMO channels was proposed in [26, 44] and later proved optimal in [27]. Intuitively, it seems reasonable to expect the same network to maximize Fisher information. However, the proof given in [27] does not apply here, both because it relies on properties of capacity which are not shared by Fisher information and because we consider arbitrary antenna noise, whereas [27] assumed uniform black-body radiation.

From (2.13), we see that Fisher information depends on the antenna noise, matching and amplifiers only through the noise covariance $\Sigma_{\mathbf{n}}$ in (2.10). We therefore write $\mathcal{I}(\mathbf{u}) = \mathcal{I}(\mathbf{u}; \Sigma_{\mathbf{n}})$ to display this dependence explicitly. The following result is proved in the Appendix A.1.

Theorem 2.1: Consider two receivers of the form illustrated in Fig. 2.1 that detect the same signal $\mathbf{s} \sim \mathcal{CN}(\boldsymbol{\mu}_{\mathbf{s}}(\mathbf{u}), \Sigma_{\mathbf{s}}(\mathbf{u}))$, but which may differ in antenna noise, matching and amplifiers. Denote the noise covariances (2.10) of these receivers by $\Sigma_{\mathbf{n}}$ and $\Sigma_{\mathbf{n}'}$, respectively. Then $\Sigma_{\mathbf{n}} \leq \Sigma_{\mathbf{n}'}$ implies $\mathcal{I}(\mathbf{u}; \Sigma_{\mathbf{n}}) \geq \mathcal{I}(\mathbf{u}; \Sigma_{\mathbf{n}'})$. \diamond

We now investigate matching networks that maximize Fisher information. From Theorem 1, this goal would be achieved by any network with the property the $\Sigma_{\mathbf{n}}$ is as small as possible in each dimension. The following theorem generalizes the result of [27, eq. 40] to an arbitrary antenna noise covariance.

Theorem 2.2: The system noise covariance in (2.10) is bounded by

$$\Sigma_{\mathbf{n}} \geq \Sigma_{\mathbf{o}} + (F_{min} - 1)4k_b T_0 \mathbf{R}_{\mathbf{A}} , \quad (2.15)$$

with equality if and only if $\bar{\mathbf{Z}}_{\mathbf{A}} = z_{opt} \mathbf{I}$, where F_{min} is the minimum amplifier noise factor (2.7) calculated at temperature T_0 , and z_{opt} is the source impedance that achieves it. \diamond

Proof: Consider the bracketed expression on the right side of (2.10). If we substitute $z_{cor} = -z_{opt} + r_{cor} + r_{opt}$ and expand the product we obtain

$$\begin{aligned} r_a \mathbf{I} + g_a (\bar{\mathbf{Z}}_{\mathbf{A}} + z_{cor} \mathbf{I}) (\bar{\mathbf{Z}}_{\mathbf{A}} + z_{cor} \mathbf{I})^H &= g_a (\bar{\mathbf{Z}}_{\mathbf{A}} - z_{opt} \mathbf{I}) (\bar{\mathbf{Z}}_{\mathbf{A}} - z_{opt} \mathbf{I})^H \\ &+ \left(g_a r_{cor} + \sqrt{g_a r_a + (g_a r_{cor})^2} \right) (\bar{\mathbf{Z}}_{\mathbf{A}} + \bar{\mathbf{Z}}_{\mathbf{A}}^H) \\ &\geq (F_{min} - 1) \bar{\mathbf{R}}_{\mathbf{A}} , \end{aligned}$$

with equality if $\bar{\mathbf{Z}}_{\mathbf{A}} = z_{\text{opt}}\mathbf{I}$. Here the first step follows from

$$r_a + g_a (r_{\text{cor}} + r_{\text{opt}})^2 - 2g_a r_{\text{opt}} (r_{\text{cor}} + r_{\text{opt}}) = 0 ,$$

and the second from (2.7) and $\bar{\mathbf{R}}_{\mathbf{A}} = (1/2)(\bar{\mathbf{Z}}_{\mathbf{A}} + \bar{\mathbf{Z}}_{\mathbf{A}}^H)$. Substituting this bound into (2.10) and observing $\mathbf{R}_{\mathbf{A}} = \mathbf{M}^{-1}\bar{\mathbf{R}}_{\mathbf{A}}\mathbf{M}^{-H}$, we obtain (2.15). \diamond

Theorems 2.1 and 2.2 imply that any matching network satisfying $\bar{\mathbf{Z}}_{\mathbf{A}} = z_{\text{opt}}\mathbf{I}$ maximizes the Fisher information (and minimizes the CRB) associated with any estimation problem of the form given in Sec. 2.1. Thus, one network is simultaneously optimal for all estimation problems. We call any lossless network that achieves $\bar{\mathbf{Z}}_{\mathbf{A}} = z_{\text{opt}}\mathbf{I}$ a *minimum-noise-factor (MNF) match*. There are many ways to construct such networks. In principle, it is desirable to implement such a network using passive, reactive elements [35, pg. 171]. One example of a lossless, reciprocal network that meets these requirements is [27, eq. 42]

$$\mathbf{Z}_{\mathbf{M}} = \begin{bmatrix} \mathbf{Z}_{M11} & \mathbf{Z}_{M12} \\ \mathbf{Z}_{M21} & \mathbf{Z}_{M22} \end{bmatrix} = j \begin{bmatrix} -\mathbf{X}_{\mathbf{A}} & (r_{\text{opt}}\mathbf{R}_{\mathbf{A}})^{1/2} \\ (r_{\text{opt}}\mathbf{R}_{\mathbf{A}})^{1/2} & x_{\text{opt}}\mathbf{I} \end{bmatrix}. \quad (2.16)$$

While MNF matching minimizes the CRB, implementing multi-port matching networks can be complicated in practice. For comparison, we therefore also consider a simpler, suboptimal single-port matching network. In minimum-noise factor *self-matching* [82], we ignore coupling between the antennas, and apply a separate matching network to each individual antenna that seeks to minimize the noise factor of that antenna in isolation. This form of matching can be described by

$$\mathbf{Z}_{SM} = \begin{bmatrix} \mathbf{Z}_{M11} & \mathbf{Z}_{M12} \\ \mathbf{Z}_{M21} & \mathbf{Z}_{M22} \end{bmatrix} = j \begin{bmatrix} -\mathbf{X}_{AS} & (r_{\text{opt}}\mathbf{R}_{AS})^{1/2} \\ (r_{\text{opt}}\mathbf{R}_{AS})^{1/2} & x_{\text{opt}}\mathbf{I} \end{bmatrix}, \quad (2.17)$$

where $\mathbf{R}_{AS} = \text{diag}(\mathbf{R}_A)$ and $\mathbf{X}_{AS} = \text{diag}(\mathbf{X}_A)$. Here $\text{diag}(\cdot)$ denotes an operation that retains only the diagonal elements of the matrix, and sets all off-diagonal elements to zero. We will compare the performance of these two matching networks in Sec. V.

2.2.3 Bayesian Cramér-Rao Bound

We have so far assumed $\mathbf{u} \in \mathbb{R}^p$ is a fixed, unknown vector. When \mathbf{u} is modeled as random with prior pdf $p(\mathbf{u})$, there is a similar lower bound on the estimation error covariance, called the Bayesian Cramér-Rao Bound (BCRB). For $p(\mathbf{y}, \mathbf{u}) = p(\mathbf{y}|\mathbf{u})p(\mathbf{u})$, Van Trees [28, pg. 72] showed¹

$$\boldsymbol{\Sigma}_{\hat{\mathbf{u}}} \triangleq E_{\mathbf{y}, \mathbf{u}} [(\hat{\mathbf{u}} - \mathbf{u})(\hat{\mathbf{u}} - \mathbf{u})^T] \geq \boldsymbol{\mathcal{I}}^{-1}, \quad (2.18)$$

for any estimator $\hat{\mathbf{u}}$ (unbiased or not), where $\boldsymbol{\mathcal{I}}$ is the Bayesian information,

$$\begin{aligned} \boldsymbol{\mathcal{I}} &\triangleq E_{\mathbf{y}, \mathbf{u}} \left[\left(\frac{\partial \ln p(\mathbf{y}, \mathbf{u})}{\partial \mathbf{u}} \right) \left(\frac{\partial \ln p(\mathbf{y}, \mathbf{u})}{\partial \mathbf{u}} \right)^T \right] \\ &= E_{\mathbf{u}} [\boldsymbol{\mathcal{I}}(\mathbf{u})] + E_{\mathbf{u}} \left[\left(\frac{\partial \ln p(\mathbf{u})}{\partial \mathbf{u}} \right) \left(\frac{\partial \ln p(\mathbf{u})}{\partial \mathbf{u}} \right)^T \right], \end{aligned} \quad (2.19)$$

and $\boldsymbol{\mathcal{I}}(\mathbf{u})$ is the Fisher information (2.13). The Bayesian information depends on antenna noise, matching networks and amplifiers only through the noise covariance $\boldsymbol{\Sigma}_{\mathbf{n}}$ in (2.10). For a given $p(\mathbf{u})$, we can therefore write $\boldsymbol{\mathcal{I}} = \boldsymbol{\mathcal{I}}(\boldsymbol{\Sigma}_{\mathbf{n}})$ to display this dependence explicitly.

Theorem 2.3: Consider two receivers of the form illustrated in Fig. 2.1 that detect the same signal \mathbf{s} , where $p_{\mathbf{s}|\mathbf{u}} \sim \mathcal{CN}(\boldsymbol{\mu}_{\mathbf{s}}(\mathbf{u}), \boldsymbol{\Sigma}_{\mathbf{s}}(\mathbf{u}))$, but which may differ in antenna noise, matching networks and amplifiers. Denote the noise covariances (2.10) of these receivers

¹Van Trees proved (2.19) for complex \mathbf{u} ; however, the real-parameter version is easily derived and has already appeared in other publications (e.g., [29]).

by $\Sigma_{\mathbf{n}}$ and $\Sigma_{\mathbf{n}'}$, respectively. Then, for any prior distribution $p(\mathbf{u})$, $\Sigma_{\mathbf{n}} \leq \Sigma_{\mathbf{n}'}$ implies $\mathcal{I}(\Sigma_{\mathbf{n}}) \geq \mathcal{I}(\Sigma_{\mathbf{n}'})$. \diamond

Proof of Theorem 2.3: In (2.19), the only term that depends on $\Sigma_{\mathbf{n}}$ is $\mathcal{I}(\mathbf{u})$. For fixed $p(\mathbf{u})$, Theorem 3 thus follows immediately from Theorem 1 and (2.19). \diamond

Theorems 2 and 3 demonstrate that MNF matching, $\bar{\mathbf{Z}}_{\mathbf{A}} = z_{\text{opt}}\mathbf{I}$, maximizes the Bayesian information, and hence minimizes the BCRB associated with any Bayesian estimation problem. Thus, MNF matching is again simultaneously optimal for all estimation problems.

2.2.4 Complex Parameters

When the parameters to be estimated are complex, say $\mathbf{u} = \mathbf{u}_r + j\mathbf{u}_i \in \mathbb{C}^p$, the bounds above can always be applied as written to the problem of estimating the *real* vector $\bar{\mathbf{u}} \triangleq (\mathbf{u}_r^T, \mathbf{u}_i^T)^T \in \mathbb{R}^{2p}$. However, the bounds can often be expressed in a more intuitive, complex form. Define the complex gradient of the log-likelihood by [74, Sec. 15.7]

$$\frac{\partial \ln p(\mathbf{y}|\mathbf{u})}{\partial \mathbf{u}^*} \triangleq \frac{1}{2} \left[\frac{\partial \ln p(\mathbf{y}|\mathbf{u})}{\partial \mathbf{u}_r} + j \frac{\partial \ln p(\mathbf{y}|\mathbf{u})}{\partial \mathbf{u}_i} \right].$$

For non-Bayesian estimation of a fixed vector $\mathbf{u} \in \mathbb{C}^p$, when the pseudo-information vanishes,

$$E_{\mathbf{y}|\mathbf{u}} \left[\left(\frac{\partial \ln p(\mathbf{y}|\mathbf{u})}{\partial \mathbf{u}^*} \right) \left(\frac{\partial \ln p(\mathbf{y}|\mathbf{u})}{\partial \mathbf{u}^*} \right)^T \right] = \mathbf{0}_{p \times p}, \quad (2.20)$$

we can express the CRB in (2.12) in the equivalent complex form

$$\Sigma_{\hat{\mathbf{u}}}(\mathbf{u}) \triangleq E_{\mathbf{y}}[(\hat{\mathbf{u}} - \mathbf{u})(\hat{\mathbf{u}} - \mathbf{u})^H] \geq \mathcal{I}^{-1}(\mathbf{u}), \quad (2.21)$$

where $\mathcal{I}(\mathbf{u})$ is the (complex) Fisher Information

$$\mathcal{I}(\mathbf{u}) \triangleq E_{\mathbf{y}|\mathbf{u}} \left[\left(\frac{\partial \ln p(\mathbf{y}|\mathbf{u})}{\partial \mathbf{u}^*} \right) \left(\frac{\partial \ln p(\mathbf{y}|\mathbf{u})}{\partial \mathbf{u}^*} \right)^H \right]. \quad (2.22)$$

Similarly, for Bayesian estimation of a complex $\mathbf{u} \in \mathbb{C}^p$ with prior pdf $p(\mathbf{u})$, if

$$E_{\mathbf{y},\mathbf{u}} \left[\left(\frac{\partial \ln p(\mathbf{y}, \mathbf{u})}{\partial \mathbf{u}^*} \right) \left(\frac{\partial \ln p(\mathbf{y}, \mathbf{u})}{\partial \mathbf{u}^*} \right)^T \right] = \mathbf{0}_{p \times p} \quad (2.23)$$

holds, we can express the BCRB (2.19) for $\bar{\mathbf{u}} \triangleq (\mathbf{u}_r^T, \mathbf{u}_i^T)^T$ in the equivalent complex form:

$$\Sigma_{\bar{\mathbf{u}}} \triangleq E_{\mathbf{y},\mathbf{u}}[(\hat{\mathbf{u}} - \mathbf{u})(\hat{\mathbf{u}} - \mathbf{u})^H] \geq \mathcal{I}^{-1}, \quad (2.24)$$

where \mathcal{I} is the (complex) Bayesian information

$$\mathcal{I} \triangleq E_{\mathbf{y},\mathbf{u}} \left[\left(\frac{\partial \ln p(\mathbf{y}, \mathbf{u})}{\partial \mathbf{u}^*} \right) \left(\frac{\partial \ln p(\mathbf{y}, \mathbf{u})}{\partial \mathbf{u}^*} \right)^H \right]. \quad (2.25)$$

2.3 Two Spatial Estimation Problems

We now apply the results of Sec. 2.2 to two important estimation problems that rely in an essential way on the physical size of the sensor array. Our aim is to explore how coupling and matching impacts spatial estimation problems in compact arrays. Each application will be explored numerically in Sec. V.

2.3.1 MIMO Channel Estimation

Most studies of MIMO capacity assume that channel fading path gains are known at the receiver [6, 7]. In practice, these gains are estimated from the received signal, often from

a known training sequence embedded in the transmitted data. Accurate estimation can be challenging when the number of path gains is large, but is important because even small estimation errors can significantly reduce capacity [30,31].

There is an extensive literature on MIMO channel estimation (e.g., [20,32] and references therein); however, relatively few papers have considered the impact of coupling. In [22,23], Liu *et al* investigated the performance of least-squares and minimum-mean-squared error (MMSE) estimators for MIMO channel estimation in the presence of coupling. Both papers suggest coupling can increase estimation error for arrays of dipoles spaced 0.2λ – 0.4λ apart.

In this section, we consider the impact of coupling and matching on the Bayesian CRB of training-based MIMO channel estimation. Our work differs from [22,23] in several ways. First, we include physical models of antenna and amplifier noise which depend on the antenna configuration, matching networks and amplifier characteristics, whereas a fixed noise model that is independent of the antennas and front end is assumed in [22,23]. Second, we focus on the impact of matching and derive optimal matching networks that minimize the BCRB, whereas matching is not considered in [22,23].

Suppose the transmitter sends a known training sequence $\mathbf{x}_1, \dots, \mathbf{x}_T \in \mathbb{C}^N$ from N uncoupled antennas to the receiver. If the path gains and matching are fixed for the duration of transmission, the corresponding observation at time t is given by (2.9) with $\mathbf{s} = \mathbf{H}\mathbf{x}_t$, so

$$\mathbf{y}_t = \mathbf{H}\mathbf{x}_t + \mathbf{n}_t, t = 1, \dots, T, \quad (2.26)$$

where $\mathbf{H} \in \mathbb{C}^{M \times N}$ is the channel fading matrix, and \mathbf{n}_t is an independent sequence of noise vectors with distribution $\mathcal{CN}(\mathbf{0}, \mathbf{\Sigma}_n)$, where $\mathbf{\Sigma}_n$ is given in (2.10). Since the transmit

antennas are identical and uncoupled, we can assume the columns of \mathbf{H} are independent, identically-distributed (i.i.d.) zero-mean, circularly-symmetric complex random vectors. If \mathbf{h}_j is the j -th column of \mathbf{H} , we denote this distribution by $\mathbf{h}_j \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Sigma}_{\mathbf{H}})$, where $\boldsymbol{\Sigma}_{\mathbf{H}} = E \{ \mathbf{h}_j \mathbf{h}_j^H \}$ is the column covariance of \mathbf{H} , and E denotes the expectation.

Arranging the sequence into a matrix, $\mathbf{X} \triangleq [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T] \in \mathbb{C}^{N \times T}$, we can write

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{N}, \quad (2.27)$$

where $\mathbf{Y} \triangleq [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_T]$ and $\mathbf{N} \triangleq [\mathbf{n}_1, \mathbf{n}_2, \dots, \mathbf{n}_T]$. We convert this to vector form using the identity $\text{vec}(\mathbf{A}\mathbf{X}\mathbf{B}) = (\mathbf{B}^T \otimes \mathbf{A})\text{vec}(\mathbf{X})$ [33]

$$\text{vec}(\mathbf{Y}) = (\mathbf{X}^T \otimes \mathbf{I}_M)\text{vec}(\mathbf{H}) + \text{vec}(\mathbf{N}), \quad (2.28)$$

where $\text{vec}(\mathbf{X})$ is the vectorization operator that stacks the columns of \mathbf{X} into a single vector and \otimes is the Kronecker product. Note $\text{vec}(\mathbf{H})$ and $\text{vec}(\mathbf{N})$ are independent, zero-mean complex Gaussian vectors with covariances $\mathbf{I}_N \otimes \boldsymbol{\Sigma}_{\mathbf{H}}$ and $\mathbf{I}_T \otimes \boldsymbol{\Sigma}_{\mathbf{n}}$ respectively.

For a linear observation model $\mathbf{y} = \mathbf{A}\mathbf{u} + \mathbf{n}$ where \mathbf{u} and \mathbf{n} are zero-mean and Gaussian, the Bayesian information is given by $\mathcal{I} = \mathbf{A}^H \boldsymbol{\Sigma}_{\mathbf{n}}^{-1} \mathbf{A} + \boldsymbol{\Sigma}_{\mathbf{u}}^{-1}$. Applying this to (2.28), with $\mathbf{u} = \text{vec}(\mathbf{H})$ and $\mathbf{n} = \text{vec}(\mathbf{N})$, we have

$$\begin{aligned} \mathcal{I} &= (\mathbf{X}^T \otimes \mathbf{I}_M)^H \boldsymbol{\Sigma}_{\text{vec}(\mathbf{N})}^{-1} (\mathbf{X}^T \otimes \mathbf{I}_M) + \boldsymbol{\Sigma}_{\text{vec}(\mathbf{H})}^{-1} \\ &= (\mathbf{X}^* \otimes \mathbf{I}_M) (\mathbf{I}_T \otimes \boldsymbol{\Sigma}_{\mathbf{n}}^{-1}) (\mathbf{X}^T \otimes \mathbf{I}_M) + (\mathbf{I}_T \otimes \boldsymbol{\Sigma}_{\mathbf{H}}^{-1}) \\ &= (\mathbf{X}^* \mathbf{X}^T \otimes \boldsymbol{\Sigma}_{\mathbf{n}}^{-1}) + (\mathbf{I}_N \otimes \boldsymbol{\Sigma}_{\mathbf{H}}^{-1}), \end{aligned} \quad (2.29)$$

which follows from the identities $(\mathbf{A} \otimes \mathbf{B})^{-1} = \mathbf{A}^{-1} \otimes \mathbf{B}^{-1}$ and $(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) =$

(AC) \otimes (BD).

For MIMO channel estimation, the training sequences associated with different antennas are often chosen to be equal-energy and orthogonal, $\mathbf{X}\mathbf{X}^H = (PT/N)\mathbf{I}_N$. Here (2.29) reduces to

$$\mathcal{I} = \mathbf{I}_N \otimes \left(\frac{PT}{N} \boldsymbol{\Sigma}_{\mathbf{n}}^{-1} + \boldsymbol{\Sigma}_{\mathbf{H}}^{-1} \right). \quad (2.30)$$

The resulting Bayesian CRB on any estimator $\hat{\mathbf{H}}$ is then

$$\boldsymbol{\Sigma}_{\hat{\mathbf{H}}} \geq \mathcal{I}^{-1} = \mathbf{I}_N \otimes \left(\frac{PT}{N} \boldsymbol{\Sigma}_{\mathbf{n}}^{-1} + \boldsymbol{\Sigma}_{\mathbf{H}}^{-1} \right)^{-1}. \quad (2.31)$$

It follows the average per-symbol mean-squared estimation error (MSE) is bounded by

$$\text{MSE}_{\hat{\mathbf{H}}} \triangleq \left(\frac{1}{MN} \right) \text{tr}(\boldsymbol{\Sigma}_{\hat{\mathbf{H}}}) \geq \frac{1}{M} \text{tr} \left(\frac{PT}{N} \boldsymbol{\Sigma}_{\mathbf{n}}^{-1} + \boldsymbol{\Sigma}_{\mathbf{H}}^{-1} \right)^{-1}. \quad (2.32)$$

2.3.2 Direction-of-Arrival Estimation

Next we consider the problem of estimating the directions-of-arrival (DoA) of N narrow-band signals incident on the array. The signals are modeled as deterministic, which is often called the “conditional” model [76]. It is well known that many other important problems in array processing take a similar form, such as estimating the parameters of a superposition of exponentials in noise, or resolving overlapping multipaths.

As noted earlier, the impact of coupling on DoA estimation has been considered by several authors. Svantesson [12] derived a CRB for DoA estimation in the presence of coupling. Using similar models, Lau and Bach-Anderson [14] showed strong coupling can significantly degrade the CRB, but losses can be significantly reduced by using multi-port

conjugate matching between the array and front end. Both papers model the effects of coupling on the received signal, but do not consider the impact of coupling and matching on physical noise sources. Instead, observation noise is modeled as spatially white, with a fixed distribution that is independent of the antennas and matching. In contrast, other studies of DoA estimation with coupling have considered physical models of antenna and amplifier noise [24, 25], but do not consider matching.

The aim of this section is to investigate the impact of coupling and matching on the CRB of DoA estimation using the physical noise models in (2.9), and to determine the performance that can be achieved when the best possible matching network is used for the given noise conditions. We assume the received signal at time t is a superposition of N signals $x_t(1), \dots, x_t(N)$ that arrive from directions ϕ_1, \dots, ϕ_N , respectively. If we define $\mathbf{x}_t = [x_t(1), \dots, x_t(N)]^T \in \mathbb{C}^{N \times 1}$, the observation at time t is given by (2.9) with $\mathbf{s} = \mathbf{G}(\boldsymbol{\phi})\mathbf{x}_t$, so

$$\mathbf{y}_t = \mathbf{G}(\boldsymbol{\phi})\mathbf{x}_t + \mathbf{n}_t, \quad t = 1, \dots, T, \quad (2.33)$$

where $\boldsymbol{\phi} = (\phi_1, \dots, \phi_N)^T \in \mathbb{R}^N$, $\mathbf{G}(\boldsymbol{\phi}) = [\mathbf{g}(\phi_1), \dots, \mathbf{g}(\phi_N)]^T \in \mathbb{C}^{M \times N}$, and $\mathbf{g}(\phi) \in \mathbb{C}^{M \times 1}$ denotes the steering vector of the array. As noted earlier, \mathbf{n}_t is an independent sequence of noise vectors with distribution $\mathcal{CN}(\mathbf{0}, \boldsymbol{\Sigma}_{\mathbf{n}})$, where $\boldsymbol{\Sigma}_{\mathbf{n}}$ is defined in (2.10). In general, $\mathbf{g}(\phi)$ depends on the structure and geometry of the array; specific examples will be considered in Sec. V.

For the particular case of spatially white observation noise, $\boldsymbol{\Sigma}_{\mathbf{n}} = \gamma \mathbf{I}$, a closed-form expression for the Fisher information in (2.13) was given in [76, eq. (2.11)]:

$$\mathcal{I}(\boldsymbol{\phi}) = \left(\frac{2T}{\gamma} \right) \text{Re} [\mathbf{H} \odot \mathbf{P}^T], \quad (2.34)$$

where \odot is the Schur product,

$$\begin{aligned} \mathbf{H} &\triangleq \mathbf{D}^H [\mathbf{I} - \mathbf{G}(\mathbf{G}^H \mathbf{G})^{-1} \mathbf{G}^H] \mathbf{D} , \\ \mathbf{D} &\triangleq \left[\frac{d\mathbf{g}(\phi_1)}{d\phi}, \dots, \frac{d\mathbf{g}(\phi_N)}{d\phi} \right] , \quad \mathbf{P} \triangleq \frac{1}{T} \sum_{t=1}^T \mathbf{x}_t \mathbf{x}_t^H . \end{aligned} \quad (2.35)$$

We can extend this result to colored noise by observing that pre-whitening the observation, say $\bar{\mathbf{y}}_t = \boldsymbol{\Sigma}_{\mathbf{n}}^{-1/2} \mathbf{y}_t$, does not change Fisher information provided $\boldsymbol{\Sigma}_{\mathbf{n}}$ is non-singular and independent of ϕ . The whitened observation is given by (2.33) with $\mathbf{g}(\phi)$ replaced by $\bar{\mathbf{g}}(\phi) \triangleq \boldsymbol{\Sigma}_{\mathbf{n}}^{-1/2} \mathbf{g}(\phi)$, and $\boldsymbol{\Sigma}_{\mathbf{n}}$ replaced by $\bar{\boldsymbol{\Sigma}}_{\mathbf{n}} \triangleq \mathbf{I}$. It follows from (2.34) that the CRB for colored noise is

$$\text{CRB}_{\hat{\phi}} \triangleq \boldsymbol{\mathcal{I}}^{-1}(\phi) = (2T \text{Re} [\bar{\mathbf{H}} \odot \mathbf{P}^T])^{-1} , \quad (2.36)$$

where $\bar{\mathbf{G}} = \boldsymbol{\Sigma}_{\mathbf{n}}^{-1/2} \mathbf{G}$, $\bar{\mathbf{D}} = \boldsymbol{\Sigma}_{\mathbf{n}}^{-1/2} \mathbf{D}$, and so

$$\begin{aligned} \bar{\mathbf{H}} &= \bar{\mathbf{D}}^H [\mathbf{I} - \bar{\mathbf{G}}(\bar{\mathbf{G}}^H \bar{\mathbf{G}})^{-1} \bar{\mathbf{G}}^H] \bar{\mathbf{D}} \\ &= \mathbf{D}^H [\boldsymbol{\Sigma}_{\mathbf{n}}^{-1} - \boldsymbol{\Sigma}_{\mathbf{n}}^{-1} \mathbf{G}(\mathbf{G}^H \boldsymbol{\Sigma}_{\mathbf{n}}^{-1} \mathbf{G})^{-1} \mathbf{G}^H \boldsymbol{\Sigma}_{\mathbf{n}}^{-1}] \mathbf{D} . \end{aligned} \quad (2.37)$$

2.4 Numerical Results

We now present numerical examples to illustrate the results of Sec. IV. Consider a receive array consisting of M parallel, half-wavelength dipoles with inter-element spacing d , where $0.05\lambda \leq d \leq \lambda$. In order to capture the effects of near-field scattering among antennas, we numerically calculate the receive pattern $\mathbf{g}(\phi)$ and the array impedance $\mathbf{Z}_{\mathbf{A}}$ in (2.1) using the Numerical Electromagnetics Code (NEC) [34], where each antenna has

a thickness of $10^{-3}\lambda$ and is divided into 25 computational segments. For comparison, we also consider an array of uncoupled dipoles with receive patterns [12, eq. 2.11]

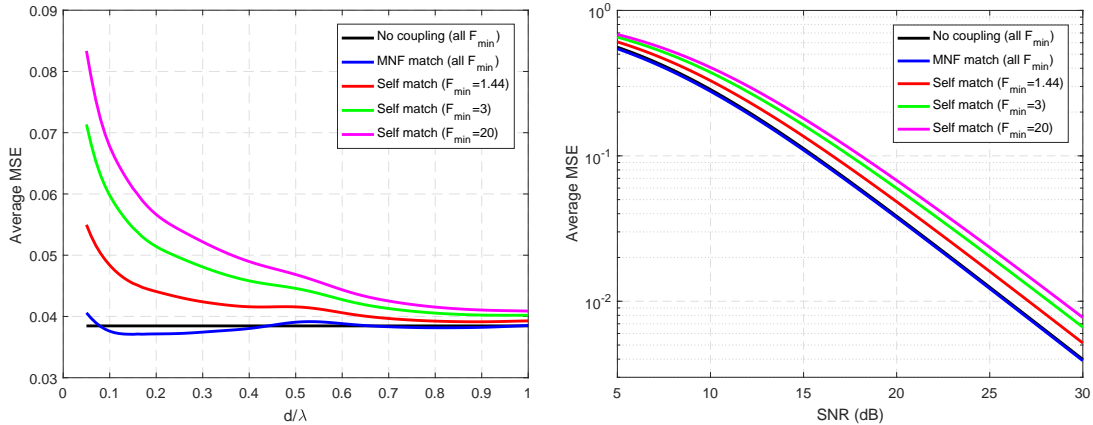
$$[\mathbf{g}(\theta, \phi)]_i = a(\theta) \exp\{-2\pi j(d/\lambda)(i-1)\sin(\theta)\cos(\phi)\}, \quad a(\theta) = \frac{\lambda \cos(\pi \cos(\theta))}{\pi \sin(\theta)} \quad (2.38)$$

where $\theta \in [0, \pi]$ and $\phi \in [0, 2\pi)$ denote elevation and azimuth, respectively. For signals that arrive in the azimuthal plane, $\theta = \pi/2$, we take the pattern to be $\mathbf{g}(\phi) \triangleq \mathbf{g}(\pi/2, \phi)$. The impedance matrix of the uncoupled array is $\text{diag}\{Z_A, Z_A, \dots, Z_A\}$, where Z_A is the impedance of an isolated dipole.

For the amplifier, we consider the example of the SiGe low-noise amplifier (LNA) in [83], designed for applications at 2.4GHz. At $T_0 = 298\text{K}$ and $R_{BIAS} = 1.2\text{ k}\Omega$, this LNA has noise parameters $F_{min} = 1.44$, $r_a = 11.57\ \Omega$, $g_a = 5.32\text{ mS}$, and $z_{cor} = 43.90\angle -1.7^\circ\ \Omega$. To gain insight into how the relative mix of antenna and amplifier noise affects results, we will also vary F_{min} by considering the noise parameters $r_a = 11.57\alpha\ \Omega$, $g_a = 5.32\alpha\text{ mS}$ for different values of $\alpha \geq 0$. Thus, $\alpha = 0$ implies $F_{min} = 1$, which corresponds to a system with only antenna noise, similar to the model considered in [43]. Increasing α increases amplifier noise; for example, $\alpha = 20.7$ gives $F_{min} = 10$, where 90% of the noise arises in the amplifiers. More generally, for a perfectly-matched dipole in isolation with signal-to-noise ratio $\text{SNR} = P/4k_bT_0Br_A F_{min}$, the fraction of system noise that arises in the antennas is F_{min}^{-1} .

2.4.1 MIMO Channel Estimation

We now calculate the bounds on MIMO channel estimation derived in Sec. IV. Consider a two-dimensional propagation model in which the received signal consists of a large number of vertically-polarized plane waves that arrive uniformly from all directions in



(a) Versus dipole spacing d for SNR=10 dB (b) Versus noise factor F_{min} for $d = 0.1\lambda$

Figure 2.2: BCRB on mean-squared channel estimation error for a 4×4 MIMO channel

the plane orthogonal to the dipoles. We assume fading is quasi-static, so the channel is constant over the duration of the training sequence. Under these assumptions, the covariance (2.26) is given by [39, eq. 33]

$$[\Sigma_{\mathbf{H}}]_{mn} = \frac{1}{2\pi} \int_0^{2\pi} g_m(\phi) g_n^*(\phi) e^{j2\pi(d/\lambda)(m-n)\cos\phi} d\phi, \quad (2.39)$$

where $g_i(\phi)$ is the open-circuit voltage induced in the i -th antenna by a zero-phase, unit-amplitude plane wave arriving from azimuthal angle ϕ . We omit the exponent $e^{j2\pi(d/\lambda)(m-n)\cos\phi}$ in (2.39) when the $g_i(\phi)$ share a common phase center. Here the patterns were computed for 120 equally-spaced angles, and the results were used to numerically integrate (2.39).

In Fig. 2.2a, we plot the lower bound (2.32) on average mean-squared channel estimation error (MSE) versus dipole spacing d for a system with SNR=10 dB, $T = 10$, and $M = N = 4$ antennas at the transmitter and receiver. We consider both self-matching (2.17) and optimal MNF matching (2.16), and various values of the noise factor F_{min} . We

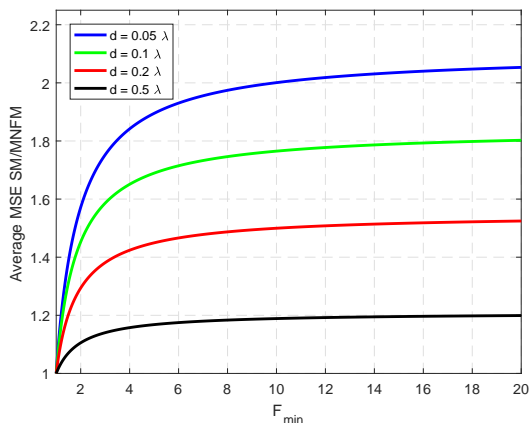


Figure 2.3: Ratio BCRB-SM/BCRB-MNF versus F_{min} for SNR=10 dB

also include the bound for an uncoupled array, calculated from (2.38). Note the BCRB of optimal MNF matching is close ($\pm 3\%$) to that of an uncoupled array for all d in the range of interest, and neither plot depends on F_{min} . However, the BCRB of self-matching is quite sensitive to F_{min} and d , and diverges significantly from optimal for small d and large F_{min} . Not surprisingly, the curves converge as d increases for all F_{min} , since coupling also vanishes. In Fig. 2.2b, we plot the same curves versus SNR for $d = 0.1\lambda$. Again, MNF matching is comparable to an uncoupled array and, for each F_{min} , the performance loss (in dB) of self-matching relative to optimal matching appears similar at most SNRs. As in [22, 23], we conclude the estimation error of self-matching can degrade significantly in the presence of strong coupling, although we consider very different signal and noise models. In contrast to [22, 23], however, we further conclude this degradation occurs only when amplifier noise is significant, and can be entirely eliminated with optimal matching.

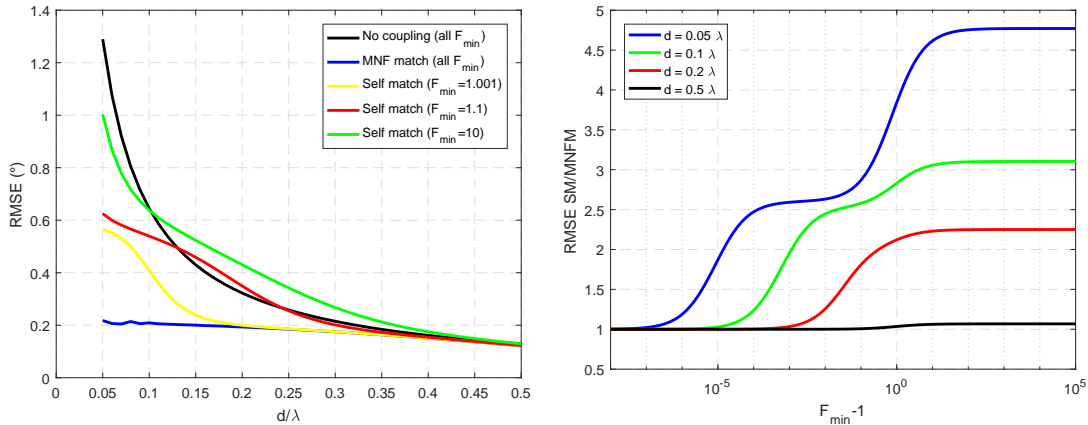
Now let BCRB-SM and BCRB-MNF denote the BCRB calculated with self-matching and MNF matching, respectively. To explore how performance varies with the mix of antenna and amplifier noise, in Fig. 2.3 we plot the ratio BCRB-SM/BCRB-MNF versus

F_{min} for SNR = 10 dB, $T = 10$ and different values of d . Note the performance loss of self-matching relative to optimal increases monotonically with F_{min} to an asymptotic value. When the noise arises exclusively in the antennas ($F_{min} = 1$), self-matching yields the same performance as MNF matching for all d . This is intuitively reasonable, since the observation model (2.9) does not depend on matching when $r_a = g_a = 0$. When amplifier noise dominates, however, self-matching increases the BCRB by 20% for the standard spacing of $d = 0.5\lambda$, and by 105% at $d = 0.05\lambda$. Even a small amount of amplifier noise appears to significantly degrade the performance of self-matching: for the values of d in Fig. 2.3, 50% of the degradation happens by $F_{min} = 2$, and 75% by $F_{min} = 4$.

2.4.2 Direction-of-Arrival Estimation

We now consider the CRB (2.36), which bounds the error in estimating the directions-of-arrival of target signals located in the azimuthal plane. We consider the same array, matching networks and amplifier model as in the last section, and we adopt the same definition of SNR.

For a single target, the CRB is scalar and depends on the signal only through the product $PT = \sum_{t=1}^T |x_1(t)|^2$. We define the root-mean-squared angular estimation error (RMSE) as $(\text{CRB}_{\hat{\phi}})^{1/2}$ where $\text{CRB}_{\hat{\phi}}$ is defined in (2.36). In Fig. 2.4a, we plot the RMSE in degrees versus d for a target at broadside ($\phi = 90^\circ$), for SNR = 10 dB, $T = 200$ and various values of F_{min} . Note the optimal MNF match is always better than no coupling, and neither depends on F_{min} . However, the performance of no coupling degrades much more as d decreases. Although our noise and matching models are different, these curves appear qualitatively similar to the results for multiport-conjugate matching and no coupling reported in [14, Fig. 3(a)]. However, the self-match behaves in a different



(a) Versus dipole spacing d for SNR=10 dB (b) Versus noise factor F_{min} for $d = 0.1\lambda$

Figure 2.4: CRB on RMS angular error for 4-element array with one target at $\phi = 90^\circ$ and $T = 200$

way, that depends on the composition of observation noise: When $F_{min} = 1$, the self-match coincides with the MNF match for all d . This is intuitive, since matching does not affect performance when antenna noise dominates [27, 43]. However, the addition of even a small amount of amplifier noise significantly alters performance, as can be seen by comparing results for $F_{min} = 1, 1.001$. This occurs because the antenna noise covariance (2.4) has some small eigenvalues for small d , and the directions associated with those eigenvalues provide very informative observations of the target. Thus, even a small increase in amplifier noise can increase the noise in those directions by orders of magnitude.

To see this sensitivity to F_{min} more clearly, Fig. 2.4b plots the ratio of the CRB for the self-match and MNF matching versus $F_{min} - 1$. In the presence of strong coupling, note the performance of self-matching for $d = 0.05\lambda$ is significantly degraded even when $F_{min} - 1 \approx 10^{-5}$. The curves also exhibit an odd staircase behavior, which can be explained as follows. Recall we vary amplifier noise via the parameter α . Let $\Sigma_{\mathbf{n}}(\alpha)$ be the noise covariance for α and note $F_{min} = 1 + 0.44\alpha$. It can be shown the first term in (2.37)

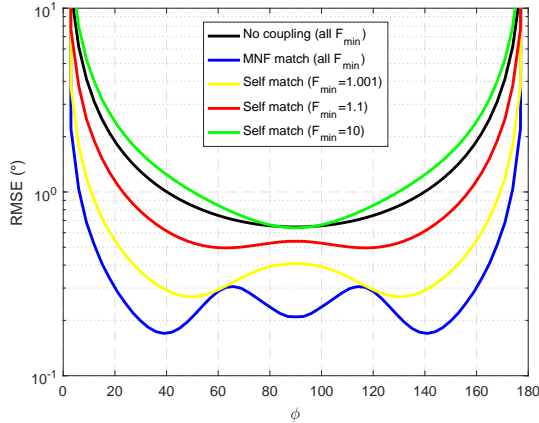


Figure 2.5: CRB on RMS angular error versus target angle ϕ for $d = 0.1\lambda$.

dominates, so $\mathcal{I}(\phi) \approx 2PT \sum_{i=1}^M |d_i|^2 (1 + \alpha \lambda_i)^{-1}$ where λ_i are the eigenvalues of the matrix $\mathbf{S} = \Sigma_{\mathbf{o}}^{-1/2} \Sigma_{\mathbf{n}}(1) \Sigma_{\mathbf{o}}^{-1/2} - \mathbf{I}$ and $\mathbf{d} = \mathbf{U} \Sigma_{\mathbf{o}}^{-1/2} \mathbf{D}$, where $\mathbf{S} = \mathbf{U}^H \mathbf{\Lambda} \mathbf{U}$ denotes the eigensystem of \mathbf{S} . From this, it is clear $\text{CRB}_{\hat{\phi}} = \mathcal{I}(\phi)^{-1}$ increases most rapidly when α is near $1/\lambda_i$ for some i , as observed in the figure.

In Fig. 2.5, we plot the RMSE versus direction-of-arrival ϕ for a single target with $d = 0.1\lambda$, SNR= 10 dB and $T = 200$. Not surprisingly, all matching techniques generally perform well near broadside ($\phi = 90^\circ$) and poorly near end-fire ($\phi = 0^\circ, 180^\circ$). As in Fig. 2.4a, the performance of self-matching degrades with increasing F_{min} for each ϕ . Interestingly, the performance gain of optimal matching depends strongly on target direction as well as F_{min} : ranging from a nearly 7-fold improvement for targets at $\pm 50^\circ$ of broadside when amplifier noise dominates, to no improvement for any ϕ when $F_{min} = 1$. For any matching network and F_{min} , define the beamwidth to be the range of ϕ for which the CRB performs within a factor of 2 of its best performance. Note MNF matching has a beamwidth of roughly 146° about the broadside for all F_{min} , whereas self-matching has a beamwidth that varies from 146° for $F_{min} = 1$ down to 110° when $F_{min} = 10$. Thus,

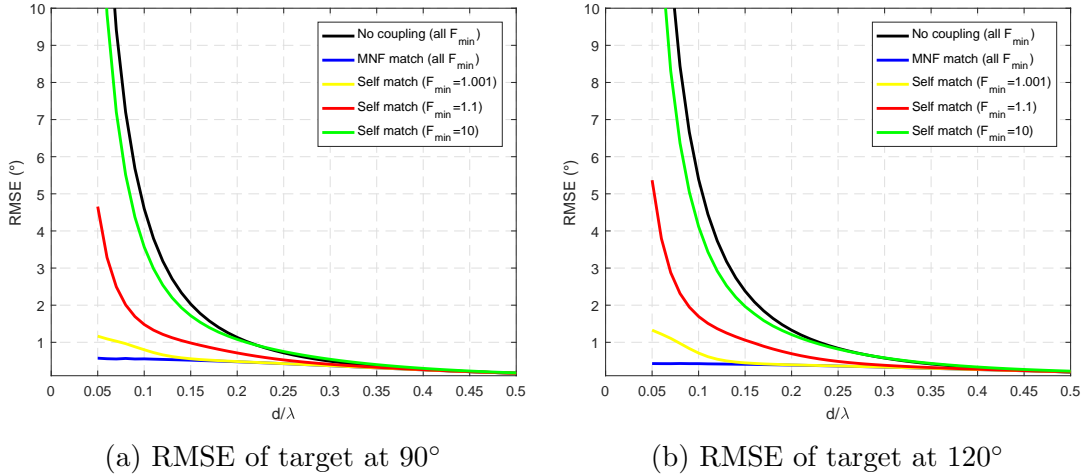


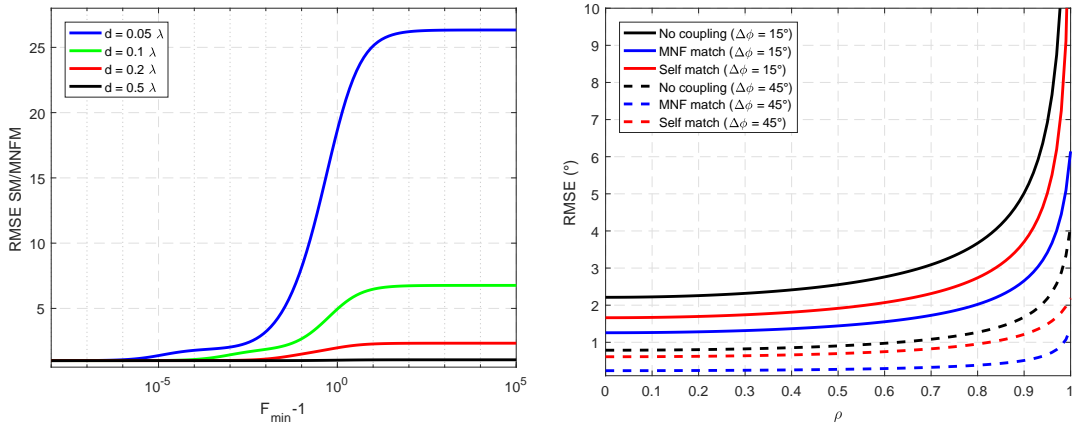
Figure 2.6: CRB on RMS angular error for targets at 90° and 120° with $\rho = 0$, SNR=10 dB.

when amplifier noise is significant, MNF matching also significantly expands the angular range over which the CRB is close to its optimal value.

Next consider two equal-energy targets located at $\phi = 90^\circ$ and 120° , respectively. If we define $\rho \triangleq \sum_{t=0}^T x_1(t)x_2^*(t)/PT$, then the matrix \mathbf{P} in (2.34) is given by

$$\mathbf{P} = P \begin{bmatrix} 1 & \rho \\ \rho^* & 1 \end{bmatrix}. \quad (2.40)$$

In Figs. 2.6a and 2.6b, we plot the RMSE of each target versus d for orthogonal signals ($\rho = 0$), SNR = 10 dB and $T = 200$. Note the RMSEs of each target are qualitatively similar. MNF matching performs significantly better than no coupling for both targets, and neither depends on F_{min} . Again, the performance of self-matching is sensitive to coupling and F_{min} . For $F_{min} = 1$, self-matching coincides with MNF matching for all d . However, a small amount of amplifier noise causes a significant degradation that varies with d . The RMSEs do not appear as sensitive to F_{min} as in the single target case, presumably



(a) Versus F_{\min} for $\rho = 0, \Delta\phi = 30^\circ$ (b) Versus ρ for $d = 0.2\lambda$ and $\Delta\phi = 15^\circ, 45^\circ$

Figure 2.7: CRB on RMSE of target at 90° with second target at $90^\circ + \Delta\phi$ and SNR=10 dB

because interference from the other target reduces the information in directions where the antenna noise covariance is small. When amplifier noise dominates, MNF matching outperforms self-matching by a factor of 20 in the presence of strong coupling.

To see how performance depends on the source of observation noise, in Fig. 2.7a we plot ratio of the RMSE for self-matching and MNF matching versus $F_{\min} - 1$. Again, the performance of self-matching degrades as the fraction of amplifier noise increases. However, the degradation is much greater for two targets than it was for one; thus, optimal matching appears to be more important in multiple target environments. Finally, in Fig. 2.7b, we plot the RMSE of a target at 90° versus target signal correlation ρ in the presence of a second target located at $90^\circ + \Delta\phi$ for different values of $\Delta\phi$ with $F_{\min} = 1.44$, SNR=10 dB and $T = 200$. Note performance varies significantly with the angle of the second target, and degrades monotonically with correlation ρ , and the worst performance occurs when the same signal arrives from both directions ($\rho = 1$).

2.5 Conclusions

In this chapter, we have considered a general class of Bayesian and non-Bayesian estimation problems that observe the signal through a compact array front-end consisting of coupled antennas, impedance matching, amplifiers, and physical sources of antenna and amplifier noise. We derived the Cramér-Rao bounds associated with these problems, and showed that minimum-noise-factor matching is universally optimal in the sense of minimizing the CRB for all problems in the class. We then considered two applications: MIMO channel estimation and direction-of-arrival estimation. For both applications, numerical results suggest the performance of conventional antenna self-matching is highly sensitive to amplifier noise factor F_{min} in the presence of strong coupling. However, MNF matching eliminates this sensitivity, and significantly outperforms self-matching when amplifier noise is non-negligible. Further, in DoA estimation, optimal matching can also significantly increase the beamwidth, or range of target angles that can be accurately estimated. These improvements can be traded for other important benefits, such as reduced SNR requirements, faster estimation times, or increased coverage area.

Chapter 3

Performance Limits of Arbitrary Compact Arrays

In the last chapter, we considered a general class of Bayesian and non-Bayesian estimation problems, in which the signal of interest is observed through a sensor front-end consisting of coupled antennas, an impedance matching network, amplifiers, and physical noise sources. We derived the Cramer-Rao Bound (CRB) of such system and showed that minimum-noise-factor matching is universally optimal in the sense of minimizing the CRB for all estimation problems in the class. These results provide a way to calculate the CRB for any sensor array in terms of the signal and noise properties, and the array radiation pattern and impedance. Thus, if we are given a particular sensor array, these results show how antenna matching and estimation algorithms can be jointly optimized to make the best use of that array, in the sense of minimizing the CRB. On a more fundamental level, it is natural to ask whether estimation theory can provide guidance *on the design of the sensor array itself*. All arrays detect information by indirectly observing the currents induced by incident signal and noise fields in some volume of conducting

material, say V . How much information is contained in the currents on V ? How does the information vary with the shape and material of V , and with the number and placement of antenna ports? Most importantly, how can we use this knowledge to design sensor arrays that efficiently capture all of the useful information within a given volume?

In this chapter, we consider the problem of estimating a collection of real parameters \mathbf{u} from an incident electromagnetic field. Our goal is to quantify the amount of useful information about \mathbf{u} that can be extracted from the currents induced in an arbitrary conducting volume V , and to determine the best possible estimation error that can be achieved by any M -port array that observes the field through V . Several authors have studied the information contained in electromagnetic fields within a volume, often to determine degrees-of-freedom [71, 72] or decompose information among incident spherical waves [64–67, 73]. In [67], the authors calculate the maximum Fisher information for DoA and polarization estimation that can be achieved by any lossless array that fits within a spherical volume. Our work differs from [67] in several ways: (1) we consider arrays of arbitrary shape and conductivity; (2) we consider the general problem of estimating the parameters of an incident Gaussian field; (3) we consider the impact of physical sources of noise, such as sky and amplifier noise; (4) we focus only narrowband estimation problems, whereas [67] considers broadband signals; and (5) we consider optimal design of the number and placement of array ports, whereas [67] only compares the Fisher information of the sphere with two conventional arrays. In [79], the author determines the maximum MIMO capacity attainable by any M -port array that observes an incident field through an arbitrary volume V . This chapter considers signal and front-end models similar to ours, but considers capacity rather than signal estimation, and so employs different tools and metrics from those considered here.

The rest of the chapter is organized as follows. In Sec. 3.1, we present our system

model and introduce the CRB for an arbitrary M -port array with Gaussian observations. In Sec. 3.2, we consider an array formed from an arbitrary conductor V , and derive formulas relating the Fisher information of the array to the properties of V and the incident electromagnetic field. In Sec. 3.3, we define the concept of aperture Fisher information, decompose the information in V among a collection of unit-power eigencurrents, and derive bounds on the Fisher information of any M -port array that observes the signal through V . Finally, we present numerical examples in Sec. 3.4 and summarize our conclusions in Sec. 3.5.

3.1 System Model

Consider a receiver that estimates a collection of real parameters $\mathbf{u} \in \mathbb{R}^P$ using observations from M antennas. The antennas may experience a certain degree of mutual coupling, so the signals and noise observed by each antenna may be correlated. Before the signals are processed, they pass through an analog front-end consisting of an impedance matching network followed by a bank of noisy amplifiers, as illustrated in Fig. 3.1. In Chapter 2, we derived a formula for the Cramér-Rao Bound for the estimation of \mathbf{u} for lossless arrays with arbitrary impedance matching and amplifiers. In this section, we extend these results to arrays with loss. In Sec. 3.2, we will use these results to study arbitrary arrays formed from a common conductor.

3.1.1 Sensor Array

Let $\mathbf{v} \in \mathbb{C}^M$ denote the complex-baseband voltage across the terminals of the array and let $\mathbf{i} \in \mathbb{C}^M$ be the corresponding current flowing into the array. The relationship between

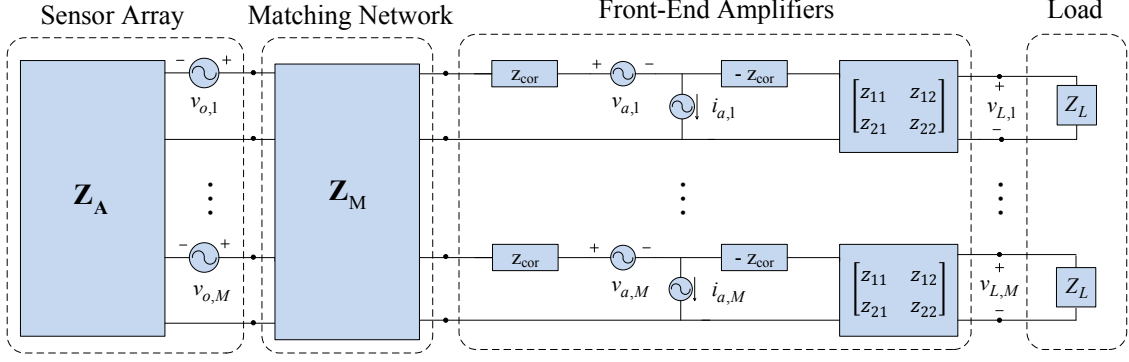


Figure 3.1: Circuit model of a sensor array front-end

the two can be described by

$$\mathbf{v} = \mathbf{Z}_A \mathbf{i} + \mathbf{v}_o, \quad (3.1)$$

where $\mathbf{Z}_A \in \mathbb{C}^{M \times M}$ is the array impedance matrix and \mathbf{v}_o is the open-circuit voltage induced by the incident field. Here, diagonal elements of \mathbf{Z}_A represent antenna self-impedances, while off-diagonal elements represent mutual coupling between antennas. We can always write

$$\mathbf{Z}_A = \mathbf{R}_A + j\mathbf{X}_A, \quad (3.2)$$

where $\mathbf{R}_A \triangleq \frac{1}{2}(\mathbf{Z}_A + \mathbf{Z}_A^H)$, $\mathbf{X}_A \triangleq \frac{1}{2j}(\mathbf{Z}_A - \mathbf{Z}_A^H)$, and $(\cdot)^H$ denotes the conjugate-transpose. If the antennas are reciprocal $\mathbf{Z}_A = \mathbf{Z}_A^T$, then $\mathbf{R}_A = \text{Re}[\mathbf{Z}_A]$ and $\mathbf{X}_A = \text{Im}[\mathbf{Z}_A]$ are both real, symmetric matrices.

In Chapter 2, we considered only lossless arrays. In this work, we explore the impact of internal losses on estimation accuracy. Thus, the array resistance matrix is modeled as a sum of a radiation resistance \mathbf{R}_{AE} and loss resistance \mathbf{R}_{AI} [87]:

$$\mathbf{R}_A = \mathbf{R}_{AE} + \mathbf{R}_{AI}. \quad (3.3)$$

The radiation resistance is closely related to the antenna radiation patterns. Let the pattern $\mathbf{E}_m(\theta, \phi)$ be defined by the far-field approximation of field produced by antenna m at (r, θ, ϕ) when the antenna is driven by current i_m ,

$$\mathbf{E}(r, \theta, \phi) = (1/r)e^{-jkr}i_m\mathbf{E}_m(\theta, \phi) , \quad (3.4)$$

where $k = 2\pi/\lambda$ and λ is the wavelength. If all patterns are expressed relative to a common phase center, then for any lossless array in free space, we have [87, eq. 38]

$$[\mathbf{R}_{\mathbf{AE}}]_{mn} = \frac{1}{\eta} \iint_{S_1} \mathbf{E}_m^T(\theta, \phi)\mathbf{E}_n^*(\theta, \phi) dS \quad (3.5)$$

where $\eta = 120\pi \Omega$ is the impedance of free space, and $\iint_{S_1} (\cdot) dS$ denotes the surface integral over the surface of the unit sphere S_1 . Note that (3.5) can be applied to both lossless and *lossy* antennas, since $\mathbf{R}_{\mathbf{A}}$ and $\mathbf{R}_{\mathbf{AE}}$ have the same antenna patterns.

Consider the problem of estimating \mathbf{u} based on indirect observations of the open-circuit voltage $\mathbf{v}_{\mathbf{o}}$ in (3.1). In general, we assume $\mathbf{v}_{\mathbf{o}}$ can be expressed as

$$\mathbf{v}_{\mathbf{o}} = \mathbf{s}_{\mathbf{A}} + \mathbf{n}_{\mathbf{o}} , \quad (3.6)$$

where $\mathbf{s}_{\mathbf{A}} \in \mathbb{C}^M$ is the target signal and $\mathbf{n}_{\mathbf{o}}$ represents noise. Here we take $\mathbf{s}_{\mathbf{A}}$ to be a circularly-symmetric Gaussian random vector [74, Sec. 15.7] with mean $\boldsymbol{\mu}_{\mathbf{s}_{\mathbf{A}}}(\mathbf{u})$ and covariance $\boldsymbol{\Sigma}_{\mathbf{s}_{\mathbf{A}}}(\mathbf{u})$, which we denote by $\mathbf{s}_{\mathbf{A}} \sim \mathcal{CN}(\boldsymbol{\mu}_{\mathbf{s}_{\mathbf{A}}}(\mathbf{u}), \boldsymbol{\Sigma}_{\mathbf{s}_{\mathbf{A}}}(\mathbf{u}))$. Specific examples of estimation problem of this form will be considered in Sec. 3.4.

The term $\mathbf{n}_{\mathbf{o}}$ in (3.6) represents antenna noise due to thermal noise from the loss resistance $\mathbf{R}_{\mathbf{AI}}$ as well as exterior noise from the environment. The noise contributed by $\mathbf{R}_{\mathbf{AI}}$ can be modeled as Johnson-Nyquist noise. Twiss [75] has shown that a resistance $\mathbf{R}_{\mathbf{AI}}$

at temperature T_0 in Kelvin will produce an open-circuit noise voltage with distribution $\mathcal{CN}(\mathbf{0}, 4k_b T_0 B \mathbf{R}_{\mathbf{AI}})$, where $\mathbf{0} \in \mathbb{C}^M$ is the all-zero vector, $k_b = 1.38 \times 10^{-23}$ J/K is Boltzmann's constant and B is the system bandwidth in Hz. We assume the exterior noise has a known distribution $\mathcal{CN}(\mathbf{0}, \mathbf{\Sigma}_e)$, where $\mathbf{\Sigma}_e$ is independent of \mathbf{u} . Thus antenna noise can be modeled as $\mathbf{n}_o \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma}_o)$, where

$$\mathbf{\Sigma}_o = 4k_b T_0 B \mathbf{R}_{\mathbf{AI}} + \mathbf{\Sigma}_e . \quad (3.7)$$

As an example, if the sensor array with radiation resistance $\mathbf{R}_{\mathbf{AE}}$ is surrounded by an isotropic black-body of temperature T_A , Twiss has also shown that the corresponding exterior noise covariance is given by $\mathbf{\Sigma}_e = 4k_b T_A B \mathbf{R}_{\mathbf{AE}}$. If $T_A = T_0$, the combined antenna noise from both sources is distributed as $\mathbf{n}_o \sim \mathcal{CN}(\mathbf{0}, 4k_b T_0 B \mathbf{R}_{\mathbf{A}})$ in this case.

3.1.2 Front-End

In Fig. 3.1, the signals observed by the estimation algorithms are the load voltages $\mathbf{v}_L = (v_{L,1}, \dots, v_{L,M})^T$. Prior to reaching the load, the signals pass through a front-end consisting of a matching network followed by a bank of independent and identical amplifiers. As in [27], we represent the amplifiers using the standard Rothe-Dahlke equivalent network, where $\mathbf{v}_a \sim \mathcal{CN}(\mathbf{0}, 4k_b T_0 B r_a \mathbf{I})$ and $\mathbf{i}_a \sim \mathcal{CN}(\mathbf{0}, 4k_b T_0 B g_a \mathbf{I})$ are independent random vectors that represent internal amplifier noise sources, and \mathbf{I} is the identity matrix. For simplicity, we call these circuits amplifiers, but any sequence of linear, noisy circuits can be represented this way [27].

In this section, we first consider the case of no matching network in Fig. 3.1. Using basic circuit analysis, we can show the load voltages in Fig. 3.1 are given by (e.g. [39,

eq. 14])

$$\mathbf{v}_L = \mathbf{K} \left[\mathbf{v}_o - \mathbf{v}_a - (\mathbf{Z}_A + z_{\text{cor}} \mathbf{I}) \mathbf{i}_a \right], \quad (3.8)$$

where $\mathbf{K} = \frac{z_{21}z_L}{z_L+z_{22}} \left(\mathbf{Z}_A + \frac{z_{11}(z_L+z_{22})-z_{12}z_{21}}{z_L+z_{22}} \mathbf{I} \right)^{-1}$. Since optimal estimation performance is not affected by an invertible transformation of the observation, we can define the observation to be

$$\begin{aligned} \mathbf{y}_A &= \mathbf{K}^{-1} \mathbf{v}_L = \mathbf{v}_o - \mathbf{v}_a - (\mathbf{Z}_A + z_{\text{cor}} \mathbf{I}) \mathbf{i}_a \\ &= \mathbf{s}_A + \mathbf{n}_A, \end{aligned} \quad (3.9)$$

where $\mathbf{n}_A \triangleq \mathbf{n}_o - \mathbf{v}_a - (\mathbf{Z}_A + z_{\text{cor}} \mathbf{I}) \mathbf{i}_a$ is the overall system noise. Intuitively, (3.9) expresses all sources in (3.8) in terms of equivalent sources at the open-circuit array terminals.

3.1.3 Cramér-Rao Bound

Consider \mathbf{u} to be fixed and real. For a given receiver, let $\hat{\mathbf{u}} = \hat{\mathbf{u}}(\mathbf{y}_A)$ be any unbiased estimator of $\mathbf{u} \in \mathbb{R}^P$ based on the observation of \mathbf{y}_A , so $E_{\mathbf{y}_A|\mathbf{u}}[\hat{\mathbf{u}}] = \mathbf{u}$, where $E_{\mathbf{y}_A|\mathbf{u}}$ denotes expectation with respect to the conditional pdf of \mathbf{y}_A given \mathbf{u} . The Cramér-Rao Bound (CRB) provides a fundamental lower bound on error covariance of $\hat{\mathbf{u}}$: [74, Sec. 15.7]

$$\Sigma_{\hat{\mathbf{u}}}(\mathbf{u}) \triangleq E_{\mathbf{y}_A|\mathbf{u}}[(\hat{\mathbf{u}} - \mathbf{u})(\hat{\mathbf{u}} - \mathbf{u})^T] \geq \mathcal{I}_A^{-1}(\mathbf{u}), \quad (3.10)$$

where $\mathcal{I}_A(\mathbf{u})$ is the Fisher information matrix (FIM). Here, we use the notation $\mathbf{A} \geq \mathbf{B}$ for Hermitian matrices \mathbf{A} and \mathbf{B} , to indicate $\mathbf{A} - \mathbf{B}$ is positive semi-definite.

For complex Gaussian observation model (3.9), the FIM is given by [74, Sec. 15.7]

$$[\mathcal{I}_{\mathbf{A}}(\mathbf{u})]_{ij} = 2\text{Re} \left[\frac{\partial \boldsymbol{\mu}_{\mathbf{s}_{\mathbf{A}}}^H}{\partial u_i} \boldsymbol{\Sigma}_{\mathbf{y}_{\mathbf{A}}}^{-1}(\mathbf{u}) \frac{\partial \boldsymbol{\mu}_{\mathbf{s}_{\mathbf{A}}}}{\partial u_j} \right] + \text{tr} \left[\boldsymbol{\Sigma}_{\mathbf{y}_{\mathbf{A}}}^{-1}(\mathbf{u}) \frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}_{\mathbf{A}}}(\mathbf{u})}{\partial u_i} \boldsymbol{\Sigma}_{\mathbf{y}_{\mathbf{A}}}^{-1}(\mathbf{u}) \frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}_{\mathbf{A}}}(\mathbf{u})}{\partial u_j} \right] \quad (3.11)$$

where tr , Re and $(\cdot)^H$ denote the trace, real part, and conjugate transpose, respectively. Hereafter, we write this formula compactly as

$$\mathcal{I}_{\mathbf{A}}(\mathbf{u}) = \mathcal{F}(\boldsymbol{\mu}_{\mathbf{y}_{\mathbf{A}}}, \boldsymbol{\Sigma}_{\mathbf{y}_{\mathbf{A}}}) . \quad (3.12)$$

The lower bound is achieved by any estimator $\hat{\mathbf{u}}$ that satisfies

$$\frac{\partial p(\mathbf{y}_{\mathbf{A}}|\mathbf{u})}{\partial \mathbf{u}} = \mathcal{I}_{\mathbf{A}}(\mathbf{u}) [\hat{\mathbf{u}}(\mathbf{y}_{\mathbf{A}}) - \mathbf{u}] , \quad (3.13)$$

in which case the estimator is called *efficient*.

3.1.4 CRB Optimal Matching

We have so far assumed no matching between the sensors and front-end in Fig. 3.1. However, several studies have shown that impedance matching can significantly enhance capacity for similar MIMO communication systems (e.g. [27] and references therein). In Chapter 2, we also showed that impedance matching can significantly improve the performance of estimation algorithms, especially when antenna coupling and amplifier noise are significant.

For $M = 1$, the noise contributed by the amplifier in (3.9) is often measured by the *noise factor* $F \triangleq \sigma_{\mathbf{n}_{\mathbf{A}}}^2 / \sigma_{\mathbf{n}_{\mathbf{o}}}^2$, which depends on antenna impedance. The noise factor takes

its minimum value (e.g. [27, Appendix])

$$F_{min} = 1 + 2 \left(g_a r_{cor} + \sqrt{g_a r_a + (g_a r_{cor})^2} \right) , \quad (3.14)$$

when the antenna impedance equals $z_{opt} = \sqrt{r_a/g_a + r_{cor}^2} - jx_{cor}$, where g_a , r_a and $z_{cor} = r_{cor} + jx_{cor}$ are internal amplifier noise parameters.

For general M , the matching network in Fig. 3.1 is described by a block impedance matrix

$$\mathbf{Z}_M \triangleq \begin{bmatrix} \mathbf{Z}_{M11} & \mathbf{Z}_{M12} \\ \mathbf{Z}_{M21} & \mathbf{Z}_{M22} \end{bmatrix} , \quad (3.15)$$

where each submatrix lies in $\mathbb{C}^{M \times M}$. Ideally the network is designed with passive, reactive elements so it is lossless, reciprocal and noiseless. The network is lossless and noiseless if [39, eq. 30] $\mathbf{Z}_M = -\mathbf{Z}_M^H$ and reciprocal if $\mathbf{Z}_M = \mathbf{Z}_M^T$ where $[\cdot]^T$ denotes the transpose.

When matching is applied, the noise in the observation model (3.9) is changed to

$$\begin{aligned} \mathbf{y}_A &= \mathbf{s}_A + \mathbf{n}_A , \\ \mathbf{n}_A &\triangleq \mathbf{n}_o - \mathbf{M}^{-1} [\mathbf{v}_a + (\bar{\mathbf{Z}}_A + z_{cor}\mathbf{I}) \mathbf{i}_a] , \end{aligned} \quad (3.16)$$

where $\mathbf{M} \triangleq \mathbf{Z}_{M21} (\mathbf{Z}_A + \mathbf{Z}_{M11})^{-1}$ and $\bar{\mathbf{Z}}_A \triangleq -\mathbf{M}\mathbf{Z}_{M12} + \mathbf{Z}_{M22}$. Since \mathbf{n}_o , \mathbf{v}_a and \mathbf{i}_a are Gaussian and independent, the total observation noise is distributed as $\mathbf{n}_A \sim \mathcal{CN}(\mathbf{0}, \Sigma_{\mathbf{n}_A})$ where

$$\Sigma_{\mathbf{n}_A} = \Sigma_o + 4k_b T_0 B \mathbf{M}^{-1} \left[r_a \mathbf{I} + g_a (\bar{\mathbf{Z}}_A + z_{cor}\mathbf{I}) (\bar{\mathbf{Z}}_A + z_{cor}\mathbf{I})^H \right] \mathbf{M}^{-H} . \quad (3.17)$$

In our model (3.17), note that impedance matching only changes the noise covariance $\Sigma_{\mathbf{n}_A}$. For any array and matching network, we write $\mathcal{I}_A(\mathbf{u}) = \mathcal{I}_A(\mathbf{u}; \Sigma_{\mathbf{n}_A})$ to display the

dependence of the FIM on the noise covariance. In Chapter 2, we showed that $\Sigma_{\mathbf{n}_A} \geq \Sigma'_{\mathbf{n}_A}$ implies $\mathcal{I}_A(\mathbf{u}; \Sigma_{\mathbf{n}_A}) \leq \mathcal{I}_A(\mathbf{u}; \Sigma'_{\mathbf{n}_A})$. For the system model of Fig. 3.1, we also showed that the noise covariance satisfies

$$\Sigma_{\mathbf{n}_A} \geq \Sigma_{\mathbf{o}} + (F_{min} - 1)4k_b T_0 B \mathbf{R}_A, \quad (3.18)$$

with equality if and only if $\bar{\mathbf{Z}}_A = z_{opt} \mathbf{I}$. We call any lossless network that achieves $\bar{\mathbf{Z}}_A = z_{opt} \mathbf{I}$ a *minimum-noise-factor (MNF) match*. Taken together, these two results imply that MNF matching is optimal in the sense of maximizing the FIM (or equivalently minimizing the CRB) for all estimation problems involving Gaussian signals.

One example of matching network that meets these requirements is [27, eq. 42]

$$\mathbf{Z}_M = \begin{bmatrix} \mathbf{Z}_{M11} & \mathbf{Z}_{M12} \\ \mathbf{Z}_{M21} & \mathbf{Z}_{M22} \end{bmatrix} = j \begin{bmatrix} -\mathbf{X}_A & (r_{opt} \mathbf{R}_A)^{1/2} \\ (r_{opt} \mathbf{R}_A)^{1/2} & x_{opt} \mathbf{I} \end{bmatrix}. \quad (3.19)$$

Note that (3.19) is a reciprocal network only if $\mathbf{Z}_A \neq \mathbf{Z}_A^T$, $\mathbf{R}_A \notin \mathbb{R}^{M \times M}$.

3.2 Arrays from Arbitrary Conductors

A sensor array collects signal information by indirectly observing the currents induced in some conducting body by the incident electromagnetic fields. In the last section, we expressed the Fisher information associated with a sensor array in terms of the array impedance, and the signal and noise components of the open-circuit voltage (3.6). In this section, we adopt a more general perspective, and consider an array formed from an arbitrary conductor of shape V . This conductor could be a single, contiguous structure such as a patch, or a collection of structures, such as an array of dipoles. We assume

the currents and fields on V are observed through a collection of M ports that couple to V . The aim of this section is to relate the Fisher information observed at these ports to the properties of the underlying conductor V and incident signal and noise fields. In Sec. 3.3, we use this relationship to determine the maximum Fisher information that can be achieved by any array formed from V , and to gain insight into how to design ports that achieve this maximum.

3.2.1 Fields and Currents on Lossy Conductors

We first review the equations relating the fields and currents in a lossy conductor, following [87]. Consider an arbitrary region V filled with a homogeneous material of conductivity $\sigma < +\infty$. The fields in V can be decomposed into the sum of an incident field due to external sources \mathbf{E}_o , and a scattered field \mathbf{E}_s induced by the volume¹ current \mathbf{J} . The relationship between these quantities is given by the constitutive equation [87, eq. 50]

$$\mathbf{J} = \sigma (\mathbf{E}_o + \mathbf{E}_s) . \quad (3.20)$$

Note the total field is a linear function of the current that depends only on conductivity, $\mathbf{Z}_I(\mathbf{J}) = \sigma^{-1}\mathbf{J}$. Using the Green's function, the scattered field can also be expressed as a linear function of current that depends only on the aperture V ,

$$\mathbf{E}_s = -\mathbf{Z}_E(\mathbf{J}) . \quad (3.21)$$

¹In a strong conductor, the fields and currents are largely confined to the surface of V and (3.22) could be viewed as surface fields and currents.

Thus, we can rewrite (3.20) as a linear equation that relates the incident field to the volume currents

$$\mathbf{Z}(\mathbf{J}) \triangleq \mathbf{Z}_{\mathbf{E}}(\mathbf{J}) + \mathbf{Z}_{\mathbf{I}}(\mathbf{J}) = \mathbf{E}_o . \quad (3.22)$$

For calculations, it is more convenient to work with discretized fields and currents, rather than the continuous quantities in (3.22). Using the Method of Moments, for example, we can express all of the fields and currents in terms of a common set of basis functions, say

$$\mathbf{J} = \mathbf{J}(s) = J_1 \mathbf{W}_1(s) + \dots + J_K \mathbf{W}_K(s) \text{ for all } s \in V . \quad (3.23)$$

Here we assume the basis is normalized in a way that preserves power calculations, so

$$-\frac{1}{2} \int_V \mathbf{E} \cdot \mathbf{J}^* dv = -\frac{1}{2} (E_1 J_1^* + \dots + E_M J_M^*) . \quad (3.24)$$

A detailed procedure to discretize these equations is discussed in [88]. This process is also automated in the software tool FEKO [81]. Once (3.22) is discretized, we can interpret \mathbf{J} and \mathbf{E}_o as vectors in \mathbb{C}^K , and \mathbf{Z} , $\mathbf{Z}_{\mathbf{E}}$ and $\mathbf{Z}_{\mathbf{I}}$ as matrices in $\mathbb{C}^{K \times K}$.

3.2.2 Embedded Ports

We normally observe the fields and currents on V through one or more ports that couple to the conductor. In this section, we consider ports that are embedded in V , in the sense that they are completely contained in V and require no additional structures to be attached, except a shielded cable with negligible scattering. Suppose V is used as an antenna array with M embedded ports, as illustrated in Fig. 3.2. Let v_m and i_m denote the voltage and current into the m -th port, respectively. To calculate the Fisher information

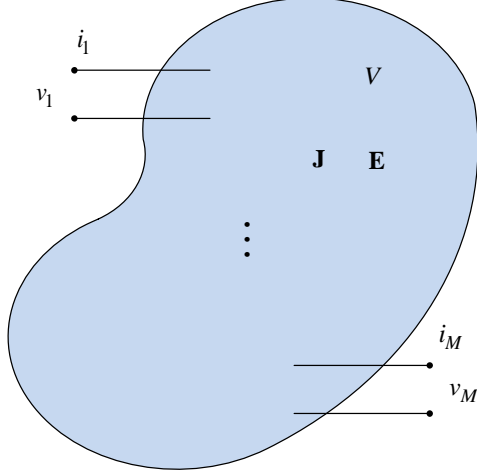


Figure 3.2: An array formed from an arbitrary conductor V

(3.11) of this array, we must find the array impedance \mathbf{Z}_A and the open-circuit voltage \mathbf{v}_o in (3.1). We now derive these quantities from the properties of the underlying conductor V and the incident field \mathbf{E}_o .

When the array is driven by port currents \mathbf{i} , the (complex) power supplied to V is [89, eq. 9.3.3]

$$P_s = -\frac{1}{2}(v_1 i_1^* + \cdots + v_M i_M^*) = -\frac{1}{2} \mathbf{v}^T \mathbf{i}^* . \quad (3.25)$$

Note each port current produces a proportional volume current on V , so the volume current produced by all of the ports can be written as

$$\mathbf{J} = \mathbf{C}_1 i_1 + \cdots + \mathbf{C}_M i_M = \mathbf{C} \mathbf{i} ,$$

where $\mathbf{C}_m \in \mathbb{C}^K$ is the current on V generated by a 1A current into port m when all other ports are open-circuit, and $\mathbf{C} \in \mathbb{C}^{K \times M}$ is the matrix whose m -th column is \mathbf{C}_m .

Using (3.24) and [89, eq. 7.3.23], the power supplied by \mathbf{J} to V can also be expressed as

$$P_s = -\frac{1}{2}\mathbf{E}^T\mathbf{J}^*, \quad (3.26)$$

where \mathbf{E} is the sum of the scattered field due to \mathbf{J} and the incident field,

$$\mathbf{E} = \mathbf{Z}\mathbf{J} + \mathbf{E}_o. \quad (3.27)$$

Equating the two power expressions, it follows

$$-\frac{1}{2}\mathbf{v}^T\mathbf{i}^* = -\frac{1}{2}[\mathbf{Z}_A\mathbf{i} + \mathbf{v}_o]^T\mathbf{i}^* = -\frac{1}{2}[\mathbf{Z}\mathbf{C}\mathbf{i} + \mathbf{E}_o]^T\mathbf{C}^*\mathbf{i}^* = -\frac{1}{2}[\mathbf{C}^H\mathbf{Z}\mathbf{C}\mathbf{i} + \mathbf{C}^H\mathbf{E}_o]^T\mathbf{i}^*. \quad (3.28)$$

Since these equations agree for all \mathbf{i} , we conclude

$$\mathbf{Z}_A = \mathbf{C}^H\mathbf{Z}\mathbf{C}, \quad \mathbf{v}_o = \mathbf{C}^H\mathbf{E}_o. \quad (3.29)$$

The incident field can be decomposed into two parts,

$$\mathbf{E}_o = \mathbf{s}_V + \mathbf{n}_e, \quad (3.30)$$

where $\mathbf{s}_V \in \mathbb{C}^K$ and $\mathbf{n}_e \in \mathbb{C}^K$ are the signal and noise fields on V , respectively. Similar to \mathbf{v}_o in (3.6), we assume the signal field is Gaussian with a known mean and covariance, $\mathbf{s}_V \sim \mathcal{CN}(\boldsymbol{\mu}_{\mathbf{s}_V}(\mathbf{u}), \boldsymbol{\Sigma}_{\mathbf{s}_V}(\mathbf{u}))$. Analogous to \mathbf{n}_o in (3.6), we assume the noise field is distributed as $\mathbf{n}_e \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Sigma}_{\mathbf{n}_e})$, where $\boldsymbol{\Sigma}_{\mathbf{n}_e} = 4k_bT_0B\mathbf{R}_I + \boldsymbol{\Sigma}_{e,V}$ where $\mathbf{R}_I = (1/2)(\mathbf{Z}_I + \mathbf{Z}_I^H)$. Here $4k_bT_0B\mathbf{R}_I$ represents noise due to internal losses, and $\boldsymbol{\Sigma}_{e,V}$ represents external noise. In particular, if the aperture is surrounded by an isotropic black-body of temper-

ature T_0 , the combined noise from both sources has covariance $\Sigma_{\mathbf{n}_e} = 4k_b T_0 B \mathbf{R}$, where $\mathbf{R} = (1/2)(\mathbf{Z} + \mathbf{Z}^H)$.

From (3.29), we can relate the properties of \mathbf{s}_A and \mathbf{n}_o in (3.6) to the incident field (3.30). From (3.6) and (3.29), we can infer $\mathbf{s}_A = \mathbf{C}^H \mathbf{s}_V$ is Gaussian with mean and covariance

$$\boldsymbol{\mu}_{\mathbf{s}_A}(\mathbf{u}) = \mathbf{C}^H \boldsymbol{\mu}_{\mathbf{s}_V}(\mathbf{u}) , \quad \Sigma_{\mathbf{s}_A}(\mathbf{u}) = \mathbf{C}^H \Sigma_{\mathbf{s}_V}(\mathbf{u}) \mathbf{C} . \quad (3.31)$$

Similarly, $\mathbf{n}_o = \mathbf{C}^H \mathbf{n}_e$ and so

$$\Sigma_{\mathbf{o}} = \mathbf{C}^H \Sigma_{\mathbf{n}_e} \mathbf{C} = \mathbf{C}^H [4k_b T_0 B \mathbf{R}_I + \Sigma_{e,V}] \mathbf{C} . \quad (3.32)$$

From (3.29), we have $\mathbf{R}_{AI} = \mathbf{C}^H \mathbf{R}_I \mathbf{C}$, so the first term in the brackets produces the first term on the right in (3.7). It follows the second term in the brackets produces the other term in (3.7),

$$\Sigma_{\mathbf{e}} = \mathbf{C}^H \Sigma_{e,V} \mathbf{C} . \quad (3.33)$$

We can now use these equations to calculate Fisher information for any incident Gaussian signal and noise fields, any aperture V , any collection of M embedded ports, and any matching network. The incident signal field is defined by its mean $\boldsymbol{\mu}_{\mathbf{s}_V}(\mathbf{u})$ and covariance $\Sigma_{\mathbf{s}_V}(\mathbf{u})$; the incident noise field is defined by the covariance $\Sigma_{e,V}$. The aperture V is described by the radiation and loss impedances, \mathbf{Z}_E and \mathbf{Z}_I , and the noise temperature T_0 . The ports are specified by the current matrix \mathbf{C} . Given these quantities, we can calculate the mean and covariances of the signal and noise components in (3.6) from (3.31) and (3.32), respectively. We can then calculate radiation and loss impedances of the array

from (3.29). For any matching network, with matching matrix \mathbf{M} in (3.17), we can then calculate the total observation noise in the array from (3.17), and the resulting Fisher information from (3.11).

As noted in Sec. 3.1, the array FIM is always maximized when MNF matching (3.19) is used. In this case, the array observation (3.17) reduces to

$$\mathbf{y}_A = \mathbf{s}_A + \mathbf{n}_A , \quad (3.34)$$

where $\mathbf{n}_A \sim \mathcal{CN}(\mathbf{0}, \Sigma_{\mathbf{n}_A})$ and $\Sigma_{\mathbf{n}_A} = \Sigma_{\mathbf{o}} + (F_{min} - 1)4k_b T_0 B \mathbf{R}_A$. From (3.31)–(3.33), note we can also write this equation as $\mathbf{y}_A = \mathbf{C}^H \mathbf{y}_V$ where \mathbf{y}_V is defined by

$$\mathbf{y}_V \triangleq \mathbf{s}_V + \mathbf{n}_V , \quad (3.35)$$

and $\mathbf{n}_V \sim \mathcal{CN}(\mathbf{0}, \Sigma_{\mathbf{n}_V})$ and

$$\Sigma_{\mathbf{n}_V} \triangleq \Sigma_{\mathbf{n}_e} + (F_{min} - 1)4k_b T_0 B \mathbf{R} . \quad (3.36)$$

Comparing with (3.30), we can view \mathbf{y}_V as a noisy observation of the incident field \mathbf{E}_o , where the last term in $\Sigma_{\mathbf{n}_V}$ represents amplifier noise referred to the volume V . This equation shows the array observation $\mathbf{y}_A = \mathbf{C}^H \mathbf{y}_V$ provides a partial view of the aperture observation \mathbf{y}_V , and the relationship between the distributions of these observations is given by

$$\boldsymbol{\mu}_{\mathbf{y}_A}(\mathbf{u}) = \mathbf{C}^H \boldsymbol{\mu}_{\mathbf{y}_V}(\mathbf{u}) , \quad \Sigma_{\mathbf{y}_A}(\mathbf{u}) = \mathbf{C}^H \Sigma_{\mathbf{y}_V}(\mathbf{u}) \mathbf{C} . \quad (3.37)$$

When optimal matching is used, we conclude the array FIM (3.11) can be expressed in

terms of the incident field as

$$\mathcal{I}_{\mathbf{A}}(\mathbf{u}) = \mathcal{F}(\boldsymbol{\mu}_{\mathbf{y}_{\mathbf{A}}}(\mathbf{u}), \boldsymbol{\Sigma}_{\mathbf{y}_{\mathbf{A}}}(\mathbf{u})) = \mathcal{F}(\mathbf{C}^H \boldsymbol{\mu}_{\mathbf{y}_{\mathbf{V}}}(\mathbf{u}), \mathbf{C}^H \boldsymbol{\Sigma}_{\mathbf{y}_{\mathbf{V}}}(\mathbf{u}) \mathbf{C}) , \quad (3.38)$$

where \mathcal{F} is given in (3.12).

3.3 Aperture Fisher Information

In this section, we investigate the maximum Fisher information that can be extracted from an arbitrary conducting structure V . As in Sec. 3.2, consider any array that observes the fields and currents on V through M embedded ports. We show the maximum Fisher information obtainable with any such array is given by a quantity called the *aperture Fisher information*, that represents all of the useful information contained in the currents and fields induced in the conducting structure. For a broad class of estimation problems, we also show how arrays can be designed, at least in principle, to achieve the aperture Fisher information or minimize diagonal elements of the CRB.

We define aperture Fisher information by way of a thought experiment: Suppose an array could be constructed from V in such a way that each port corresponds to a basis function in (3.23), in the sense that applying a 1A current to port m produces a volume current equal to the m -th basis function in (3.23). It follows this virtual array has $M = K$ ports and the current matrix is $\mathbf{C} = \mathbf{I}$. If optimal matching is used with this array, (3.38) implies the resulting Fisher information would be

$$\mathcal{I}_V(\mathbf{u}) \triangleq \mathcal{F}(\boldsymbol{\mu}_{\mathbf{y}_{\mathbf{V}}}(\mathbf{u}), \boldsymbol{\Sigma}_{\mathbf{y}_{\mathbf{V}}}(\mathbf{u})) . \quad (3.39)$$

We call this quantity the *aperture Fisher information matrix* (AFI). From (3.39), the AFI

appears to depend on the basis (3.23) used to describe the fields and currents on V . In general, adding basis functions increases the number of observations and could increase $\mathcal{I}_V(\mathbf{u})$. Presumably, one could bound the AFI above by an analogous quantity involving the continuous fields on V . However, in this chapter, we assume the continuous fields are discretized in a way that captures all of the relevant signal and noise information in the aperture; more precisely, \mathbf{y}_V is a sufficient statistic for estimating \mathbf{u} on the basis of the continuous fields on V . With this assumption, AFI will not depend on the basis selected. In fact, if adding basis functions increases AFI, then the original basis was clearly not sufficient to capture all of the useful information in the aperture.

3.3.1 Bounds on Arbitrary Arrays

Intuitively, it seems reasonable to expect that the Fisher information contained in the fields and currents of V is an upper bound to the Fisher information of any array constructed from V . The following theorem makes this intuition precise, and is the main result of this chapter. Here the notation $\mathbf{A} \leq \mathbf{B}$ for Hermitian matrices \mathbf{A} and \mathbf{B} means $\mathbf{B} - \mathbf{A}$ is positive semi-definite.

Theorem 3.1: Let V be any conductor excited by the incident Gaussian signal and noise fields (3.30), and let $\mathcal{I}_V(\mathbf{u})$ denote the corresponding aperture Fisher information (3.39). Let $\mathcal{I}_{\mathbf{A}}(\mathbf{u})$ denote the Fisher information (3.38) of any array formed by embedding M ports in V . Then $\mathcal{I}_{\mathbf{A}}(\mathbf{u}) \leq \mathcal{I}_V(\mathbf{u})$. \diamond

Before proving the theorem, a few remarks are in order. First, Theorem 1 suggests that $\mathcal{I}_V(\mathbf{u})$ is an upper bound on the Fisher information of any array that observes the signal through V , in the sense of the matrix partial ordering. Intuitively, $\mathcal{I}_V(\mathbf{u})$ can be

viewed as the maximum Fisher information that any array can extract from V . Second, we are usually more interested in the Cramér-Rao bound, $\text{CRB}_{\mathbf{A}}(\mathbf{u}) = \mathcal{I}_{\mathbf{A}}^{-1}(\mathbf{u})$, rather than Fisher information. From [86, Cor. 7.7.4(a)], if both matrices are invertible, Theorem 1 implies

$$\text{CRB}_V(\mathbf{u}) \triangleq \mathcal{I}_V^{-1}(\mathbf{u}) \leq \mathcal{I}_{\mathbf{A}}^{-1}(\mathbf{u}) = \text{CRB}_{\mathbf{A}}(\mathbf{u}) . \quad (3.40)$$

In this sense, $\text{CRB}_V(\mathbf{u})$ is the minimum CRB that any array formed from V can achieve. Finally, note the diagonal element $[\text{CRB}_{\mathbf{A}}(\mathbf{u})]_{ii}$ is a lower bound on the error of any unbiased estimator of parameter u_i . Note (3.40) implies $\text{CRB}_{\mathbf{A}}(\mathbf{u}) - \text{CRB}_V(\mathbf{u})$ is positive semi-definite. It follows that each diagonal element of this matrix is non-negative [86, Cor. 7.7.4(b)]; hence for each i

$$[\text{CRB}_V(\mathbf{u})]_{ii} \leq [\text{CRB}_{\mathbf{A}}(\mathbf{u})]_{ii} . \quad (3.41)$$

For any array formed from V , the diagonal elements of $\mathcal{I}_V^{-1}(\mathbf{u})$ are therefore a lower bound on the error of any unbiased estimator of the corresponding parameter of \mathbf{u} .

Proof of Theorem 3.1: Since MNF matching maximizes $\mathcal{I}_{\mathbf{A}}(\mathbf{u})$, it is enough to prove the theorem for optimal matching. For Hermitian matrices \mathbf{A} and \mathbf{B} , note $\mathbf{A} \leq \mathbf{B}$ holds if and only if $\mathbf{w}^H \mathbf{A} \mathbf{w} \leq \mathbf{w}^H \mathbf{B} \mathbf{w}$ for all $\mathbf{w} \in \mathbb{C}^p$ [86, pg. 402]. From (3.38) and (3.39), to prove the theorem it therefore suffices to show for all \mathbf{u}, \mathbf{C} and \mathbf{w} that

$$\mathbf{w}^H \mathcal{F}(\mathbf{C}^H \boldsymbol{\mu}_{\mathbf{y}_V}, \mathbf{C}^H \boldsymbol{\Sigma}_{\mathbf{y}_V} \mathbf{C}) \mathbf{w} \leq \mathbf{w}^H \mathcal{F}(\boldsymbol{\mu}_{\mathbf{y}_V}, \boldsymbol{\Sigma}_{\mathbf{y}_V}) \mathbf{w} , \quad (3.42)$$

Here we suppress the dependence of $\boldsymbol{\mu}_{\mathbf{y}_V}(\mathbf{u})$ and $\boldsymbol{\Sigma}_{\mathbf{y}_V}(\mathbf{u})$ on \mathbf{u} for compactness.

From (3.11), we observe

$$\begin{aligned}
\mathbf{w}^H \mathcal{F}(\boldsymbol{\mu}_{\mathbf{y}_V}, \boldsymbol{\Sigma}_{\mathbf{y}_V}) \mathbf{w} &= \sum_{ij} w_i^* w_j \left\{ \frac{\partial \boldsymbol{\mu}_{\mathbf{y}_V}^H}{\partial u_i} \boldsymbol{\Sigma}_{\mathbf{y}_V}^{-1} \frac{\partial \boldsymbol{\mu}_{\mathbf{y}_V}}{\partial u_j} + \left(\frac{\partial \boldsymbol{\mu}_{\mathbf{y}_V}^H}{\partial u_i} \boldsymbol{\Sigma}_{\mathbf{y}_V}^{-1} \frac{\partial \boldsymbol{\mu}_{\mathbf{y}_V}}{\partial u_j} \right)^* \right. \\
&\quad \left. + \text{tr} \left[\boldsymbol{\Sigma}_{\mathbf{y}_V}^{-1} \frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}_V}}{\partial u_i} \boldsymbol{\Sigma}_{\mathbf{y}_V}^{-1} \frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}_V}}{\partial u_j} \right] \right\} \\
&= \mathbf{x}^H \boldsymbol{\Sigma}_{\mathbf{y}_V}^{-1} \mathbf{x} + \bar{\mathbf{x}}^H \boldsymbol{\Sigma}_{\mathbf{y}_V}^{-1} \bar{\mathbf{x}} + \text{tr} \left[\boldsymbol{\Sigma}_{\mathbf{y}_V}^{-1} \mathbf{A}^H \boldsymbol{\Sigma}_{\mathbf{y}_V}^{-1} \mathbf{A} \right] , \tag{3.43}
\end{aligned}$$

where

$$\mathbf{x} \triangleq \sum_j w_j \frac{\partial \boldsymbol{\mu}_{\mathbf{y}_V}}{\partial u_j} , \quad \bar{\mathbf{x}} \triangleq \sum_j w_j^* \frac{\partial \boldsymbol{\mu}_{\mathbf{y}_V}}{\partial u_j} , \quad \mathbf{A} \triangleq \sum_j w_j \frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}_V}}{\partial u_j} .$$

Similarly, we find

$$\begin{aligned}
\mathbf{w}^H \mathcal{F}(\mathbf{C}^H \boldsymbol{\mu}_{\mathbf{y}_V}, \mathbf{C}^H \boldsymbol{\Sigma}_{\mathbf{y}_V} \mathbf{C}) \mathbf{w} &= \mathbf{x}^H \mathbf{C} (\mathbf{C}^H \boldsymbol{\Sigma}_{\mathbf{y}_V} \mathbf{C})^{-1} \mathbf{C}^H \mathbf{x} + \bar{\mathbf{x}}^H \mathbf{C} (\mathbf{C}^H \boldsymbol{\Sigma}_{\mathbf{y}_V} \mathbf{C})^{-1} \mathbf{C}^H \bar{\mathbf{x}} \\
&\quad + \text{tr} \left[(\mathbf{C}^H \boldsymbol{\Sigma}_{\mathbf{y}_V} \mathbf{C})^{-1} \mathbf{C}^H \mathbf{A}^H \mathbf{C} (\mathbf{C}^H \boldsymbol{\Sigma}_{\mathbf{y}_V} \mathbf{C})^{-1} \mathbf{C}^H \mathbf{A} \mathbf{C} \right] \\
&= \mathbf{x}^H \boldsymbol{\Sigma}^{-1} \mathbf{x} + \bar{\mathbf{x}}^H \boldsymbol{\Sigma}^{-1} \bar{\mathbf{x}} + \text{tr} \left[\boldsymbol{\Sigma}^{-1} \mathbf{A}^H \boldsymbol{\Sigma}^{-1} \mathbf{A} \right] , \tag{3.44}
\end{aligned}$$

where the last step follows from the trace identity $\text{tr}[\mathbf{BC}] = \text{tr}[\mathbf{CB}]$ and

$$\boldsymbol{\Sigma}^{-1} \triangleq \mathbf{C} (\mathbf{C}^H \boldsymbol{\Sigma}_{\mathbf{y}_V} \mathbf{C})^{-1} \mathbf{C}^H . \tag{3.45}$$

Consider the matrix $\mathbf{U} \triangleq \boldsymbol{\Sigma}_{\mathbf{y}_V}^{1/2} \mathbf{C} (\mathbf{C}^H \boldsymbol{\Sigma}_{\mathbf{y}_V} \mathbf{C})^{-1/2}$, where $\boldsymbol{\Sigma}_{\mathbf{y}_V}^{-1/2}$ denotes the Hermitian square root [86, Thrm. 7.2.6] of $\boldsymbol{\Sigma}_{\mathbf{y}_V}$. Note that $\mathbf{U}^H \mathbf{U} = \mathbf{I}$, so the rows of \mathbf{U} are orthonormal. Since $\mathbf{U}^H \mathbf{U}$ and $\mathbf{U} \mathbf{U}^H$ must have the same eigenvalues, it follows that all

of the non-zero eigenvalues of $\mathbf{U}\mathbf{U}^H$ are one. Hence,

$$\mathbf{I} \geq \mathbf{U}\mathbf{U}^H = \boldsymbol{\Sigma}_{\mathbf{y}_V}^{1/2} \mathbf{C} (\mathbf{C}^H \boldsymbol{\Sigma}_{\mathbf{y}_V} \mathbf{C})^{-1} \mathbf{C}^H \boldsymbol{\Sigma}_{\mathbf{y}_V}^{1/2}, \quad (3.46)$$

or equivalently $\boldsymbol{\Sigma}^{-1} \leq \boldsymbol{\Sigma}_{\mathbf{y}_V}^{-1}$. From this, it follows immediately that

$$\mathbf{x}^H \mathbf{C} (\mathbf{C}^H \boldsymbol{\Sigma}_{\mathbf{y}_V} \mathbf{C})^{-1} \mathbf{C}^H \mathbf{x} \leq \mathbf{x}^H \boldsymbol{\Sigma}_{\mathbf{y}_V}^{-1} \mathbf{x}, \quad (3.47)$$

$$\bar{\mathbf{x}}^H \mathbf{C} (\mathbf{C}^H \boldsymbol{\Sigma}_{\mathbf{y}_V} \mathbf{C})^{-1} \mathbf{C}^H \bar{\mathbf{x}} \leq \bar{\mathbf{x}}^H \boldsymbol{\Sigma}_{\mathbf{y}_V}^{-1} \bar{\mathbf{x}}. \quad (3.48)$$

For any complex matrix \mathbf{D} with M columns, $\boldsymbol{\Sigma}^{-1} \leq \boldsymbol{\Sigma}_{\mathbf{y}_V}^{-1}$ also implies each diagonal element of $\mathbf{D}^H [\boldsymbol{\Sigma}_{\mathbf{y}_V}^{-1} - \boldsymbol{\Sigma}^{-1}] \mathbf{D}$ is positive, so

$$\text{tr} [\mathbf{D}^H \boldsymbol{\Sigma}_{\mathbf{y}_V}^{-1} \mathbf{D}] \geq \text{tr} [\mathbf{D}^H \boldsymbol{\Sigma}^{-1} \mathbf{D}]. \quad (3.49)$$

We can use this to bound the right-most term in (A.1):

$$\begin{aligned} \text{tr} [\boldsymbol{\Sigma}_{\mathbf{y}_V}^{-1} \mathbf{A}^H \boldsymbol{\Sigma}_{\mathbf{y}_V}^{-1} \mathbf{A}] &= \text{tr} \left[\left(\mathbf{A} \boldsymbol{\Sigma}_{\mathbf{y}_V}^{-1/2} \right)^H \boldsymbol{\Sigma}_{\mathbf{y}_V}^{-1} \mathbf{A} \boldsymbol{\Sigma}_{\mathbf{y}_V}^{-1/2} \right] \\ &\geq \text{tr} \left[\left(\mathbf{A} \boldsymbol{\Sigma}_{\mathbf{y}_V}^{-1/2} \right)^H \boldsymbol{\Sigma}^{-1} \mathbf{A} \boldsymbol{\Sigma}_{\mathbf{y}_V}^{-1/2} \right] \\ &= \text{tr} \left[\left(\mathbf{A}^H \boldsymbol{\Sigma}^{-1/2} \right)^H \boldsymbol{\Sigma}_{\mathbf{y}_V}^{-1} \mathbf{A}^H \boldsymbol{\Sigma}^{-1/2} \right] \\ &\geq \text{tr} \left[\left(\mathbf{A}^H \boldsymbol{\Sigma}^{-1/2} \right)^H \boldsymbol{\Sigma}^{-1} \mathbf{A}^H \boldsymbol{\Sigma}^{-1/2} \right] \\ &= \text{tr} \left[(\mathbf{C}^H \boldsymbol{\Sigma}_{\mathbf{y}_V} \mathbf{C})^{-1} \mathbf{C}^H \mathbf{A}^H \mathbf{C} (\mathbf{C}^H \boldsymbol{\Sigma}_{\mathbf{y}_V} \mathbf{C})^{-1} \mathbf{C}^H \mathbf{A} \mathbf{C} \right], \end{aligned} \quad (3.50)$$

where each inequality follows from (A.3), and each equality from $\text{tr} [\mathbf{BC}] = \text{tr} [\mathbf{CB}]$. Combining (A.1), (3.44), (A.2) and (A.4), we have proved (3.42). Thus we conclude $\mathcal{I}_{\mathbf{A}}(\mathbf{u}) \leq \mathcal{I}_V(\mathbf{u})$. \diamond

3.3.2 Optimal Array Design: $\Sigma_{\mathbf{y}_V}$ Independent of Parameters

For any array constructed from V , the error covariance of any unbiased estimator of u_i is bounded below in the Cramér-Rao bound by $[\mathcal{I}_{\mathbf{A}}^{-1}(\mathbf{u})]_{ii}$. This raises several natural questions: What is the minimum value of $[\mathcal{I}_{\mathbf{A}}^{-1}(\mathbf{u})]_{ii}$ over all possible M -port arrays constructed from V ? What is the minimum value of $\text{tr} [\mathcal{I}_{\mathbf{A}}^{-1}(\mathbf{u})]$? How do the answers to these questions vary with M ? Can we characterize the current matrices \mathbf{C} of arrays that achieve these minima?

In this chapter, we investigate these questions for a broad class of estimation problems in which either the mean $\boldsymbol{\mu}_{\mathbf{y}_V}$ or the covariance $\Sigma_{\mathbf{y}_V}$ of the incident field do not depend on \mathbf{u} . In this section, we first suppose $\Sigma_{\mathbf{y}_V}$ does not depend on \mathbf{u} . Note this includes as a special case the *conditional model* of direction-of-arrival estimation considered in [76], where the signal is considered to be nonrandom ($\Sigma_{\mathbf{s}_V} = \mathbf{0}$ and $\Sigma_{\mathbf{s}_A} = \mathbf{0}$). Suppose $\mathcal{I}_{\mathbf{A}}(\mathbf{u})$ is the FIM of an M -port array constructed from V with current matrix \mathbf{C} . As noted after Theorem 3.1, the diagonal elements of the CRB are bounded by the AFI,

$$[\mathcal{I}_{\mathbf{A}}^{-1}(\mathbf{u})]_{ii} \geq [\mathcal{I}_V^{-1}(\mathbf{u})]_{ii} , \quad (3.51)$$

with equality when each of the K basis functions is a port. However, the number ports in an array is usually far less than the number of basis functions needed to adequately describe the fields and currents on V . For a given \mathbf{u} and $M \ll K$, under what conditions on \mathbf{C} can we achieve equality in (3.51)? A partial answer is provided by the following theorem.

Theorem 3.2: Suppose $\Sigma_{\mathbf{y}_V}$ is non-singular and independent of \mathbf{u} , and define

$$\mathbf{Y} \triangleq \mathbf{Y}(\mathbf{u}) = \Sigma_{\mathbf{y}_V}^{-1/2} \frac{\partial \mu_{\mathbf{y}_V}}{\partial \mathbf{u}}. \quad (3.52)$$

If $M \geq \text{rank}[\mathbf{Y}]$ for $\mathbf{u} \in \mathbb{C}^p$, there exists an M -port current matrix, $\mathbf{C} \in \mathbb{C}^{K \times M}$, such that $\mathcal{I}_{\mathbf{A}}(\mathbf{u}) = \mathcal{I}_V(\mathbf{u})$. Equality is achieved by any current matrix \mathbf{C} that satisfies

$$\mathbf{U}\mathbf{U}^H\mathbf{Y} = \mathbf{Y}. \quad (3.53)$$

where $\mathbf{U} \triangleq \Sigma_{\mathbf{y}_V}^{1/2} \mathbf{C} (\mathbf{C}^H \Sigma_{\mathbf{y}_V} \mathbf{C})^{-1/2}$. In particular, equality is achieved for $\mathbf{C} = \Sigma_{\mathbf{y}_V}^{-1/2} \mathbf{V}$, where \mathbf{V} is any matrix whose columns are an orthonormal basis for the columns of \mathbf{Y} . \diamond

Before proving the theorem, we make several observations: (a) Since $\mathbf{Y}(\mathbf{u})$ is a $K \times p$ matrix, it follows $\text{rank}[\mathbf{Y}] \leq p$ for all \mathbf{u} . Thus, we can achieve equality in (3.51) for any \mathbf{u} whenever the number of ports is greater than or equal to the number of parameters to be estimated. (b) Note equality in Theorem 2 is guaranteed *only for one specific* \mathbf{u} ; Theorem 2 does not guarantee existence of a \mathbf{C} that achieves equality for all \mathbf{u} simultaneously. For example, if \mathbf{u} represents a direction-of-arrival, Theorem 2 provides conditions under which we can achieve the best possible sensitivity in one particular direction, \mathbf{u} , but we cannot necessarily optimize sensitivity simultaneously in all directions. (c) We can optimize sensitivity in multiple directions by choosing the columns of \mathbf{U} to satisfy (3.53) for multiple values of \mathbf{u} . For example, to optimize performance in two directions, say \mathbf{u}_1 and \mathbf{u}_2 , we can choose the columns of \mathbf{U} to be a basis for the columns of both $\mathbf{Y}(\mathbf{u}_1)$ and $\mathbf{Y}(\mathbf{u}_2)$. (d) When the observation noise is standard thermal noise, the observation covariance $\Sigma_{\mathbf{y}_V}$ is proportional to the resistance matrix, say $\Sigma_{\mathbf{y}_V} = \alpha \mathbf{R}$, where $\mathbf{R} = (1/2)[\mathbf{Z} + \mathbf{Z}^H]$. In this case, the current matrix $\mathbf{C} = \Sigma_{\mathbf{y}_V}^{-1/2} \mathbf{V}$ has a simple, intuitive

interpretation. Here $\mathbf{C}^H \boldsymbol{\Sigma}_{\mathbf{y}_V} \mathbf{C} = \alpha \mathbf{I}$. Recalling that the columns of \mathbf{C} represent the currents generated by each port, this condition implies that the ports are uncoupled and the array observation noise is spatially white, $\boldsymbol{\Sigma}_{\mathbf{y}_A} = \alpha \mathbf{I}$.

Proof of Theorem 2: If $\boldsymbol{\Sigma}_{\mathbf{y}_V}$ is independent of \mathbf{u} , then $\frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}_V}}{\partial u_i} = \mathbf{0}$ for each i . From (3.52), the AFI (3.39) can be written as

$$\mathcal{I}_V(\mathbf{u}) = 2\text{Re} \left[\frac{\partial \boldsymbol{\mu}_{\mathbf{y}_V}^H}{\partial \mathbf{u}} \boldsymbol{\Sigma}_{\mathbf{y}_V}^{-1} \frac{\partial \boldsymbol{\mu}_{\mathbf{y}_V}}{\partial \mathbf{u}} \right] = 2\text{Re} [\mathbf{Y}^H \mathbf{Y}] , \quad (3.54)$$

where \mathbf{U} is defined in (3.53). Similarly, the array Fisher information can be written as

$$\begin{aligned} \mathcal{I}_A(\mathbf{u}) &= 2\text{Re} \left[\frac{\partial \boldsymbol{\mu}_{\mathbf{y}_A}^H}{\partial \mathbf{u}} \boldsymbol{\Sigma}_{\mathbf{y}_A}^{-1} \frac{\partial \boldsymbol{\mu}_{\mathbf{y}_A}}{\partial \mathbf{u}} \right] \\ &= 2\text{Re} \left[\frac{\partial \boldsymbol{\mu}_{\mathbf{y}_V}^H}{\partial \mathbf{u}} \mathbf{C}^H (\mathbf{C} \boldsymbol{\Sigma}_{\mathbf{y}_V} \mathbf{C}^H)^{-1} \mathbf{C} \frac{\partial \boldsymbol{\mu}_{\mathbf{y}_V}}{\partial \mathbf{u}} \right] = 2\text{Re} [\mathbf{Y}^H \mathbf{U} \mathbf{U}^H \mathbf{Y}] . \end{aligned} \quad (3.55)$$

Note $\mathbf{U} \in \mathbb{C}^{K \times M}$ satisfies $\mathbf{U}^H \mathbf{U} = \mathbf{I}_M$, and so has orthonormal columns. From (3.55), it is clear that $\mathbf{U} \mathbf{U}^H \mathbf{Y} = \mathbf{Y}$ implies $\mathcal{I}_A(\mathbf{u}) = \mathcal{I}_V(\mathbf{u})$. Moreover, (3.53) is achieved by any \mathbf{U} whose columns are an orthonormal basis for the columns of \mathbf{Y} , which is possible for all $M \geq \text{rank}[\mathbf{Y}]$. Finally, it is easy to verify $\mathbf{C} = \boldsymbol{\Sigma}_{\mathbf{y}_V}^{-1/2} \mathbf{V}$ satisfies (3.53) whenever the columns of \mathbf{V} are an orthonormal basis for the columns of \mathbf{Y} . \diamond

3.3.3 Optimal Array Design: $\boldsymbol{\mu}_{\mathbf{y}_V}$ Independent of Parameters

We now reconsider the questions raised in the last section for estimation problems in which the mean of the incident field, $\boldsymbol{\mu}_{\mathbf{y}_V}$, does not depend on \mathbf{u} . For example, this includes as a special case the *unconditional model* of direction-of-arrival estimation considered in [76], where the signal is considered to be random.

Suppose $\mathcal{I}_{\mathbf{A}}(\mathbf{u})$ is the FIM of an M -port array with current matrix \mathbf{C} constructed from V . As noted earlier, the diagonal entries of the CRB are bounded below by the aperture FI

$$[\mathcal{I}_{\mathbf{A}}^{-1}(\mathbf{u})]_{ii} \geq [\mathcal{I}_V^{-1}(\mathbf{u})]_{ii} , \quad (3.56)$$

with equality when each of the K basis functions is a port. For a given \mathbf{u} and $M \ll K$, what is the minimum value of $[\mathcal{I}_{\mathbf{A}}^{-1}(\mathbf{u})]_{ii}$ and what current matrices \mathbf{C} achieve it? A partial answer is provided by the following theorem.

Theorem 3.3: Suppose $\Sigma_{\mathbf{y}_V}$ is non-singular and $\boldsymbol{\mu}_{\mathbf{y}_V}$ is independent of \mathbf{u} . For any $\mathbf{u} \in \mathbb{C}^p$, let

$$\mathbf{A}_i \triangleq \mathbf{A}_i(\mathbf{u}) = \Sigma_{\mathbf{y}_V}^{-1/2} \frac{\partial \Sigma_{\mathbf{y}_V}}{\partial \mathbf{u}} \Sigma_{\mathbf{y}_V}^{-1/2} , \quad i = 1, \dots, p . \quad (3.57)$$

Let $\hat{\mathbf{A}}_i$ denote the orthogonal projection of \mathbf{A}_i onto the subspace spanned by the other matrices in the inner product $\langle \mathbf{A}, \mathbf{B} \rangle \triangleq \text{tr}[\mathbf{A}^H \mathbf{B}]$. For any array with $M \leq K$ ports, we have

$$[\mathcal{I}_{\mathbf{A}}^{-1}(\mathbf{u})]_{ii} \geq (\sigma_{1,i}^2 + \dots + \sigma_{M,i}^2)^{-1} , \quad (3.58)$$

where $\sigma_{1,i} \geq \sigma_{2,i} \geq \dots \geq \sigma_{K,i}$ are the singular values of the Hermitian matrix $\mathbf{A}_i - \hat{\mathbf{A}}_i$ in decreasing order. Let $\mathbf{V} \in \mathbb{C}^{K \times M}$ be the matrix with columns equal to the M orthonormal singular vectors which correspond to the singular values in (3.58). Then equality holds in (3.58) for any $\mathbf{C} \in \mathbb{C}^{K \times M}$ such that $\mathbf{U} \triangleq \Sigma_{\mathbf{y}_V}^{1/2} \mathbf{C} (\mathbf{C}^H \Sigma_{\mathbf{y}_V} \mathbf{C})^{-1/2}$ satisfies

$$\mathbf{U} \mathbf{U}^H \mathbf{V} = \mathbf{V} .$$

In particular, the current matrix $\mathbf{C} = \boldsymbol{\Sigma}_{\mathbf{yV}}^{-1/2} \mathbf{V}$ satisfies this condition. \diamond

Before proving the theorem, let us make several observations: (a) An immediate consequence of Theorem 3 for $M = K$ is

$$[\mathcal{I}_V^{-1}(\mathbf{u})]_{ii} = \|\mathbf{A}_i - \hat{\mathbf{A}}_i\|^{-2}, \quad (3.59)$$

where $\|\mathbf{A}\|^2 \triangleq \langle \mathbf{A}, \mathbf{A} \rangle$ is the Frobenius norm. It follows we can achieve equality in (3.56) whenever $M \geq \text{rank}[\mathbf{A}_i - \hat{\mathbf{A}}_i]$. (b) The conditions for minimizing $[\mathcal{I}_V^{-1}(\mathbf{u})]_{ii}$ apply to one specific \mathbf{u} and i ; the theorem does not guarantee the existence of a \mathbf{C} that achieves equality simultaneously for all \mathbf{u} and i . In general, there will be a tradeoff between minimizing $[\mathcal{I}_V^{-1}(\mathbf{u})]_{ii}$ for different values of \mathbf{u} and i . We conjecture that the best tradeoff between the error covariances $[\mathcal{I}_V^{-1}(\mathbf{u})]_{ii}$ and $[\mathcal{I}_V^{-1}(\mathbf{u})]_{jj}$ is related to the solutions of the generalized eigenvalue problem

$$[\mathbf{A}_i - \hat{\mathbf{A}}_i] \mathbf{x} = \lambda [\mathbf{A}_j - \hat{\mathbf{A}}_j] \mathbf{x}. \quad (3.60)$$

Proof of Theorem 3: If $\boldsymbol{\mu}_{\mathbf{yV}}$ does not depend on \mathbf{u} , then $\frac{\partial \boldsymbol{\mu}_{\mathbf{yV}}}{\partial \mathbf{u}} = \mathbf{0}$ and the AFI (3.39) is

$$\begin{aligned} [\mathcal{I}_V(\mathbf{u})]_{ij} &= \text{tr} \left[\boldsymbol{\Sigma}_{\mathbf{yV}}^{-1} \frac{\partial \boldsymbol{\Sigma}_{\mathbf{yV}}}{\partial u_i} \boldsymbol{\Sigma}_{\mathbf{yV}}^{-1} \frac{\partial \boldsymbol{\Sigma}_{\mathbf{yV}}}{\partial u_j} \right] = \text{tr} \left[\boldsymbol{\Sigma}_{\mathbf{yV}}^{-1/2} \frac{\partial \boldsymbol{\Sigma}_{\mathbf{yV}}}{\partial u_i} \boldsymbol{\Sigma}_{\mathbf{yV}}^{-1} \frac{\partial \boldsymbol{\Sigma}_{\mathbf{yV}}}{\partial u_j} \boldsymbol{\Sigma}_{\mathbf{yV}}^{-1/2} \right] \\ &= \text{tr} [\mathbf{A}_i \mathbf{A}_j] = \langle \mathbf{A}_i, \mathbf{A}_j \rangle, \end{aligned} \quad (3.61)$$

where \mathbf{A}_i is given in (3.57) and $\langle \mathbf{A}, \mathbf{B} \rangle = \text{tr} [\mathbf{A}^H \mathbf{B}]$ is the inner product on $\mathbb{C}^{K \times K}$. Here the second step follows from $\text{tr}[\mathbf{AB}] = \text{tr}[\mathbf{BA}]$, and the last by observing \mathbf{A}_i is Hermitian.

Proceeding in a similar way, we have

$$\begin{aligned}
[\mathcal{I}_{\mathbf{A}}(\mathbf{u})]_{ij} &= \text{tr} \left[\mathbf{C} (\mathbf{C}^H \boldsymbol{\Sigma}_{\mathbf{y}\mathbf{v}} \mathbf{C})^{-1} \mathbf{C}^H \frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}\mathbf{v}}}{\partial u_i} \mathbf{C} (\mathbf{C}^H \boldsymbol{\Sigma}_{\mathbf{y}\mathbf{v}} \mathbf{C})^{-1} \mathbf{C}^H \frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}\mathbf{v}}}{\partial u_j} \right] \\
&= \text{tr} \left[(\mathbf{C}^H \boldsymbol{\Sigma}_{\mathbf{y}\mathbf{v}} \mathbf{C})^{-1/2} \mathbf{C}^H \frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}\mathbf{v}}}{\partial u_i} \mathbf{C} (\mathbf{C}^H \boldsymbol{\Sigma}_{\mathbf{y}\mathbf{v}} \mathbf{C})^{-1} \mathbf{C}^H \frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}\mathbf{v}}}{\partial u_j} \mathbf{C} (\mathbf{C}^H \boldsymbol{\Sigma}_{\mathbf{y}\mathbf{v}} \mathbf{C})^{-1/2} \right] \\
&= \text{tr} [\mathbf{U}^H \mathbf{A}_i^H \mathbf{U} \mathbf{U}^H \mathbf{A}_j \mathbf{U}] \quad , \tag{3.62}
\end{aligned}$$

where \mathbf{U} is defined in the theorem. Note $\mathbf{U}^H \mathbf{U} = \mathbf{I}$, so the rows of \mathbf{U} are orthonormal. Further note $\mathbf{P} \triangleq \mathbf{U} \mathbf{U}^H$ is Hermitian and idempotent, $\mathbf{P}^2 = \mathbf{P}$, and so is a projection matrix of rank M . Using the trace identity, we can write

$$[\mathcal{I}_{\mathbf{A}}(\mathbf{u})]_{ij} = \text{tr} [\mathbf{A}_i^H \mathbf{U} \mathbf{U}^H \mathbf{A}_j \mathbf{U} \mathbf{U}^H] = \text{tr} [\mathbf{A}_i^H \mathbf{P}^2 \mathbf{A}_j \mathbf{P}^2] = \langle \mathbf{P} \mathbf{A}_i \mathbf{P}, \mathbf{P} \mathbf{A}_j \mathbf{P} \rangle \tag{3.63}$$

It suffices to prove (3.58) for $i = 1$. To calculate $[\mathcal{I}_{\mathbf{V}}^{-1}(\mathbf{u})]_{11}$, partition the AFI into block form:

$$\mathcal{I}_{\mathbf{V}}(\mathbf{u}) = \begin{bmatrix} V_{11} & \mathbf{V}_{21}^H \\ \mathbf{V}_{21} & \mathbf{V}_{22} \end{bmatrix} \quad , \tag{3.64}$$

where $V_{11} \in \mathbb{C}$, $\mathbf{V}_{21} \in \mathbb{C}^{(p-1) \times 1}$, $\mathbf{V}_{22} \in \mathbb{C}^{(p-1) \times (p-1)}$. From the matrix inversion lemma [86, pg. 18], we have

$$[\mathcal{I}_{\mathbf{V}}^{-1}(\mathbf{u})]_{11} = (V_{11} - \mathbf{V}_{21}^H \mathbf{V}_{22}^{-1} \mathbf{V}_{21})^{-1} = (V_{11} - \mathbf{w}^H \mathbf{V}_{22} \mathbf{w})^{-1} \tag{3.65}$$

where $\mathbf{w} \triangleq (w_2, \dots, w_p) = \mathbf{V}_{22}^{-1} \mathbf{V}_{21}$. Since $\mathbf{V}_{22} \mathbf{w} = \mathbf{V}_{21}$, (3.61) implies

$$\sum_{j=2}^p \langle \mathbf{A}_j, \mathbf{A}_i \rangle w_j = \langle \mathbf{A}_1, \mathbf{A}_i \rangle \quad , \quad i = 2, \dots, p \quad . \tag{3.66}$$

These can be recognized as the *normal equations* [90, Thrm. 7.4.6] that define the projection of the matrix \mathbf{A}_1 onto the subspace \mathcal{L}_1 of $\mathbb{C}^{K \times K}$ spanned by $\mathbf{A}_2, \dots, \mathbf{A}_p$. In particular,

$$\hat{\mathbf{A}}_1 \triangleq \sum_{j=2}^p w_j \mathbf{A}_j$$

is the orthogonal projection of \mathbf{A}_1 onto \mathcal{L}_1 and

$$\langle \mathbf{A}_1 - \hat{\mathbf{A}}_1, \mathbf{A}_j \rangle = 0, \quad i = 2, \dots, p. \quad (3.67)$$

From these last two equations, it follows $\mathbf{w}^H \mathbf{V}_{22} \mathbf{w} = \|\hat{\mathbf{A}}_1\|^2$ and

$$[\mathcal{I}_V^{-1}(\mathbf{u})]_{11} = \|\mathbf{A}_1 - \hat{\mathbf{A}}_1\|^{-2} \quad (3.68)$$

where $\|\mathbf{A}\|^2 = \langle \mathbf{A}, \mathbf{A} \rangle$ is the Frobenius norm. When $\mathbf{A}_2, \dots, \mathbf{A}_p$ are linearly independent, $\hat{\mathbf{A}}_1$ can also be expressed directly in terms of \mathbf{A}_1 in vectorized form as

$$\text{vec}[\hat{\mathbf{A}}_1] = \mathbf{B} (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H \text{vec}[\mathbf{A}_1], \quad (3.69)$$

where $\text{vec}[\mathbf{A}_1]$ is the vectorization operator [33] that stacks the columns of \mathbf{A}_1 into a single column vector of dimension K^2 and $\mathbf{B} \in \mathbb{C}^{K^2 \times p}$ is the matrix whose j -th column is $\text{vec}[\mathbf{A}_j]$.

Proceeding in the same way, we can show

$$[\mathcal{I}_A^{-1}(\mathbf{u})]_{11} = \|\mathbf{P} \mathbf{A}_1 \mathbf{P} - \hat{\mathbf{A}}'_1\|^{-2} \quad (3.70)$$

where $\hat{\mathbf{A}}'_1$ is the projection of $\mathbf{P} \mathbf{A}_1 \mathbf{P}$ onto the subspace \mathcal{L}_2 spanned by $\mathbf{P} \mathbf{A}_2 \mathbf{P}, \dots, \mathbf{P} \mathbf{A}_p \mathbf{P}$. Since the basis vectors of \mathcal{L}_2 are projections of those of \mathcal{L}_1 , it follows $\mathcal{L}_2 \subseteq \mathcal{L}_1$. Hence,

$\hat{\mathbf{A}}'_1$ can be calculated by first projecting \mathbf{A}_1 onto \mathcal{L}_1 to obtain $\hat{\mathbf{A}}_1$, and then projecting $\hat{\mathbf{A}}_1$ onto \mathcal{L}_2 , which yields $\hat{\mathbf{A}}'_1 = \mathbf{P}\hat{\mathbf{A}}_1\mathbf{P}$. It follows that

$$[\mathcal{I}_{\mathbf{A}}^{-1}(\mathbf{u})]_{11} = \|\mathbf{P}(\mathbf{A}_1 - \hat{\mathbf{A}}_1)\mathbf{P}\|^{-2} . \quad (3.71)$$

Recalling that $\mathbf{P}^2 = \mathbf{P} \leq \mathbf{I}$, we can write

$$\begin{aligned} \|\mathbf{P}(\mathbf{A}_1 - \hat{\mathbf{A}}_1)\mathbf{P}\|^2 &= \text{tr} \left[\mathbf{U}^H(\mathbf{A}_1 - \hat{\mathbf{A}}_1)^H \mathbf{U} \mathbf{U}^H (\mathbf{A}_1 - \hat{\mathbf{A}}_1) \mathbf{U} \right] \\ &\leq \text{tr} \left[\mathbf{U}^H(\mathbf{A}_1 - \hat{\mathbf{A}}_1)^H (\mathbf{A}_1 - \hat{\mathbf{A}}_1) \mathbf{U} \right] . \end{aligned} \quad (3.72)$$

The last trace equals the sum of the M eigenvalues of $\mathbf{U}^H(\mathbf{A}_1 - \hat{\mathbf{A}}_1)^H (\mathbf{A}_1 - \hat{\mathbf{A}}_1) \mathbf{U}$. From the Poincaré Separation Theorem [86, Cor. 4.3.16], these eigenvalues are less than or equal to the M largest eigenvalues of $(\mathbf{A}_1 - \hat{\mathbf{A}}_1)^H (\mathbf{A}_1 - \hat{\mathbf{A}}_1)$, or equivalently

$$\|\mathbf{P}(\mathbf{A}_1 - \hat{\mathbf{A}}_1)\mathbf{P}\|^2 \leq \sigma_{1,1}^2 + \cdots + \sigma_{M,1}^2 \quad (3.73)$$

where $\sigma_{1,1} \geq \sigma_{2,1} \geq \cdots \geq \sigma_{M,1}$ are the first M singular values of $\mathbf{A}_1 - \hat{\mathbf{A}}_1$. From (3.72), we see that equality is achieved when the columns of \mathbf{U} are the corresponding unit singular vectors. Thus, we conclude that

$$[\mathcal{I}_{\mathbf{A}}^{-1}(\mathbf{u})]_{11} \geq (\sigma_{1,1}^2 + \cdots + \sigma_{1,M}^2)^{-1} , \quad (3.74)$$

thereby proving the theorem. ◇

3.4 Numerical Examples

In this section, we illustrate our results for three different estimation problems. For each problem, we consider the three arrays depicted in Fig. 3.3, which are formed from the same underlying conductor V . Here V is a copper wire of conductivity $\sigma = 5.96 \times 10^7$ S/m with length $L = 0.5$ m and radius $r = 0.5$ mm operating at a frequency of $f = 300$ MHz. The three arrays differ only in the number and locations of ports: $M = 1$ is a dipole with one port in the center; $M = 2$ consists of two ports located 0.125m from each end; $M = 3$ contains all three ports. For this V , the impedance \mathbf{Z} and radiation and loss resistances, $\mathbf{R}_{\mathbf{E}}$ and $\mathbf{R}_{\mathbf{I}}$, were obtained numerically for $K = 99$ triangular basis functions using FEKO [81], a method-of-moments electromagnetic solver. As in [91], we model each port as a “delta-gap” – a gap so small it does not alter the impedance of V . With this assumption, we can calculate the coupling matrix \mathbf{C} of each array in Fig. 3.3 via $\mathbf{C}^H = \mathbf{Y}_{rc}^{-1} \mathbf{Y}_c^T$. Here $\mathbf{Y} = \mathbf{Z}^{-1}$ is the admittance matrix of V , \mathbf{Y}_c is the submatrix of \mathbf{Y} containing only columns corresponding to port locations, and \mathbf{Y}_{rc} is a submatrix of \mathbf{Y}_c containing only rows corresponding to port locations. We further assume that each array is connected to a bank of M amplifiers with noise factor F_{min} using the optimal MNF matching network (3.19).

In the sections that follow, each estimation example can be specified in terms of the incident signal and noise field models and the parameters to be estimated.

3.4.1 Estimating the Temperature of Black-Body Radiation

First consider the example of an array surrounded by a black-body enclosure, where the goal is to estimate the temperature of the enclosure. Let $\mathbf{R}_{\mathbf{E}}$ and $\mathbf{R}_{\mathbf{I}}$ denote the radiation and loss resistances of V , respectively. The “signal” component of the incident

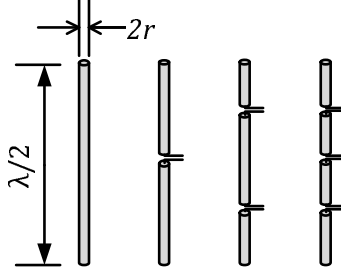


Figure 3.3: Three arrays formed from the wire V .

field (3.30) is induced by the enclosure, and is distributed as $\mathbf{s}_V \sim \mathcal{CN}(\mathbf{0}, 4k_b T_E B \mathbf{R}_E)$, where k_b is Boltzmann's constant, B is the bandwidth in Hz, and T_E is the temperature of the enclosure. There is no external noise, so the noise \mathbf{n}_V in (3.36) is due entirely to the loss resistance and amplifiers, and so

$$\Sigma_{\mathbf{n}_V} = 4k_b T_0 B \mathbf{R}_I + (F_{min} - 1) 4k_b T_0 B \mathbf{R}, \quad (3.75)$$

where T_0 is the temperature of V . If the parameter to be estimated is $u \triangleq 1 + (T_E - T_0)/(T_0 F_{min})$, the aperture observation (3.35) has mean $\boldsymbol{\mu}_{\mathbf{y}_V} = \mathbf{0}$ and covariance $\Sigma_{\mathbf{y}_V} = 4k_b T_0 F_{min} B (u \mathbf{R}_E + \mathbf{R}_I)$.

Since $\boldsymbol{\mu}_{\mathbf{y}_V}$ is independent of u , the results of Theorem 3 apply: For any M -port array formed from V , the error covariance of any unbiased estimator of u is bounded by the scalar CRB (3.58)

$$\mathcal{I}_{\mathbf{A}}^{-1}(\mathbf{u}) \geq (\sigma_{1,1}^2 + \cdots + \sigma_{M,1}^2)^{-1}, \quad (3.76)$$

where $\sigma_{1,1} \geq \sigma_{2,1} \geq \dots \geq \sigma_{K,1}$ are the ordered singular values of

$$\mathbf{A}_1 - \hat{\mathbf{A}}_1 = \mathbf{A}_1 = (u\mathbf{R}_E + \mathbf{R}_I)^{-1/2} \mathbf{R}_E (u\mathbf{R}_E + \mathbf{R}_I)^{-1/2}. \quad (3.77)$$

Here $\mathbf{A}_1, \hat{\mathbf{A}}_1$ are defined in Theorem 3, and $\hat{\mathbf{A}}_1$ vanishes because there is only one parameter. In particular, when $M = K$ (3.76) reduces to the aperture CRB (3.59).

This bound has a simple physical interpretation in terms of the efficiencies of certain eigencurrents on V . Consider the generalized eigenvalue problem

$$\eta [\mathbf{R}_E + \mathbf{R}_I] \mathbf{J} = \mathbf{R}_E \mathbf{J}. \quad (3.78)$$

If $\mathbf{R}_E + \mathbf{R}_I$ is positive definite, there exist K eigencurrents of this equation such that

$$\mathbf{J}_i^H [\mathbf{R}_E + \mathbf{R}_I] \mathbf{J}_k = \delta_{ik}, \quad \mathbf{J}_i^H \mathbf{R}_E \mathbf{J}_k = \eta_k \delta_{ik}. \quad (3.79)$$

Intuitively, $\mathbf{J}_1, \dots, \mathbf{J}_K$ are unit-power eigencurrents that jointly diagonalize the radiation and loss resistance matrices. Note, the eigenvalue η_k can be written as

$$\eta_k = \frac{\mathbf{J}_k^H \mathbf{R}_E \mathbf{J}_k}{\mathbf{J}_k^H \mathbf{R}_E \mathbf{J}_k + \mathbf{J}_k^H \mathbf{R}_I \mathbf{J}_k}. \quad (3.80)$$

If we drive V with current \mathbf{J}_k , note $\mathbf{J}_k^H \mathbf{R}_E \mathbf{J}_k$ is the power radiated and $\mathbf{J}_k^H \mathbf{R}_I \mathbf{J}_k$ is the power dissipated in internal losses. Thus the eigenvalue $0 \leq \eta_k \leq 1$ represents the fraction of power that is radiated, which can be recognized as the *radiation efficiency* of eigencurrent \mathbf{J}_k [92, pg. 872]. From (3.79), we can show \mathbf{J}_k is also an eigenvector of $u\mathbf{R}_E + \mathbf{R}_I$

with corresponding eigenvalue $u\eta_k + 1 - \eta_k$, and an eigenvector of \mathbf{A}_1 with eigenvalue

$$\sigma_{k,1} = \frac{\eta_k}{u\eta_k + 1 - \eta_k}. \quad (3.81)$$

Since \mathbf{A}_1 is semi-definite and Hermitian, the singular values coincide with the eigenvalues. For any M -port array formed from V , we conclude (3.76) can also be expressed as

$$\mathcal{I}_{\mathbf{A}}^{-1}(\mathbf{u}) \geq \left[\sum_{i=1}^M \left(\frac{\eta_k}{u\eta_k + 1 - \eta_k} \right)^2 \right]^{-1}, \quad (3.82)$$

where η_1, \dots, η_M are the radiation efficiencies of the M most-efficient eigencurrents of (3.78).

This example illustrates that our results can lead to trivial results for certain estimation problems involving perfect conductors ($\sigma = +\infty$). To see this, note $\mathbf{R}_{\mathbf{I}}$ vanishes when V is a perfect conductor. If $\mathbf{R}_{\mathbf{E}}$ has full rank, then every \mathbf{J} is a solution of (3.78) with $\eta = 1$, so $\eta_1 = \dots = \eta_K = 1$. The aperture CRB is given by (3.76) with $M = K$, which yields

$$\mathcal{I}_V^{-1}(u) \triangleq \frac{u^2}{K}, \quad (3.83)$$

where K is the dimension of $\mathbf{R}_{\mathbf{E}}$. Here the AFI grows with the number of basis functions in (3.23) for any V , invalidating our assumption that the basis (3.23) captures all of the useful information contained in the continuous fields and currents on V . For this problem, the aperture CRB based on the continuous fields is zero, a trivial result. A similar phenomenon was observed for a different information measure in [79], which showed that MIMO aperture capacity can be infinite for lossless conductors.

Now consider an example with finite conductivity: the three arrays in Fig. 3.3 are

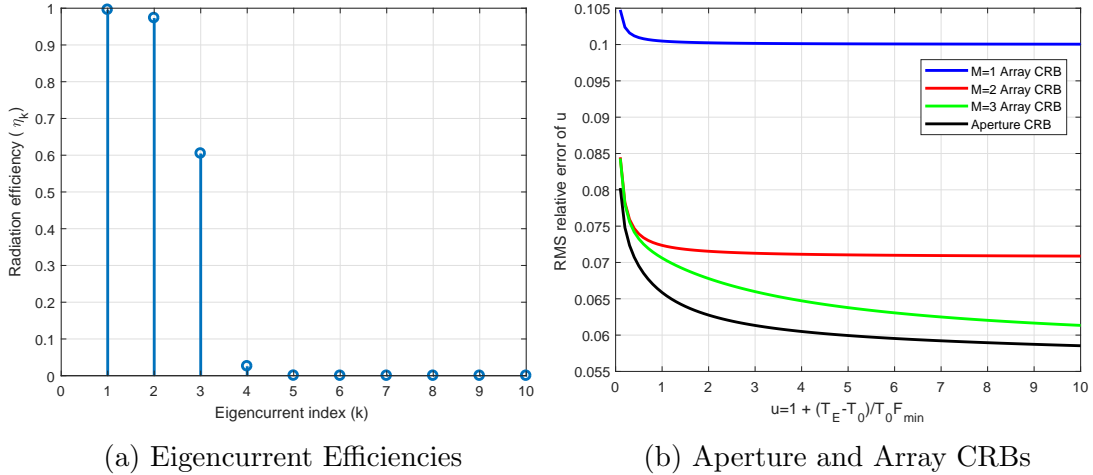


Figure 3.4: Eigencurrent efficiencies and CRBs for temperature estimation example.

formed from a copper wire of conductivity $\sigma = 5.96 \times 10^7$ S/m. The eigencurrent efficiencies η_k of this wire were calculated by solving (3.79) in MATLAB and are plotted in Fig. 3.4a. Note the efficiencies fall off rapidly with k , and all but the first three eigencurrents are quite lossy. Thus, in this example, introducing a small loss reduces the number of efficient eigencurrents from $K = 99$ to three. In Fig. 3.4b, we plot the ratio of square-root of the aperture CRB to u , i.e. $u^{-1}\mathcal{I}_V^{-1/2}(u)$, which provides a lower bound on the root-mean-square (RMS) relative estimation error of u . For comparison, we also plot the array CRBs for the three arrays depicted in Fig. 3.3. Note the CRB of the $M = 1$ dipole is 31-70% larger than the aperture CRB, so the dipole fails to capture much of the useful information about u that is contained in the fields and currents on V . Similarly, the CRB for two ports is 6-20% above the aperture bound, and the CRB for three ports is 3-8% above. Here we see that additional ports can significantly reduce the CRB, and $M = 3$ ports performs close to the aperture CRB. This is intuitively reasonable, since Fig. 3.4a shows most of the information is contained in the first three eigencurrents.

The examples above illustrate that even small internal losses can significantly increase aperture CRB. When $\sigma < +\infty$, the higher-order eigenmodes of Theorem 3 typically become increasingly lossy, and most of the Fisher information resides in the first few eigenmodes. These eigenmodes can usually be well modeled by a sufficiently fine MoM basis. For this reason, we will include internal losses in all of the remaining examples in this chapter.

3.4.2 Estimating Direction-of-arrival of a Known Signal

Suppose the wire V in the previous example is now used to estimate direction-of-arrival (DoA) of an incident plane wave. We consider a scenario similar to [84], where a far-field source transmits a signal with known polarization to the receiver. Let θ and ϕ denote spherical coordinates centered at the midpoint of the wire, as illustrated in Fig. 3.5. The incident signal field is a sequence of N plane waves arriving from direction θ, ϕ of the form

$$\begin{bmatrix} E_{\theta,t} \\ E_{\phi,t} \end{bmatrix} = x s_t \begin{bmatrix} \cos \gamma \\ \sin \gamma e^{j\psi} \end{bmatrix}, \quad t = 1, \dots, N, \quad (3.84)$$

where s_t is the signal, $x = x_r + jx_i$ is a complex path gain that encodes the signal amplitude and phase, and polarization is determined by the Johnson vector [85, eq. 8] on the right, where $\tan \gamma$ and ψ are the magnitude and angle of the polarization ratio $E_{\phi,t}/E_{\theta,t}$. Here we assume s_t, γ and ψ are known and we want to estimate x, θ, ϕ . Let $\mathbf{P}_i(\theta, \phi) = [p_{\theta,i}(\theta, \phi), p_{\phi,i}(\theta, \phi)]^T$ denote the radiation pattern of the i -th basis function in (3.23). The field \mathbf{s}_V in (3.30) induced on V at time t by the incident field (3.85) can now

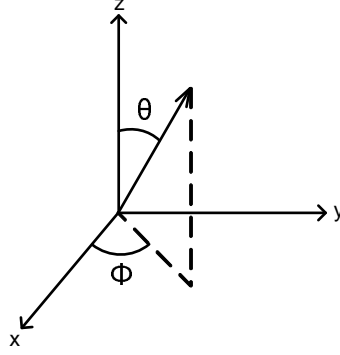


Figure 3.5: Spherical coordinates

be written as

$$\mathbf{s}_V(t) = x s_t \mathbf{g}(\theta, \phi), \quad t = 1, \dots, N, \quad (3.85)$$

where $\mathbf{g}(\theta, \phi) \in \mathbb{C}^{K \times 1}$ is defined by

$$g_i(\theta, \phi) \triangleq \mathbf{P}_i^T(\theta, \phi) \begin{bmatrix} \cos \gamma \\ \sin \gamma e^{j\psi} \end{bmatrix}.$$

We model the antenna noise in (3.30) by the Twiss model, so $\mathbf{n}_e(t) \sim \mathcal{CN}(\mathbf{0}, 4k_b T_0 B \mathbf{R})$, $t = 1, 2, \dots, N$. With these assumptions, the aperture observation model (3.37) has mean $\boldsymbol{\mu}_{\mathbf{y}_V(t)} = x s_t \mathbf{g}(\theta, \phi)$ and covariance $\boldsymbol{\Sigma}_{\mathbf{y}_V} = 4k_b T_0 F_{min} B \mathbf{R}$.

Assuming perfect time synchronization, we consider the problem of using the arrays in Fig. 3.3 to estimate the parameters $\mathbf{u} = [x_r, x_i, \theta, \phi]$. Since $\boldsymbol{\Sigma}_{\mathbf{y}_V}$ does not depend on \mathbf{u} , the aperture and array Fisher information are given by (3.54) and (3.55), respectively. As in (A.2) and (A.3) of [84], the sum of the AFI for all N symbols can be expressed in

block matrix form as

$$\mathcal{I}_V = \begin{bmatrix} \mathcal{I}_{V\mathbf{x}\mathbf{x}} & \mathcal{I}_{V\psi\mathbf{x}}^T \\ \mathcal{I}_{V\psi\mathbf{x}} & \mathcal{I}_{V\psi\psi} \end{bmatrix}, \quad (3.86)$$

where $\mathbf{x} = [x_r, x_i]$ and $\psi = [\theta, \phi]$. Assuming the observation noise is temporally white, each sub-matrix is given by

$$\mathcal{I}_{V\mathbf{x}\mathbf{x}} = 2PN\mathbf{g}(\theta, \phi)^H \Sigma_{\mathbf{y}\mathbf{v}}^{-1} \mathbf{g}(\theta, \phi) \mathbf{I}_2 \quad (3.87)$$

$$\mathcal{I}_{V\psi\mathbf{x}} = 2PN \operatorname{Re} \left\{ [1, i] \otimes x^* \frac{\partial \mathbf{g}(\theta, \phi)^H}{\partial \psi} \Sigma_{\mathbf{y}\mathbf{v}}^{-1} \mathbf{g}(\theta, \phi) \right\} \quad (3.88)$$

$$\mathcal{I}_{V\psi\psi} = 2PN|x|^2 \operatorname{Re} \left\{ \frac{\partial \mathbf{g}(\theta, \phi)^H}{\partial \psi} \Sigma_{\mathbf{y}\mathbf{v}}^{-1} \frac{\partial \mathbf{g}(\theta, \phi)}{\partial \psi} \right\} \quad (3.89)$$

where $\Sigma_{\mathbf{y}\mathbf{v}} = 4k_b T_0 F_{min} B \mathbf{R}$ and $P = (1/N) \sum_{t=1}^N |s(t)|^2$ is the average power of the signal.

First we consider a simple one-dimensional DoA estimation problem. Suppose a known sequence of $N = 100$ symbols s_1, \dots, s_N is transmitted with a path gain of $x = 1$ in a right-hand circularly-polarized wave, so $\gamma = \pi/4$ and $\psi = -\pi/2$ in (3.85). This signal is sensed using the arrays in Fig. 3.3, where each is centered at the origin of the coordinate system in Fig. 3.5 and oriented along the z -axis. Since each array is symmetric about the z -axis, the radiation pattern (3.85) does not depend on ϕ and the unknown parameters are $\mathbf{u} = [x_r, x_i, \theta]$. The Fisher information is calculated using (3.86) by setting $\psi = [\theta]$.

In Fig. 3.6, we plot $([\mathcal{I}_V^{-1}]_{\theta, \theta})^{1/2}$, the aperture bound on the RMS estimation error of θ , versus θ for a signal-to-noise ratio of $SNR \triangleq P/(4k_b T_0 F_{min} B) = 10$. Also shown are the corresponding bounds for the $M = 2$ and 3 arrays in Fig. 3.3. The $M = 1$ case is omitted, because the FIM is singular, implying that there is no meaningful lower bound

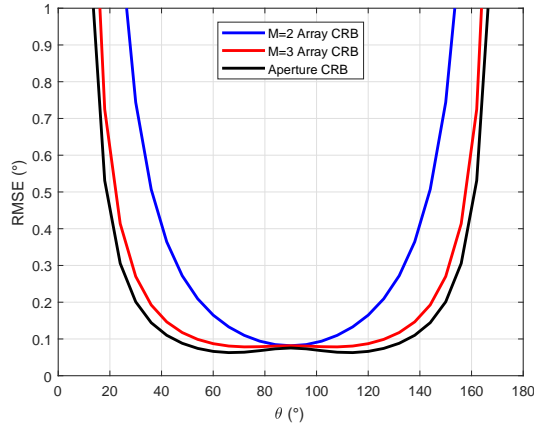


Figure 3.6: CRB on RMSE versus target angle θ using half-wavelength wire.

on the estimation error [80]. Note all of the CRBs are lowest near broadside ($\theta = 90^\circ$) and grow rapidly near endfire ($\theta = 0^\circ$ or 180°). As expected, the aperture CRB of V is a lower bound on the CRB of each array formed from V . Note the CRB of the $M = 2$ array coincides with the aperture bound near broadside, and so achieves the best possible performance of any array formed from the wire V ; however, the RMS error of this array is more than 3.7 times larger than the aperture bound for $\theta \leq 30^\circ$ and $\theta \geq 60^\circ$. Interestingly, the $M = 3$ array is relatively close to the aperture bound for nearly all θ , which suggests that most of the information contained in the fields and currents of V can be observed with only three ports. If we define the beamwidth of an array to be the range of θ for which the CRB performs within a factor of 2 of its best performance, then both the aperture bound and $M = 3$ array have beamwidths of roughly 100° , whereas the $M = 2$ array has a beamwidth of only about 60° . Thus, using three ports not only lowers the CRB, but also increases the angular range over which the CRB is close to its optimal value.

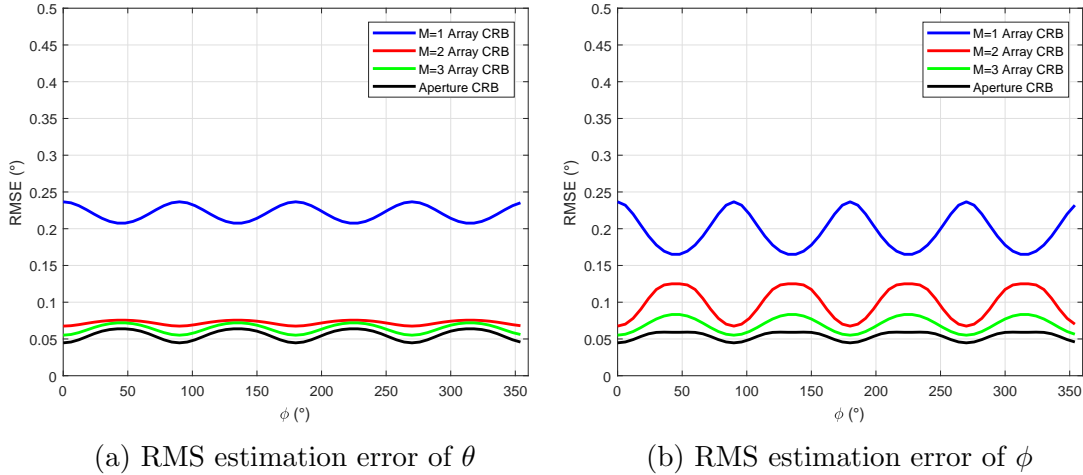


Figure 3.7: CRB on RMSE versus arrival angle ϕ for signal arriving in $\theta = 90^\circ$ plane.

3.4.3 Two-dimensional DoA Estimation using Tripoles

Next consider the problem of estimating both θ and ϕ using three, mutually-orthogonal wire antennas. A conventional tripole array [85] is formed from three $M = 1$ dipoles, each centered at the origin and aligned with the x , y , and z axes in Fig. 3.5. Here we call this an $M = 1$ tripole, although it actually has three ports. Similarly, the $M = 2$ and $M = 3$ tripoles are defined by replacing each dipole with the $M = 2$ and $M = 3$ arrays in Fig. 3.3, respectively. Thus, the $M = 2$ and $M = 3$ tripoles have six and nine ports, respectively. In this section, we take the incident signal to be the same $N = 100$, circularly-polarized wave considered in Fig. 3.6, except now we estimate both angles-of-arrival, θ and ϕ .

First suppose the incident signal arrives in the plane $\theta = 90^\circ$. We plot the aperture bound $([\mathcal{I}_V^{-1}]_{\theta,\theta})^{1/2}$ on the RMS estimation error of θ versus the signal angle-of-arrival ϕ in Fig. 3.7a for SNR= 10 dB. Also shown are the corresponding array CRBs for the $M = 1, 2$ and 3 tripoles. For all three arrays, note the RMS error of θ is relatively constant over the angle-of-arrival ϕ . Interestingly, the CRB of the conventional $M = 1$ tripole array is

3.3 to 5.3 times worse than the aperture CRB. However, the $M = 2$ and $M = 3$ arrays perform close to the aperture bound for all ϕ . Thus, in this example, adding extra ports to conventional tripoles can significantly improve the performance of DoA estimation. Similarly, we plot $([\mathcal{I}_V^{-1}]_{\phi,\phi})^{1/2}$ versus ϕ under the same conditions in Fig. 3.7b. Note the estimation error of ϕ varies more with angle-of-arrival. The conventional $M = 1$ tripole performs 2.8 to 5.3 times worse than the aperture bound for all ϕ . Adding a second port to each wire can significantly reduce the error, but the improvement varies widely with ϕ . We note the $M = 3$ array performs close to the aperture bound for all ϕ .

Finally, Figs. 3.8a and 3.8b plot the same quantities as Figs. 3.7a and 3.7b for a target signal arriving in the $\theta = 60^\circ$ conical plane. Once again, we find the conventional $M = 1$ tripole lies significantly above the aperture bound, and fails to capture all of the useful information in V . Adding an extra port can significantly reduce the error, but the improvement depends on the arrival angle ϕ . Once again, the $M = 3$ tripole performs close to the aperture bound for all ϕ .

All of the results above suggest that tripole antennas formed from conventional centered dipoles fail to capture all of the useful information contained in the fields and currents on these wires. Performance can often be significantly improved by introducing additional ports on each of the dipoles. Our results further suggest that array and aperture CRBs can be used as design criteria to guide the choice of the number and placement of ports for any conductor V . Although we have considered only simple wire antennas in this section, the concepts presented here are general and can be applied to other conductors of interest, such as patch antennas.

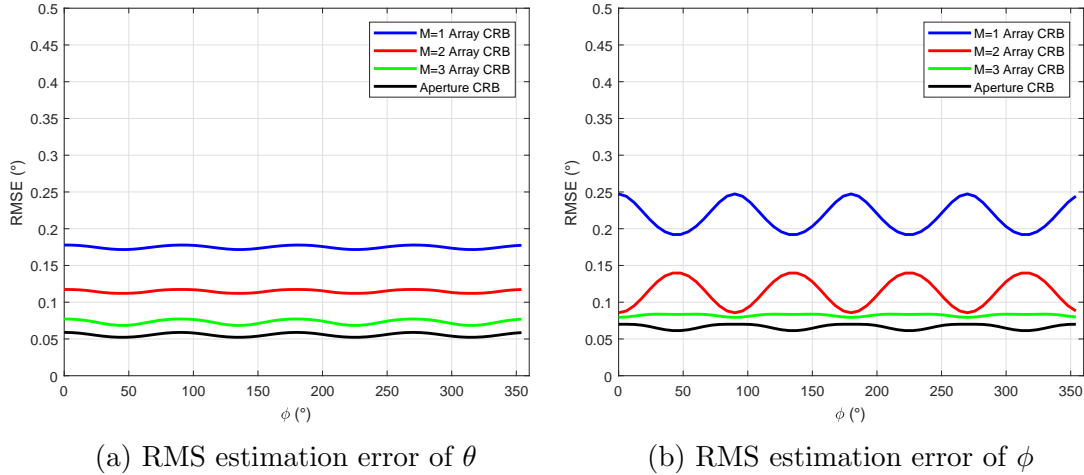


Figure 3.8: CRB on RMSE versus arrival angle ϕ for signal arriving in $\theta = 60^\circ$ plane.

3.5 Conclusion

In this chapter, we have considered antenna arrays built from an arbitrary conductor V , and showed how array properties are changed by V and antenna ports. For Gaussian incident electromagnetic field, we derived the Fisher information for any array based on physical models of antenna and amplifier noise. We further proved this Fisher information is always bounded above by the *aperture Fisher information*, which represents all the information contained in the currents of V . For general estimation problems, we showed how the currents on V can be decomposed into a series of unit-power eigencurrents, and derived what antenna ports can observe these eigencurrents and capture the most Fisher information. We then considered two applications, radiation temperature estimation and direction-of-arrival estimation, using several arrays built from wire structures. For both applications, our results suggest that conventional dipole and tripole antennas do not take full use of their underlying structures, and the system performance can be significantly improved by using more antenna ports. Further, in temperature estimation, we showed

that even a small material loss in V may have a large impact on the aperture Fisher information.

Chapter 4

Impact of Antenna Aperture on Fisher Information

In previous chapters, we studied the general problem of estimating the parameters of an incident Gaussian electromagnetic field using a sensor array that observes the field through the currents in an arbitrary conductor V . We characterized the maximum Fisher information that can be achieved with this array using a front-end of the form described in Chapter 2. We proved this Fisher information is always bounded above by a quantity we call aperture Fisher information (AFI), which intuitively represents all of the useful information contained in the currents of V . In this chapter, we study how the properties of V itself changes the AFI, and use these results as a benchmark to evaluate the performance of several widely-used antenna arrays.

The rest of this chapter is organized as follows. In Sec. 4.1, we derive the relationship between the AFI of an aperture and its substructures, and give numerical examples to illustrate how AFI varies with the size and shape of the aperture. In Sec. 4.2, we study the performance of several conventional antenna arrays, and examine how well

each array captures the information contained in the physical space it occupies. Finally, we summarize the main contribution of this chapter in Sec. 4.3.

4.1 Antenna Shape

In this section, we study how the aperture Fisher information changes with antenna shape. Several authors have considered optimizing the shape of a multi-port antenna system, usually in order to enhance directivity or improve impedance matching and cross-polarization between ports [77,78]. Most of these studies involve a computationally-intensive search among the substructures of some parent structure V . When one of the design objectives of shape optimization is to minimize the error incurred in estimating the parameters of an incident Gaussian field, the tools developed in this dissertation can be used to focus a search on the most promising substructures. In [79], Kundu showed that the aperture ergodic capacity of any structure V is an upper bound on the aperture capacity of any substructure. In a similar way, we now show the aperture Fisher information of any structure V is an upper bound on the AFI of any substructure. Thus, the AFI provides a way to quantify the loss of information associated with each substructure in a way that is independent of the number and placement of ports. This may lead to more efficient searches, since we can then focus attention on substructures that preserve most of the aperture Fisher information contained in the initial structure V .

Suppose the parent structure V has aperture impedance \mathbf{Z} and induced current \mathbf{J} as in (3.22), where both are represented relative to the basis $\mathbf{W}_j(s)$, $j = 1, \dots, K$. Similarly,

the substructure $\bar{V} \in V$ can be described by a similar equation

$$\bar{\mathbf{E}} = \bar{\mathbf{Z}}(\bar{\mathbf{J}}) + \bar{\mathbf{E}}_o . \quad (4.1)$$

where $\bar{\mathbf{E}}, \bar{\mathbf{J}} \in \mathbb{C}^{\bar{K}}$ represent fields and currents on \bar{V} .

In general, we may choose to represent $\bar{\mathbf{E}}$ and $\bar{\mathbf{J}}$ with different basis functions, say $\bar{\mathbf{W}}_j(s)$, $j = 1, \dots, \bar{K}$, where $s \in \bar{V}$. Since \bar{V} is a substructure of V , however, each basis function of \bar{V} can also be expressed in terms of the basis functions of V as $\bar{\mathbf{W}}_j(s) = \sum_i W_{ij} \mathbf{W}_i(s)$ for all s in \bar{V} . It follows that $\bar{\mathbf{J}}$ on \bar{V} represents the same current as $\mathbf{J} = \mathbf{W}\bar{\mathbf{J}}$ on V , and similarly $\mathbf{E} = \mathbf{W}\bar{\mathbf{E}}$, where $\mathbf{W} = \{W_{ij}\}$ represents a simple change of basis.

Since (complex) power calculations will produce the same results in both bases, we have

$$\bar{\mathbf{J}}^H \bar{\mathbf{Z}} \bar{\mathbf{J}} = \bar{\mathbf{J}}^H \bar{\mathbf{E}} = \mathbf{J}^H \mathbf{E} = \bar{\mathbf{J}}^H \mathbf{W}^H \mathbf{Z} \mathbf{W} \bar{\mathbf{J}} \quad (4.2)$$

$$\bar{\mathbf{J}}^H \bar{\mathbf{E}}_o = \mathbf{J}^H \mathbf{E}_o = \bar{\mathbf{J}}^H \mathbf{W}^H \mathbf{E}_o . \quad (4.3)$$

Since the equalities hold for all $\bar{\mathbf{J}}$, it follows $\bar{\mathbf{Z}} = \mathbf{W}^H \mathbf{Z} \mathbf{W}$, $\bar{\mathbf{E}}_o = \mathbf{W}^H \mathbf{E}_o$.

In Chapter 3, we showed that the AFI of an aperture V can be calculated from an equivalent observation model \mathbf{y}_V , given in (3.39). Let $\mathbf{y}_{\bar{V}}$ denote the corresponding observation model for \bar{V} in V . Taken together, the equations above imply $\mathbf{y}_{\bar{V}}$ has the same distribution as $\mathbf{W}^H \mathbf{y}_V$. Following the derivation in (3.35)-(3.37) with \mathbf{C} replaced by \mathbf{W} , we find $\mathbf{y}_{\bar{V}}$ is a Gaussian observation with

$$\mu_{\mathbf{y}_{\bar{V}}} = \mathbf{W}^H \mu_{\mathbf{y}_V} , \quad \Sigma_{\mathbf{y}_{\bar{V}}} = \mathbf{W}^H \Sigma_{\mathbf{y}_V} \mathbf{W} \quad (4.4)$$

and the corresponding Fisher information is given by

$$\mathcal{I}_{\bar{V}}(\mathbf{u}) = \mathcal{F}\left(\boldsymbol{\mu}_{\mathbf{y}_{\bar{V}}}(\mathbf{u}), \boldsymbol{\Sigma}_{\mathbf{y}_{\bar{V}}}(\mathbf{u})\right) = \mathcal{F}\left(\mathbf{W}^H \boldsymbol{\mu}_{\mathbf{y}_V}(\mathbf{u}), \mathbf{W}^H \boldsymbol{\Sigma}_{\mathbf{y}_V}(\mathbf{u}) \mathbf{W}\right). \quad (4.5)$$

Thus, the relationship between the aperture observation \mathbf{y}_V and the substructure observation $\mathbf{y}_{\bar{V}}$ is similar to the relationship between the aperture and an embedded array (3.37). It follows that we can prove a result analogous to Theorem 3.1 here.

Theorem 4.1: Under the same radiation environment, suppose structure V has aperture Fisher information $\mathcal{I}_V(\mathbf{u})$, and structure $\bar{V} \subseteq V$ has aperture Fisher information $\mathcal{I}_{\bar{V}}(\mathbf{u})$. Then $\mathcal{I}_{\bar{V}}(\mathbf{u}) \leq \mathcal{I}_V(\mathbf{u})$ for any $\bar{V} \subseteq V$. \diamond

Proof of Theorem 4.1: For any \mathbf{W} , observe (4.5) is a special case of (3.38) corresponding to $\mathbf{C} = \mathbf{W}$. From Theorem 3.1, we know that $\mathcal{I}_V(\mathbf{u}) \geq \mathcal{I}_{\bar{V}}(\mathbf{u})$ always holds.

The theorem above has two immediate, intuitive implications: first, it implies that removing metal from any conducting structure cannot increase Fisher information. Thus, aperture Fisher information cannot decrease as the physical dimensions of an antenna increase. It follows that longer, thicker wires typically capture more information than shorter, thinner ones; and, bigger patches are more informative than smaller ones. Second, the theorem provides a convenient way to evaluate how well an array samples the information contained within the space it occupies. Antenna systems for wireless applications are often designed to fit within a given form factor. For example, a cellular antenna must fit within the form factor of a cellular handset, and the physical dimensions of base station arrays are often limited by concerns about wind shear. The theorem above suggests that the AFI of a conductor with the physical dimensions of the handset is an upper bound to the AFI of any antenna that fits within the handset. Thus, we can use the handset AFI as a benchmark to measure how efficiently any cellular antenna samples

the information in the interior of the handset.

In the remainder of this section, we present several numerical examples to illustrate how AFI depends on the size and shape of the aperture for two different estimation problems: noise temperature estimation and direction-of-arrival estimation. We focus first on simple apertures, such as planar squares and spheres. In Sec. 4.2, we will further use these apertures to study the information contained in the space occupied by certain conventional arrays, such as uniform linear arrays and vector sensors.

4.1.1 Black-Body Radiation Temperature Estimation

To explore how estimation performance scales with aperture size, we again consider the radiation temperature problem described earlier. As shown in Sec. 3.4.1, the aperture observation \mathbf{y}_V is a zero-mean Gaussian vector with covariance $\mathbf{\Sigma}_{\mathbf{y}_V} = 4k_b T_0 B(u\mathbf{R}_E + \mathbf{R}_I)$, and the goal is to estimate the parameter u . For the aperture V , we consider an $l \times l \times 1$ cm copper patch (conductivity $\sigma = 5.96 \times 10^7$ S/m), where the side length l varies from 0.05λ to 0.5λ . Here λ is the wavelength of 30 MHz. In this example, \mathbf{R}_E and \mathbf{R}_I were calculated using FEKO with 380 RWG basis functions.

As noted earlier, the CRB (3.82) for this problem depends only on the eigencurrent efficiencies. The efficiencies were calculated by solving (3.80) in MATLAB. In Fig. 4.1a and 4.1b, we plot the efficiencies of the first 100 eigencurrents for square patches with side length $l = 0.1\lambda$ and 0.25λ , respectively. The number of efficient modes ($\eta_k > 0.1$) is relatively large: 10 for $l = 0.1\lambda$ and 16 for $l = 0.25\lambda$. This suggests that a large number of ports may be required to capture all of the information in this aperture. Note that when the radiation temperature equals the receiver temperature ($T_E = T_o$ and $u = 1$), the aperture Fisher information equals the summation of the squared eigencurrent efficiencies,

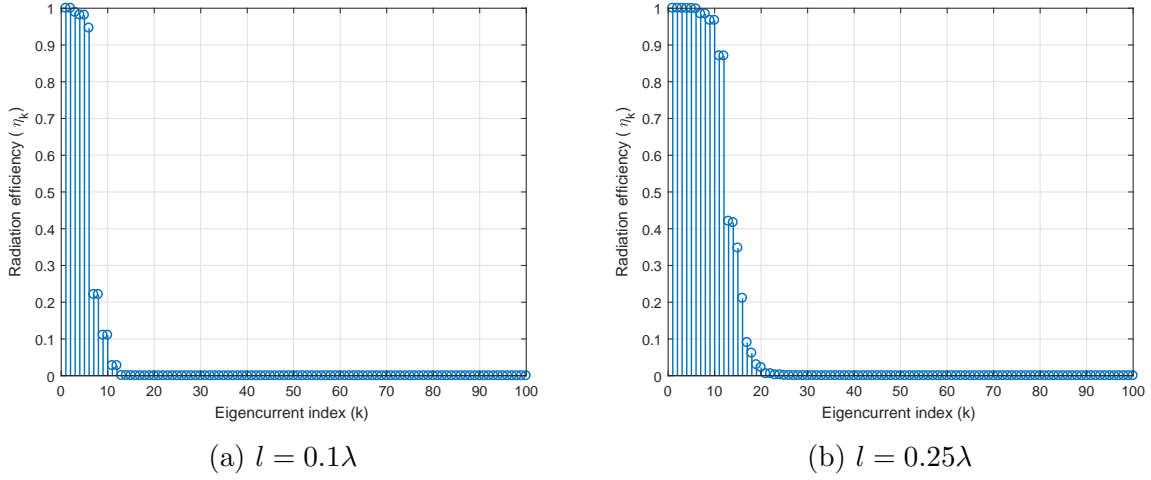


Figure 4.1: Eigencurrent efficiencies of a Square

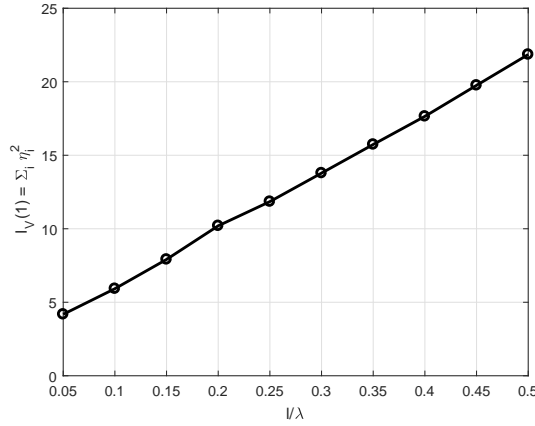


Figure 4.2: $\mathcal{I}_V(1)$ versus side length l for a square aperture

$\mathcal{I}_V(1) = \sum_i^K \eta_i^2$. We plot this value in Fig. 4.2, which appears to be increasing with l linearly within the range of l we considered, $\mathcal{I}_V(1) \approx 2.22 + 39.27l/\lambda$. In Fig. 4.3, we next plot the lower bound on the relative error implied by the aperture CRB, i.e. $u^{-1}\mathcal{I}_V^{-1/2}(u)$, for square patches with side length ranging from 0.05λ to 0.5λ . The estimation error decreases quickly as the size of the square grows. For example, increasing the size of the

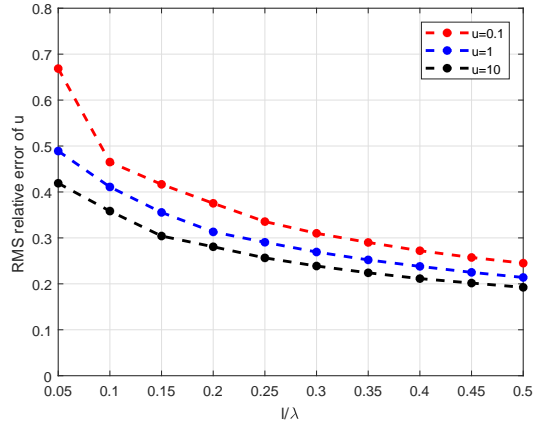


Figure 4.3: Aperture CRB on Temperature Estimation using a Square

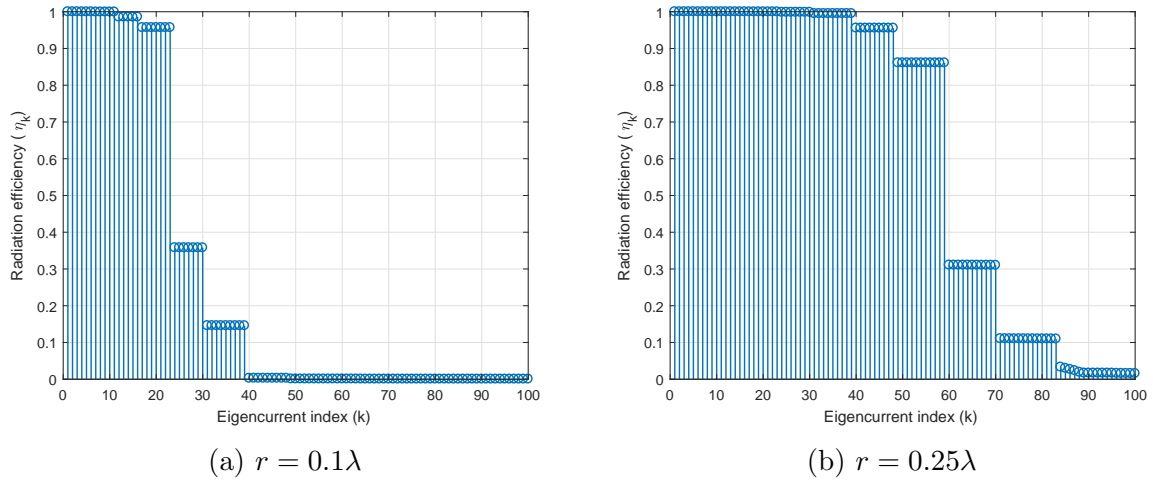


Figure 4.4: Eigencurrent efficiencies of a sphere

square from $l = 0.05\lambda$ to $l = 0.5\lambda$ changes the estimation error from 0.489 to 0.213 in the $u = 1$ case, which is a factor of 2.33 decrease.

Next, we recalculate the quantities in Fig. 4.1, Fig. 4.2, and Fig. 4.3 for a spherical copper conductor with radius ranging from $r = 0.05\lambda$ to $r = 0.5\lambda$, where λ is the wavelength of 30 MHz. \mathbf{R}_E and \mathbf{R}_I were calculated in FEKO again, using 1263 basis functions.

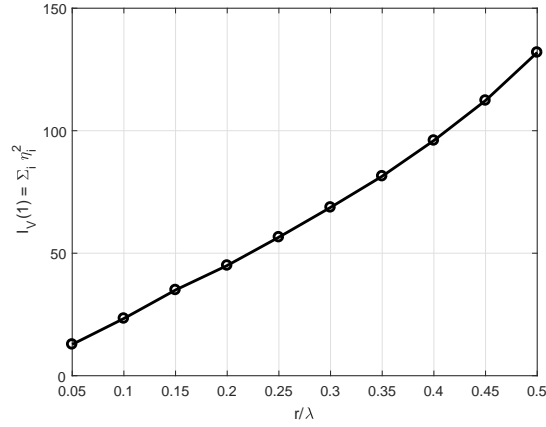


Figure 4.5: $\mathcal{I}_V(1)$ versus radius r for a spherical aperture

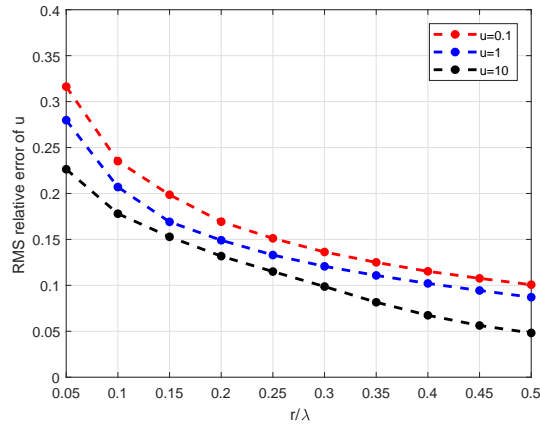


Figure 4.6: Aperture CRB on Temperature Estimation using a sphere

In Fig. 4.4a and 4.4b, we plot the radiation efficiencies of the first 100 eigencurrents for $r = 0.1\lambda$ and 0.25λ , respectively. Again, note the large number of efficient modes (39 for $r = 0.1\lambda$ and 83 for $r = 0.25\lambda$), which implies a large number of ports would be needed to capture all of the available information. Interestingly, several groups of eigencurrents have exactly the same efficiency - we conjecture this is caused by the high degree of symmetry in \mathbf{R}_E and \mathbf{R}_I , which is expected for a spherical structure. We plot the summation of the

squared eigencurrent efficiencies, $\mathcal{I}_V(1) = \sum_i^K \eta_i^2$, corresponding to different r values in Fig. 4.5. Note that $\mathcal{I}_V(1)$ now grows with r faster than linearly, possibly because we are now considering a 3-dimensional structure, rather than a 2-dimensional one. In Fig. 4.6, we plot the CRB on relative error, $u^{-1}\mathcal{I}_V^{-1/2}(u)$, where $\mathcal{I}_V(u)$ is given in (3.82).

4.1.2 One-dimensional DoA Estimation

Next, we consider the one-dimensional DoA estimation problem illustrated in Sec. 3.4.2. As mentioned before, this problem has been studied in numerous papers, where the receiver array is usually a ULA of dipoles (e.g. [76] and [14]). For conventional ULAs with dipoles spaced at least 0.5λ apart, the performance usually increases with the number of elements. However, our results in Sec. 2.4.2 show that, when the optimal matching network is applied, DoA estimation accuracy of an array remains constant as the neighboring distance decreases. This raises an interesting question: if the overall size of a dipole ULA is fixed, will estimation accuracy continue to increase as we pack more antennas into this region, or does performance saturate at some point past which additional gains are marginal? According to Theorem 4.1, the performance of the dipole array is bounded above by the AFI of the planar square that encloses the array. In a sense, the ULA samples the electromagnetic field over this square region, and we can increase the “spatial sampling frequency” by adding more dipoles.

Consider an array of $N = 2$ to 5 half-wavelength copper dipoles with radius 0.5 cm, that occupy a vertical square $0.5\lambda \times 0.5\lambda$ aperture, as illustrated in Fig. 4.7. For a ULA of N dipoles, the inter-element distance is then $d = 0.5\lambda/(N - 1)$. The enclosing aperture is a $0.5\lambda \times 0.5\lambda$ copper square of thickness 1cm. For this problem, we calculate \mathbf{R}_E , \mathbf{R}_I and the antenna pattern $\mathbf{g}(\theta, \phi)$ in (3.85) in FEKO, using 380 RWG basis functions to

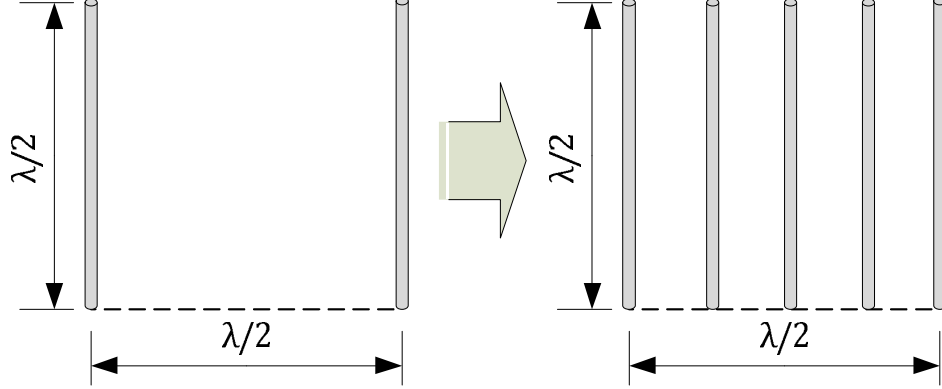


Figure 4.7: A uniform linear array of 2 to 5 dipoles

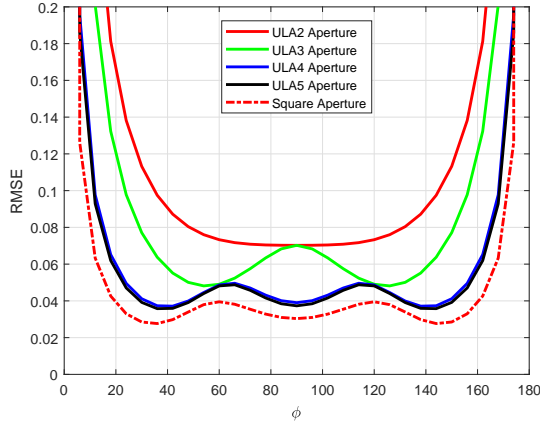


Figure 4.8: Aperture CRB on ϕ estimation using a ULA

model the square, and 99 basis functions to model each dipole, and we model sky noise using the Twiss model, so $\Sigma_e = 4k_b T_0 B \mathbf{R}_{\mathbf{AE}}$ in (3.7). Suppose a target located on the plane $\theta = 90^\circ$ emits a vertically-polarized signal at 30 MHz that consists of $N = 100$ symbols with a path gain of $x = 1$, so the received observation is given by (3.85) with $\gamma = 0$. We assume the receiver knows that the target is located on the horizontal plane ($\theta = 90^\circ$), and only estimates the ϕ angle.

The aperture Fisher information for these estimation problems is calculated from

(3.87) to (3.89), where $\boldsymbol{\psi} = \phi$. In Fig. 4.8, we plot aperture CRBs on the RMS estimation error of ϕ , $([\mathbf{I}_V^{-1}]_{\phi,\phi})^{1/2}$, for signals coming from different ϕ directions for a signal-to-noise ratio $SNR \triangleq P/(4k_bT_0F_{min}B)$ of 10 dB. The figure shows that all CRBs are relatively low near the broadside ($\phi = 90^\circ$) and grow rapidly near the end-fire direction ($\phi = 0^\circ$ or 180°). Note that using more dipoles reduces the aperture CRB, but the improvement appears to saturate at 4 dipoles; since 5 dipoles yields nearly the same bound to within the thickness of the plot lines. The minimum RMSE is 0.070° for the two-dipole array, 0.048° for the three-dipole array, 0.037° for the four-dipole array, and 0.036° for the five-dipole array. There is a relatively small performance gap (0.011° or less over $\phi = 30^\circ$ to $\phi = 150^\circ$) between the five-dipole array and the square structure. One possible explanation is that the radiation/receiving efficiency of dipole apertures are relatively low compared with the aperture, since dipoles have small cross sectional area therefore their loss impedance is large.

4.1.3 Two-dimensional DoA Estimation

Over the past few decades, many authors have considered two-dimensional DoA estimation using a variety of antenna structures, such as dipole arrays and vector sensors consisting of co-located loops and dipoles. The tools developed in this paper provide a way to study how well these antennas collect the information contained in the physical spaces they occupy. For example, a tripole antenna [85] consists of three mutually-orthogonal dipoles; a 6-element vector sensor [58] adds three mutually orthogonal loops to the tripole. Both of these antenna systems can be viewed as sampling the incident field within a physical space that could be approximated by an enclosing sphere.

Theorem 4.1 implies that using a tripole array rather than the sphere will generally

entail a loss of Fisher information. Intuitively, part of this loss is due to the size and shape of the tripole (i.e., the aperture) and part of the loss is due to the number and placement of ports on this aperture. In this section, we study how much information is lost due to the aperture. In the next section, we study how much information is lost due to the number and placement of ports.

We study three apertures: 1) a copper sphere with radius r ; 2) a copper tripole consisting of three wires of length $2r$, each centered at the origin and aligned with the x,y, and z-axes, respectively; and 3) a copper vector sensor consisting of three, mutually-orthogonal wires with length $2r$ and three, mutually orthogonal loops with radius r , where each loop is centered at the origin and orthogonal to one of the wires. For all three apertures, we vary r from 0.05λ to 0.5λ to study how their performance varies with size. Note that the tripole is a substructure of the vector sensor; therefore, according to Theorem 4.1, the AFI of a tripole cannot be greater than the AFI of a vector sensor of the same size.

We calculate $\mathbf{R}_{\mathbf{E}}$, $\mathbf{R}_{\mathbf{I}}$, and $\mathbf{g}(\theta, \phi)$ of these structures in FEKO. We use 1263 RWG basis functions to model the sphere, 297 basis functions to model the tripole, and 633 basis functions to model the vector sensor. We also model sky noise using the Twiss model. Suppose a target located in the (θ, ϕ) direction emits right-hand circularly-polarized signals at 30 MHz that consists of $N = 100$ symbols with a path gain of $x = 1$, so the received observation is given by (3.85) with $\gamma = \pi/4$ and $\psi = -\pi/2$. The SNR is set to be 10 dB. We assume the receiver knows the transmitted waveform and estimates both the θ and ϕ angles.

The aperture Fisher information for these estimation problems is calculated from (3.87) to (3.89), where $\boldsymbol{\psi} = [\theta, \phi]$. For a target located on the horizontal plane ($\theta = 90^\circ$), we plot the aperture CRBs on the RMS estimation error of θ and ϕ , $([\mathcal{I}_V^{-1}]_{\theta,\theta})^{1/2}$ and

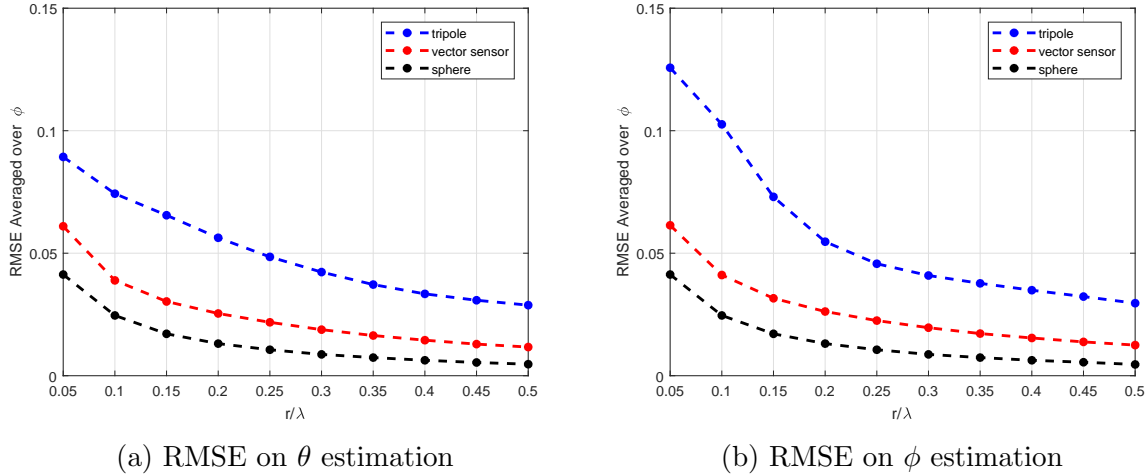


Figure 4.9: Aperture CRB on DoA estimation

$([\mathcal{I}_V^{-1}]_{\phi,\phi})^{1/2}$, averaged over all ϕ angles, in Fig. 4.9a and Fig. 4.9b, respectively. Consistent with Theorem 4.1, note that the aperture bounds decrease monotonically with size, and the sphere performs better than its substructure, the vector sensor. Similarly, the vector sensor outperforms its substructure, the tripole. For example, at size $r = 0.05\lambda$, the aperture bound on θ estimation is 0.0288° for a tripole, 0.0117° for a vector sensor, and 0.0047° for a sphere. Over the range of r we consider, the performance gap between a tripole and its enclosing sphere is significant in general; however, the ratio between the RMSE of a vector sensor and a sphere can be as small as 1.5 when $r = 0.05\lambda$. The aperture bound of a sphere decreases the most as its size increases: when the aperture size increases from $r = 0.05\lambda$ to $r = 0.5\lambda$, the average RMSE on θ estimation is reduced by a factor of 3.10 for a tripole, 5.21 for a vector sensor and 8.79 for a sphere; for ϕ estimation, these values are 4.25, 4.91 and 8.98, respectively.

These results suggest that tripoles and vector sensors do not capture most of the useful information contained in the physical spaces they occupy, except possibly for 6-

element vector sensors that are extremely small ($r < 0.1\lambda$). This loss is entirely due to size and shape of the metal conductors, and not the number and placement of ports.

4.2 Conventional Sensor Arrays

In the last section, we studied the information lost by using tripoles and vector sensors relative to an enclosing sphere, in the context of 2-dimensional DoA estimation. These results illustrate how estimation performance is limited by the size and shape of the conductors. In this section, we now consider how much additional information is lost due to the number and placement of ports on these conductors. In particular, we consider several conventional arrays studied in Sec 4.1 and explore how the estimation accuracy changes with combinations of antenna ports.

4.2.1 Dipole ULA

One of the most common antenna arrays is the center-fed dipole ULA, where each dipole is one half-wavelength long, and has only one port in the center. The results of Sec. 3.4 show that using multiple ports on one dipole can significantly improve its performance in some estimation applications. For the one-dimensional DoA estimation problem we considered in Sec. 4.1.2, can we significantly increase the estimation accuracy by using more ports on the same wire?

We consider the same dipole ULA apertures illustrated in Fig. 4.7. For each dipole, we consider the three sets of delta-gap ports shown in Fig. 3.3: one port in the center, two ports located 0.125λ from each end, and all three ports. As discussed in Sec. 3.4, the current matrix for delta-gap ports is $\mathbf{C}^H = \mathbf{Y}_{rc}^{-1}\mathbf{Y}_c^T$, where \mathbf{Y}_c is the submatrix of $\mathbf{Y} = \mathbf{Z}^{-1}$ containing only columns corresponding to port locations, and \mathbf{Y}_{rc} is a submatrix

of \mathbf{Y}_c containing only rows corresponding to port locations. We model this array in FEKO using the same basis functions as in Sec. 4.1.2, and calculate the corresponding array FIM by substituting \mathbf{R}_E , \mathbf{R}_I , $\mathbf{g}(\theta, \phi)$, and \mathbf{C} into (3.38).

As in Sec 4.1.2, we consider estimation of the azimuth ϕ for a target located on the horizontal plane ($\theta = 90^\circ$), which sends $N = 100$ known symbols that are vertically-polarized with a path gain of $x = 1$. The signal SNR is chosen to be 10 dB. Again, we use the CRB on the RMSE of ϕ estimation, $([\mathcal{I}_V^{-1}]_{\phi, \phi})^{1/2}$, as the performance metric. In Fig. 4.10, we compare the aperture CRB and the array CRB for these ULAs when each dipole has only one port in the center. We can see that these two bounds are generally close: for example, on the $\phi = 90^\circ$ direction, the array bound is only $26.5 \pm 0.5\%$ higher than the aperture bound for all ULAs we considered. This indicates that the conventional center-fed dipole arrays are actually observing most of the AFI of the wire conductors. A comparison of the array and aperture CRBs for 2 and 3 ports per wire is plotted in Fig. 4.11a and Fig. 4.11b, respectively. We can see that two ports give virtually the same array CRB as one port placed between them, and yet using three ports on each dipole essentially achieves the aperture bound (less than 10% higher RMSE for all ULAs).

4.2.2 Tripole

Next, we study the array CRB of a tripole for the two-dimensional DoA estimation application. We consider the same three sets of delta-gap ports on each wire: one port on the center of each wire, two ports located $0.25l$ from each end (where $l = 2r$ is the length of each wire), and all three ports. The total number of antenna ports for the three cases are 3, 6, and 9, respectively. We use the same set of basis functions as in Sec. 4.1.3 and calculate \mathbf{R}_E , \mathbf{R}_I , $\mathbf{g}(\theta, \phi)$, and \mathbf{C} for the array FIM calculation using (3.38).

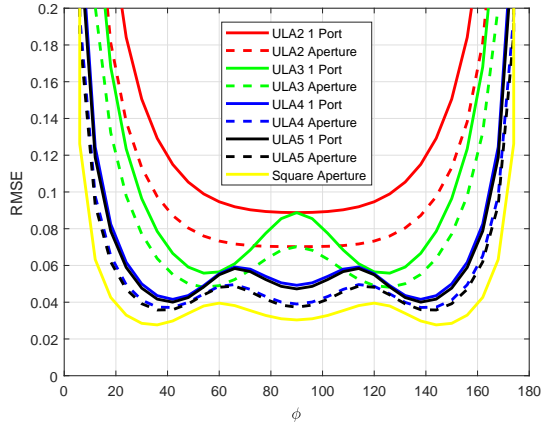
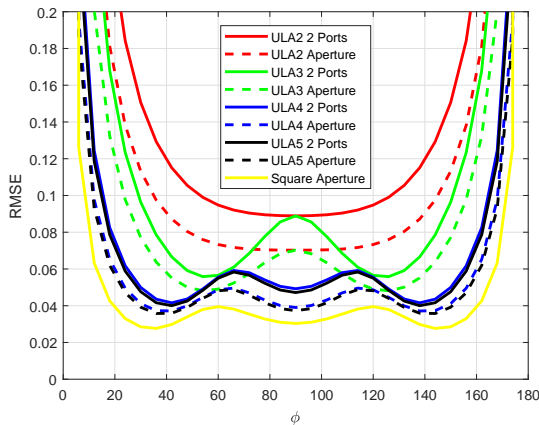
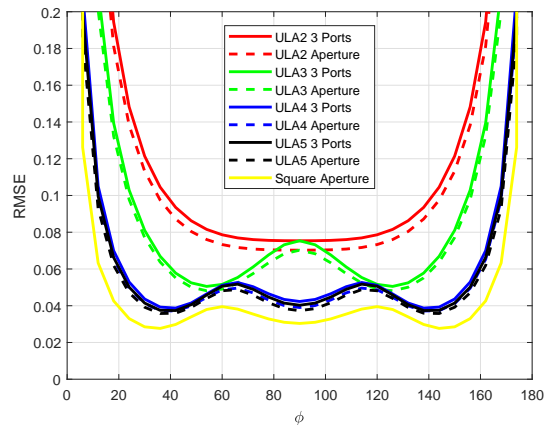


Figure 4.10: CRBs for 1 port per wire



(a) CRBs for 2 ports per wire



(b) CRBs for 3 ports per wire

Figure 4.11: Array CRB versus Aperture CRB for ϕ estimation

We consider the same scenario as in Sec. 4.1.3, where a target sends $N = 100$ right-hand circularly-polarized symbols with a path gain of $x = 1$ at $SNR = 10$ dB from the direction (θ, ϕ) . The receiver estimates both θ and ϕ angles of the target. In Fig. 4.12a and Fig. 4.12b, we plot the CRBs on the RMS estimation error of θ and ϕ for a target located in the horizontal plane ($\theta = 90^\circ$), using a $r = 0.1\lambda$ tripole with different port

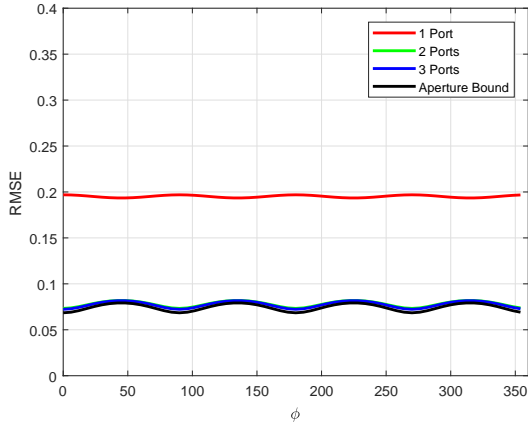
combinations. For the estimation of θ , we can see the estimation accuracy does not vary significantly with ϕ : over all ϕ directions, using 1 port on each wire gives an RMSE of about 0.195° , yet using 2 or 3 ports reduces it to around 0.075° , which essentially achieves the aperture bound. The RMSE on ϕ estimation is different: using 1 port on each wire gives an RMSE around 0.190° over all ϕ directions, yet the CRB varies significantly with ϕ when more ports are used. For example, with 3 ports on each wire, the array CRB ranges from 0.171° on $\phi = 45^\circ$ to 0.072° on $\phi = 90^\circ$.

In Fig. 4.12a and Fig. 4.12b, we plot the array CRB using different port combinations on a $r = 0.25\lambda$ tripole. The results are similar to $r = 0.1\lambda$ tripole in general: all CRBs on θ estimation are relatively constant over all ϕ directions, yet the CRB on ϕ estimation varies significantly over ϕ . One port on each wire does not exploit this aperture fully; using two or more ports can reduce the estimation error by up to a factor of 3 in certain directions.

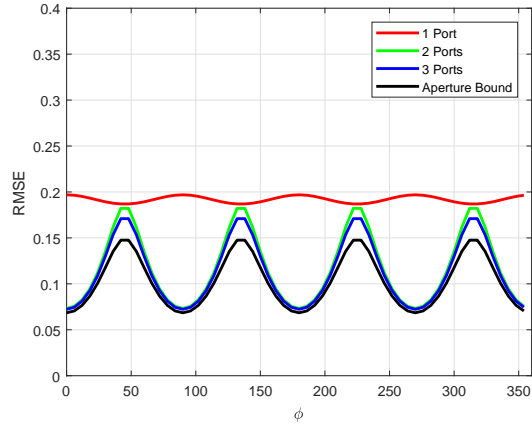
We have shown that every aperture has lower aperture CRB than its substructures; however, the results above show that this is not generally true of arrays with fixed port locations. For example, we note the $r = 0.25\lambda$ tripole with one port per wire contains the corresponding $r = 0.1\lambda$ array as a substructure. However, from Fig. 4.12 and 4.13, note the CRB of the $r = 0.25\lambda$ array is 0.236° on $\phi = 90^\circ$ whereas the $r = 0.1\lambda$ tripole is 0.197° . Of course, this does not violate Theorem 4.1, which applies only to the aperture Fisher information and not to array Fisher information.

4.2.3 Vector Sensor

Finally, we consider a vector sensor that consists of three mutually-orthogonal dipoles and three mutually-orthogonal loops. For this aperture, we consider four combinations of

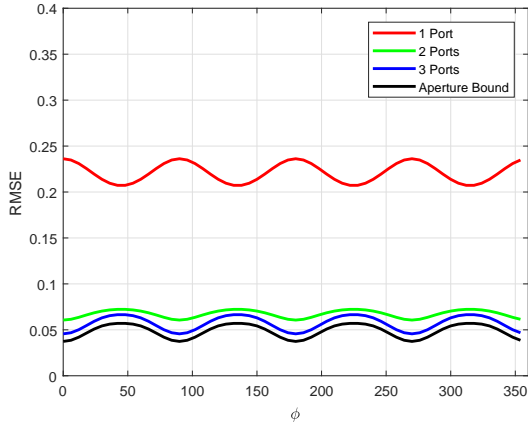


(a) RMSE on θ estimation

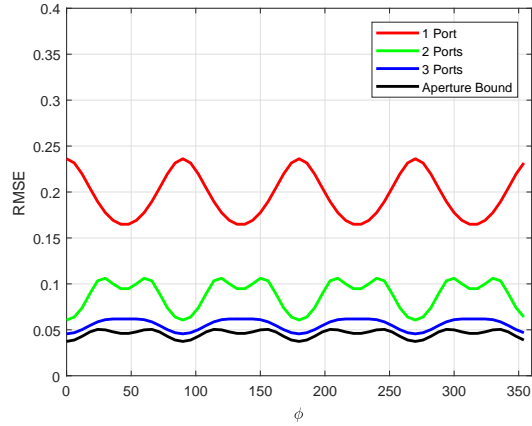


(b) RMSE on ϕ estimation

Figure 4.12: Array CRB for a tripole, $r = 0.1\lambda$



(a) RMSE on θ estimation



(b) RMSE on ϕ estimation

Figure 4.13: Array CRB for a tripole, $r = 0.25\lambda$

antenna ports as examples: 1) 1 port on each dipole and 1 port on each loop (6 ports in total); 2) 2 ports on each dipole and 2 ports on each loop (12 ports in total); 3) 3 ports on each dipole and 4 ports on each loop (21 ports in total); 4) 3 ports on each dipole and 8 ports on each loop (33 ports in total). The ports on dipoles have same locations as in the last subsection, and the ports on loops are uniformly distributed. If we consider one

dipole and one loop as a basic composing unit for a vector sensor, then the positions of the ports on each unit are illustrated by red dots Fig. 4.14. Note that when assigning loop ports, we avoid locations where the loop connects with other loops or dipoles. \mathbf{R}_E , \mathbf{R}_I , \mathbf{g} , and \mathbf{C} are calculated in FEKO using the same basis functions considered in Sec. 4.1.3.

For a vector sensor whose three wires are aligned along the x, y, and z axes, the port locations are not circularly symmetric for the 6 and 12 ports cases in Fig. 4.14, which causes the DoA estimation accuracy to vary significantly over ϕ . To remove this asymmetry, we firstly rotate the structure along the z-axis by 45° clockwise, then rotate it along the y-axis by 54.74° counter-clockwise.

We consider the two-dimensional DoA estimation problem under the same scenario as in Sec. 4.2.2, where a target sends $N = 100$ right-hand circularly-polarized symbols with a path gain of $x = 1$ at $SNR = 10$ dB from direction (θ, ϕ) . The receiver estimates both θ and ϕ of the target. In Fig. 4.15a and Fig. 4.15b, we plot the array bound on the RMS estimation error of θ and ϕ for a target located in the horizontal plane ($\theta = 90^\circ$), using a $r = 0.1\lambda$ vector sensor with all four port combinations. The estimation accuracy of both θ and ϕ are almost constant over all ϕ directions. We can see that when the most conventional feeding method (1 port on each dipole and 1 port on each loop) is used, the RMSE is 0.150° on θ estimation and 0.137° on ϕ estimation. However, if we increase the total number of ports to 12, these two values can be decreased significantly to 0.080° and 0.109° , respectively. Using more modes can further decrease the array bound, and almost all of the AFI can be observed using 33 ports.

Under the same radiation environment, the corresponding array bound of a $r = 0.25\lambda$ vector sensor on θ and ϕ estimation is plotted in Fig. 4.16. For each combination of ports, the estimation accuracy of θ and ϕ are similar over all ϕ directions: the RMSE error is around 0.10° for 6 ports, 0.057° for 12 ports, 0.047° for 21 ports, and 0.025° for 33 ports.

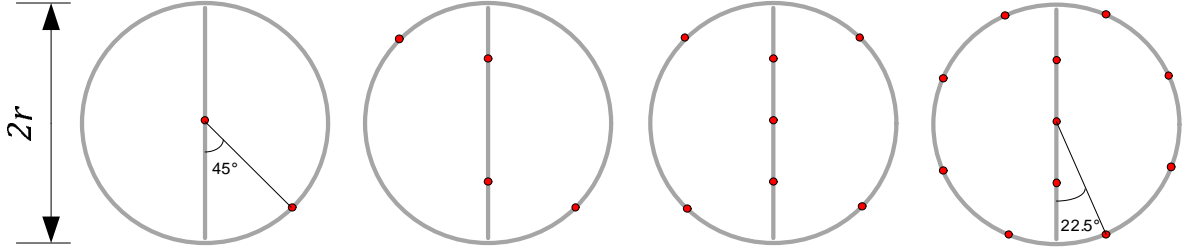


Figure 4.14: Antenna ports on vector sensor

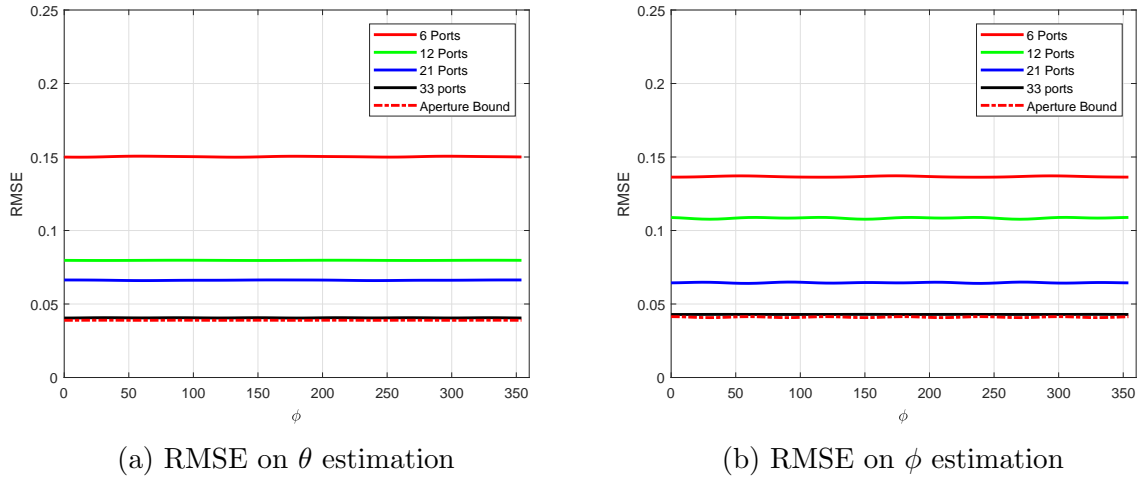
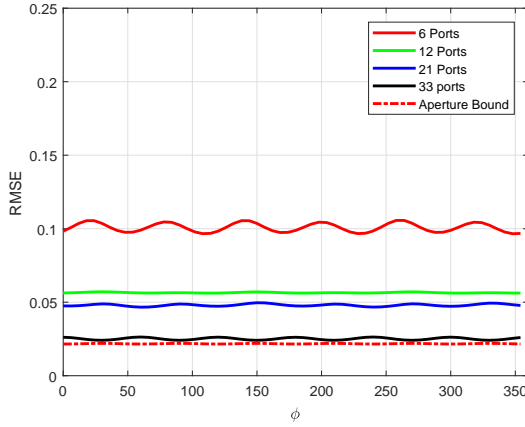


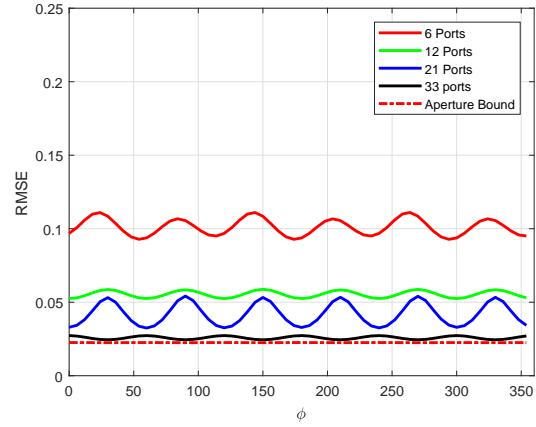
Figure 4.15: Array CRB for a vector sensor, $r = 0.1\lambda$

Similarly, using only 6 ports leads to a relatively high estimation error, and the aperture bound can be achieved when 33 ports are used.

From Figs. 4.15 and 4.16, we see that the conventional placement of ports on a vector sensor (6 ports in total) does not capture all of the information contained in the structure. In particular, more accurate DoA estimates may be obtained by adding more antenna ports. Furthermore, note that a large number of ports may be required to achieve the aperture bound of a vector sensor, even if it is electrically small.



(a) RMSE on θ estimation



(b) RMSE on ϕ estimation

Figure 4.16: Array CRB for a vector sensor, $r = 0.25\lambda$

4.3 Conclusion

In this chapter, we evaluated the aperture Fisher information and array Fisher information of several antenna structures for three different estimation applications. We proved that the AFI of a structure is always greater than or equal to the AFI of its substructures. Using square and spherical apertures as examples, we calculated how the AFI increases with the aperture size for radiation intensity and DoA estimation applications. We then studied the performance of several substructures of squares and spheres, which have been considered in the literature, and evaluated their ability to observe the information contained in the spaces they occupy. For one-dimensional DoA estimation and a fixed square aperture, we showed that placing dipoles closer than $\lambda/2$ apart can significantly increase AFI. However, performance tends to saturate at about $\lambda/6$ spacing. Furthermore, we demonstrated that in this application, most of the AFI can be observed by using only one port on each dipole, which is the most conventional way of using a dipole ULA. We also studied the DoA estimation ability of tripoles and vector sensors. Our results

show that the conductors used in them do not capture most of the AFI contained in the physical spaces they occupy, as evidenced by comparing these structures to an enclosing sphere. Moreover, the conventional placement of ports on tripoles and vector sensors induces additional loss in Fisher information, which can be significantly reduced by adding more antenna ports.

Chapter 5

Conclusion and Future Work

In this concluding chapter, we summarize the main contributions of this dissertation and suggest topics for future work.

5.1 Summary of Dissertation

In Chapter 1, we briefly reviewed existing literature on array signal processing, especially those who consider compact arrays. In conventional arrays, antennas are separated by a relatively large distance. However, deploying arrays on platforms with size limitation requires placing antennas close together. This causes antennas to interact with each other and can profoundly change the array properties. To improve the performance of compact arrays, impedance matching network can be applied to decouple antennas, increase power transfer, and so on. Although several authors have analyzed compact arrays from different perspectives, no one has yet considered the impact of impedance matching on the performance of estimators based on physical noise models. The lack of a comprehensive study on the performance of compact arrays in estimation applications has motivated

our work in Chapter 2.

In Chapter 2, we studied a general class of Bayesian and non-Bayesian estimation problems, in which the signal of interest is observed through a sensor front-end consisting of coupled antennas, an impedance matching network, and a bank of uncoupled amplifiers. Considering physical noise sources, we derived the CRB associated with each problem and explored how it varies with impedance matching and the physical characteristics of the signals and noise. We showed that MNF matching is universally optimal in the sense of minimizing the CRB for all estimation problems in the class. To demonstrate our results, we studied MIMO channel estimation and DoA estimation problems using a 4-dipole ULA with various antenna separating distance, and calculated the CRB corresponding to different matching networks and amplifier noise levels. Our results suggest that the CRB is generally sensitive to the physical sources of observation noise. However, optimal matching can eliminate this sensitivity and significantly reduce estimation error in the presence of strong coupling.

While the result of Chapter 2 allows us to calculate the performance bound of any antenna arrays, it is still unclear how to design the optimal array for estimation problems. In Chapter 3, instead of studying any array in particular, we considered the general problem of estimating the parameters of an incident Gaussian electromagnetic field using sensor arrays built from an arbitrary conductor. We characterized the maximum Fisher information that can be observed by such arrays when a front-end consisting of a matching network followed by a bank of amplifiers are used. We proved this Fisher information is always bounded above by a quantity we call aperture Fisher information, which intuitively represents all of the useful information contained in the currents of this conductor. For a broad class of estimation problems, we showed that aperture Fisher information can be associated with a system of unit-power eigencurrents on this conductor, which

decompose the volume currents into a sequence of independent, scalar signal observations. We showed the Fisher information of any M -port array is bounded above by the information in the best M eigencurrents, and derived conditions under which equality can be attained. Through numerical examples involving wire arrays, we found that certain conventional arrays miss most of the useful information contained in their volume currents, and illustrated how arrays can be modified to capture this missing information.

After analyzing all components in a receiver of the form in Chapter 3, we studied the conductor which the array is built from in Chapter 4. We proved that the AFI of a structure is always greater than or equal to the AFI of its substructures. Using squares and spheres as examples, we studied how the AFI increases with aperture size for radiation intensity and DoA estimation problems. We then used the results as benchmarks to evaluate the performance of several substructures of squares and spheres, which have been studied in the literature. For example, our results on one-dimensional DoA estimation problem using an ULA of half-wavelength dipoles showed that placing dipoles at less than $\lambda/2$ spacing can significantly reduce the estimation error. However, the performance tends to saturate at about $\lambda/6$ neighboring distance. Furthermore, we showed that using only one port on each dipole, which is the most conventional way of using such arrays, observes most of the AFI. We also studied the DoA estimation ability of three-dimensional antenna arrays including tripoles, vector sensors, and their enclosing spheres. We concluded that most AFI contained in the electromagnetic field within the sphere cannot be observed by tripoles and vector sensors. Moreover, the conventional ports placements on tripoles and vector sensors induced additional loss in Fisher information. Adding more antenna ports can significantly improve their estimation accuracy.

5.2 Future Work

During the development of this work, we encountered several interesting questions that are yet to be answered. In this section, we highlight some of these questions that may serve as future research directions.

Optimal Single-port Matching Network

In Chapter 2, we derived the properties of the optimal matching network for estimation applications. However, the actual implementation of a multi-port matching network can be difficult, especially under strong antenna coupling conditions. On the contrary, single-port matching networks are sub-optimal yet easy to build. In this dissertation, we studied one type of single port matching technique called self match, which ignores the coupling effect and seeks to minimize the noise factor of each antenna in isolation. However, it is unclear if other single port matching networks can significantly outperform self match. It is therefore of practical interest to look for the optimal single port matching network.

Similar questions have been considered by researchers for wireless communication applications. [42] shows that for a MIMO system whose receive antennas are coupled, applying different single port matching networks can change the channel capacity significantly. Furthermore, they derived conditions for optimal single port matching for a receiving array consists of 2 half-wavelength dipoles with small neighboring distance. It seems possible to derive the optimal single port matching network for estimation applications as well, especially for arrays with certain geometric symmetries.

Limits of Estimation in Broadband Applications

This dissertation systematically studied the performance of a compact array under narrowband conditions. However, radio systems usually work at non-zero bandwidth in practical applications. As frequency varies, nearly all properties of the sensor array may change, including antenna impedance, radiation pattern, noise covariance, etc.

Nearly all chapters of this dissertation can be extended to incorporate broadband analysis. For the matching network discussed in Chapter 2, studies including [93] and [94] have shown that perfect impedance match may be impossible to achieve over a positive bandwidth using passive networks. According to our derivation, an effective matching network should take the frequency-dependent antenna impedance into consideration to achieve the possibly “Minimum-Noise-Factor Broadband Match”, or even jointly consider the antenna impedance and Fisher Information on different frequency to achieve a “Maximum-Fisher-Information Broadband Match”.

Similarly, our studies in Chapter 3 and 4 can also include broadband conditions. Frequency variation of the antenna impedance and signal and noise characteristics are likely to change the aperture Fisher information over certain bandwidth. Consequently, the current matrix of ports and Fisher information eigencurrents may be frequency dependent, which adds new challenges to ports and aperture design.

REFERENCES

- [1] H. Krim and M. Viberg, "Two decades of array signal processing research: The parametric approach," *IEEE Signal Processing Magazine*, vol. 13, no. 4, pp. 67–94, Jul. 1996.
- [2] H. L. Van Trees, *Optimum Array Processing (Part IV of Detection, Estimation, and Modulation Theory)*. New York: Wiley, 2002.
- [3] B. D. Van Veen and K. M. Buckley, "Beamforming: a versatile approach to spatial filtering," *IEEE ASSP Magazine*, vol. 5, no. 2, pp. 4–24, Apr. 1988.
- [4] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276–280, Mar. 1986.
- [5] A.-J. van der Veen, M. C. Vanderveen and A. Paulraj, "Joint angle and delay estimation using shift-invariance techniques," *IEEE Transactions on Signal Processing*, vol. 46, no. 2, pp. 405–418, Feb. 1998.
- [6] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, no. 3, pp. 311–335, Mar. 1998.
- [7] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, Nov. 1999.
- [8] I. Gupta and A. Ksienski, "Effect of mutual coupling on the performance of adaptive arrays," *IEEE Transactions on Antennas and Propagation*, vol. 31, no. 5, pp. 785–791, Sep. 1983.
- [9] H. Steyskal and J. S. Herd, "Mutual coupling compensation in small array antennas," *IEEE Transactions on Antennas and Propagation*, vol. 38, no. 12, pp. 1971–1975, Dec. 1990.
- [10] C. C. Yeh, M. L. Leou and D. R. Ucci, "Bearing estimations with mutual coupling present," *IEEE Transactions on Antennas and Propagation*, vol. 37, no. 10, pp. 1332–1335, Oct. 1989.
- [11] B. Friedlander and A. J. Weiss, "Direction finding in the presence of mutual coupling," *IEEE Transactions on Antennas and Propagation*, vol. 39, no. 3, pp. 273–284, Mar. 1991.

- [12] T. Svantesson, "Direction finding in the presence of mutual coupling," Licentiate Thesis, Department of Signals and Systems, Chalmers University of Technology, Sweden, 1999.
- [13] T. Svantesson, "Antennas and propagation from a signal processing perspective," PhD Thesis, Department of Signals and Systems, Chalmers University of Technology, Sweden, 2001.
- [14] B. K. Lau and J. B. Andersen, "Direction-of-arrival estimation for closely coupled arrays with impedance matching," *2007 International Conference on Information, Communications and Signal Processing*, pp. 1–5, Dec. 2007.
- [15] M. Lin and L. Yang, "Blind calibration and DOA estimation with uniform circular arrays in the presence of mutual coupling", *IEEE Antennas and Wireless Propagation Letters*, vol. 5, no. 1, pp. 315–318, Dec. 2006.
- [16] C. K. E. Lau, R. S. Adve and T. K. Sarkar, "Minimum norm mutual coupling compensation with applications in direction of arrival estimation," *IEEE Transactions on Antennas and Propagation*, vol. 52, no. 8, pp. 2034–2041, Aug. 2004.
- [17] Z. Ye and C. Liu, "On the resiliency of MUSIC direction finding against antenna sensor coupling," *IEEE Transactions on Antennas and Propagation*, vol. 56, no. 2, pp. 371–380, Feb. 2008.
- [18] S. Hwang, S. Burintramart, T. K. Sarkar and S. R. Best, "Direction of arrival (DOA) estimation using electrically small tuned dipole antennas," *IEEE Transactions on Antennas and Propagation*, vol. 54, no. 11, pp. 3292–3301, Nov. 2006.
- [19] E. De Carvalho and D. T. M. Slock, "Cramer-Rao bounds for semi-blind, blind and training sequence based channel estimation", *Signal Processing Advances in Wireless Communications, First IEEE Signal Processing Workshop on*, pp. 129–132, Apr. 1997.
- [20] M. Biguesh and A. B. Gershman, "Training-based MIMO channel estimation: a study of estimator tradeoffs and optimal training signals", *IEEE Transactions on Signal Processing*, pp. 884–893, Vol. 54, no. 3, Mar. 2006.
- [21] C. Shin, R. W. Heath and E. J. Powers, "Blind channel estimation for MIMO-OFDM systems", *IEEE Transactions on Vehicular Technology*, vol. 56, no. 2, pp. 670–685, Mar. 2007.
- [22] X. Liu, S. Lu, M. E. Bialkowski and H. T. Hui, "MMSE channel estimation for MIMO system with receiver equipped with a circular array antenna," *2007 Asia-Pacific Microwave Conference*, Bangkok, Thailand, pp. 1-4, Dec. 2007

- [23] X. Liu and M. E. Bialkowski, "Effect of antenna mutual coupling on MIMO channel estimation and capacity," *International Journal of Antennas and Propagation*, vol. 2010, Article ID. 306173, 2010.
- [24] A. Kisliansky, R. Shavit and J. Tabrikian, "Direction-of-arrival estimation in the presence of noise coupling in antenna arrays," *IEEE Transactions on Antennas and Propagation*, vol. 55, no. 7, pp. 1940–1947, Jul. 2007.
- [25] C. H. Niow and H. T. Hui, "Improved noise modeling with mutual coupling in receiving antenna arrays for direction-of-arrival estimation," *IEEE Transactions on Wireless Communications*, vol. 11, no. 4, pp. 1616–1621, Apr. 2012.
- [26] M. L. Morris and M. A. Jensen, "Improved network analysis of coupled antenna diversity performance," *IEEE Transactions on Wireless Communications*, vol. 4, no. 4, pp. 1928–1934, Jul. 2005.
- [27] C. P. Domizioli and B. L. Hughes, "Front-end design for compact MIMO receivers: A communication theory perspective," *IEEE Transactions on Communications*, vol. 60, no. 10, pp. 2938–2949, Oct. 2012.
- [28] H. L. Van Trees, *Detection, Estimation and Modulation Theory, vol. 1*. New York: John Wiley, 1968.
- [29] J. Dauwels, "Computing Bayesian Cramér-Rao bounds," *Proc. IEEE International Symposium on Information Theory*, Adelaide, Australia, pp. 425–429, Sept. 4–9, 2005
- [30] B. Hassibi and B. Hochwald, "How much training is needed in multiple-antenna wireless links?," *IEEE Transactions on Information Theory*, vol. 49, no. 4, pp. 951–963, Apr. 2003.
- [31] T. Yoo and A. Goldsmith, "Capacity and power allocation for fading MIMO channels with channel estimation error," *IEEE Transactions on Information Theory*, vol. 52, no. 5, pp. 2203–2214, May. 2006.
- [32] M. Dong and L. Tong, "Optimal design and placement of pilot symbols for channel estimation," *IEEE Transactions on Signal Processing*, vol. 50, no. 12, pp. 3055–3069, Dec. 2002.
- [33] J. W. Brewer, "Kronecker products and matrix calculus in systems theory," *IEEE Transactions on Circuits and Systems*, vol. 25, no. 9, pp 772–781, Sep. 1978.
- [34] G. J. Burke and A. J. Poggio, "Numerical electromagnetics code (NEC) - method of moments," *Technical Document 11, Naval Ocean Systems Center*, San Diego, CA, Jan. 1981.

- [35] D. M. Pozar, *Microwave Engineering*, 3rd edition. Hoboken, New Jersey: Wiley, 2009.
- [36] C. Desoer, “The maximum power transfer theorem for n-ports”, *IEEE Transactions on Circuit Theory*, vol. 20, no. 3, pp. 328–330, May 1973.
- [37] J. W. Wallace and M. A. Jensen, “Mutual coupling in MIMO wireless systems: a rigorous network theory analysis,” *IEEE Transactions on Wireless Communications*, vol. 3, no. 4, pp. 1317–1325, Jul. 2004.
- [38] K. F. Warnick and M. A. Jensen, “Optimal noise matching for mutually coupled arrays”, *IEEE Transactions on Antennas and Propagation*, vol. 55, no. 6, pp. 1726–1731, Jun. 2007.
- [39] C. P. Domizioli and B. L. Hughes, “Noise correlation in compact diversity receivers,” *IEEE Transactions on Communications*, vol. 58, no. 5, pp. 1426–1436, May 2010.
- [40] B. K. Lau, J. B. Andersen, G. Kristenson and A. F. Molisch, “Impact of matching network on the capacity of compact MIMO systems,” in *Proceedings of the Nordic Antenna Symposium*, Linköping, Sweden, May 30-Jun. 1, 2006.
- [41] B. K. Lau, J. B. Andersen, A. F. Molisch and G. Kristensson, “Antenna matching for capacity maximization in compact MIMO systems,” *2006 3rd International Symposium on Wireless Communication Systems*, Valencia, Spain, Sep. 2006.
- [42] Y. Fei, Y. Fan, B. K. Lau and J. S. Thompson, “Optimal single-port matching impedance for capacity maximization in compact MIMO arrays,” *IEEE Transactions on Antennas and Propagation*, vol. 56, no. 11, pp. 3566–3575, Nov. 2008.
- [43] M. J. Gans, “Channel capacity between antenna arrays – part I: sky noise dominates,” *IEEE Transactions on Communications*, vol. 54, no. 9, pp. 1586–1592, Sep. 2006.
- [44] M. J. Gans, “Channel capacity between antenna arrays - part II: amplifier noise dominates,” *IEEE Transactions on Communications*, vol. 54, no. 11, pp. 1983–1992, Nov. 2006.
- [45] P. S. Taluja and B. L. Hughes, “Bandwidth limitations and broadband matching for coupled multi-antenna systems,” *2011 IEEE Global Telecommunications Conference - GLOBECOM 2011*, Kathmandu, Nepal, 5-9 Dec. 2011.
- [46] A. Nehorai and E. Paldi, “Vector-sensor array processing for electromagnetic source localization,” *IEEE Transactions on Signal Processing*, vol. 42, no. 2, pp. 376–398, Feb. 1994.

- [47] B. Hochwald and A. Nehorai, “Identifiability in array processing models with vector-sensor applications,” *IEEE Transactions on Signal Processing*, vol. 44, no. 1, pp. 83–95, Jan. 1996.
- [48] L. L. Monte, B. Elnour, D. Erricolo and A. Nehorai, “Design and realization of a distributed vector sensor for polarization diversity applications,” *International Waveform Diversity and Design Conference*, pp. 358–361, Pisa, Italy, 4-8 Jun. 2007.
- [49] X. Gong, Z. Liu, Y. Xu and M. I. Ahmad, “Direction-of-arrival estimation via twofold mode-projection,” *Signal Processing*, vol. 89, no. 5, pp. 831–842, May 2009.
- [50] L. L. Monte, B. Elnour and D. Erricolo, “Distributed 6D vector antennas design for direction of arrival applications,” *International Conference on Electromagnetics in Advanced Applications*, Torino, Italy, Sep. 2007.
- [51] K. T. Wong, “Direction finding/polarization estimation-dipole and/or loop triad(s),” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 37, no. 2, pp. 679–684, Apr. 2001.
- [52] K. T. Wong, X. Yuan, “ ‘Vector cross-product direction-finding’ with an electromagnetic vector-sensor of six orthogonally oriented but spatially noncollocating dipoles/loops,” *IEEE Transactions on Signal Processing*, Vol. 59, no. 1, pp. 160–171, Jan. 2011.
- [53] J. Tabrikian, R. Shavit and D. Rahamim, “An efficient vector sensor configuration for source localization,” *IEEE Signal Processing Letters*, Vol. 11, no. 8, pp. 690–693, Aug. 2004.
- [54] X. Yuan, K. T. Wong, Z. Xu and K. Agrawal, “Various compositions to form a triad of collocated dipoles/loops, for direction finding and polarization estimation,” *IEEE Sensors Journal*, vol. 12, no. 6, pp. 1763–1771, Jun. 2012.
- [55] H. Xin and J. Ding, “An improved two-antenna direction of arrival (DOA) technique inspired by human ears,” *IEEE Antennas and Propagation Society International Symposium*, San Diego, CA, USA, Jul. 2008.
- [56] N. Behdad, M. Li and M. Al-Joumayly, “Biologically-inspired antenna arrays based on the hearing mechanism of the parasitoid fly *Ormia Ochracea*,” *IEEE International Symposium on Antennas and Propagation (APSURSI)*, pp. 1526–1529, Spokane, WA, USA, 3–8 Jul. 2011.
- [57] W. Martinsen, “Giselle: a mutually orthogonal triple twin-loop ground symmetrical broadband receiving antenna for the HF band,” Command, Control, Communications and Intelligence Division DSTO Defence Science and Technology Organisation, July 2009.

- [58] M. Slater, “Design and analysis of direction-of-arrival estimation systems using electrically small antenna arrays,” Doctoral Thesis, Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, IL, USA, 2012.
- [59] N. M. Abbasi, “An Experimental Investigation into the Feasibility of MIMO Techniques within the HF Band”, PhD Thesis, University of Leicester, Leicester, UK.
- [60] R. Garbacz and R. H. Turpin, “A generalized expansion for radiated and scattered fields,” *IEEE Transactions on Antennas and Propagation*, vol. AP-19, no. 3, pp. 348–358, May 1971.
- [61] R. F. Harrington and J. R. Mautz, “Theory of characteristic modes for conducting bodies,” *IEEE Transactions on Antennas and Propagation*, vol. AP-19, no. 5, pp. 622–628, Sep. 1971.
- [62] R. F. Harrington and J. R. Mautz, “Computation of characteristic modes for conducting bodies,” *IEEE Transactions on Antennas and Propagation*, vol. AP-19, no. 5, pp. 629–639, Sep. 1971.
- [63] R. F. Harrington, J. R. Mautz and Y. Chang, “Characteristic modes for dielectric and magnetic bodies,” *IEEE Transactions on Antennas and Propagation*, vol. AP-20, no. 2, pp. 194–198, Mar. 1972.
- [64] M. Gustafsson and S. Nordebo, “Characterisation of MIMO antennas using spherical vector waves,” *IEEE Transactions on Antenna and Propagation*, vol. 54, no. 9, pp. 2679–2682, Sep. 2006.
- [65] M. Gustafsson and S. Nordebo, “On the spectral efficiency of a sphere,” *Progress In Electromagnetics Research*, vol. 67, pp. 275–296, 2007.
- [66] A. A. Glazunov, M. Gustafsson, A. F. Molisch and F. Tufvesson, “Physical modeling of multiple-input multiple-output antennas and channels by means of the spherical vector wave expansion,” *IET Microwaves, Antennas and Propagation*, vol. 4, no. 6, pp. 778–791, Jun. 2010.
- [67] S. Nordebo, M. Gustafsson and J. Lundback, “Fundamental limitations for DOA and polarization estimation with applications in array signal processing”, *IEEE Transactions on Signal Processing*, vol. 54, no. 10, pp. 4055–4061, Oct. 2006.
- [68] G. A. Akers and C. F. Corbin, “The effect of the number of modes and feed locations in angle-of-arrival estimation using a multimode antenna,” *IEEE National Aerospace and Electronics Conference*, pp. 65–69, Fairborn, OH, USA, 14–16 Jul. 2010.

- [69] H. Li, Z. T. Miers and B. K. Lau, “Design of orthogonal MIMO handset antennas based on characteristic mode manipulation at frequency bands below 1 GHz,” *IEEE Transactions on Antennas and Propagation*, vol. 62, no. 5, pp. 2756–2766, May 2014.
- [70] S. K. Chaudhury, H. J. Chaloupka and A. Ziroff, “Multiport antenna systems for MIMO and diversity,” *Proceedings of the Fourth European Conference on Antennas and Propagation*, Barcelona, Spain, 12-16 Apr. 2010.
- [71] A. S. Y. Poon, R. W. Brodersen and D. N. C. Tse, “Degrees of freedom in multiple antenna channels: a signal space approach,” *IEEE Transactions on Information Theory*, vol. 51, no. 2, pp. 523-536, Feb. 2005.
- [72] O. M. Bucci, C. Gennarelli and C. Savarese, “Representation of electromagnetic fields over arbitrary surfaces by a finite and nonredundant number of samples,” *IEEE Transactions on Antenna and Propagation*, vol. 46, no. 3, pp. 351-359, Mar. 1998.
- [73] M. Migliore, “An intuitive electromagnetic approach to MIMO communication systems,” *IEEE Antennas and Propagation Magazine*, vol. 48, no. 3, pp. 128-137, Jun. 2006.
- [74] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, New Jersey: Prentice Hall, 1993.
- [75] R. Q. Twiss, “Nyquist’s and Thevenin’s theorems generalized for nonreciprocal linear networks,” *Journal of Applied Physics*, vol. 26, no. 5, pp. 599–602, May. 1955.
- [76] P. Stoica and A. Nehorai, “Performance study of conditional and unconditional direction-of-arrival estimation,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, no. 10, pp. 1783–1795, Oct. 1990.
- [77] J. L. T. Ethier and D. A. McNamara, “Antenna shape synthesis without prior specification of the feedpoint locations,” *IEEE Transactions on Antennas and Propagation*, vol. 62, no. 10, pp. 4919–4934, Oct. 2014.
- [78] B. Yang and J. Adams, “Systematic shape optimization of symmetric MIMO antennas using characteristic modes,” *IEEE Transactions on Antennas and Propagation*, vol. PP, no. 99, Aug. 2015.
- [79] L. Kundu, “Information-theoretic limits on MIMO antennas,” Doctoral Thesis, Department of Electrical and Computer Engineering, North Carolina State University, USA, Sep. 2016.
- [80] P. Stoica and T. L. Marzetta, “Parameter estimation problems with singular information matrices,” *IEEE Transactions on Signal Processing*, vol. 49, no. 1, Jan. 2001.

- [81] A. Elsherbeni, P. Nayeri and C. J. Reddy, “Antenna analysis and design using FEKO electromagnetic simulation software,” *SciTech Publishing*, Edison, NJ, 2014.
- [82] M. L. Morris and M. A. Jensen, “Network model for MIMO systems with coupled antennas and noisy amplifiers,” *IEEE Transactions on Antennas and Propagation*, vol. 53, no. 1, pp. 545–552, Jan. 2005.
- [83] <https://www.maximintegrated.com/en/products/comms/wireless-rf/MAX2644.html>
- [84] A. Dogandzic and A. Nehorai, “Cramér-Rao bounds for estimating range, velocity, and direction with an active array,” *IEEE Transactions on Signal Processing*, vol. 49, no. 6, pp. 1122–1137, Jun. 2001.
- [85] R. T. Compton Jr., “The tripole antenna: an adaptive array with full polarization flexibility,” *IEEE Transactions on Antennas and Propagation*, vol. 29, no. 6, pp. 944–952, Nov. 1981.
- [86] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, UK: Cambridge University Press, 1990.
- [87] R. F. Harrington, J. R. Mautz and Y. Chang, “Characteristics modes for dielectric and magnetic bodies,” *IEEE Transactions on Antenna and Propagation*, vol. AP-20, no. 2, pp. 194-198, Mar. 1972.
- [88] R. F. Harrington and J. R. Mautz, “Computation of characteristics modes for conducting bodies,” *IEEE Trans. Ant. and Prop.*, vol. AP-19, no. 5, pp. 629–639, Sep. 1971.
- [89] D. H. Staelin, A. Morgenthaler and J. Au, *Electromagnetic Waves*, 1st edition, Prentice Hall, 1998.
- [90] S. R. Garcia and R. A. Horn, *A Second Course on Linear Algebra*, Cambridge: Cambridge University Press, 2017.
- [91] R. F. Harrington, “Matrix methods for field problems,” *Proceedings of the IEEE*, vol. 55, no. 2, pp. 136–149, Feb. 1967.
- [92] C. A. Balanis, *Antenna Theory Analysis and Design*, 2nd ed. New York: Wiley, 1997.
- [93] H. W. Bode, *Network Analysis and Feedback Amplifier Design*. Van Nostrand: New York, 1945.
- [94] R. M. Fano, “Theoretical limitations on the broad-band matching of arbitrary impedances,” *Journal of the Franklin Institute*, vol. 249, pp. 57-83, Jan. 1950.

APPENDIX

Appendix A

Proof of theorem 2.1

First observe $\mathcal{I}(\mathbf{u}; \Sigma_{\mathbf{n}}) \geq \mathcal{I}(\mathbf{u}; \Sigma_{\mathbf{n}'})$ only if $\mathcal{I}(\mathbf{u}; \Sigma_{\mathbf{n}}) - \mathcal{I}(\mathbf{u}; \Sigma_{\mathbf{n}'})$ is positive semi-definite.

It thus suffices [86, pg. 402] to show $\mathbf{w}^H [\mathcal{I}(\mathbf{u}; \Sigma_{\mathbf{n}}) - \mathcal{I}(\mathbf{u}; \Sigma_{\mathbf{n}'})] \mathbf{w} \geq 0$ for all $\mathbf{w} \in \mathbb{C}^p$.

From (2.13), we observe

$$\begin{aligned} \mathbf{w}^H \mathcal{I}(\mathbf{u}; \Sigma_{\mathbf{n}}) \mathbf{w} &= \sum_{ij} w_i^* w_j \left\{ \frac{\partial \boldsymbol{\mu}_{\mathbf{s}}^H}{\partial u_i} \Sigma_{\mathbf{y}}^{-1}(\mathbf{u}) \frac{\partial \boldsymbol{\mu}_{\mathbf{s}}}{\partial u_j} + \left(\frac{\partial \boldsymbol{\mu}_{\mathbf{s}}^H}{\partial u_i} \Sigma_{\mathbf{y}}^{-1}(\mathbf{u}) \frac{\partial \boldsymbol{\mu}_{\mathbf{s}}}{\partial u_j} \right)^* \right. \\ &\quad \left. + \text{tr} \left[\Sigma_{\mathbf{y}}^{-1}(\mathbf{u}) \frac{\partial \Sigma_{\mathbf{s}}(\mathbf{u})}{\partial u_i} \Sigma_{\mathbf{y}}^{-1}(\mathbf{u}) \frac{\partial \Sigma_{\mathbf{s}}(\mathbf{u})}{\partial u_j} \right] \right\} \\ &= \mathbf{x}^H \Sigma_{\mathbf{y}}^{-1}(\mathbf{u}) \mathbf{x} + \bar{\mathbf{x}}^H \Sigma_{\mathbf{y}}^{-1}(\mathbf{u}) \bar{\mathbf{x}} + \text{tr} \left[\Sigma_{\mathbf{y}}^{-1}(\mathbf{u}) \mathbf{A}^H \Sigma_{\mathbf{y}}^{-1}(\mathbf{u}) \mathbf{A} \right] , \quad (\text{A.1}) \end{aligned}$$

where the first step follows from $\Sigma_{\mathbf{y}}(\mathbf{u}) = \Sigma_{\mathbf{s}}(\mathbf{u}) + \Sigma_{\mathbf{n}}$, and the second from

$$\mathbf{x} \triangleq \sum_j w_j \frac{\partial \boldsymbol{\mu}_{\mathbf{s}}}{\partial u_j} , \quad \bar{\mathbf{x}} \triangleq \sum_j w_j^* \frac{\partial \boldsymbol{\mu}_{\mathbf{s}}}{\partial u_j} , \quad \mathbf{A} \triangleq \sum_j w_j \frac{\partial \Sigma_{\mathbf{s}}(\mathbf{u})}{\partial u_j} ,$$

and observing $\partial \Sigma_{\mathbf{s}}(\mathbf{u}) / \partial u_j$ is Hermitian and $\bar{\mathbf{x}}^H \Sigma_{\mathbf{y}}^{-1}(\mathbf{u}) \bar{\mathbf{x}}$ is real. Note \mathbf{x} , $\bar{\mathbf{x}}$, and \mathbf{A} do not depend on $\Sigma_{\mathbf{n}}$.

Defining $\Sigma_{\mathbf{y}'}(\mathbf{u}) \triangleq \Sigma_{\mathbf{s}}(\mathbf{u}) + \Sigma_{\mathbf{n}'}$, we note $\Sigma_{\mathbf{n}} \leq \Sigma_{\mathbf{n}'}$ implies $\Sigma_{\mathbf{y}}(\mathbf{u}) \leq \Sigma_{\mathbf{y}'}(\mathbf{u})$ and hence $\Sigma_{\mathbf{y}}^{-1}(\mathbf{u}) \geq \Sigma_{\mathbf{y}'}^{-1}(\mathbf{u})$. It follows

$$\mathbf{x}^H \Sigma_{\mathbf{y}}^{-1}(\mathbf{u}) \mathbf{x} \geq \mathbf{x}^H \Sigma_{\mathbf{y}'}^{-1}(\mathbf{u}) \mathbf{x}, \quad \bar{\mathbf{x}}^H \Sigma_{\mathbf{y}}^{-1}(\mathbf{u}) \bar{\mathbf{x}} \geq \bar{\mathbf{x}}^H \Sigma_{\mathbf{y}'}^{-1}(\mathbf{u}) \bar{\mathbf{x}}. \quad (\text{A.2})$$

For any complex matrix \mathbf{D} with M columns, $\Sigma_{\mathbf{y}}^{-1}(\mathbf{u}) \geq \Sigma_{\mathbf{y}'}^{-1}(\mathbf{u})$ also implies each diagonal element of $\mathbf{D}^H [\Sigma_{\mathbf{y}}^{-1}(\mathbf{u}) - \Sigma_{\mathbf{y}'}^{-1}(\mathbf{u})] \mathbf{D}$ is positive, so

$$\text{tr} [\mathbf{D}^H \Sigma_{\mathbf{y}}^{-1}(\mathbf{u}) \mathbf{D}] \geq \text{tr} [\mathbf{D}^H \Sigma_{\mathbf{y}'}^{-1}(\mathbf{u}) \mathbf{D}]. \quad (\text{A.3})$$

Let $\Sigma_{\mathbf{y}}^{-1/2}(\mathbf{u})$ denote the Hermitian square root [86, Thrm. 7.2.6] of $\Sigma_{\mathbf{y}}(\mathbf{u})$. We can bound the right-most term in (A.1) by

$$\begin{aligned} \text{tr} [\Sigma_{\mathbf{y}}^{-1}(\mathbf{u}) \mathbf{A}^H \Sigma_{\mathbf{y}}^{-1}(\mathbf{u}) \mathbf{A}] &= \text{tr} \left[\left(\mathbf{A} \Sigma_{\mathbf{y}}^{-1/2}(\mathbf{u}) \right)^H \Sigma_{\mathbf{y}}^{-1}(\mathbf{u}) \mathbf{A} \Sigma_{\mathbf{y}}^{-1/2}(\mathbf{u}) \right] \\ &\geq \text{tr} \left[\left(\mathbf{A} \Sigma_{\mathbf{y}}^{-1/2}(\mathbf{u}) \right)^H \Sigma_{\mathbf{y}'}^{-1}(\mathbf{u}) \mathbf{A} \Sigma_{\mathbf{y}}^{-1/2}(\mathbf{u}) \right] \\ &= \text{tr} \left[\left(\mathbf{A}^H \Sigma_{\mathbf{y}'}^{-1/2}(\mathbf{u}) \right)^H \Sigma_{\mathbf{y}}^{-1}(\mathbf{u}) \mathbf{A}^H \Sigma_{\mathbf{y}'}^{-1/2}(\mathbf{u}) \right] \\ &\geq \text{tr} \left[\left(\mathbf{A}^H \Sigma_{\mathbf{y}'}^{-1/2}(\mathbf{u}) \right)^H \Sigma_{\mathbf{y}'}^{-1}(\mathbf{u}) \mathbf{A}^H \Sigma_{\mathbf{y}'}^{-1/2}(\mathbf{u}) \right] \\ &= \text{tr} [\Sigma_{\mathbf{y}'}^{-1}(\mathbf{u}) \mathbf{A}^H \Sigma_{\mathbf{y}'}^{-1}(\mathbf{u}) \mathbf{A}], \end{aligned} \quad (\text{A.4})$$

where each inequality follows from (A.3), and each equality from the identity $\text{tr} [\mathbf{BC}] = \text{tr} [\mathbf{CB}]$. Combining (A.1), (A.2) and (A.4), we have proved $\mathbf{w}^H [\mathcal{I}(\mathbf{u}; \Sigma_{\mathbf{n}}) - \mathcal{I}(\mathbf{u}; \Sigma_{\mathbf{n}'})] \mathbf{w} \geq 0$. Since $\mathbf{w} \in \mathbb{C}^p$ was arbitrary, conclude $\mathcal{I}(\mathbf{u}; \Sigma_{\mathbf{n}}) \geq \mathcal{I}(\mathbf{u}; \Sigma_{\mathbf{n}'})$. \diamond