ABSTRACT

DOMIZIOLI, CARLO P. Noise Analysis and Low-Noise Design for Compact Multi-Antenna Receivers: A Communication Theory Perspective. (Under the direction of Dr. Brian L. Hughes).

Multiple-input, multiple-output (MIMO) systems combine the deployment of multiple antennas at both the transmitter and receiver with sophisticated signal processing to improve the performance of wireless communications. As with any communication system, developing an accurate yet mathematically tractable channel model is essential to analyzing the performance of actual systems. Prior studies of MIMO channel modeling have provided detailed models for fading correlation – either due to the propagation environment or through mutual coupling between the antennas – and how this correlation affects performance. On the other hand, relatively little attention has been paid to the noise correlation.

In this dissertation we consider noise analysis and low-noise design for compact MIMO receivers. We begin by analyzing the performance of several MIMO communication schemes in the presence of fading and noise correlation. It is shown that fading and noise correlation have opposite effects on performance, so properly accounting for both in the channel model is crucial to accurately predicting performance. Next we develop a circuit model for compact multi-antenna receivers that includes noise generated by the antennas, front-end amplifiers, and other components. Through analytical and numerical examples we demonstrate that the noise may be correlated, and that different noise sources may impact performance in profoundly different ways. Finally, we derive low-noise design theorems for MIMO front-ends from communication-theoretic principles.

Noise Analysis and Low-Noise Design for Compact Multi-Antenna Receivers:
A Communication Theory Perspective

by
Carlo Peter Domizioli

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

Electrical Engineering

Raleigh, North Carolina

2009

APPROVED BY:

Dr. J. Keith Townsend	Dr. Stephen L. Campbell
Dr. Kevin G. Gard	Dr. Gianluca Lazzi
Dr. Brian L. Hughes Chair of Advisory Committee	

BIOGRAPHY

Carlo P. Domizioli received the B.S. degree in electrical engineering from Tennessee Technological University in May 2005. In August 2005 he began working towards his PhD in electrical engineering at North Carolina State University under the advisement of Prof. Brian L. Hughes. During his career at NC State, Carlo studied the modeling and impact of receiver noise sources in multi-antenna wireless communication systems, resulting in the publication of several conference and journal papers. He was also involved in experimental work on underwater optical communications. Upon completion of his PhD studies in August 2009, Carlo began his professional career as a communication systems engineer for Northrop Grumman in Fairfax, VA.

ACKNOWLEDGMENTS

I would first like to thank my advisor, Prof. Brian Hughes, for his support and guidance throughout my PhD career. I have benefited tremendously from not only his technical expertise, but also from observing how he clearly communicates those technical concepts – a combination of attributes found too seldom among engineers. In addition, I always appreciated the careful balance he maintained in suggesting interesting research avenues to explore, while also allowing me the freedom to pursue problems of my own choosing.

The multidisciplinary approach of this dissertation was made possible through numerous discussions with Profs. Kevin Gard and Gianluca Lazzi, whose respective expertise in RF design and electromagnetics allowed me to expand my research area into these fields. Ultimately, none of this research would be possible without the funding we received from the National Science Foundation, in particular, grants CCF-0312686, CCF-0515164, and CCF-0728803.

One of the benefits of a PhD program is the opportunity to surround oneself with students of different backgrounds, yet similar ambitions. In this regard I would like to thank my fellow research group members, Yuhan Dong and Pawandeep Singh, for the years of collaboration and camaraderie. I am also thankful to one of our undergraduate students, Kiran Karra, for developing the Python scripts used in our numerical electromagnetics computations. I would also like to thank William Cox and Jim Simpson for all that I learned from our many useful discussions during experimental work on underwater optical communications.

Finally, I owe the most gratitude to my wife Rachel, whose constant support and encouragement sustained me through the many highs and lows of graduate school. Being a "PhD widow" is no doubt trying on the strongest of marriages, but her love and dedication allowed me the opportunity to pursue my ambitions while we built our life together. This dissertation is dedicated to her.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

Introduction

In this dissertation we consider noise analysis and low-noise design for multiple-input, multiple-output (MIMO) wireless communication systems. We approach this topic from a communication theory perspective by first extending the MIMO channel model to include noise correlation (Chapter 2), then specializing the noise model to include detailed circuit models of various receiver components (Chapter 3), and finally, by developing theorems for low-noise design from communication-theoretic principles (Chapter 4). In this chapter we introduce the MIMO channel model, discuss some prior literature, and present an outline of the dissertation.

MIMO systems combine the deployment of multiple antennas at both the transmitter and receiver with sophisticated signal processing to improve the performance of wireless communication systems. The interest in MIMO was sparked over a decade ago by the pioneering studies of Telatar [82] and Foschini and Gans [29], which predicted that large spectral efficiencies could be achieved by using antenna arrays to communicate over

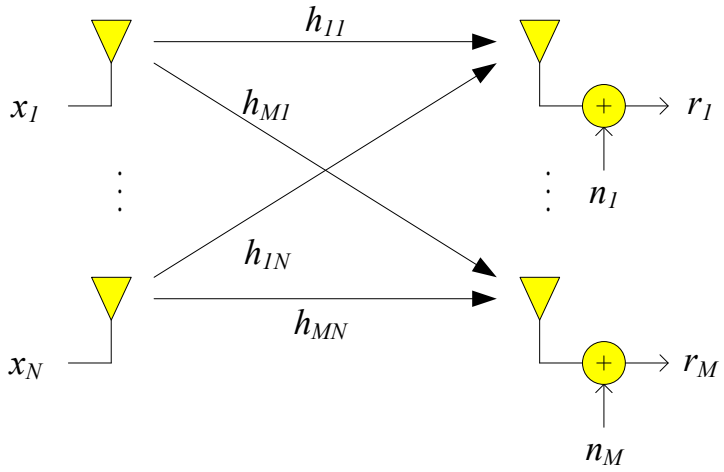


Figure 1.1: MIMO channel model.

fading multipath channels. These results spurred an unprecedented surge in scholarly research, and led to the development of commercial wireless systems. For example, recent revisions to international standards for wireless local area networks (WLANs) and metropolitan area networks (WMANs) – IEEE 802.11n and 802.16e, respectively[1] – both use MIMO signaling to vastly improve data rates over previous standards. As with any communication system, developing an accurate yet mathematically tractable channel model is essential to analyzing the performance of actual systems.

A channel model for a MIMO system with N transmit and M receive antennas is shown in Fig. 1.1 (cf., [33, Ch. 10], [83, Ch. 7]). The details and underlying assumptions of this model will be developed in Chapter 2; here we briefly discuss some key concepts. The signals $x_1, \ldots, x_N$ sent over the transmit antennas are corrupted by two channel

[1]The 802.11 and 802.16 standards define the physical (PHY) and medium access control (MAC) layers of WiFi and WiMAX wireless networks, respectively.

impairments: fading and noise. The fading path gains h_{ij} model the random power fluctuations that result from constructive and destructive interference of the scattered radio waves. As we shall see in the next section, there is a rich literature on statistical models for the fading path gains. On the other hand, the modeling and impact of the noise n_i has received relatively little attention. These topics, in addition to optimal low-noise design principles, form the main contribution of this dissertation.

1.1 Prior Work

Early studies on MIMO modeled the fading path gains h_{ij} as independent random variables. While this assumption may be valid if the antennas are separated by several wavelengths, in compact mobile devices such large spacings may be impractical. This observation led to a flood of research activity that provided detailed models for how the fading path gains may become correlated – either through the propagation environment or through *mutual coupling* between the antennas – and how this correlation affects performance. In this section we provide a brief summary of prior work in this area. We focus rather specifically on papers that provide analytical models for fading correlation and mutual coupling, and study their impact on MIMO channel capacity. Before proceeding, we note that the impact of fading correlation and mutual coupling has long been studied in diversity receivers [51], adaptive arrays [34], and various other multi-antenna systems. Therefore, a comprehensive discussion of the literature on fading correlation would take us too far afield, so we limit our coverage to selected papers that fit well within the scope of this dissertation.

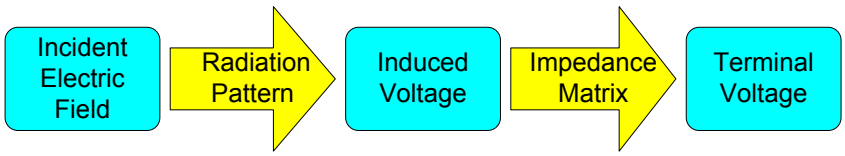


Figure 1.2: The incident electric field and terminal voltage are related by the array radiation pattern and impedance matrix.

One of the initial studies on MIMO capacity with fading correlation was provided in 2000 by Shiu *et al.* [75], which employed a classic ray-tracing propagation model previously developed for receive diversity. A key contribution of this work was the introduction of the *separable correlation model*, which models the correlation between two fading path gains as the product of two quantities: the transmit and receive correlations. Subsequent studies focused on various mathematical techniques for computing capacity under the separable correlation model. In [16] asymptotic results in random matrix theory were used to evaluate capacity for transmit and receive correlation as the number of antennas approached infinity. Other approaches, including simplifications of existing closed-form expressions and asymptotic (in SNR or the number of antennas) analyses may be found in several MIMO special issues published in 2003 [71], [72], [39]. Some of these results are neatly summarized in the monograph on random matrix theory by Tulino and Verdu [84]. Alternate models for fading correlation include rank-deficient channels [13], [32] and virtual channel representations [73].

While the above studies provide models for the propagation environment, they do not account for all of the potential interactions that may occur between the antennas. As illustrated in Fig. 1.2, the voltage induced in an antenna array by an incident electric

field depends on the array *radiation pattern*, and this induced voltage is related to the terminal (observed) voltage by the array *impedance matrix*. For certain arrays, e.g., dipole antennas separated by several wavelengths, these relationships are trivial and the terminal voltage is simply a scalar multiple of the incident field. However, in compact arrays several complications may occur. First, the radiation pattern of each antenna may become distorted due to its close proximity to other antennas. Second, current flowing in one antenna may induce a voltage in a neighboring antenna, i.e., the antennas are coupled. This *mutual coupling* may be modeled by a non-diagonal impedance matrix, cf., [4, Ch. 8].

Several researchers were quick to notice the lack of antenna modeling in the above fading correlation studies; however, there was some initial disagreement about the impact of mutual coupling on capacity. Earlier studies [79], [52], and [92] showed that mutual coupling may decrease the correlation between fading path gains, so the authors concluded that mutual coupling may *increase* capacity. This result was somewhat surprising as mutual coupling is typically regarded as detrimental (e.g., the last section of [29]). Subsequent studies considered the impact of mutual coupling on both the power and correlation of the fading path gains, observing that an impedance mismatch between the coupled array and its terminating impedance results in a power loss. It was shown that the penalty associated with this power loss may outweigh the aforementioned benefit of decorrelating, in which case mutual coupling indeed degrades capacity. However, as demonstrated in [91], by optimizing the impedance match for maximum power transfer [21], one can not only mitigate this power loss, but provide complete decoupling of the array. On the other hand, [50] shows that this decoupling may sharply reduce band-

width. Additional theoretical and experimental works on mutual coupling in MIMO systems may be found, for example, in the 2006 special issue [43].

The studies presented so far have introduced increasingly detailed models of the fading correlation; however, little attention has been paid to the noise. In fact, most of the papers cited above assume a spatially white noise model, only mentioning noise correlation in reference to interference (cf., [56]). However, for closely-spaced antennas even noise generated in the receiver may become correlated. For example, noise in a front-end amplifier may disperse across neighboring receiver chains through coupled antennas. Some recent work has improved the noise modeling in multi-antenna receivers. A non-trivial noise model for a MIMO receiver was considered in [58], where the authors introduced a standard amplifier model into their previous model for mutual coupling [91]. They demonstrated that matching networks optimized for noise figure outperformed those designed for maximum power transfer. More recently, Gans [30], [31] evaluated the impact of antenna thermal noise and (spatially white) amplifier noise individually. It was shown that matching networks offer no improvement in capacity when antenna noise is dominant, while they have a tremendous impact on capacity when amplifier noise is dominant.

In summary, over the last decade the modeling of fading correlation in MIMO has progressed from the simple independent fading model used in the pioneering studies, to separable models from which transmit and receive correlation may be derived, and finally to models that account for the array radiation pattern and impedance matrix. Recent studies have also included more realistic noise models for the antennas and amplifiers. As a consequence of this progress, the study of compact MIMO transceivers has become increasingly multidisciplinary, combining elements from communication theory,

electromagnetics, and RF circuit design.

1.2 Overview of Dissertation

As evidenced by the above literature review, far more attention has been devoted to fading correlation in MIMO than noise correlation. Since fading and noise play equal roles in determining most system performance metrics, a channel model that does not correctly represent noise cannot be expected to correctly predict performance. In this dissertation we focus on noise analysis, noise modeling, and low-noise design for compact MIMO receivers.

We begin in the first two sections of Chapter 2 with an overview of some of the basic assumptions underlying the channel model used in our work and most of the prior literature. The remainder of the chapter is devoted to the analysis of some common MIMO communication schemes in the presence of fading and noise correlation. It is shown that fading and noise correlation have opposite effects on performance, so properly accounting for both in the channel model is crucial to accurately predicting performance.

In Chapter 3 we develop a noise model for compact multi-antenna receivers. The model includes noise generated by the antennas, front-end amplifiers, and other receiver components. Using this model we show that noise in a compact multi-antenna receiver may be spatially correlated. Expressions relating noise correlation to properties of the antennas and amplifiers are derived and the impact of noise correlation on the outage probability of an optimal diversity combiner is studied. Examples illustrating the relationship between mutual coupling and noise correlation demonstrate how different noise

sources may impact performance in profoundly different ways. Other topics, including the impact of matching, amplifier unilaterality, and directional fading and sky noise, are also considered.

In Chapter 4 we combine the communications schemes presented Chapter 2 with a generalization of the noise model introduced in Chapter 3. Our main task is to develop low-noise design principles for compact multi-antenna receivers. We proceed by first examining the performance metrics of several MIMO communication schemes, from which we determine a common optimality criterion. The result is a natural generalization of the well-known concept of minimum noise figure design for *single-antenna* receivers. With this optimality criteria we are able to prove the optimality of a matching network that was only conjectured to be optimal in prior studies. We also show that two amplifiers with the same noise figure may perform very differently in the presence of coupled antennas.

Finally, in Chapter 5 we summarize our main findings and make some suggestions for future work. Appendix A provides details for the numerical electromagnetics simulations we used to compute the antenna array radiation pattern and impedance matrix in Chapters 3 and 4, and also highlights some of the different assumptions underlying this technique and that used in some of the related literature.

Chapter 2

MIMO Communication with Correlated Fading and Noise

We begin our study by introducing noise correlation into the MIMO channel model and examining its impact on the performance of several MIMO communication schemes. In addition to providing background material, the main purpose of this chapter is to demonstrate that fading and noise correlation play equal roles in determining most system performance metrics. This observation motivates the detailed noise models considered in the remainder of this dissertation.

A stochastic wireless channel model specifies the fading and noise statistics as a function of time, frequency, and space. Since the focus of this dissertation is on spatial correlation, we shall take a simplistic view of both the time and frequency coherence of fading and noise. These topics are introduced in the context of single-input, single-output (SISO) systems in the first section. In the next section we present a well-known model

for the spatial correlation of fading. The spatial coherence of noise is one of the central topics of this thesis, so we postpone a detailed discussion of this subject to later chapters.

The purpose of these first two sections is to place the model used throughout this dissertation within the context of a much broader class of wireless channel models, so we are concise and selective in our presentation. More complete introductions to wireless channels are available in the many textbooks on the subject. A classic treatment is Jakes' text [44], which is based on research conducted at Bell Laboratories during their development of cellular systems in the 1960s and early 1970s. More modern treatments are available in the introductory chapters of the texts by Rappaport [69], Goldsmith [33], and Tse and Viswanath [83]. Texts devoted specifically to MIMO include [64], [6], and [62], with the latter providing the strongest emphasis on channel modeling. The introductory communications text by Proakis [67] also includes a chapter on wireless channels.

After introducing the channel model, we next consider the form several popular MIMO communication schemes take under spatially correlated fading and noise. Each scheme is introduced within the context of its requirements on both the channel knowledge and maximum codeword length. We finish the chapter by evaluating the performance of each communication scheme under a simple exponential correlation model, which, while not based on any specific physical model, lends itself to simple interpretation and captures the general effect of correlated fading and noise. It is shown that fading and noise correlation have quite opposite effects on performance, so properly accounting for both of them in a channel model is crucial to accurately predicting performance.

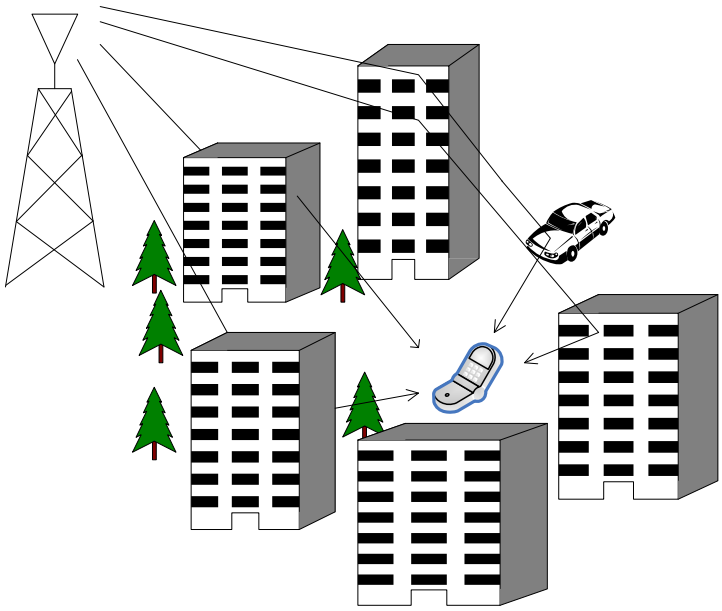


Figure 2.1: Illustration of wireless propagation in an urban environment.

2.1 SISO Channel Model

In this section we present a channel model for SISO systems, which is the basis for the MIMO channel used throughout this dissertation. Most of the results of this section can be found in wireless communication textbooks; we shall follow [83, Ch. 2].

A typical wireless propagation environment is illustrated in Fig. 2.1. A complex baseband signal $x(t)$ transmitted over the channel scatters throughout the environment and arrives at the receiver as a large number of *multipath components*. Assuming the signal is further corrupted at the receiver by additive white Gaussian noise (AWGN) with power spectral density $\frac{N_0}{2}$, we may model the received signal by

$$r(t) = \sum_i \alpha_i(t) e^{-j2\pi f_c \tau_i(t)} x(t - \tau_i(t)) + n(t) \ , \tag{2.1}$$

where $\alpha_i(t)$ and $\tau_i(t)$ are the attenuation and propagation delay of the i^{th} multipath component at time t and $n(t)$ is complex baseband AWGN. The propagation delay affects both the complex baseband signal and its carrier $e^{j2\pi f_c t}$, where f_c is the carrier frequency in Hz. We can view the wireless channel (2.1) as a time-varying linear filter channel, $r(t) = \int h(t,u)x(u)du + n(t)$, where the impulse response is

$$h(t,u) = \sum_i \alpha_i(t) e^{-j2\pi f_c \tau_i(t)} \delta(u - t + \tau_i(t)) \ , \tag{2.2}$$

and $\delta(t)$ is the Dirac-delta function.

Two important time-domain characterizations of the channel model (2.1) are the *delay spread* and *coherence time*. The delay spread T_d is the maximum time delay between two multipath components[1],

$$T_d \triangleq \max_{i,j,t} |\tau_i(t) - \tau_j(t)| \ . \tag{2.3}$$

The coherence time T_c is roughly defined as the largest time interval for which both $\alpha_i(t)$ and $\tau_i(t)$ are approximately constant. A common convention is to take the coherence time as the reciprocal of the largest difference between the Doppler shifts $f_c \frac{d}{dt}\tau_i$ of each multipath, referred to as the *Doppler spread*

$$D_s \triangleq \max_{i,j,t} f_c \left| \frac{d}{dt}(\tau_i - \tau_j) \right| \ . \tag{2.4}$$

Typical values for T_d and T_c are on the order of microseconds and milliseconds, respectively, and channels for which $T_d \ll T_c$ are referred to as *underspread*. From (2.2) we see that an impulse transmitted at any time $u \in [0, T_c]$ over an underspread channel will be

[1]It is assumed that the propagation delays and attenuations are continuously differentiable functions defined on some finite interval.

received as

$$\sum_i \alpha_i(0) e^{-j2\pi f_c \tau_i(0)} \delta(u - t + \tau_i(0)) \ . \tag{2.5}$$

We may therefore model an underspread wireless channel by a sequence of linear, time-invariant (LTI) channels that change every T_c seconds.

Practical wireless channels are constrained in bandwidth, so we may modify the infinite-bandwidth impulse response (2.5) to reflect a bandpass bandwidth B by cascading it with an ideal lowpass filter. The resulting channel is

$$h_B(t) = \sum_i \alpha_i e^{-j2\pi f_c \tau_i} \mathrm{sinc}\left[B(t - \tau_i)\right] \ , \tag{2.6}$$

where $\mathrm{sinc}\, t \triangleq \frac{\sin \pi t}{\pi t}$ and we have dropped the time arguments from the attenuations and propagation delays. The reciprocal of T_d is often referred to as the *coherence bandwidth*, and will be denoted by B_c. If $B \ll B_c$, then $B\tau_i \approx B\tau_1$ for all i, and we may write

$$h_B(t) \approx \mathrm{sinc}\left[B(t - \tau_1)\right] \sum_i \alpha_i e^{-j2\pi f_c \tau_i} \ . \tag{2.7}$$

This is referred to as a *frequency-flat* channel since the Fourier transform of (2.7) is constant over $[-\frac{B}{2}, \frac{B}{2}]$. On the other hand, if $B \gg B_c$ we are able to resolve the individual multipaths in (2.6) and the channel is said to be *frequency-selective*. Although we shall use a frequency-flat channel throughout this dissertation, this does not imply we are only considering low-bandwidth applications. By using orthogonal frequency division multiplexing (OFDM), one can divide a frequency-selective wideband channel into a set of frequency-flat narrowband channels. For example, the IEEE 802.16e WMAN standard [42] uses OFDM to divide a 10 MHz channel into 10.9 kHz subchannels.

The complex baseband signal $x(t)$ has bandwidth $\frac{B}{2}$, so by the Nyquist sampling theorem it is completely specified by samples $x[k] \triangleq x(k/B)$ taken every $\frac{1}{B}$ seconds.

Thus we may model the underspread, frequency-flat channel in discrete time by the *block fading* model

$$r[k] = hx[k] + n[k] \ , \quad k = 0, 1, \ldots, K-1 \ . \tag{2.8}$$

Here $h \triangleq \sum_i \alpha_i e^{-j2\pi f_c \tau_i}$ is the *fading path gain*, which remains constant over the block length[2] $K \triangleq \lfloor T_c B \rfloor$, and recalling the noise model we see that $n[k] \triangleq n(k/B)$ is an independent and identically distributed (i.i.d.) sequence of zero-mean, circularly symmetric, complex Gaussian (ZMCSCG) random variables with variance $N_0 B$, denoted by $n[k] \sim \mathcal{CN}(0, N_0 B)$.

The fading path gain h in (2.8) depends on the amplitude and propagation delay of each multipath component. Usually detailed information about each multipath is not available, so it is fruitful to model h as a random variable. In most practical systems $f_c \tau_i \gg 1$, so the phase $2\pi f_c \tau_i$ of each multipath component is well-modeled by a uniformly-distributed random variable on $[0, 2\pi]$. Assuming a *rich scattering environment* in which there are a large number of i.i.d. multipaths, by the central limit theorem $h \sim \mathcal{CN}(0, \sum_i \mathrm{E}[\alpha_i^2])$, where $\mathrm{E}[\cdot]$ denotes the expectation. This is referred to as *Rayleigh fading*, since $|h|$ has a Rayleigh distribution. Other distributions for h may be derived for channels with non-identically distributed multipath components, e.g., Rician fading for propagation environments with a line-of-sight (LOS) component, but we shall restrict our attention to Rayleigh fading.

Throughout this dissertation no explicit appeal to the time domain is required, so we

[2]Recalling our earlier discussion on typical values for delay spread and coherence time, and the assumptions for underspread ($T_d \ll T_c$), frequency-flat ($B \ll B_c = \frac{1}{T_d}$) channels, we see that block lengths on the order of a hundred are typical.

drop the time indices in (2.8) and simply write the SISO channel model as

$$r = hx + n \ , \tag{2.9}$$

where $h \sim \mathcal{CN}(0,1)$ and $n \sim \mathcal{CN}(0, N_0 B)$. We have normalized the fading path gain power to unity by absorbing the signal attenuation into the transmit power $P \triangleq \mathrm{E}|x|^2$. In writing (2.9) it is implied that h remains fixed over K consecutive uses of the channel and n is i.i.d. for each use of the channel.

2.2 MIMO Channel Model

Now consider a MIMO system with N transmit and M receive antennas, which we hereafter refer to as an $N \times M$ *system*. We assume the channel between each transmit and receive antenna pair is identical to the SISO model (2.9), so the MIMO channel may be expressed in vector form as

$$\mathbf{r} = \mathbf{H}\mathbf{x} + \mathbf{n} \ , \tag{2.10}$$

where $\mathbf{r} \in \mathcal{C}^M$ is the received signal, $\mathbf{H} \in \mathcal{C}^{M \times N}$ is the *channel matrix*, $\mathbf{x} \in \mathcal{C}^N$ is the transmitted signal, and $\mathbf{n} \in \mathcal{C}^M$ is noise. The transmit power is $P = \mathrm{tr}(\mathbf{\Sigma_x})$, where $\mathbf{\Sigma_x} \triangleq \mathrm{E}[\mathbf{x}\mathbf{x}^\dagger]$ is the *transmit correlation matrix*, $\mathrm{tr}(\cdot)$ is the trace, and the $\dagger$ superscript denotes the conjugate-transpose. We shall use $*$ and T superscripts to denote conjugation and transposition, respectively, so $\mathbf{A}^\dagger = (\mathbf{A}^*)^T$. The elements of the channel matrix and noise vector are jointly ZMCSCG, so it remains to specify the correlation between these elements, i.e., the *spatial correlation*. In this section we focus on the channel matrix correlations; in later chapters we will determine the *noise correlation matrix* $\mathbf{\Sigma_n} \triangleq \mathrm{E}[\mathbf{n}\mathbf{n}^\dagger]$

for specific receiver models.

Correlation between the elements h_{ij} of $\mathbf{H}$ may be expressed by the $MN \times MN$ block matrix

$$\mathbf{\Psi} \triangleq \begin{bmatrix} \mathrm{E}[\mathbf{h}_1\mathbf{h}_1^\dagger] & \cdots & \mathrm{E}[\mathbf{h}_1\mathbf{h}_N^\dagger] \\ \vdots & \ddots & \vdots \\ \mathrm{E}[\mathbf{h}_N\mathbf{h}_1^\dagger] & \cdots & \mathrm{E}[\mathbf{h}_N\mathbf{h}_N^\dagger] \end{bmatrix} , \tag{2.11}$$

where $\mathbf{h}_i$ denotes the i^{th} column of $\mathbf{H}$, so the $(k,l)^{\text{th}}$ element of $\mathrm{E}[\mathbf{h}_i\mathbf{h}_j^\dagger]$ is $\mathrm{E}[h_{ki}h_{lj}^*]$. Under the *separable correlation model* (cf., [62, Ch. 3]) the correlations satisfy

$$\mathrm{E}[h_{ki}h_{lj}^*] = [\mathbf{\Sigma}_\mathbf{h}^\mathbf{T}]_{ij} \cdot [\mathbf{\Sigma}_\mathbf{h}^\mathbf{R}]_{kl} , \tag{2.12}$$

where $\mathbf{\Sigma}_\mathbf{h}^\mathbf{T}$ and $\mathbf{\Sigma}_\mathbf{h}^\mathbf{R}$ are the *transmit* and *receive correlation matrices*, respectively, and $[\mathbf{A}]_{ij}$ denotes the $(i,j)^{\text{th}}$ element of a matrix $\mathbf{A}$. The concept of separable correlations was introduced in [75] for a ray-tracing propagation model, and is frequently assumed throughout the MIMO literature. Intuitively, the underlying assumption is that the correlation between signals sent from two transmit antennas is independent of the receive antenna from which they are observed and, conversely, the correlation between signals observed by two receive antennas is independent of the transmit antenna from which the signal was sent. This is typically a good approximation in rich scattering environments with no LOS between the transmitter and receiver [75]. By substituting (2.12) into (2.11), we obtain

$$\mathbf{\Psi} = \mathbf{\Sigma}_\mathbf{h}^\mathbf{T} \otimes \mathbf{\Sigma}_\mathbf{h}^\mathbf{R} , \tag{2.13}$$

where $\otimes$ denotes the Kronecker product. For this reason the assumption (2.12) is often referred to as the *Kronecker correlation model* (cf., [62, pg. 82]).

The exact form of $\mathbf{\Sigma_h^T}$ and $\mathbf{\Sigma_h^R}$ depends on the antennas and surrounding propagation environment. Generally speaking, these matrices become diagonal as both the antenna separations and the number of scatterers around the array increase. As an example, consider the downlink of the system illustrated in Fig. 2.1. The mobile receiver is surrounded by many scatterers, so antenna separations of only a few wavelengths may result in $\mathbf{\Sigma_h^R} \approx \mathbf{I}$, where $\mathbf{I}$ is the identity matrix. On the other hand, the base station is located at a higher elevation away from scatterers, so antenna separations of 10s of wavelengths may be needed for $\mathbf{\Sigma_h^T} \approx \mathbf{I}$ [75]. Of course, space constraints are much less stringent for a base station tower than for a compact mobile device, so it may still be easier to design a cellular system for less transmit correlation than receive correlation.

The main topic of this dissertation is fading and noise correlation in compact MIMO *receivers*, so $\mathbf{\Sigma_h^R}$ will typically be non-diagonal. Therefore, to simplify the analysis we will assume $\mathbf{\Sigma_h^T} = \mathbf{I}$, i.e., there is no transmit correlation. Since there will be no further need to distinguish between transmit and receive correlation, we shall simplify notation by removing the superscript from $\mathbf{\Sigma_h^R}$. By substituting $\mathbf{\Sigma_h^T} = \mathbf{I}$ and $\mathbf{\Sigma_h^R} = \mathbf{\Sigma_h}$ into (2.13), from (2.11) we obtain $\mathrm{E}[\mathbf{h}_i \mathbf{h}_j^\dagger] = \mathbf{\Sigma_h} \delta_{ij}$, where δ_{ij} is the Kronecker-Delta function. Thus for no transmit correlation the columns of $\mathbf{H}$ in (2.10) are i.i.d. ZMCSCG random vectors with correlation matrix $\mathbf{\Sigma_h}$, denoted by $\mathbf{h}_i \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma_h})$. Specific formulas relating $\mathbf{\Sigma_h}$ to the antennas and propagation model will be provided in the next chapter.

For concreteness we summarize the developments of these last two sections by restating the assumptions of our channel model

$$\mathbf{r} = \mathbf{H}\mathbf{x} + \mathbf{n} \ . \tag{2.14}$$

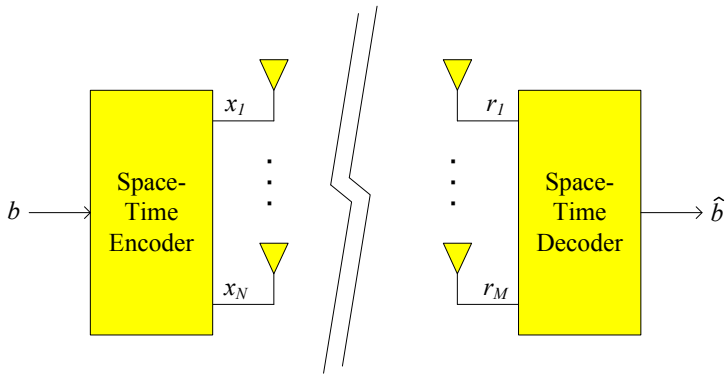


Figure 2.2: Digital communication over the MIMO channel.

The columns $\mathbf{h}_i$ of the $M \times N$ channel matrix $\mathbf{H}$ are i.i.d. $\mathcal{CN}(\mathbf{0}, \mathbf{\Sigma_h})$. The value of $\mathbf{H}$ is drawn according to this distribution, and remains fixed over K consecutive uses of the channel, i.e., during the transmission of one *block*. The power of the transmitted signal $\mathbf{x}$ is $P = \mathrm{tr}(\mathbf{\Sigma_x})$. The noise is $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma_n})$ and drawn i.i.d. for each use of the channel, so we say that the noise is *temporally white*. In general the noise correlation matrix $\mathbf{\Sigma_n}$ is non-diagonal, so we say the noise is *spatially colored*. If $\mathbf{\Sigma_n} = N_0 B\mathbf{I}$ we say that the noise is *spatially white*, and if in addition $\mathbf{\Sigma_h} = \mathbf{I}$, we shall refer to the model as the *i.i.d. fading and noise* channel.

2.3 MIMO Communication Schemes

Consider the transmission of a binary sequence b over the MIMO channel (2.14), as illustrated in Fig. 2.2. The binary sequence is mapped to the transmitted signal $\mathbf{x} = [x_1 \dots x_N]^T$ by a *space-time encoder*, and an estimate $\hat{b}$ of the data is formed from the

received signal $\mathbf{r} = [r_1 \cdots r_M]^T$ by a *space-time decoder*. In practice, temporal encoding and modulation of the binary sequence may be performed by separate devices either before or after spatial encoding (cf., [33, Fig. 10.10]), however, for our purposes we may view encoding and modulation across both space and time as a single operation. For clarity we have suppressed the D/A and A/D converters and RF front-ends required to convert the complex baseband, discrete-time channel in Fig. 2.2 to a practical analog channel. The modeling and impact of receiver front-ends will be the focus of subsequent chapters.

Different types of space-time encoders and decoders have been suggested for different constraints on the channel knowledge and codeword length. For example, the seminal works on MIMO channel capacity [82], [29], spatial multiplexing [96], space-time coding [80], and transmit beamforming [54] all make different assumptions about the channel knowledge and codeword length. In this section we evaluate the channel capacity and the performance of these systems for the MIMO channel (2.14) with correlated fading and noise. We first present a brief discussion of channel knowledge and codeword length for the block fading channel.

2.3.1 Channel Knowledge and Codeword Length

Consider a typical sample function of a fading path gain h_{ij}, illustrated in Fig. 2.3 with some of the transmitted blocks superimposed. The blocks contain *training symbols* that are used to estimate the channel matrix $\mathbf{H}$. If the receiver obtains a perfect estimate, we say that we have *channel state information at the receiver* (CSIR). Under

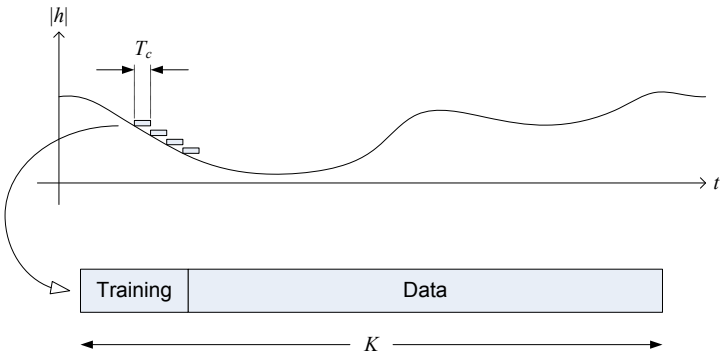


Figure 2.3: A block fading channel.

certain conditions channel state information may also be available at the transmitter, and we have *full CSI*. We shall always assume CSIR is available, so the distinguishing feature in channel knowledge is whether or not there is CSI available to the transmitter. One particular scenario that allows for transmitter CSI is a time-division duplex (TDD) system, where communication in both the forward and reverse (the roles of the transmitter and receiver are reversed) links occur at different times within each data block. In certain scenarios the channel is *reciprocal*, i.e., the channel matrix of the reverse link is $\mathbf{H}^T$, so channel estimates obtained at the receiver of the forward link may be taken as transmitter CSI for the reverse link, and vice-versa. As an example, in the IEEE 802.16e WMAN standard [42], data blocks are divided equally between the uplink and downlink channels, so both *open loop* (CSIR) and *closed loop* (full CSI) MIMO schemes are specified.

Decoding a temporally encoded signal usually introduces a time delay since the receiver may have to wait for the entire codeword to arrive before it is decoded. Therefore, a

high-level delay constraint on the data will impose a maximum length on the codewords. For example, in real-time voice applications delays of more than a few hundred milliseconds may be intolerable, while for file transfers much laxer constraints are possible. If the delay constraint is on the order of the coherence time T_c, coding over one or only a few consecutive blocks is possible, and as illustrated in Fig. 2.3, the channel matrix $\mathbf{H}$ may not change significantly. If $\mathbf{H}$ remains constant during the *entire* codeword transmission, we have a *slow fading* channel. On the other hand, if the delay constraint is orders of magnitude larger than T_c, coding over a long sequence of blocks is possible. If the sequence of channel matrices associated with these blocks is stationary and ergodic, one can show that the capacity of this channel is identical to that of a hypothetical channel in which $\mathbf{H}$ is temporally white [82] – this is referred to as *fast fading*. In practice, one could construct a fast fading channel from (2.14) by performing appropriate interleaving and de-interleaving at its input and output (cf., [83, Ch. 5]).

In summary, we can consider four distinct scenarios in which we have some combination of CSIR or full CSI with slow or fast fading. Each case will give rise to the various communication schemes we discuss below. The performance of many SISO communication schemes (2.9) is parameterized by the signal-to-noise ratio (SNR),

$$\sigma \triangleq P \cdot \frac{\mathrm{E}|h|^2}{\mathrm{E}|n|^2} = \frac{P}{N_0 B} \ . \tag{2.15}$$

The MIMO communication schemes we present below are parameterized by the following extension of SNR to the vector channel (2.14), which we will refer to as the *SNR matrix*:

$$\mathbf{\Sigma} \triangleq P \cdot \mathbf{\Sigma}_{\mathbf{h}}^{1/2} \mathbf{\Sigma}_{\mathbf{n}}^{-1} \mathbf{\Sigma}_{\mathbf{h}}^{1/2} \ , \tag{2.16}$$

where $\mathbf{A}^{1/2}$ denotes the positive definite square root of a positive definite matrix $\mathbf{A}$

(cf., [40, pg. 406]). We will have much to say about the role of the SNR matrix in later chapters, for now we just observe that it appears in each of the below metrics.

2.3.2 Ergodic Capacity

The capacity of a channel is defined as the maximum possible rate at which one can communicate with an arbitrarily low probability of error [74]. For any fixed h, the SISO channel (2.9) is an AWGN channel with SNR $|h|^2\sigma$, for which the capacity is (cf., [19, Ch. 9])

$$C_{\text{AWGN}} = \log(1 + |h|^2\sigma) \ . \tag{2.17}$$

With this result in mind, it is useful to think of a slow fading SISO channel as an AWGN channel with the SNR randomly drawn from a Rayleigh distribution. The Rayleigh probability density function (pdf) has a non-zero mass around zero, so from (2.17) it should not be surprising to learn that the capacity of the slow fading SISO channel is zero (cf., [83, pg.188]). More generally, in [82] it was shown that the capacity of the slow fading MIMO channel (2.14) is zero. As an alternative to capacity, communication engineers have found it useful to consider the *outage capacity* of slow fading MIMO channels, defined as the *probability* that one can communicate error-free at a certain rate [29]. However, we shall not consider outage capacity, opting instead to examine specific schemes designed for slow fading channels.

Due to the stationarity and ergodicity of the fast fading channel, one may average over many realizations of the channel and obtain a non-zero capacity (cf., [83, pg. 200]). To distinguish this from the outage capacity of a slow fading channel, the capacity of

a fast fading channel is often referred to as the *ergodic capacity*. The original MIMO ergodic capacity derivations for i.i.d. fading and noise with CSIR [82] and full CSI [46] show that codewords of the form $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma_x})$ are optimal. Moreover, inspection of these analyses reveals that Gaussian signaling is still optimal (i.e., mutual information is maximized) for spatially correlated fading and noise, so finding the capacity of (2.14) reduces to a transmit covariance matrix optimization. For CSIR we obtain

$$C_{\mathrm{R}} = \max_{\mathbf{\Sigma_x}:\ 0 \le \mathrm{tr}(\mathbf{\Sigma_x}) \le 1} \mathrm{E}\left[\log\det\left(\mathbf{I} + \mathbf{\Sigma_x}\mathbf{H}_w^\dagger \mathbf{\Sigma} \mathbf{H}_w\right)\right] \ , \tag{2.18}$$

where $\mathbf{H}_w$ denotes a channel matrix with i.i.d. $\mathcal{CN}(0,1)$ entries. The capacity with full CSI is

$$C_{\mathrm{F}} = \max_{\rho:\ \mathrm{E}[\rho] \le 1} \mathrm{E}\left[\max_{\mathbf{\Sigma}_{\mathbf{x}|\mathbf{H}_w}:\ 0 \le \mathrm{tr}(\mathbf{\Sigma}_{\mathbf{x}|\mathbf{H}_w}) \le \rho} \log\det\left(\mathbf{I} + \mathbf{\Sigma}_{\mathbf{x}|\mathbf{H}_w}\mathbf{H}_w^\dagger \mathbf{\Sigma} \mathbf{H}_w\right)\right] \ , \tag{2.19}$$

where $\mathbf{\Sigma}_{\mathbf{x}|\mathbf{H}_w} \triangleq \mathrm{E}[\mathbf{x}\mathbf{x}^\dagger|\mathbf{H}_w]$ is the transmit covariance conditioned on $\mathbf{H}_w$.

Under the assumption of i.i.d. fading and noise, Telatar [82] showed that $\mathbf{\Sigma_x} = \frac{1}{N}\mathbf{I}$ is optimal for CSIR. It is clear that (2.18) still benefits from the symmetry and convexity properties[3] used in his proof, so we conclude that

$$C_{\mathrm{R}} = \mathrm{E}\left[\log\det\left(\mathbf{I} + \frac{1}{N}\mathbf{H}_w^\dagger \mathbf{\Sigma} \mathbf{H}_w\right)\right] \ . \tag{2.20}$$

The solution to (2.19) is an obvious modification to the well-known *space-time waterfilling* distribution (cf., [83, pg. 346]), obtained by aligning the transmit covariance eigenvectors with those of $\mathbf{H}_w^\dagger \mathbf{\Sigma} \mathbf{H}_w$. Denoting the ordered eigenvalues of a Hermitian matrix $\mathbf{A}$ by

[3] $\mathbf{\Sigma_n}^{-1/2}\mathbf{H}\mathbf{U} \sim \mathbf{\Sigma_n}^{-1/2}\mathbf{H}$ for any unitary $\mathbf{U}$ and the argument of (2.18) preserves the positive definiteness of any $\mathbf{\Sigma_x}$. Note that our assumption that the columns of $\mathbf{H}$ are i.i.d. is essential.

$\lambda_1(\mathbf{A}) \geq \cdots \geq \lambda_N(\mathbf{A})$, the full CSI capacity is

$$C_{\mathrm{F}} = \sum_{i=1}^{\min\{N,M\}} \mathrm{E}\left[\log\left(1 + P_i^* \lambda_i(\mathbf{H}_w^\dagger \boldsymbol{\Sigma} \mathbf{H}_w)\right)\right] \ , \tag{2.21}$$

where the waterfilling power allocations are given by

$$P_i^* = \left[\mu - \frac{1}{\lambda_i(\mathbf{H}_w^\dagger \boldsymbol{\Sigma} \mathbf{H}_w)}\right]^+ \ , \tag{2.22}$$

$[x]^+$ denotes the positive part of x, and the water level μ satisfies $\sum_i \mathrm{E}[P_i^*] = 1$.

2.3.3 Spatial Multiplexing

The CSIR capacity derivation suggests that sending independent data streams over each transmit antenna is desirable. Optimal decoding requires joint maximum a-posteriori (MAP) detection of the received signals, which may become computationally infeasible for even moderate N and signal constellations. The V-BLAST receiver proposed in [96] uses a linear estimator $\tilde{\mathbf{x}} = \mathbf{A}\mathbf{r}$, $\mathbf{A} \in \mathcal{C}^{N \times M}$, in conjunction with individual MAP detection and successive cancelation. Both *zero forcing* (ZF) and *linear minimum-mean square error* (MMSE) are considered; here we briefly evaluate the capacity of these receivers under our model (2.14) by making simple modifications to the i.i.d. fading and noise case (cf., [83, Sec. 8.3]).

The ZF transformation assumes the transmitted signal $\mathbf{x}$ can be any complex vector, and forms the maximum-likelihood (ML) estimate [90, Ch. 5],

$$\tilde{\mathbf{x}} = \arg\max_{\mathbf{x} \in \mathcal{C}^N} \ln f_{\mathbf{r}|\mathbf{x}}(\mathbf{r}|\mathbf{x}) = \left(\boldsymbol{\Sigma}_{\mathbf{n}}^{-1/2}\mathbf{H}\right)^+ \boldsymbol{\Sigma}_{\mathbf{n}}^{-1/2}\mathbf{r} \ , \tag{2.23}$$

where $f_{\mathbf{r}|\mathbf{x}}(\mathbf{r}|\mathbf{x})$ is the conditional pdf of $\mathbf{r}$ given $\mathbf{x}$ and the $+$ superscript denotes the Moore-Penrose generalized inverse. Using the identity $(\boldsymbol{\Sigma}_{\mathbf{n}}^{-1/2}\mathbf{H})^+\boldsymbol{\Sigma}_{\mathbf{n}}^{-1/2} = \mathbf{H}^+$ and ob-

serving that $\mathbf{H}$ is full rank with probability one, we may express the ZF transformation as

$$\mathbf{A}_{\text{ZF}} = \begin{cases} (\mathbf{H}^\dagger\mathbf{H})^{-1}\mathbf{H}^\dagger \ , & N \leq M \\ \mathbf{H}^\dagger(\mathbf{H}\mathbf{H}^\dagger)^{-1} \ , & N > M \end{cases} \ . \tag{2.24}$$

Thus noise correlation does not change the structure of the ZF receiver. For clarity, let $N \leq M$ so that, conditioned on $\mathbf{H}$, the ZF filter output equals $\mathbf{x}$ plus noise with covariance matrix $(\mathbf{H}_w^\dagger \boldsymbol{\Sigma} \mathbf{H}_w)^{-1}$. We now have N parallel Gaussian channels, so for individual MAP detection of the received signals the ZF capacity is

$$C_{\text{ZF}} = \sum_{i=1}^{N} \text{E}\left[\log\left(1 + \frac{1}{N[(\mathbf{H}_w^\dagger \boldsymbol{\Sigma} \mathbf{H}_w)^{-1}]_{ii}}\right)\right] \ , \tag{2.25}$$

where we use the notation $[\mathbf{A}]_{ij}$ to denote the $(i,j)^{\text{th}}$ element of a matrix $\mathbf{A}$. From the channel coding theorem we know that the error probability in decoding each stream may be made as close to zero as desired (cf., [19, Ch. 9]), so it is beneficial to subtract the i^{th} correctly decoded signal from $\mathbf{r}$ before performing ZF on stream $i+1$. This is ZF with successive cancelation (ZF-SC). Observe that for correctly-decoded [4] SC the input to the ZF filter for stream i is only affected by interference from streams $i+1, \ldots, N$, so the capacity of V-BLAST with ZF-SC (for $N \leq M$) is given by

$$C_{\text{ZF-SC}} = \sum_{i=1}^{N} \text{E}\left[\log\left(1 + \frac{1}{N[(\mathbf{H}_w^{i\dagger} \boldsymbol{\Sigma} \mathbf{H}_w^i)^{-1}]_{ii}}\right)\right] \ , \tag{2.26}$$

where $\mathbf{H}_w^i$ denotes the matrix obtained by deleting the first $i-1$ columns of $\mathbf{H}_w$.

The linear MMSE transformation also assumes that $\mathbf{x} \in \mathcal{C}^N$, and forms the MMSE

[4] An important consideration in implementing SC is the impact of subtracting *incorrectly* decoded data [90, Ch. 7]; however, we shall not explore this possibility.

estimate $\tilde{\mathbf{x}} = \mathbf{A}_{\text{MMSE}}\mathbf{r}$, where [90, Ch. 6]

$$\begin{aligned}\mathbf{A}_{\text{MMSE}} &= \arg\min_{\mathbf{A}\in\mathcal{C}^{N\times M}} \text{tr E}[(\mathbf{Ar}-\mathbf{x})(\mathbf{Ar}-\mathbf{x})^\dagger] \\ &= \frac{P}{N}\mathbf{H}^\dagger\left(\frac{P}{N}\mathbf{H}\mathbf{H}^\dagger + \mathbf{\Sigma}_\mathbf{n}^{-1}\right)^{-1} . \end{aligned} \tag{2.27}$$

Proceeding as before, we assume correctly-decoded SC so that $\tilde{x}_i$ is only affected by interference from $x_{i+1}, \ldots, x_N$. Thus the i^{th} estimated signal of the MMSE-SC is

$$\tilde{x}_i = [\mathbf{A}^i_{\text{MMSE}}\mathbf{H}^i]_{ii}x_i + \sum_{j>i}[\mathbf{A}^i_{\text{MMSE}}\mathbf{H}^i]_{ij}x_j + z_i \ , \tag{2.28}$$

where $\mathbf{H}^i$ denotes the matrix obtained by deleting the first $i-1$ columns of $\mathbf{H}$, $\mathbf{A}^i_{\text{MMSE}}$ is given by (2.27) with $\mathbf{H}$ replaced by $\mathbf{H}^i$, and $z_i \sim \mathcal{CN}(0, [\mathbf{A}^i_{\text{MMSE}}\mathbf{\Sigma}_\mathbf{n}\mathbf{A}^{i\dagger}_{\text{MMSE}}]_{ii})$. The capacity is therefore

$$C_{\text{MMSE-SC}} = \sum_{i=1}^{N} \text{E}\left[\log\left(1 + \frac{\frac{P}{N}|[\mathbf{A}^i_{\text{MMSE}}\mathbf{H}^i]_{ii}|^2}{\frac{P}{N}\sum_{j>i}|[\mathbf{A}^i_{\text{MMSE}}\mathbf{H}^i]_{ij}|^2 + [\mathbf{A}^i_{\text{MMSE}}\mathbf{\Sigma}_\mathbf{n}\mathbf{A}^{i\dagger}_{\text{MMSE}}]_{ii}}\right)\right] . \tag{2.29}$$

The expression for MMSE-SC capacity is considerably more complicated than the ZF-SC capacity (2.26) because the MMSE filter does not completely remove the subchannel interference. However, it is well-known [88] that the MMSE-SC receiver is a capacity-achieving architecture[5], so we may simply conclude

$$C_{\text{MMSE-SC}} = C_{\text{R}} = \text{E}\left[\log\det\left(\mathbf{I} + \frac{1}{N}\mathbf{H}_w^\dagger\mathbf{\Sigma}\mathbf{H}_w\right)\right] . \tag{2.30}$$

The optimality of the MMSE-SC receiver will be confirmed in the next section by evaluating (2.29) and (2.30).

[5]This may be demonstrated by applying the chain rule of mutual information and noting that $[\mathbf{A}_{\text{MMSE}}\mathbf{r}]_i = \frac{P}{N}\mathbf{h}_i^\dagger\left(\frac{P}{N}\mathbf{H}\mathbf{H}^\dagger + \mathbf{\Sigma}_\mathbf{n}^{-1}\right)^{-1}\mathbf{r}$ is a sufficient statistic for detecting x_i in Gaussian noise $\sum_{j\neq i}\mathbf{h}_j x_j + \mathbf{n}$ (cf., [83, pg. 362]).

2.3.4 Space-Time Block Codes

Now let us consider a slow fading channel in which $\mathbf{H}$ remains constant throughout the transmission. Here coding over many channel realizations is not possible and the spatial multiplexing system described above is subject to outage. In this case it may be desirable to improve reliability (at the expense of data rate) by introducing some spatial redundancy in the transmitted signal through the use of space-time codes. Specifically, consider the transmission of RL data symbols during L channel uses with a rate-R linear space-time block code (STBC) (cf., [49])

$$\mathbf{R} = \mathbf{H}\mathbf{X} + \mathbf{N} \ , \tag{2.31}$$

where the i^{th} column of $\mathbf{R}$, $\mathbf{X}$, and $\mathbf{N}$ denotes the received signal, transmitted signal, and noise during the i^{th} channel use, respectively. For equiprobable codewords the ML detector

$$\hat{\mathbf{X}}_{\text{ML}} = \arg \min_{\mathbf{X}_i \in \mathcal{X}} \|\mathbf{\Sigma}_{\mathbf{n}}^{-1/2}(\mathbf{R} - \mathbf{H}\mathbf{X}_i)\|_{\text{F}}^2 \tag{2.32}$$

minimizes the probability of error, where $\|\cdot\|_{\text{F}}^2$ is the Frobenius norm and $\mathcal{X}$ is the code space. An upper bound for the probability of error can be obtained by considering the *pairwise error probability* (PEP) P_{ij} between codewords $\mathbf{X}_i$ and $\mathbf{X}_j$, defined as the error probability of the hypothetical ML binary decision $\mathbf{X} \in \{\mathbf{X}_i, \mathbf{X}_j\}$. From (2.32) we obtain

$$\text{P}_{ij} = \text{E}\left[\text{Q}\left(\sqrt{\frac{1}{2}\text{tr}\left[(\mathbf{X}_i - \mathbf{X}_j)^\dagger \mathbf{H}_w^\dagger \mathbf{\Sigma} \mathbf{H}_w (\mathbf{X}_i - \mathbf{X}_j)\right]}\right)\right] \ . \tag{2.33}$$

In our numerical examples we shall focus on orthogonal STBCs [81], in which case $(\mathbf{X}_i - \mathbf{X}_j)(\mathbf{X}_i - \mathbf{X}_j)^\dagger = d_{ij}^2 \mathbf{I}$, where d_{ij} is the Euclidian distance between the corresponding vectors of RL data symbols. The PEP upper bound on the probability of error contains

the sum over all pairs $(\mathbf{X}_i, \mathbf{X}_j)$, which is usually dominated by the terms with codewords corresponding to the smallest Euclidean distance $d_{\min}$ between a single pair of data symbols. To derive the relationship between $d_{\min}$ and the power constraint, let us consider BPSK modulation. Here the symbol constellation is $\pm\frac{d_{\min}}{2}$, and to satisfy the power constraint implied in (2.33) we must have $\frac{1}{L}\mathrm{E}[\mathrm{tr}(\mathbf{X}\mathbf{X}^\dagger)] = \frac{1}{L}\cdot N\cdot RL\cdot\frac{d_{\min}^2}{4} \leq 1$. Therefore $d_{\min} = \frac{2}{\sqrt{NR}}$ and the PEP is

$$\mathrm{P}_{ij} = \mathrm{E}\left[\mathrm{Q}\left(\sqrt{\frac{2}{NR}\mathrm{tr}\left(\mathbf{H}_w^\dagger\mathbf{\Sigma}\mathbf{H}_w\right)}\right)\right] \ . \tag{2.34}$$

2.3.5 Beamforming

If channel knowledge can also be obtained at the transmitter, the reliability of the slow fading channel may be improved relative to STBCs (again at the expense of data rate) by sending a single data symbol x over the channel. The instantaneous SNR is

$$\gamma(\mathbf{w}_T, \mathbf{w}_R) = P\frac{|\mathbf{w}_R^\dagger\mathbf{H}\mathbf{w}_T|^2}{\mathbf{w}_R^\dagger\mathbf{\Sigma}_\mathbf{n}\mathbf{w}_R} \ , \tag{2.35}$$

where $\mathbf{w}_T, \mathbf{w}_R$ are the beamforming weights applied to the transmit and receive array, respectively. For spatially white noise, the optimal weights are given by the MIMO maximum-ratio combiner (MIMO-MRC) [54]. The derivation for correlated fading and noise is similar,

$$\max_{\substack{\mathbf{w}_R,\mathbf{w}_T\\ \|\mathbf{w}_T\|=1}} \gamma(\mathbf{w}_T, \mathbf{w}_R) = \max_{\substack{\mathbf{w}_T\\ \|\mathbf{w}_T\|=1}} P\|\mathbf{\Sigma}_\mathbf{n}^{-1/2}\mathbf{H}\mathbf{w}_T\|^2 = P\lambda_1(\mathbf{H}^\dagger\mathbf{\Sigma}_\mathbf{n}^{-1}\mathbf{H}) \ , \tag{2.36}$$

which follows from the Cauchy-Schwarz inequality and Rayleigh-Ritz theorem. The outage probability of this system is

$$P_{\text{out}} = \Pr\left\{\lambda_1(\mathbf{H}_w^\dagger\mathbf{\Sigma}\mathbf{H}_w) \leq \tau\right\} \ , \tag{2.37}$$

where τ is a non-negative threshold.

2.4 Performance in Correlated Fading and Noise

We conclude this chapter by evaluating the communication schemes presented in the previous section for some specific examples of the fading and noise correlation matrices $\mathbf{\Sigma_h}$ and $\mathbf{\Sigma_n}$. In particular, we consider an *exponential correlation model* (cf., [55]) where

$$[\mathbf{\Sigma_h}]_{ij} = \rho_h^{|i-j|} \qquad \text{and} \qquad [\mathbf{\Sigma_n}]_{ij} = N_0 B \rho_n^{|i-j|} \tag{2.38}$$

for some correlation coefficients $\rho_h, \rho_n \in [0, 1)$. To compute the SNR matrix (2.16) we need only specify the value of each correlation coefficient and the SISO SNR (2.15), which we take as 10 dB. In later chapters we shall derive correlation matrices from physical models of the fading and noise; for now the exponential correlation model will serve as a useful tool for examining the impact of fading and noise correlation, and also provides us an opportunity to discuss how the performance metrics are computed throughout this dissertation.

We shall use two techniques to compute the performance metrics: Monte-Carlo simulations and numerical integration. Other techniques, e.g., bounds and asymptotic analyses for capacity, may be found in MIMO special issues [71], [72], [39]. In the Monte-Carlo approach we use a random number generator to produce an i.i.d. sequence of $\mathbf{H}_w$ matrices, from which the means and distributions specified by each metric may be approximated by sample means and histograms. A rigorous discussion of the sample size needed to produce a desired confidence interval would take us too far afield, so we simply follow standards established in the MIMO literature (e.g., [29]) and use 10,000 simulations.

If a large number of data points are desired, e.g., over a range of ρ_h and ρ_n, the Monte-Carlo approach may be time consuming. As an alternative, some of the metrics may be evaluated using numerical integration. The CSIR capacity (2.20) may be expressed as

$$\begin{aligned} C_R &= \sum_{i=1}^{\min\{N,M\}} \mathrm{E}\left[\log\left(1+\frac{1}{N}\lambda_i(\mathbf{H}_w^\dagger \mathbf{\Sigma}\mathbf{H}_w)\right)\right] \\ &= \min\{N,M\}\int_0^\infty \log\left(1+\frac{1}{N}\lambda\right) f(\lambda)d\lambda \ , \end{aligned} \tag{2.39}$$

where $f(\lambda)$ is the pdf of an unordered, positive eigenvalue of [6] $\mathbf{H}_w^\dagger \mathbf{\Sigma}\mathbf{H}_w$. An expression for this pdf, when the eigenvalues of $\mathbf{\Sigma}$ are distinct, is available in [3], and we may evaluate the above integral numerically. Similarly, the full CSI capacity (2.21) is

$$C_F = \min\{N,M\}\int_{1/\mu}^\infty \log(\mu\lambda) f(\lambda)d\lambda \ , \tag{2.40}$$

where the water level μ satisfies

$$\min\{N,M\}\int_{1/\mu}^\infty \left(\mu - \frac{1}{\lambda}\right) f(\lambda)d\lambda = 1 \ . \tag{2.41}$$

The MIMO-MRC outage probability (2.37) is simply the cumulative distribution function (cdf) of the largest eigenvalue of $\mathbf{H}_w^\dagger \mathbf{\Sigma}\mathbf{H}_w$, which is also available in the literature, e.g., [63], [97]. As one example, for $N = M$ and distinct SNR matrix eigenvalues $\lambda_i(\mathbf{\Sigma})$ we obtain

$$P_{\text{out}} = \frac{\det \mathbf{A}}{\prod_{\substack{i=1\\ j>i}}^N \lambda_i^i(\mathbf{\Sigma})\left[1-\lambda_i(\mathbf{\Sigma})/\lambda_j(\mathbf{\Sigma})\right]} \ , \tag{2.42}$$

where

$$[\mathbf{A}]_{ij} = \frac{\lambda_i^j(\mathbf{\Sigma})}{(j-1)!}\int_0^{\tau/\lambda_i(\mathbf{\Sigma})} u^{j-1}e^{-u}du \tag{2.43}$$

[6] This is known as a *complex central Wishart matrix*, cf., [84, Ch. 2].

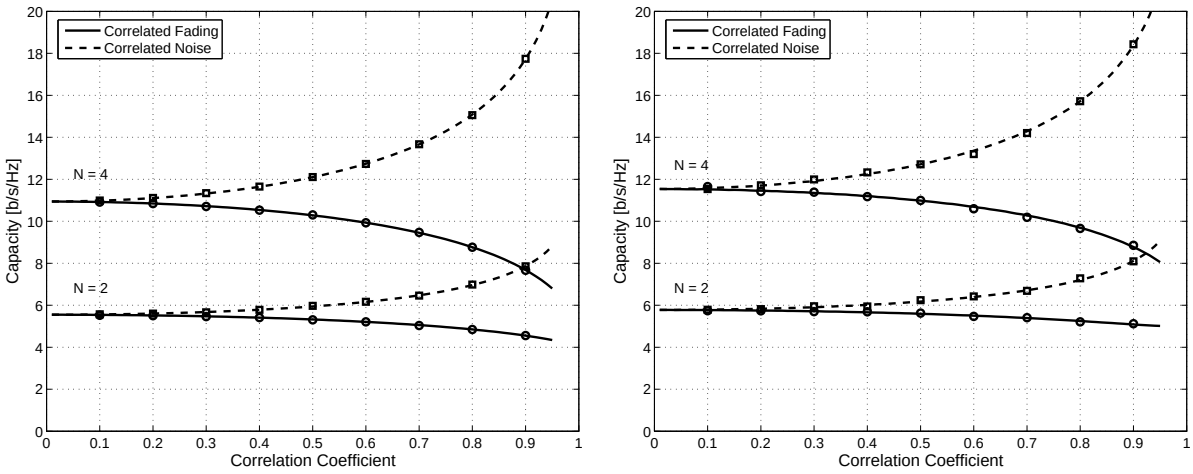


Figure 2.4: (a) CSIR and (b) full CSI capacity of $N \times N$ systems with correlated fading and noise. The lines are computed using numerical integration, and the points using Monte-Carlo simulations.

may be evaluated numerically. The STBC pairwise error probability (2.33) for orthogonal codes is

$$\begin{aligned} \mathrm{P}_{ij} &= \mathrm{E}\left[\mathrm{Q}\left(\sqrt{\frac{d_{ij}^2}{2}\mathrm{tr}(\mathbf{H}_w^\dagger \mathbf{\Sigma} \mathbf{H}_w)}\right)\right] \\ &= \frac{1}{\pi}\int_0^{\pi/2} \prod_{i=1}^{M} \left(\frac{\sin^2 u}{\sin^2 u + \frac{d_{ij}^2}{4}\lambda_i(\mathbf{\Sigma})}\right)^N du \ . \end{aligned} \tag{2.44}$$

The above result follows by observing that $\mathrm{tr}\left(\mathbf{H}_w^\dagger \mathbf{\Sigma} \mathbf{H}_w\right) \sim \sum_{i,j} |[\mathbf{H}_w]_{ij}|^2 \lambda_j(\mathbf{\Sigma})$, so the above error probability is identical to that of an NM-branch optimal diversity combiner with independent, non-identical SNRs, which is evaluated in [76, Sec. 9.2].

The ergodic capacity of 2×2 and 4×4 MIMO channels with CSIR and full CSI is shown as a function of the fading and noise correlation coefficient in Fig. 2.4. The correlated fading curves were calculated for $\rho_n = 0$ and $\rho_h \in (0, 0.95]$ in (2.38), and

conversely for correlated noise. The lines are computed using numerical integration of (2.39) and (2.40) at 0.01 increments of the correlation coefficient, and the points are computed using Monte-Carlo simulations of (2.20) and (2.21) at 0.1 increments. The logarithm is base-2, so the capacity units are bits per second per Hertz (b/s/Hz). There are several observations to make from these results. First, Monte-Carlo simulations and numerical integration appear to give consistent results, so we may use the more expedient technique of numerical integration in future computations. Second, the benefit of full CSI is almost negligible for all values of fading and noise correlation. However, we note that this conclusion applies specifically to the case we have considered where the SNR is large and $N \leq M$; if either the SNR is small or $N > M$, full CSI may offer a large improvement in capacity (cf., [83, Sec. 5.4.6]). The final and most important observation we make from these plots is that fading and noise correlation have quite opposite effects. Fading correlation degrades capacity, while capacity increases with noise correlation.

To gain some insight into the distinctive behavior of fading and noise correlation, let us examine the CSIR ergodic capacity of an $N \times 2$ system for limiting values of ρ_h and ρ_n. For correlated fading and $\rho_n = 0$, the SNR matrix has eigenvalues $\lambda_1(\mathbf{\Sigma}) = (1+\rho_h)\sigma$ and $\lambda_2(\mathbf{\Sigma}) = (1-\rho_h)\sigma$, where σ is the SISO SNR (2.15). Thus as $\rho_h \to 1$, we see that $\lambda_1(\mathbf{H}_w^\dagger \mathbf{\Sigma} \mathbf{H}_w) \to 2\sigma\|\mathbf{h}_w\|^2$ and $\lambda_2(\mathbf{H}_w^\dagger \mathbf{\Sigma} \mathbf{H}_w) \to 0\|\mathbf{h}_w\|^2$, where $\mathbf{h}_w \in \mathcal{C}^N$ is $\mathcal{CN}(\mathbf{0}, \mathbf{I})$ and $\|\cdot\|$ is the Euclidean norm. Therefore, from (2.39) we conclude that the CSIR capacity of an $N \times 2$ system with spatially white noise and $\rho_h = 1$ is $\mathrm{E}[\log(1 + 2\frac{\sigma}{N}\|\mathbf{h}_w\|^2)]$. From (2.20) we see that this is identical to the CSIR capacity of an $N \times 1$ system with twice the transmit power. Since $\rho_h = 1$ implies the rows of $\mathbf{H}$ are equal with probability one, we may obtain the aforementioned $N \times 1$ channel by performing the operation $r_1 + r_2$.

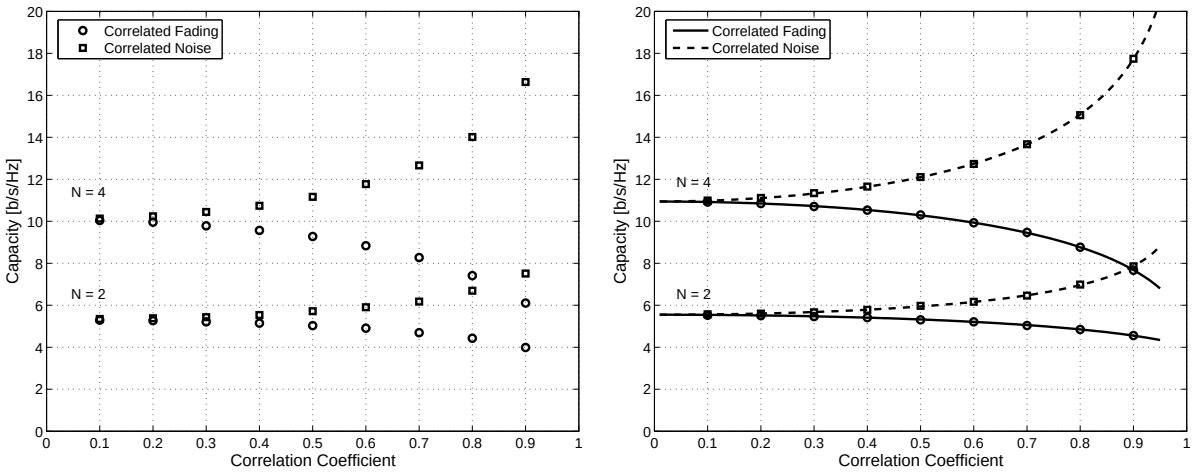


Figure 2.5: (a) ZF-SC and (b) MMSE-SC V-BLAST capacity of $N \times N$ systems with correlated fading and noise.

The simplicity of this optimal array processing reflects the fact that the receiver provides an *array gain* (of 3 dB) but no *multiplexing gain* (cf., [83, Sec. 8.2]). On the other hand, for correlated noise and $\rho_h = 0$ we have $\lambda_1(\mathbf{\Sigma}) = \frac{\sigma}{1-\rho_n}$ and $\lambda_2(\mathbf{\Sigma}) = \frac{\sigma}{1+\rho_n}$, so as $\rho_n \to 1$ the largest SNR matrix eigenvalue $\lambda_1(\mathbf{H}_w^\dagger \mathbf{\Sigma} \mathbf{H}_w) \to \infty \|\mathbf{h}_w\|^2$, and the capacity is infinite. In this case $n_1 = n_2$ with probability one, so we may create a noise-free channel by performing the operation $r_1 - r_2$.

We note that the penalty of fading correlation and benefit of noise correlation are well known for many other communication systems. For example, early studies of receive diversity [65] noted the deleterious impact of fading correlation, and Van Trees' classic text includes an extensive and elegant discussion [86, Sec. 4.3] on how optimal detectors and estimators exploit temporally correlated noise by subtracting an estimate of the noise from the received signal. The observations we made in the preceding paragraph are just

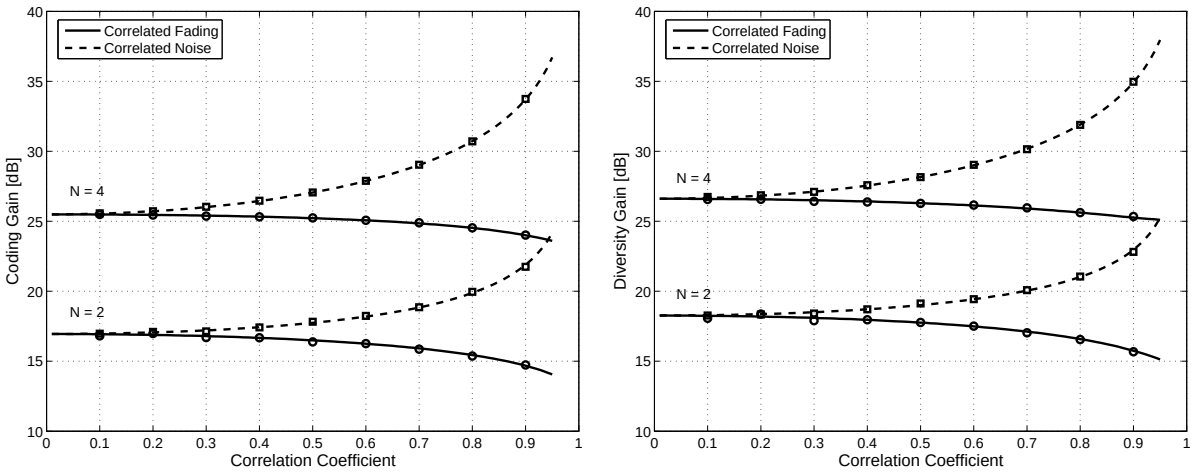


Figure 2.6: (a) Orthogonal STBC coding gain at 10^{-3} error probability and (b) MIMO-MRC diversity gain at 1% outage probability for $N \times N$ systems with correlated fading and noise.

special cases of these well-established concepts, and we shall see further evidence of this behavior in the next few examples and throughout the rest of this dissertation.

The capacity of the ZF-SC and MMSE-SC V-BLAST systems are shown in Fig. 2.5. Only Monte-Carlo simulations are provided for the ZF-SC since, to the best of our knowledge, no compact integral expression for (2.26) is available. To verify the optimality of MMSE-SC claimed in (2.30), the points in Fig. 2.5(b) were computed using Monte-Carlo simulations of (2.29) and the lines using numerical integration of the CSIR capacity (2.39). The results are indeed in agreement.

Finally, coding and diversity gains for orthogonal STBCs (2.34) and MIMO-MRC (2.37) are shown in Fig. 2.6. Here the coding gain is defined as the SNR difference between the pairwise error probability and $N = 1$ curves at an error probability of 10^{-3}, and diversity gain is similarly defined for the outage curves at a 1% probability of outage.

In Fig. 2.6(a) we use a BPSK constellation with a rate-1 code (e.g., the Alamouti code [2]) for $N = 2$ and a rate-1/2 code for $N = 4$, so $NR = 2$ for each curve.

Chapter 3

Noise Correlation in Compact Multi-Antenna Receivers

The numerical results of the previous chapter demonstrate the dramatic impact noise correlation may have on the performance of MIMO communication schemes, which motivates the study of noise models for compact multi-antenna receivers. In this chapter we evaluate the performance of an optimal diversity combiner in the presence of antenna mutual coupling and various receiver noise sources. Our receiver noise model is divided into three components: the antennas, front-end amplifiers, and everything else "downstream" from the amplifiers. This approach will allow us to characterize the unique properties of each source of noise, and to determine how each is impacted by antenna mutual coupling. Prior approaches have only modeled noise from either the antennas [30], [48] or amplifiers [59], [31], and a detailed study of how mutual coupling and the receiver structure affect noise correlation is lacking. For example, both [59] and [31] present results for spe-

cific amplifiers, but there is little discussion of how certain parameters of the amplifier (e.g., unilaterality, gain, noise figure) affect noise correlation and performance. It is the goal of this chapter to clearly articulate the characteristics of each receiver noise source and study their impact on optimal diversity performance.

We begin with an overview of our receiver model and, in order to orient the reader to receiver noise modeling, include a discussion of typical noise budgets for a single-antenna receiver. This analysis suggests that for many receivers no single source of noise is dominant. Using this model we find an expression for the signal collected by the combiner, and formulas for the outage probability of an optimal combiner are derived. In contrast to [59] and [31] (which use scattering and transmission parameters, respectively), we use an impedance formulation to describe the receiver, for which the familiar concepts of voltage and current may be more accessible to the reader unfamiliar with microwave circuit analysis. In the subsequent section we attempt to connect our model with the typical i.i.d. noise assumption. It is shown that in the presence of coupled antennas, strong assumptions on both the receiver structure and dominant source of noise may be needed to justify the white noise model. After introducing matching networks into our model, numerical results are presented. Several scenarios of interest are presented, including a detailed study of receiver noise correlation, a parametric sweep of various receiver parameters and noise sources, and the impact of directional fading and sky noise. From these results and the developed theory we conclude that noise correlation in compact multi-antenna receivers depends on the interaction between the dominant sources of noise and the coupled antennas – a result that may have profound implications for the selection of receiver components. For example, in uncoupled systems only the

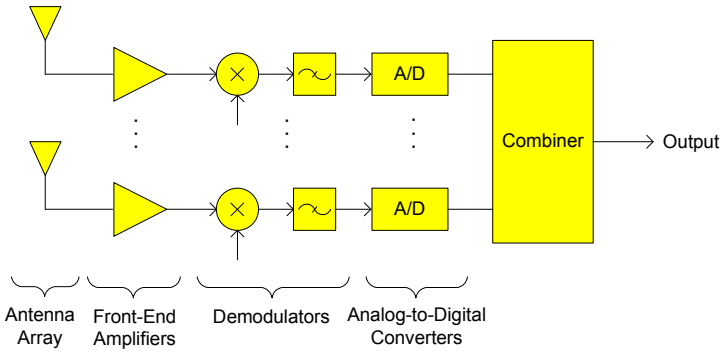


Figure 3.1: Basic diagram of a post-detection diversity receiver.

total noise power generated by a front-end amplifier is of interest, while in the presence of coupled antennas one should be concerned with how the noise is distributed between the input and output ports.

3.1 Receiver Model

A typical post-detection diversity receiver is illustrated in Fig. 3.1. The signals collected by each antenna are amplified and demodulated prior to A/D conversion and diversity combining. We assume coupling occurs only between antennas; there is no coupling between the amplifiers, demodulators or A/D converters in each branch. We further assume a narrowband system, so all impedances are constant over the system bandwidth and all signals can be expressed in complex baseband form. A circuit model of this M-branch receiver is illustrated in Fig. 3.2. Both the antenna array and front-end amplifiers are represented by equivalent noisy circuit models. As will be seen later, noise generated by receiver components "downstream" from the front-end amplifiers is consid-

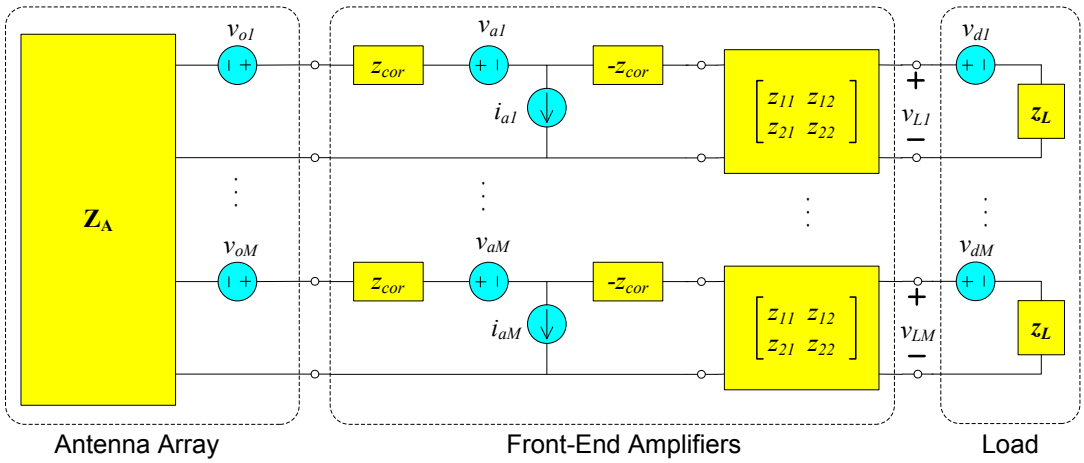


Figure 3.2: Circuit model of a post-detection diversity receiver.

erably less affected by antenna coupling, so we shall lump them together as a "load" on the amplifiers. Below we discuss these circuit models in more detail.

3.1.1 Antenna Array

The function of an antenna is to convert an incident electromagnetic field into a voltage across the antenna terminals. When the antennas of an array are closely spaced, the terminal voltage of each antenna depends not only on the field at that antenna but also on the currents flowing through neighboring antennas. The relationship between the terminal voltages and currents, arranged in M-dimensional column vectors $\mathbf{v}$ and $\mathbf{i}$, respectively, can be modeled as

$$\mathbf{v} = \mathbf{Z_A}\mathbf{i} + \mathbf{v_o} \ , \tag{3.1}$$

where $\mathbf{Z_A}$ is an $M \times M$ impedance matrix and $\mathbf{v_o}$ is an M-dimensional vector of open-circuit (induced) voltages. Here $[\mathbf{Z_A}]_{nn}$ is the *self-impedance* of antenna n, and $[\mathbf{Z_A}]_{nm}$

is the *mutual impedance* between antennas n and m. Approximate formulas for these impedances are available for thin dipoles [4, Ch. 8]; other types of antennas may be evaluated by numerical techniques.

We assume a frequency-flat, Rayleigh fading environment, so

$$\mathbf{v_o} = \mathbf{h}x + \mathbf{n_o} \ , \tag{3.2}$$

where x is the transmitted symbol, $\mathbf{h}$ is a vector of fading path gains, and $\mathbf{n_o}$ is noise. The transmit power is $P \triangleq \mathrm{E}[|x|^2]$, where $\mathrm{E}[\,\cdot\,]$ denotes the expectation. For Rayleigh fading $\mathbf{h}$ has a zero-mean, circularly-symmetric, complex Gaussian distribution, denoted by $\mathbf{h} \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma_h})$, where $\mathbf{\Sigma_h} = \mathrm{E}[\mathbf{h}\mathbf{h}^\dagger]$ is the fading correlation matrix.

For perfectly conducting antennas, $\mathbf{n_o}$ is the voltage induced in the array by noise sources in the surrounding environment. These sources may include thermal radiation, cosmic background, and interference from other electronic devices. We will use the classical model for spherically isotropic thermal radiation in which the array is surrounded by a blackbody enclosure at temperature T_A K (cf., [5, Ch. 6]). The noise voltage is then [85] $\mathbf{n_o} \sim \mathcal{CN}(\mathbf{0}, 4kT_AB\mathbf{R_A})$, where $\mathbf{R_A} \triangleq \frac{1}{2}(\mathbf{Z_A} + \mathbf{Z_A^\dagger})$, $\dagger$ denotes the conjugate-transpose, $k = 1.38 \times 10^{-23}$ J/K is Boltzmann's constant and B is the bandwidth in Hz. Gaussian interference may be modeled by a simple generalization of this model that accounts for its spatial distribution (the above model is *isotropic*) [66, pg. 642]; an example will be provided in Sec. 3.6.6. Note that if the array is imbedded in a linear medium, the reciprocity theorem of electromagnetics [4, pg. 144] implies $\mathbf{Z_A}$ is symmetric, and so $\mathbf{R_A} = \mathrm{Re}[\mathbf{Z_A}]$.

3.1.2 Front-End Amplifiers

An amplifier is usually well-modeled within its dynamic range as a linear, noisy two-port network, so it can be represented by a Thevenin equivalent network with impedance matrix

$$\begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix} \tag{3.3}$$

and open-circuit noise voltages n_1 and n_2 at the input and output ports, respectively. Many noise sources, such as thermal and shot noise, are well-modeled [5, pg. 42] as $n_i \sim \mathcal{CN}(0, \sigma_i^2)$, with $\rho_{12} \triangleq \frac{\mathrm{E}[n_1 n_2^*]}{\sigma_1 \sigma_2} \neq 0$ in general. Thus amplifier noise may be characterized by the *noise parameters* $\{\sigma_1, \sigma_2, \rho_{12}\}$. However, these particular parameters are rarely used in practice. Below we briefly present some of the more common noise metrics relevant to this study.

The Rothe-Dahlke [70] noise model, shown for each amplifier in Fig. 3.2, is typically used to model amplifiers instead of the Thevenin equivalent model discussed above. The voltage and current sources at the input port, denoted by v_a and i_a, are related to the parameters above by

$$v_a = n_1 + \frac{z_{\mathrm{cor}} - z_{11}}{z_{21}} n_2 \ , \qquad i_a = -\frac{n_2}{z_{21}} \tag{3.4}$$

where $z_{\mathrm{cor}} = r_{\mathrm{cor}} + j x_{\mathrm{cor}} = z_{11} - z_{21} \rho_{12} \frac{\sigma_1}{\sigma_2}$ is the *correlation impedance* defined so that $\mathrm{E}[v_a i_a^*] = 0$. A common convention is to normalize the variances of these noise sources by the thermal noise power generated by a 1 Ω resistor at standard temperature $T_0 = 290$ K. Thus, the *equivalent noise resistance* of the voltage source and the *equivalent noise*

conductance of the current source are given by

$$r_a = \frac{\mathrm{E}\left[|v_a|^2\right]}{4kT_0B} \ , \ g_a = \frac{\mathrm{E}\left[|i_a|^2\right]}{4kT_0B} \ . \tag{3.5}$$

The amplifiers in Fig. 3.2 are identical and uncoupled, so the joint distributions of the M amplifier voltage and current sources are $\mathbf{v_a} \sim \mathcal{CN}(\mathbf{0}, 4kT_0Br_a\mathbf{I})$ and $\mathbf{i_a} \sim \mathcal{CN}(\mathbf{0}, 4kT_0Bg_a\mathbf{I})$, where $\mathbf{I}$ is the identity matrix.

An important amplifier noise metric is the *noise factor*, F_{amp}, defined as the ratio of the output noise power to the noise power contributed by a source alone, when a source impedance $z_s = r_s + jx_s$ at standard temperature is connected to the input port, i.e., [70]

$$F_{\mathrm{amp}} = 1 + \frac{1}{r_s}\left(r_a + g_a\left|z_s + z_{\mathrm{cor}}\right|^2\right). \tag{3.6}$$

The noise factor is useful because it relates the input and output signal-to-noise ratios (SNRs) of the amplifier, given in dB by $\mathrm{SNR_{out}} = \mathrm{SNR_{in}} - \mathrm{NF}$, where $\mathrm{NF} = 10\log_{10} F_{\mathrm{amp}}$ is the *noise figure*. Note that F_{amp} attains its minimum value F_{min} when $z_s = z_{\mathrm{opt}}$ where [70]

$$F_{\mathrm{min}} = 1 + 2\left(g_a r_{\mathrm{cor}} + \sqrt{g_a r_a + \left(g_a r_{\mathrm{cor}}\right)^2}\right) \tag{3.7}$$

$$z_{\mathrm{opt}} = \sqrt{r_a/g_a + r_{\mathrm{cor}}^2} - jx_{\mathrm{cor}} \ . \tag{3.8}$$

Note the parameters $\{F_{\mathrm{min}}, g_a, z_{\mathrm{opt}}\}$ are typically used to characterize the noise statistics of low-noise amplifiers (LNAs),[1] and are related to the noise factor by

$$F_{\mathrm{amp}} = F_{\mathrm{min}} + \frac{g_a}{r_s}|z_s - z_{\mathrm{opt}}|^2 \ . \tag{3.9}$$

[1] A variation of these parameters based on the admittance version of the Rothe-Dahlke model are $\{F_{\mathrm{min}}, R_a, Y_{\mathrm{opt}}\}$. For microwave amplifiers a reflection coefficient Γ_{opt} is usually given in place of Y_{opt} [66, pg. 558].

3.1.3 Load

Since the receiver chains downstream from the amplifiers in Fig. 3.1 are assumed to be electrically isolated, the noise in each branch is independent and the impedances are uncoupled. Each branch consists of filters, mixers and other analog devices that generate noise. Here we reference all downstream noise to the amplifier output and assume that each branch may be modeled by an impedance z_L and a zero-mean Gaussian noise voltage v_{di} with equivalent noise resistance r_d, as shown in Fig. 3.2. We denote the noise factor of the downstream components by F_{dow}.

3.2 Noise in Single-Antenna Receivers

Prior studies of noise in multi-antenna receivers have assumed either antenna [30] or amplifier [59], [31] noise is dominant, and downstream noise is usually neglected. In this section we examine this assumption by considering typical noise budgets for a *single-antenna* receiver, which suggest that for many receivers no single source of noise is dominant. The results of this section may be inferred from any text on receiver design (e.g., [66, Ch. 10-13]).

Consider the model of Fig. 3.2 with $M = 1$. As discussed in Sec. 3.1.2, for $T_A = T_0$ the SNR drop across the receiver is equal to its noise factor F_{Rx}. For $T_A \neq T_0$, it can be shown (cf., [66, pg. 641]) that the SNR drop across the receiver is $1 + \frac{T_0}{T_A}(F_{\text{Rx}} - 1)$, so the SNR at the load is

$$\sigma = \frac{P}{4kT_0Br_A} \cdot \frac{1}{\frac{T_A}{T_0} + F_{\text{Rx}} - 1} \ . \tag{3.10}$$

If $G_{\text{amp}} = |z_{21}/(z_A + z_{11})|^2$ is the amplifier gain, then the receiver noise factor is related to the amplifier and downstream noise factors by (cf., [66, pg. 495])

$$F_{\text{Rx}} = F_{\text{amp}} + \frac{F_{\text{dow}} - 1}{G_{\text{amp}}} \, . \tag{3.11}$$

We may now use (3.10) and (3.11) to make the following assertions for single-antenna receivers:

- Assuming typical receiver noise figures of $3 - 10$ dB, in order for antenna noise to dominate we would need T_A to be at least $10 - 100$ times larger than the standard temperature.

- On the other hand, antenna noise of purely thermal origin typically results in $T_A/T_0 \in [-3, 0]$ dB [66, Fig. 13.6], so receiver noise figures of at least $7 - 10$ dB would be required in order to neglect thermal antenna noise.

- Many microwave amplifiers have $F_{\text{amp}}G_{\text{amp}} \in [10, 20]$ dB [66, Tbl. 10.2], in which case the downstream noise figure[2] may need to be as low as $1 - 10$ dB to be considered negligible.

From this discussion we conclude that one must typically include all three noise sources. As an example, consider the receiver in [66, Ex. 10.2]. The antenna temperature is 150 K, the LNA has a 2 dB noise figure and 10 dB gain, and the downstream components consist of a filter with 1 dB loss and a mixer with 4 dB noise figure. Here no noise source

[2] In contrast to the LNA, some downstream components are lossy, noisy devices such as filters and mixers, so noise figures in the $5 - 10$ dB range are not uncommon.

can be neglected since

$$\frac{T_A}{T_0} = 0.517 \ , \quad F_{\text{amp}} - 1 = 0.585 \ , \quad \frac{F_{\text{dow}} - 1}{G_{\text{amp}}} = 0.216 \ .$$

3.3 Outage Probability of an Optimal Combiner

We now derive the outage probability of an optimal diversity combiner for the system in Fig. 3.2. We begin by finding an expression for the voltage across the load, $\mathbf{v_L}$, and then use this expression to determine the optimal combiner. We assume throughout that all impedance matrices are invertible, so every non-zero current vector input produces a non-zero voltage vector.

The M uncoupled amplifiers may be collectively viewed as a $2M$-port network with input and output open-circuit voltages $\mathbf{v_a} + (z_{\text{cor}} - z_{11})\mathbf{i_a}$ and $-z_{21}\mathbf{i_a}$, respectively, and impedance matrix

$$\begin{bmatrix} z_{11}\mathbf{I} & z_{12}\mathbf{I} \\ z_{21}\mathbf{I} & z_{22}\mathbf{I} \end{bmatrix} \ , \tag{3.12}$$

where each $\mathbf{I}$ is $M \times M$. Relating the voltages and currents at the input and output of this network to those of the antenna array, we can show that the antenna array plus amplifiers comprise an M-port network with open-circuit voltage $\mathbf{u}$ and impedance matrix $\mathbf{Z}$, where

$$\begin{aligned} \mathbf{u} &= \mathbf{C}(\mathbf{v_o} - \mathbf{v_a}) - \left[(z_{\text{cor}} - z_{11})\,\mathbf{C} + z_{21}\mathbf{I}\right]\mathbf{i_a} \\ \mathbf{Z} &= -z_{12}\mathbf{C} + z_{22}\mathbf{I} \ , \end{aligned} \tag{3.13}$$

and $\mathbf{C} = z_{21}\left(\mathbf{Z_A} + z_{11}\mathbf{I}\right)^{-1}$. By voltage division between this network and the load, we

obtain

$$\begin{aligned}\mathbf{v_L} &= z_L\left(\mathbf{Z}+z_L\mathbf{I}\right)^{-1}\mathbf{u}+\mathbf{Z}\left(\mathbf{Z}+z_L\mathbf{I}\right)^{-1}\mathbf{v_d}\\ &= \mathbf{G}\Big[\mathbf{v_o}-\mathbf{v_a}-\left(\mathbf{Z_A}+z_{\rm cor}\mathbf{I}\right)\mathbf{i_a}+\mathbf{K}\mathbf{v_d}\Big],\end{aligned} \tag{3.14}$$

where

$$\mathbf{G} = \frac{z_{21}z_L}{z_L+z_{22}}\left(\mathbf{Z_A}+\frac{z_{11}(z_L+z_{22})-z_{12}z_{21}}{z_L+z_{22}}\mathbf{I}\right)^{-1} \tag{3.15}$$

$$\mathbf{K} = \frac{z_{22}}{z_{21}z_L}\left(\mathbf{Z_A}+\frac{z_{11}z_{22}-z_{12}z_{21}}{z_{22}}\mathbf{I}\right)\ . \tag{3.16}$$

Since we have assumed $\mathbf{G}$ is invertible, finding an optimal combiner for $\mathbf{v_L}$ is equivalent to finding one for $\mathbf{G}^{-1}\mathbf{v_L}$. Since the latter expression is simpler, we take the observed signal to be

$$\mathbf{r} \triangleq \mathbf{G}^{-1}\mathbf{v_L} = \mathbf{h}x+\mathbf{n}\ , \tag{3.17}$$

where the second equality follows by substituting (3.2) into (3.14), and defining the noise as

$$\mathbf{n} \triangleq \mathbf{n_o}-\mathbf{v_a}-\left(\mathbf{Z_A}+z_{\rm cor}\mathbf{I}\right)\mathbf{i_a}+\mathbf{K}\mathbf{v_d}\ . \tag{3.18}$$

Recalling that $\mathbf{n_o}\sim\mathcal{CN}(\mathbf{0},4kT_ABR_A)$, $\mathbf{v_a}\sim\mathcal{CN}(\mathbf{0},4kT_0Br_a\mathbf{I})$, $\mathbf{i_a}\sim\mathcal{CN}(\mathbf{0},4kT_0Bg_a\mathbf{I})$ and $\mathbf{v_d}\sim\mathcal{CN}(\mathbf{0},4kT_0Br_d\mathbf{I})$ are mutually independent, we obtain $\mathbf{n}\sim\mathcal{CN}(\mathbf{0},\mathbf{\Sigma_n})$ where

$$\mathbf{\Sigma_n} = 4kT_0B\left[\frac{T_A}{T_0}\mathbf{R_A}+r_a\mathbf{I}+g_a\left(\mathbf{Z_A}+z_{\rm cor}\mathbf{I}\right)\left(\mathbf{Z_A}+z_{\rm cor}\mathbf{I}\right)^\dagger+r_d\mathbf{K}\mathbf{K}^\dagger\right]. \tag{3.19}$$

We now derive the optimal linear combiner for the observation (3.17). For a combiner of the form $y=\mathbf{w}^\dagger\mathbf{r}$, the *instantaneous SNR* at the combiner output is

$$\gamma(\mathbf{w}) = P\frac{\left|\mathbf{w}^\dagger\mathbf{h}\right|^2}{\mathrm{E}\left|\mathbf{w}^\dagger\mathbf{n}\right|^2} = P\frac{\mathbf{w}^\dagger\mathbf{h}\mathbf{h}^\dagger\mathbf{w}}{\mathbf{w}^\dagger\mathbf{\Sigma_n}\mathbf{w}}\ . \tag{3.20}$$

For i.i.d. noise (e.g., $\mathbf{\Sigma_n} = \mathbf{I}$) it is well-known [8] that the maximum-ratio combiner (MRC) $\mathbf{w} \propto \mathbf{h}$ uniquely maximizes $\gamma(\mathbf{w})$. It is a simple exercise [22] to show that $\mathbf{w} \propto \mathbf{\Sigma_n^{-1}h}$ maximizes (3.20) for any noise correlation. The resulting optimal SNR is[3]

$$\gamma^o \;=\; \max_{\mathbf{w}} \; \gamma(\mathbf{w}) \;=\; P \cdot \mathbf{h}^\dagger \mathbf{\Sigma_n^{-1}} \mathbf{h} \; . \tag{3.21}$$

The performance of receive diversity systems is often measured by the outage probability, defined as $P_{\text{out}}(\tau) = \Pr\{\gamma^o \leq \tau\}$, where τ is a non-negative threshold. A closed-form expression for the outage can be obtained using the results of [51]: If $\lambda_1(\mathbf{\Sigma}) > \cdots > \lambda_M(\mathbf{\Sigma}) > 0$ are the distinct eigenvalues of the *SNR matrix*

$$\mathbf{\Sigma} \triangleq P \cdot \mathbf{\Sigma_h^{1/2}} \mathbf{\Sigma_n^{-1}} \mathbf{\Sigma_h^{1/2}} \; , \tag{3.22}$$

and $\mathbf{A}^{1/2}$ is the positive-definite square root of a positive definite matrix $\mathbf{A}$ [40, pg. 406], then[4]

$$P_{\text{out}}(\tau) = \sum_{j=1}^{M} \frac{\lambda_j^{M-1}(\mathbf{\Sigma}) \left(1 - e^{-\tau/\lambda_j(\mathbf{\Sigma})}\right)}{\prod_{i \neq j} (\lambda_j(\mathbf{\Sigma}) - \lambda_i(\mathbf{\Sigma}))} \; . \tag{3.23}$$

The restriction of distinct eigenvalues on $\mathbf{\Sigma}$ includes a large class of $\mathbf{\Sigma_h}$ and $\mathbf{\Sigma_n}$; however, an important exception occurs when $\mathbf{\Sigma_h}$ and $\mathbf{\Sigma_n}$ are proportional. One example of proportional $\mathbf{\Sigma_h}$ and $\mathbf{\Sigma_n}$ occurs when both the fading and noise are i.i.d.; a second example will be presented in Section 3.6.4. In this case, the SNR matrix (3.22) reduces to $\mathbf{\Sigma} = \sigma \mathbf{I}$, where σ is the single-antenna SNR (3.10), and we may use Brennan's outage formula [8]:

$$P_{\text{out}}(\tau) = \frac{1}{(M-1)!} \int_0^{\tau/\sigma} t^{M-1} e^{-t} dt \; . \tag{3.24}$$

[3] Optimization problems of this form are encountered in many other areas; adaptive array processing is one related example [41, pg. 60].

[4] This follows from the Appendix of [51] by writing $\gamma^o = \mathbf{h_w^\dagger \Sigma h_w}$, where $\mathbf{h_w} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$.

3.4 Spatially White Noise

Most prior studies of receive diversity assume spatially white noise. In this section we examine some sufficient conditions for uncorrelated noise in our model. The total noise voltage across the load is $\mathbf{Gn}$, where $\mathbf{G}$ and $\mathbf{n}$ are defined in (3.15) and (3.18). Thus the correlation matrices of noise contributions from the antennas, amplifiers, and downstream components are given by

$$\mathbf{\Sigma}_{\text{ant}} = 4kT_A B\mathbf{G}\mathbf{R}_{\mathbf{A}}\mathbf{G}^{\dagger} \tag{3.25}$$

$$\mathbf{\Sigma}_{\text{amp}} = 4kT_0 B\mathbf{G}\left[r_a\mathbf{I} + g_a(\mathbf{Z}_{\mathbf{A}} + z_{cor}\mathbf{I})(\mathbf{Z}_{\mathbf{A}} + z_{cor}\mathbf{I})^{\dagger}\right]\mathbf{G}^{\dagger} \tag{3.26}$$

$$\mathbf{\Sigma}_{\text{dow}} = 4kT_0 B\frac{r_d}{|z_L + z_{22}|^2}(z_{22}\mathbf{I} - z_{12}\mathbf{G})(z_{22}\mathbf{I} - z_{12}\mathbf{G})^{\dagger}\ . \tag{3.27}$$

When the antennas are uncoupled, $\mathbf{Z}_{\mathbf{A}}$ is diagonal and from (3.15) we see that each of the above correlation matrices is diagonal, and so the noise is spatially white. Not surprisingly, mutual coupling is therefore necessary in our model for the noise to be correlated. It is not, however, sufficient: the noise can be white when the antennas are coupled. Since the amplifier and downstream noise sources are independent, isolating these sources from the coupled antennas should result in spatially white noise. Two examples of this are given below.

An amplifier that provides zero gain from its output to input port (the reverse signal path) is said to be *unilateral* [66, pg. 524]. This condition implies $z_{12} = 0$ in Fig. 3.2 and results in a diagonal $\mathbf{\Sigma}_{\text{dow}}$ in (3.27). Since most practical front-end amplifiers are approximately unilateral ($z_{12} \approx 0$), we expect that a spatially white noise model is justified if there is a dominant source of noise downstream from the amplifiers. Later we will illustrate this result for a specific LNA.

Other assumptions may also result in white noise. In [31] the dominant noise sources are assumed to be unilateral front-end amplifiers with the property that the total noise current at the amplifier output is independent of the source impedance. To see the implications of this assumption, observe that when an impedance z_s is connected to the input port of the Rothe-Dahlke amplifier model (cf. Sec 3.1.2) with $z_{12} = 0$, the noise current at the output is

$$i_{\text{out}} = \frac{z_{21}}{(z_{11} + z_s)(z_{22} + z_L)} \left[v_a + (z_{\text{cor}} + z_s)\, i_a \right] \ . \tag{3.28}$$

This expression is independent of z_s only if $v_a = i_a(z_{11} - z_{\text{cor}})$. From (3.4), this is equivalent to assuming that the input port is noise free, $n_1 = 0$, which implies $r_a = 0$ and $z_{\text{cor}} = z_{11}$. It is easy to show that these conditions, together with the assumption $z_{12} = 0$, do indeed imply that $\mathbf{\Sigma}_{\text{amp}}$ in (3.26) is diagonal, as asserted in [31]. However, it appears that many practical amplifiers (e.g., [57]) do not satisfy these conditions, and so the amplifier noise may be correlated.

These observations suggest that, when the antennas are coupled, strong assumptions on the front-end amplifiers and downstream components may be needed to justify a white noise model. Even when the noise sources in each branch are independent (e.g., amplifier noise), these sources can interact through the coupled antennas to produce correlated noise in the observed signal. White noise can arise when the amplifiers completely isolate these sources from the antennas, but many amplifiers do not satisfy the conditions needed for this to occur.

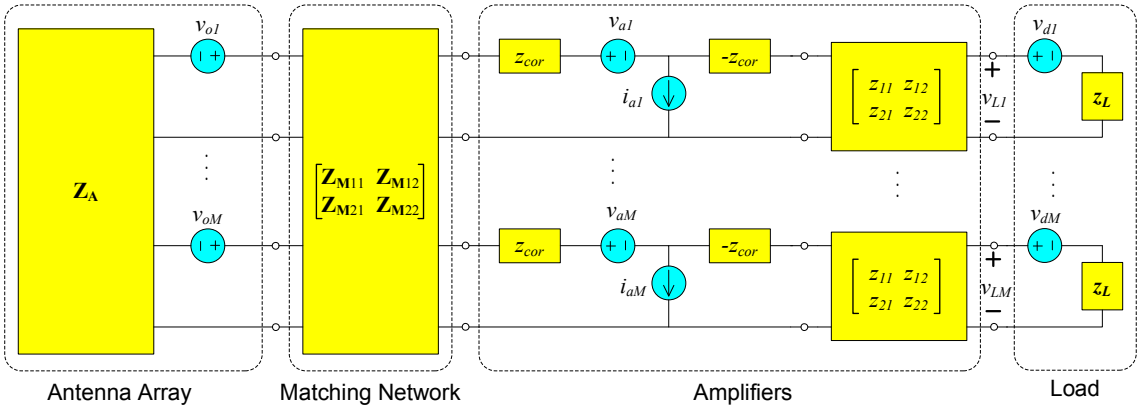


Figure 3.3: Receiver model with matching.

3.5 Matching Networks

The transfer of signal and noise power between two networks may be controlled using impedance matching techniques (cf., [66, Ch. 5]). Ideally a matching network is formed with passive, reactive elements so it is noiseless, lossless, and reciprocal. Consider a $2M$-port matching network inserted between the antennas and front-end amplifiers in Fig. 3.2 with impedance matrix

$$\begin{bmatrix} \mathbf{Z_{M11}} & \mathbf{Z_{M12}} \\ \mathbf{Z_{M21}} & \mathbf{Z_{M22}} \end{bmatrix} , \tag{3.29}$$

as shown in Fig. 3.3. If $\mathbf{v}_1, \mathbf{v}_2$ and $\mathbf{i}_1, \mathbf{i}_2$ denote the voltages and currents at the input and output of the network, respectively, the matrices in (3.29) are defined by the circuit equations

$$\begin{aligned} \mathbf{v}_1 &= \mathbf{Z_{M11}}\mathbf{i}_1 + \mathbf{Z_{M12}}\mathbf{i}_2 \\ \mathbf{v}_2 &= \mathbf{Z_{M21}}\mathbf{i}_1 + \mathbf{Z_{M22}}\mathbf{i}_2 \ , \end{aligned} \tag{3.30}$$

where each $\mathbf{Z}_{\mathbf{M}nm}$ is an $M \times M$ matrix. The network is lossless (no power is dissipated within it) provided the following conditions are satisfied [38, pg. 13]: $\mathbf{Z}_{\mathbf{M}11} = -\mathbf{Z}_{\mathbf{M}11}^{\dagger}$, $\mathbf{Z}_{\mathbf{M}22} = -\mathbf{Z}_{\mathbf{M}22}^{\dagger}$, $\mathbf{Z}_{\mathbf{M}21} = -\mathbf{Z}_{\mathbf{M}12}^{\dagger}$. Most passive networks are also reciprocal, in which case $\mathbf{Z}_{\mathbf{M}11} = \mathbf{Z}_{\mathbf{M}11}^{T}$, $\mathbf{Z}_{\mathbf{M}22} = \mathbf{Z}_{\mathbf{M}22}^{T}$, and $\mathbf{Z}_{\mathbf{M}21} = \mathbf{Z}_{\mathbf{M}12}^{T}$, where the T superscript denotes transpose.

From the standpoint of the rest of the receiver, the antennas plus matching comprise a noisy linear network which can be represented by a Thevenin equivalent circuit with [38]

$$\mathbf{v}_{\mathbf{o}}' = \mathbf{M}\mathbf{v}_{\mathbf{o}} \quad , \quad \mathbf{Z}_{\mathbf{A}}' = -\mathbf{M}\mathbf{Z}_{\mathbf{M}12} + \mathbf{Z}_{\mathbf{M}22} \ , \tag{3.31}$$

where $\mathbf{M} = \mathbf{Z}_{\mathbf{M}21}\left(\mathbf{Z}_{\mathbf{A}} + \mathbf{Z}_{\mathbf{M}11}\right)^{-1}$. Thus the results of Section 3.3 still apply with $\mathbf{\Sigma_h}$ replaced by $\mathbf{M}\mathbf{\Sigma_h}\mathbf{M}^{\dagger}$ and $\mathbf{Z_A}$ replaced by $\mathbf{Z}_{\mathbf{A}}'$. In particular, the SNR matrix (3.22) becomes[5]

$$\mathbf{\Sigma} = \frac{P}{4kT_0B}\mathbf{M}\mathbf{\Sigma_h}\mathbf{M}^{\dagger}\left[\frac{T_A}{T_0}\mathbf{R}_{\mathbf{A}}' + r_a\mathbf{I} + g_a\left(\mathbf{Z}_{\mathbf{A}}' + z_{\mathrm{cor}}\mathbf{I}\right)\left(\mathbf{Z}_{\mathbf{A}}' + z_{\mathrm{cor}}\mathbf{I}\right)^{\dagger} + r_d\mathbf{K}'\mathbf{K}'^{\dagger}\right]^{-1} , \tag{3.32}$$

where $\mathbf{R}_{\mathbf{A}}' \triangleq \frac{1}{2}(\mathbf{Z}_{\mathbf{A}}' + \mathbf{Z}_{\mathbf{A}}'^{\dagger})$ and $\mathbf{K}'$ is given by (3.16) with $\mathbf{Z_A}$ replaced by $\mathbf{Z}_{\mathbf{A}}'$.

3.6 Numerical Results

We now present numerical results for several examples of our model. In each case we use (3.23) and (3.32) to calculate the diversity gain at 1% outage for $M = 2, 4$ antennas, where diversity gain is defined as the difference in SNR between a given outage curve and

[5]For convenience we have expressed the SNR matrix as $P \cdot \mathbf{\Sigma_h}\mathbf{\Sigma_n}^{-1}$, which has the same eigenvalues (and therefore results in the same outage probability) as the SNR matrix defined in (3.22).

the $M = 1$ curve at a fixed probability. We begin with a description of the parameters[6] common to each example.

3.6.1 System Parameters

We consider a uniform linear array (ULA) of half-wavelength dipoles with inter-element spacings $0.01\lambda \leq d \leq \lambda$, where λ is the wavelength. The received electric field is composed of a large number (in the sense of the central limit theorem for Rayleigh fading) of plane waves with independent and uniformly-distributed polarization angles[7] and phases. In a spherical coordinate system with θ and ϕ denoting the zenith and azimuth, respectively, let $p(\theta, \phi) \sin\theta d\theta d\phi$ denote the probability that a plane wave arrives within solid angle $\sin\theta d\theta d\phi$. Then the entries of the fading correlation matrix are given by (cf., [17])

$$[\mathbf{\Sigma_H}]_{nm} = \int_0^{2\pi} \int_0^{\pi} g_n(\theta,\phi) g_m^*(\theta,\phi) e^{j\frac{2\pi}{\lambda}d(m-n)\sin\theta\cos\phi} p(\theta,\phi)\sin\theta d\theta d\phi \ , \tag{3.33}$$

where the antenna pattern $g_n(\theta, \phi)$ is defined as the voltage induced in the n^{th} dipole by a zero-phase, unit-amplitude, $\hat{\theta}$-polarized plane wave with angle-of-arrival (AoA) (θ, ϕ) (cf., Appendix A). In the first few examples we shall consider the standard *2D scattering* model $p(\theta, \phi) = \frac{1}{2\pi}\delta\left(\theta - \frac{\pi}{2}\right)$. We shall initially take $T_A = T_0$, and consider the impact of antenna temperature in a later example.

Approximate expressions for the impedance matrix $\mathbf{Z_A}$ and antenna patterns $g_n(\theta, \phi)$

[6]Diversity gain is unchanged for scalar multiples of the fading and noise correlation matrices such as P and B, so values for these parameters are not provided.

[7]Since dipoles respond only to the field component parallel to its axis, we may equivalently consider co-polarized waves with i.i.d. amplitudes $\alpha \in (0, 1)$, where $\Pr[\alpha \leq x] = 1 - \frac{2}{\pi}\cos^{-1} x$. Regardless, the average power loss resulting from random amplitudes may be absorbed into P.

are given in [4, Ch. 8] for infinitesimally thin wire dipoles. To account for scattering among the antennas, however, we evaluated these quantities numerically for finite-thickness dipoles using the Numerical Electromagnetics Code (NEC) [9], a method-of-moments based program. Details of the NEC simulations and results are provided in Appendix A. Each wire dipole was $10^{-3}\lambda$ thick and divided into 25 computational segments. For each value of d, $g_n(\theta, \phi)$ was computed over 11.25° increments of azimuth and zenith and the results used to numerically approximate (3.33). Note that, in the absence of scattering among the antennas, each dipole could be considered omnidirectional, $g_n(\frac{\pi}{2}, \phi) = 1$, and for 2D scattering (3.33) reduces to Clarke's [17] formula $J_0\left(2\pi \frac{d}{\lambda}(m-n)\right)$, where $J_0(x)$ is the zeroth-order Bessel function of the first kind.

The amplifier selected for this study is a low-cost SiGe, heterojunction bipolar transistor LNA [57] designed for use in the cellular band. In high-gain mode with $R_{\text{bias}} = 510\ \Omega$ and $f_c = 900$ MHz, its impedance matrix and Rothe-Dahlke noise parameters are:

$$\begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix} = \begin{bmatrix} 35.7\angle -82.0^\circ & 2.74\angle 91.8^\circ \\ 325\angle 119^\circ & 46.1\angle -23.3^\circ \end{bmatrix} \Omega$$

$$r_a = 9.45\ \Omega, \quad g_a = 3.24\ \text{mS}, \ z_{\text{cor}} = 35.3\angle -114^\circ\ \Omega$$

This amplifier is nearly unilateral ($z_{21} \gg z_{12}$) and the minimum noise figure is quite low (1.04 dB), ideal characteristics for front-end amplifiers. In practice the noise figure may be slightly higher due to impedance mismatch and other implementation issues, but still low enough (a few dBs) that amplifier noise cannot be regarded as dominant over thermal antenna noise.

We calculate the downstream equivalent noise resistance r_d from a textbook example

[66, pg. 654] in which the downstream components are a mixer with a 6 dB conversion loss and 7 dB noise figure and an intermediate-frequency amplifier with a 1.1 dB noise figure. The downstream noise figure is then ~ 7.6 dB. Converting this to an equivalent noise resistance is not straightforward since the noise factor (3.6) depends on additional parameters. For the purpose of this example, we assume the downstream components take the composite form of a unilateral amplifier with a single input-referred noise source, and that the above noise figures were calculated with the standard $Z_0 = 50\ \Omega$ input impedance. The equivalent noise resistance is then

$$r_d = Z_0(F_{\text{dow}} - 1) \approx 240\ \Omega\ . \tag{3.34}$$

3.6.2 Multiport and Self Matching

We now evaluate outage for two matching networks considered in [59]. In *optimal multiport matching* for minimum noise figure, the network (3.31) is chosen so that $\mathbf{Z}'_{\mathbf{A}} = z_{\text{opt}}\mathbf{I}$, where $z_{\text{opt}} = r_{\text{opt}} + jx_{\text{opt}}$ is the source impedance (3.8) that minimizes the amplifier noise factor F_{amp}. A lossless, reciprocal matching network that accomplishes this task is

$$\begin{bmatrix} \mathbf{Z}_{\mathbf{M}11} & \mathbf{Z}_{\mathbf{M}12} \\ \mathbf{Z}_{\mathbf{M}21} & \mathbf{Z}_{\mathbf{M}22} \end{bmatrix} = j \begin{bmatrix} -\mathbf{X}_{\mathbf{A}} & (r_{\text{opt}}\mathbf{R}_{\mathbf{A}})^{1/2} \\ (r_{\text{opt}}\mathbf{R}_{\mathbf{A}})^{1/2} & x_{\text{opt}}\mathbf{I} \end{bmatrix} \tag{3.35}$$

where $\mathbf{X}_{\mathbf{A}} \triangleq \frac{1}{2j}(\mathbf{Z}_{\mathbf{A}} - \mathbf{Z}_{\mathbf{A}}^{\dagger})$.

The matching network (3.35) may be difficult to realize in practice. A simpler, sub-optimum strategy is to apply to each receive antenna the two-port matching network that achieves the minimum noise figure for that antenna in isolation. This is called *self*

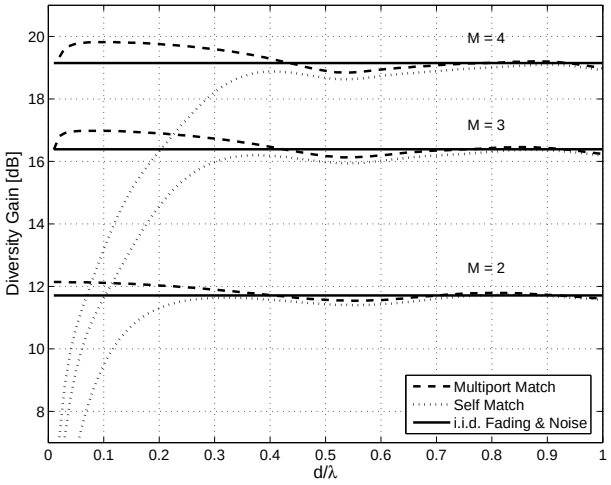


Figure 3.4: Diversity gain at 1% outage for multiport and self matching.

matching for minimum noise figure and is accomplished by the network

$$\begin{bmatrix} \mathbf{Z}_{\mathbf{M}11} & \mathbf{Z}_{\mathbf{M}12} \\ \mathbf{Z}_{\mathbf{M}21} & \mathbf{Z}_{\mathbf{M}22} \end{bmatrix} = j \begin{bmatrix} -x_A\mathbf{I} & \sqrt{r_{\text{opt}}r_A}\mathbf{I} \\ \sqrt{r_{\text{opt}}r_A}\mathbf{I} & x_{\text{opt}}\mathbf{I} \end{bmatrix} \tag{3.36}$$

where $z_A = r_A + jx_A$ is the self-impedance of each antenna in isolation.

Diversity gains for both multiport and self matching are shown in Fig. 3.4, with the i.i.d. fading and noise case (3.24) included for reference. At close inter-element spacings there is a dramatic difference between the two matching networks – gain for self-matching drops sharply for small d, while multiport matching does not. For example, there is little benefit (< 1 dB for $M = 3, 4$) in using more than two self-matched antennas for a *fixed* receiver aperture of 0.2λ, while there is apparently a significant benefit for multiport matching.

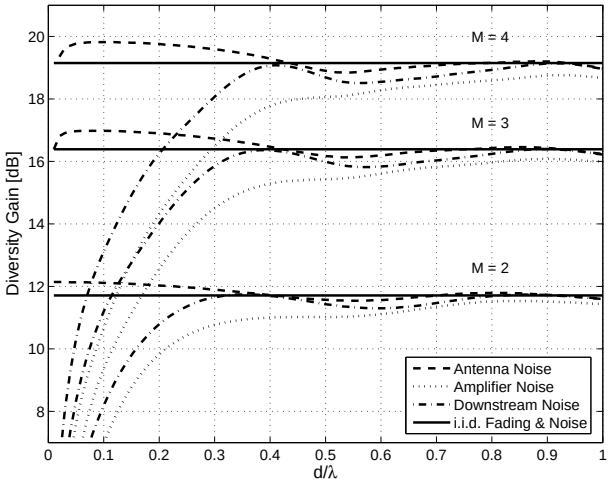


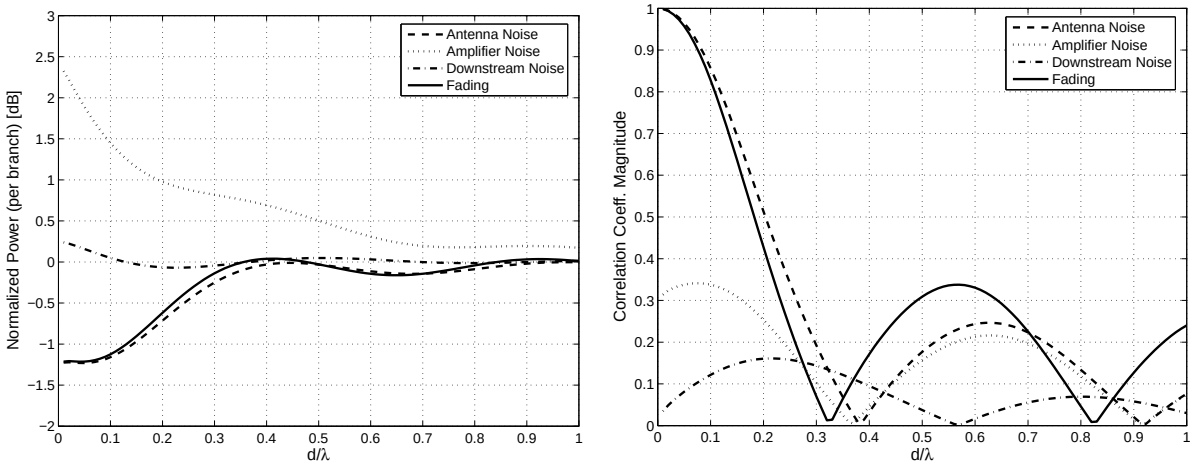
Figure 3.5: Diversity gain for various noise sources and self-matching.

3.6.3 Noise Sources

We now examine the impact of each noise source by considering antenna, amplifier, and downstream noise individually. For multiport matching $\mathbf{Z}'_{\mathbf{A}} = z_{\text{opt}}\mathbf{I}$, so from (3.32)

$$\mathbf{\Sigma} = \sigma r_A \mathbf{R}_{\mathbf{A}}^{-1/2} \mathbf{\Sigma}_{\mathbf{h}} \mathbf{R}_{\mathbf{A}}^{-1/2} \ , \tag{3.37}$$

where σ is the SNR of a single-antenna system (3.10) with $T_A = T_0$ and $z_A = z_{\text{opt}}$. Since diversity gain is unchanged by scalar multiples of $\mathbf{\Sigma}$, the impact of each noise source is identical for multiport matching. For self-matching, however, there is a strong dependence on the relative strength of each noise source. This is illustrated in Fig. 3.5, where diversity gain curves are shown for systems in which each source is dominant. Note that noise from the amplifiers has the most detrimental affect, while antenna thermal noise appears to be more benign. A system with all three noise sources (e.g., the previous example) would lie between these extremes.



(a) (b)

Figure 3.6: Output signal and noise (a) power and (b) correlation in a dual-diversity system with self-matching.

For $M = 2$, we can gain insight by examining how the power and correlation of each noise source varies with antenna spacing d. These parameters can be calculated from the matrices (3.25)-(3.27) with $\mathbf{Z_A}$ replaced by $\mathbf{Z'_A}$. For example, the antenna noise power in each branch is $\mathcal{P} = [\mathbf{\Sigma}_{\text{ant}}]_{11} = [\mathbf{\Sigma}_{\text{ant}}]_{22}$, and the correlation between the noise in each branch is $\rho = [\mathbf{\Sigma}_{\text{ant}}]_{12}/[\mathbf{\Sigma}_{\text{ant}}]_{11}$. In Fig. 3.6, we plot $\mathcal{P}$ and ρ for each noise source, and for the fading path gains (3.33). In this plot, each power is normalized by the corresponding power for $M = 1$. From Fig. 3.6(a), note that amplifier noise power increases as the antennas move closer. This explains the negative impact of amplifier noise in Fig. 3.5, and is caused by each antenna recapturing amplifier noise radiated from its neighbor. This is illustrated in Fig. 3.7, where the noise power an amplifier contributes to its own branch is compared to the power it contributes to the *other* branch. In Fig. 3.6(b) we see that both fading and antenna noise become highly correlated as the antennas move

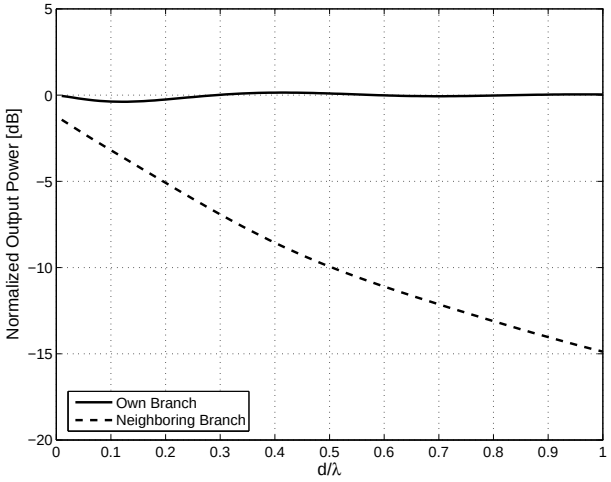


Figure 3.7: Amplifier power coupling.

closer. Since correlated noise is generally desirable in communications, exploiting this correlation in our model mitigates the undesirable impact of fading correlation for small d. Finally, observe that downstream noise power is relatively constant and correlation low for most d, similar to white noise. Indeed, setting $\mathbf{\Sigma_n} = \mathbf{I}$ and repeating the self-match simulations in Fig. 3.4 produced curves (not shown) almost indistinguishable from the downstream noise curves in Fig. 3.5. As discussed in Sec. 3.4, this is due to the amplifier unilaterality that isolates downstream components from the coupled antennas. If the amplifier was perfectly unilateral ($z_{12} = 0$), both downstream noise curves in Fig. 3.6 would be exactly zero at all spacings.

3.6.4 3D Scattering

At the end of Section 3.3 we mentioned that systems with proportional fading and noise correlation matrices have the same outage probability as a system with i.i.d. fading

and noise. In this section we examine two cases of this phenomenon, both of which follow from a well-known result in electromagnetics that we briefly present below.

Consider operating the array in transmit mode by applying a current $\mathbf{i} = [i_1 \cdots i_M]^T$ to its terminals. By superposition and reciprocity, the electric field at a large distance r from the array (i.e., in the radiation zone) is given by $\hat{\theta}\frac{1}{r}e^{-j\frac{2\pi}{\lambda}r}\sum_n i_n g_n(\theta,\phi)e^{j\frac{2\pi}{\lambda}d(1-n)\sin\theta\cos\phi}$, where $\hat{\theta}$ is the unit vector oriented along the θ direction. The array is lossless, so the real power delivered to the antennas must be equal to the radiated power:

$$\mathbf{i}^\dagger \mathbf{R_A} \mathbf{i} = \frac{1}{\eta_0}\int_0^{2\pi}\int_0^{\pi}\left|\sum_n i_n g_n(\theta,\phi)e^{j\frac{2\pi}{\lambda}d(1-n)\sin\theta\cos\phi}\right|^2 \sin\theta d\theta d\phi \ , \tag{3.38}$$

where $\eta_0 \approx 120\pi\ \Omega$ is the impedance of free space. Since this holds for all $\mathbf{i}$ we conclude that

$$[\mathbf{R_A}]_{nm} = \frac{1}{\eta_0}\int_0^{2\pi}\int_0^{\pi} g_n(\theta,\phi)g_m^*(\theta,\phi)e^{j\frac{2\pi}{\lambda}d(m-n)\sin\theta\cos\phi}\sin\theta d\theta d\phi \ . \tag{3.39}$$

This result is well-known in the antenna community (sometimes expressed in terms of scattering parameters, cf. [77]), and has been used in the context of communications to relate fading correlation and mutual impedances [89]. In particular, for a *3D scattering* model in which the received electric field arrives from each direction with equal probability, i.e., $p(\theta,\phi) = \frac{1}{4\pi}$, from (3.33) and (3.39) we see that $\mathbf{\Sigma_h} = \frac{\eta_0}{4\pi}\mathbf{R_A}$. Here we are concerned with two consequences of this result observed in[8] [30], [31] that also hold for our model:

[8]Recall that the receiver model in these references is a special case of Fig. 3.2 in which the amplifiers are unilateral with no input noise (see our discussion in Section 3.4) and downstream noise is neglected.

1. For 3D scattering and multiport matching (cf. Sec. 3.6.2), from (3.37) we see that[9] $\mathbf{\Sigma} = \frac{\sigma\eta_0}{4\pi}\mathbf{I}$.

2. For 3D scattering and *any* lossless matching network, when antenna noise is dominant (i.e., we can zero r_a, g_a, and r_d without appreciably changing $\mathbf{\Sigma_n}$), from (3.32) we have

$$\mathbf{\Sigma} \;=\; \frac{P}{4kT_0B}\mathbf{M}\mathbf{\Sigma_h}\mathbf{R_A}^{-1}\mathbf{M}^{-1} \;=\; \frac{P\eta_0}{16\pi kT_0B}\mathbf{I}, \tag{3.40}$$

   where we observed $\mathbf{R'_A} = \mathbf{M}\mathbf{R_A}\mathbf{M}^\dagger$ for lossless matching, cf. (3.31).

In both cases the diversity gain is identical to that of a system with i.i.d. fading and noise, regardless of inter-element spacing. This result is unsettling since it suggests that, at least in theory, one may realize unbounded performance improvements by packing an arbitrarily large number of antennas into a small space. However, we show below that for small d both observations are extremely sensitive to other system parameters.

Let us begin with observation 1. In this study we have assumed a narrowband, frequency-flat system, so the array impedance and radiation pattern were computed at a single frequency f_c (the value of f_c is immaterial since the antenna parameters may be normalized by λ, e.g., we are using a $\frac{\lambda}{2}$-length dipole). To test whether these results are valid in a useable frequency band around f_c, we recalculated the transformed array impedance matrix $\mathbf{Z'_A}$ and induced voltage $\mathbf{v'_o}$ with the multiport matching tuned to f_c and the array parameters calculated at $0.999f_c$, $0.99f_c$, and $0.95f_c$. The resulting

[9] In Fig. 3.4, where 2D scattering is assumed, the multiport curves are quite close to the i.i.d. fading and noise curves. Here the dipole radiation pattern attenuates many of the AoAs in the support set of $p(\theta, \phi)$ exclusive to 3D scattering.

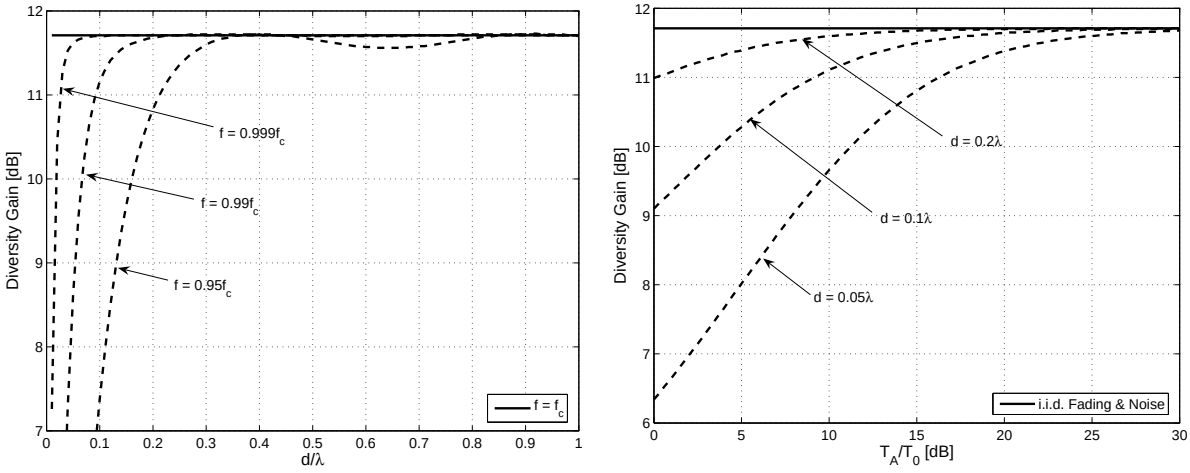


Figure 3.8: Diversity gain with 3D fading. (a) Multiport matching at various multiples of the center frequency f_c. (b) Antenna noise of varying strength.

diversity gain is shown in Fig. 3.8(a). Performance is equal to an i.i.d. fading and noise system at exactly f_c, as predicted, but drops off dramatically from this frequency for small d. A similar conclusion is reached in [50] for a decoupling matching network optimized for maximum power transfer in a system with i.i.d. noise. As noted in [31], this sharp bandwidth reduction is an artifact of supergain behavior. The problems associated with superdirective arrays have been well-chronicled in the antenna design literature (cf., [4, pg. 345]) and their implications for MIMO are considered in [60].

Now consider observation 2. Since it is clear from (3.32) that $\mathbf{\Sigma}$ is non-diagonal when we include receiver noise (i.e., r_a, g_a, and r_d are nonzero), it is interesting to consider how large we must make the antenna temperature in order to approximate an i.i.d. system. To answer this question we re-evaluated the self-matching system in Sec. 3.6.2 with 3D scattering and $T_A \neq T_0$. The resulting diversity gain is shown in Fig. 3.8(b) as a function

of antenna temperature for several inter-element spacings. Note that for $d = 0.2\lambda$, T_A only needs to be a few hundred Kelvin above standard temperature to approach the i.i.d. curve, but for $d = 0.05\lambda$ this same result requires a tremendously large temperature. Such large temperatures are usually non-thermal in origin (cf., [66, Fig. 13.6]), e.g., from a strong source of Gaussian interference.

3.6.5 Downstream Noise

As discussed in Sec. 3.4, downstream noise becomes correlated by traversing the reverse signal path (output to input) of the amplifiers and coupling through the antennas. In this example we quantify how amplifier unilaterality and antenna mutual coupling alter the impact of downstream noise on the self-matching receiver of Sec. 3.6.2. We consider 3D scattering and set $T_A = T_0$.

In Fig. 3.9 we plot diversity gain as a function of downstream noise figure (3.34), parameterized by various degrees of amplifier unilaterality and inter-element spacings. To make the amplifier [57] more (less) unilateral we performed the transformation $\{z_{12}, z_{21}\} \rightarrow \{\frac{1}{c}z_{12}, cz_{21}\}$ ($\{z_{12}, z_{21}\} \rightarrow \{cz_{12}, \frac{1}{c}z_{21}\}$), where $c = \sqrt{10}$, which alters the relative strength of the noise from the antennas (3.25) and amplifiers (3.26) while keeping the downstream noise (3.27) fixed. From this plot we see that when the antennas are strongly coupled and the amplifier is less unilateral, downstream noise has a dramatic impact on diversity gain, even for relatively low-noise devices (e.g., $F_{\text{dow}} \sim 3$ dB). For example, at $d = 0.05\lambda$ increasing the downstream noise figure from 0 to 10 dB reduces diversity gain by almost 2 dB for the "less unilateral" amplifier, while there is little change in diversity gain for the

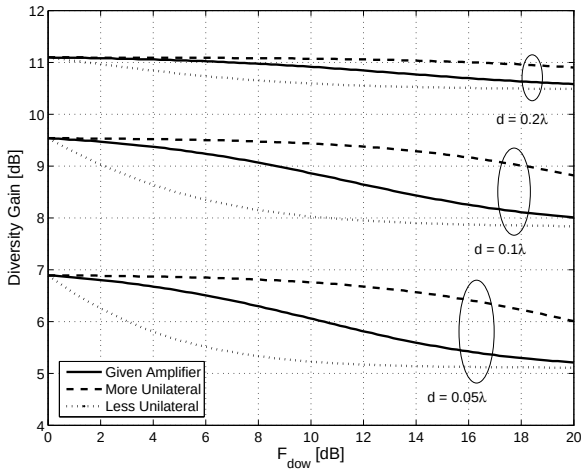


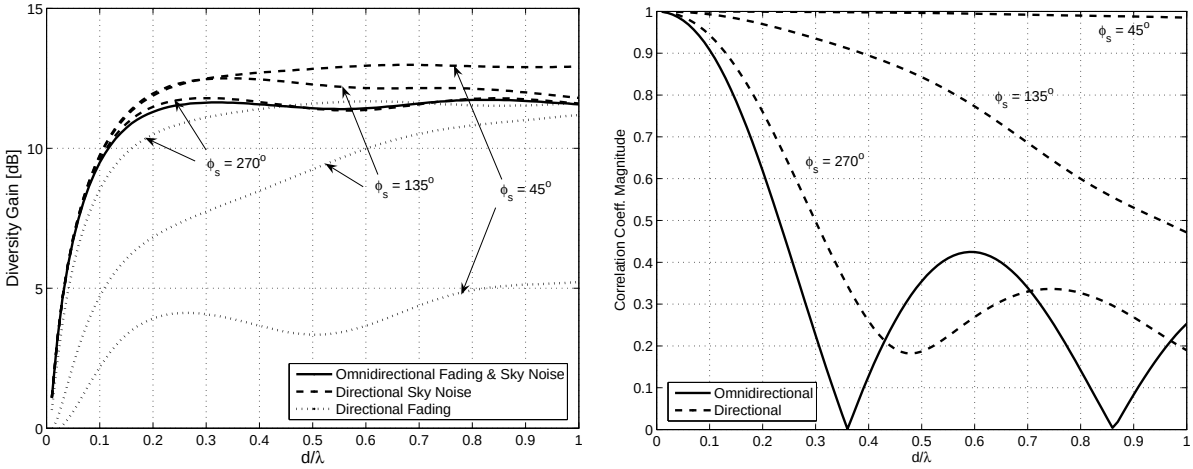
Figure 3.9: Diversity gain as a function of downstream noise figure at various antenna spacings and amplifier unilaterality.

"more unilateral" amplifier. On the other hand, at $d = 0.2\lambda$ varying both the downstream noise figure and amplifier unilaterality yields negligible change in performance.

3.6.6 Directional Fading and Sky Noise

In Fig. 3.8(b) we observed that large antenna temperatures were required to render receiver noise negligible. Such large temperatures are usually non-thermal in nature and arise from a strong source of Gaussian interference. In contrast to spherically isotropic thermal noise, interference is typically directional, so in this example we consider directional fading and sky noise.

Let us model directional fading by (3.33) with $p(\theta, \phi) = \delta\left(\theta - \frac{\pi}{2}\right) u(\phi)$, where $u(\phi)$ is uniform over $\left(-\frac{\phi_s}{2}, \frac{\phi_s}{2}\right)$ and $\phi_s \in [0, 2\pi]$ is the *angular spread*. We may replace the thermal antenna noise model used throughout this chapter with directional sky noise of



(a) (b)

Figure 3.10: (a) Diversity gain and (b) correlation for directional fading and sky noise with various angular spreads.

the same form as (3.33) by taking $\mathbf{n_o} \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma_o})$, where

$$[\mathbf{\Sigma_o}]_{nm} = \frac{4kT_ABr_A}{\phi_s} \int_{-\phi_s/2}^{\phi_s/2} g_n(\pi/2, \phi) g_m^*(\pi/2, \phi) e^{j\frac{2\pi}{\lambda}d(m-n)\cos\phi} d\phi \ , \tag{3.41}$$

which is normalized so that the sky noise power of a single-antenna receiver is $4kT_ABr_A$. We continue with the self-matching receiver model, so the SNR matrix is given by (3.32) with the $4kT_AB\mathbf{R'_A}$ term replaced by (3.41). For the purpose of this example we simply take $T_A = T_0$.

Diversity gain for various angular spreads of fading and sky noise are shown in Fig. 3.10(a). In the directional fading curve the sky noise is omnidirectional ($\phi_s = 2\pi$), and the opposite applies for the directional sky noise curve. The correlation coefficient corresponding to various angular spreads is provided in Fig. 3.10(b) for reference. Note that directional fading is generally much more harmful than directional sky noise is beneficial. For example, reducing the angular spread of fading from 360° to 45° results in an

almost 7 dB drop in diversity gain over the entire range of inter-element spacings, while reducing the sky noise angular spread by the same amount only improves performance by about 1 dB. This can be explained by observing that directional fading changes the entire signal correlation matix in (3.32), whereas directional sky noise affects only the first term of the noise correlation (i.e., amplifier and downstream noise are unchanged). Of course, if antenna noise were dominant the benefit of directional sky noise would increase.

3.7 Conclusion

A model for a compact diversity receiver was presented that articulates noise from the antennas, front-end amplifiers, and downstream components. Using this model we showed that receiver noise is generally correlated between diversity branches, in contrast the common assumption of spatially white noise. In fact, our results suggest that, in the presence of strongly coupled antennas, some stringent assumptions on the dominant source of noise and receiver structure may be needed to justify a white noise model.

We considered several examples of receivers to illustrate the characteristics of different noise sources. These examples suggest that the sources behave in profoundly different ways:

- Thermal noise from the antennas has the least detrimental effect on diversity gain since it becomes highly correlated at close antenna spacings.

- Amplifier noise can reduce diversity gain more than any other noise source since it can spread through neighboring receiver chains as the antennas become strongly

coupled.

- Noise generated downstream from the amplifiers is usually well isolated from the coupled antennas and thus behaves quite similar to spatially white noise. Its impact on diversity gain is between that of antenna and amplifier noise.

Multiport matching may significantly alter the above conclusions since one may decouple the antennas, however, at close antenna separations this may result in supergain effects that sharply reduce the system bandwidth.

While the above results are specific to the examples considered, some of the insights learned from this study may apply more broadly. Specifically, a spatially white noise model may not be accurate, even if the noise sources in each branch are independent and uncoupled. Each noise source is observed by the combiner (or any post-detection processer) only after interacting with the coupled antennas, which acts to correlate both the signal and noise components in each diversity branch. This effect is more significant for noise sources closer to the antennas, and is dependent on the unilaterality of the front-end amplifiers. We conclude that, in compact multi-antenna receivers, the impact of various noise sources depends on receiver parameters (e.g., front-end amplifier unilaterality) that are usually not considered in a standard single-antenna receiver noise budget. This is a consequence of the fact that both the power *and correlation* of each noise source are important. It is therefore vitally important to carefully identify different sources of noise, as they may have profoundly different effects on performance.

Chapter 4

Optimal Front-End Design for Compact MIMO Receivers

A critical component of modern wireless receivers is the front-end, which converts the weak RF signal collected by the antennas to a baseband signal suitable for A/D conversion. It is well-known (cf., [66]) that a SISO front-end degrades the SNR by a positive constant referred to as the *noise factor* (or *noise figure* in dB), so one usually designs the front-end for minimum noise factor. For MIMO systems with independent and identically distributed (i.i.d.) noise in each receiver chain, it is clear that one can simply apply SISO minimum noise figure design to each chain. However, when the noise is spatially colored, e.g., through mutual coupling in the antenna array or receiver, the optimal design choice is no longer clear.

In this chapter we address the problem of optimal MIMO front-end design from a communication theory perspective. Our approach will mirror the logical development

of SISO low-noise design, in which one first parameterizes system performance by SNR, then finds that designing the front-end for minimum noise figure maximizes SNR. After introducing a circuit model for the antenna array and front-end, we observe that the MIMO performance metrics derived in Sec. 2.3 depend on the positive-definiteness of the SNR matrix, which is in turn dependent on a quantity we refer to as the *noise factor matrix*. Our main result is a theorem that provides a sufficient condition for MIMO low-noise design, and is optimal for a large class of front-ends and communication schemes.

In Sec. 4.3 we apply our result to the problem of matching network design for a bank of uncoupled amplifiers. While prior studies in this area have certainly provided insight to this problem, only heuristic arguments [58], [31] of capacity-optimal matching have been provided. Matching that maximizes the output SNR of a phased array was derived in [94], however, it was only conjectured to be optimal from a channel capacity viewpoint. With the aid of the low-noise design principles developed in the prior sections we are easily able to derive matching networks that maximize capacity, and moreover, are optimal for a variety of different communication schemes.

Numerical results are presented in Sec. 4.4 that verify the optimal matching result and demonstrate that the existing SISO low-noise design principles cannot reliably predict performance in the presence of coupled antennas. With regard to the latter, we provide a bank of two SISO front-ends with the same noise factor that perform quite differently in the presence of coupled antennas – a result not predicted by standard two-port noise theory. A brief review of some relevant concepts from two-port noise modeling is provided in Appendix B.

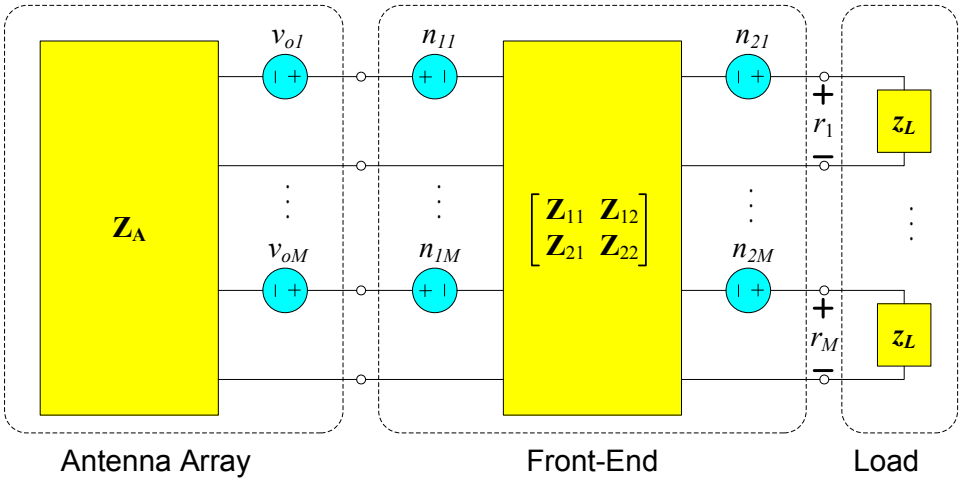


Figure 4.1: Receiver model. The antenna array and front-end are modeled with Thevenin equivalent networks.

4.1 Receiver Model

Consider a MIMO system with N transmit and M receive antennas. The signals at each receive antenna are usually amplified and down-converted before A/D conversion and subsequent digital signal processing. The analog circuits that accomplish this goal are collectively referred to as the *front-end*. In this chapter we will take a more general viewpoint than is typically found in the communications literature by only constraining the antenna array and front end to be linear networks with frequency-flat transfer function and noise power spectral density. Under these conditions we may model each device with a Thevenin equivalent network (cf., [38]), as shown in Fig. 4.1. Below we discuss these networks in more detail.

4.1.1 Antenna Array

Let $\mathbf{v}, \mathbf{i} \in \mathbb{C}^M$ denote the complex-baseband voltage across and current flowing into the antenna array terminals, respectively. The circuit equation for the array is

$$\mathbf{v} = \mathbf{Z_A}\mathbf{i} + \mathbf{v_o} \ , \tag{4.1}$$

where $\mathbf{Z_A}$ is the antenna *impedance matrix* and $\mathbf{v_o}$ is the *open-circuit* (induced) *voltage*, as shown in Fig. 4.1. Throughout this chapter we shall write impedance matrices (uniquely) as $\mathbf{Z} = \mathbf{R} + j\mathbf{X}$, where $\mathbf{R} \triangleq \frac{1}{2}(\mathbf{Z} + \mathbf{Z}^\dagger)$, $\mathbf{X} \triangleq \frac{1}{2j}(\mathbf{Z} - \mathbf{Z}^\dagger)$, and the $\dagger$ superscript denotes the conjugate-transpose. For one-port devices this reduces to $z = r + jx$, where r and x are the device resistance and reactance. For a uniform linear array (ULA), $\mathbf{Z_A} = \mathbf{R_A} + j\mathbf{X_A}$ is a function[1] of the inter-element spacing d. Mutual coupling between the antennas is reflected by the off-diagonal elements of $\mathbf{Z_A}$.

Assuming a frequency-flat propagation environment, the open-circuit voltage is given by

$$\mathbf{v_o} = \mathbf{H}\mathbf{x} + \mathbf{n_o} \ , \tag{4.2}$$

where $\mathbf{x} \in \mathbb{C}^N$ is the transmitted signal, $\mathbf{H} \in \mathbb{C}^{M \times N}$ is the channel matrix, and $\mathbf{n_o} \in \mathbb{C}^M$ is noise induced in the antennas. We assume Rayleigh fading and a Kronecker correlation model [75] with no transmit correlation, so the columns of $\mathbf{H}$ are independent and identically distributed (i.i.d.) zero-mean, circularly-symmetric, complex Gaussian random vectors, denoted by $\mathbf{h}_i \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma_h})$, where $\mathbf{\Sigma_h} \triangleq \mathrm{E}[\mathbf{h}_i \mathbf{h}_i^\dagger]$ is the fading correlation matrix and $\mathrm{E}[\,\cdot\,]$ denotes the expectation. This may be an appropriate model, for

[1] Approximate formulas are available for infinitesimally-thin wire dipoles; other antennas may be evaluated using numerical techniques [4, Ch. 8].

example, of the downlink from a base station with widely spaced antennas to a compact mobile device in a rich scattering environment. Throughout the chapter we shall impose the transmit power constraint $\mathrm{tr}(\mathbf{\Sigma_x}) \leq P$, where $\mathbf{\Sigma_x} \triangleq \mathrm{E}[\mathbf{x}\mathbf{x}^\dagger]$ is the signal correlation matrix.

For perfectly conducting antennas, $\mathbf{n_o}$ is the voltage induced in the array by noise sources in the surrounding environment. These sources may include thermal radiation, cosmic background, and interference from other electronic devices. Here we focus on thermal noise. Strictly speaking, Nyquist's [61] well-known formula for thermal noise applies to an antenna when it is enclosed by a black-body radiator [5, pg. 111]. This formula was extended to antenna arrays by Twiss [85], who showed that the noise voltage can be modeled as $\mathbf{n_o} \sim \mathcal{CN}(\mathbf{0}, 4kT_0B\mathbf{R_A})$, where $k = 1.38 \times 10^{-23}$ J/K is Boltzmann's constant, $T_0 = 290$ K is the standard temperature, and B is the bandwidth in Hz. Note that with the array imbedded in a linear medium, the reciprocity theorem of electromagnetics (cf., [4, pg. 144]) applies so that $\mathbf{Z_A}$ is symmetric, $\mathbf{R_A} = \mathrm{Re}[\mathbf{Z_A}]$, and $\mathbf{X_A} = \mathrm{Im}[\mathbf{Z_A}]$.

4.1.2 Front-End

Most prior studies of noise in MIMO receivers have been restricted to uncoupled amplifiers [58], [31]. Here we take a more general viewpoint and assume the front-end may have arbitrary coupling and noise correlation between its ports. Let $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{C}^M$ and $\mathbf{i}_1, \mathbf{i}_2 \in \mathbb{C}^M$ denote the voltages and currents at the input and output terminals of

the front-end in Fig. 4.1, respectively. These circuit quantities are related by

$$\begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{Z}_{11} & \mathbf{Z}_{12} \\ \mathbf{Z}_{21} & \mathbf{Z}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{i}_1 \\ \mathbf{i}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{n}_1 \\ \mathbf{n}_2 \end{bmatrix} , \tag{4.3}$$

where each $\mathbf{Z}_{ij} \in \mathbb{C}^{M \times M}$, and the noise voltages $\mathbf{n}_1, \mathbf{n}_2$ are zero-mean Gaussian with correlation matrices $\mathbf{\Sigma}_{ij} \triangleq \mathrm{E}[\mathbf{n}_i \mathbf{n}_j^\dagger]$.

Solving the circuit in Fig. 4.1 for the output voltage, we obtain

$$\mathbf{r} = \left[z_L \left(\mathbf{Z}_{22} + z_L \mathbf{I} - \mathbf{G}\mathbf{Z}_{12}^\dagger \right)^{-1} \mathbf{G} \right] \left(\mathbf{v_o} - \mathbf{n}_1 + \mathbf{G}^{-1}\mathbf{n}_2 \right) , \tag{4.4}$$

where $\mathbf{G} = \mathbf{Z}_{21} \left(\mathbf{Z_A} + \mathbf{Z}_{11} \right)^{-1}$ and we have assumed the impedance matrices are non-singular, i.e., every nonzero current input produces a nonzero terminal voltage. The performance metrics presented in the next section are invariant to invertible transformations applied to the received signal, so we may instead use the simpler expression

$$\mathbf{r} = \mathbf{H}\mathbf{x} + \mathbf{n} , \tag{4.5}$$

where $\mathbf{n} \triangleq \mathbf{n_o} - \mathbf{n}_1 + \mathbf{G}^{-1}\mathbf{n}_2$ is the overall system noise. Intuitively, (4.5) refers all voltages to the open-circuited antenna terminals.

4.2 Low-Noise Design and the Noise Factor Matrix

For single antenna receivers, it is well-known that designing the receiver front-end for minimum noise figure is optimal (cf., Appendix B). This follows from observing that 1) most performance metrics are optimized by maximizing the SNR and 2) minimizing the receiver noise figure maximizes the SNR. Thus we can design a receiver front-end for

minimum noise figure and be assured this is an optimal choice for a variety of applications. In this section we develop a similar notion for multi-antenna receivers. We proceed by demonstrating that, roughly speaking, 1) the performance metrics in Section 2.3 are increasing functions of the SNR matrix and 2) the SNR matrix may be maximized by minimizing a matrix we refer to as the *noise factor matrix*. We shall use the notation $\mathbf{A} \geq \mathbf{B}$ ($\mathbf{A} > \mathbf{B}$) to indicate that $\mathbf{A} - \mathbf{B}$ is nonnegative (positive) definite.

4.2.1 Form of Performance Metrics

We claim that the performance metrics in Section 2.3 may be expressed as the mean or cumulative distribution function (cdf) of the random variable

$$z = g(\mathbf{H}_w^\dagger \mathbf{\Sigma} \mathbf{H}_w) \ , \tag{4.6}$$

where for any $\mathbf{A} \geq \mathbf{B} > \mathbf{0}$ the functional g satisfies

1. $g(\mathbf{A}) \geq g(\mathbf{B})$
2. $\mathbf{A} = \mathbf{B} \Leftrightarrow g(\mathbf{C}^\dagger \mathbf{A} \mathbf{C}) = g(\mathbf{C}^\dagger \mathbf{B} \mathbf{C})$ for all full-rank $\mathbf{C} \in \mathcal{C}^{M \times N}$.

This claim is verified in Appendix C; for now we prove the following key result.

Theorem 1 *Consider two otherwise identical systems with SNR matrices* $\mathbf{\Sigma}_1 \geq \mathbf{\Sigma}_2$ *and let*

$$z_1 = g(\mathbf{H}_w^\dagger \mathbf{\Sigma}_1 \mathbf{H}_w) \ , \quad z_2 = g(\mathbf{H}_w^\dagger \mathbf{\Sigma}_2 \mathbf{H}_w) \ , \tag{4.7}$$

where g *satisfies both of the above properties. Letting* m_i *and* $F_i(x)$ *denote the mean and cdf of* z_i*, we have*

$$m_1 \geq m_2 \quad \textit{and} \quad F_1(x) \leq F_2(x) \ \forall \ x \in \mathcal{R} \ , \tag{4.8}$$

with equality if and only if $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2$.

Proof: Let $\mathbf{\Sigma}_1 \geq \mathbf{\Sigma}_2 > \mathbf{0}$. Since $\mathbf{H}_w^\dagger \mathbf{\Sigma}_1 \mathbf{H}_w \geq \mathbf{H}_w^\dagger \mathbf{\Sigma}_2 \mathbf{H}_w$ for any fixed $\mathbf{H}_w \in \mathcal{C}^{M \times N}$, we may apply Property 1 to conclude that $z_1 \geq z_2$ with probability one. Clearly $m_1 \geq m_2$, and by partitioning $\Pr[z_2 \leq x]$ we see that

$$F_2(x) = \Pr[z_2 \leq x, z_2 > z_1] + \Pr[z_2 \leq x, z_2 \leq z_1, z_1 \leq x] + \Pr[z_2 \leq x, z_2 \leq z_1, z_1 > x] \ .$$

The first term is zero and the second term is $F_1(x)$, so $F_2(x) \geq F_1(x)$. The forward part of the equality condition is trivial, so it remains to verify the converse. Let $F_1(x) = F_2(x)$ so that $m_1 = m_2$ and $z_1 - z_2$ is zero-mean. We have already seen that $z_1 - z_2$ is a nonnegative random variable, so we must have $z_1 = z_2$ with probability one. Since the support of $\mathbf{H}_w$ is the set of full-rank $M \times N$ matrices, the desired result follows immediately from Property 2. ■

4.2.2 The Noise Factor Matrix

The noise factor of a two-port network is the ratio of the output noise power to the noise power contributed by a thermal source alone (cf., Appendix B). This concept may be extended to multiport networks by defining the *noise factor matrix* as

$$\begin{aligned} \mathbf{F} &\triangleq \left(\mathbf{\Sigma_n}\big|_{\mathbf{z=0}}\right)^{-1/2} \mathbf{\Sigma_n} \left(\mathbf{\Sigma_n}\big|_{\mathbf{z=0}}\right)^{-1/2} \\ &= \mathbf{I} + \frac{1}{4kT_0B} \mathbf{R_A}^{-1/2} \mathrm{E}\left[\mathbf{z}\mathbf{z}^\dagger\right] \mathbf{R_A}^{-1/2}, \end{aligned} \tag{4.9}$$

where $\mathbf{z} \triangleq \mathbf{G}^{-1}\mathbf{n}_2 - \mathbf{n}_1$ is noise from the front-end. The two-port noise factor is useful since it equals the SNR drop across the network. Thus, from a system design perspective,

the noise factor provides a sufficient description of a device and we may ignore its internal details. As we shall see, $\mathbf{F}$ enjoys similar properties. Indeed, (3.22) may be expressed as

$$\mathbf{\Sigma} = \frac{P}{4kT_0B}\mathbf{\Sigma}_{\mathbf{h}}^{1/2}\mathbf{R}_{\mathbf{A}}^{-1/2}\mathbf{F}^{-1}\mathbf{R}_{\mathbf{A}}^{-1/2}\mathbf{\Sigma}_{\mathbf{h}}^{1/2}. \tag{4.10}$$

The noise factor of a SISO front end should be made as small as possible so that the output SNR (and thereby the performance metric) is maximized, as discussed in Appendix B. By using (4.10) and Theorem 1, we can readily state the following analogous result for MIMO receivers.

Theorem 2 *Consider two otherwise identical MIMO systems with front-end noise factor matrices* $\mathbf{F}_1 \leq \mathbf{F}_2$*. Then the CSIR* (2.20)*, full CSI* (2.21)*, and V-BLAST ZF-SC* (2.25) *and MMSE-SC* (2.29) *capacities, STBC pairwise error probabilities* (2.33)*, and MIMO-MRC outage probabilities* (2.37) *of these systems satisfy*

- $C_R^{(1)} \geq C_R^{(2)}$
- $C_F^{(1)} \geq C_F^{(2)}$
- $C_{ZF}^{(1)} \geq C_{ZF}^{(2)}$
- $C_{MMSE}^{(1)} \geq C_{MMSE}^{(2)}$
- $P_{ij}^{(1)} \leq P_{ij}^{(2)}$
- $P_{out}^{(1)} \leq P_{out}^{(2)}$,

with equality if and only if $\mathbf{F}_1 = \mathbf{F}_2$.

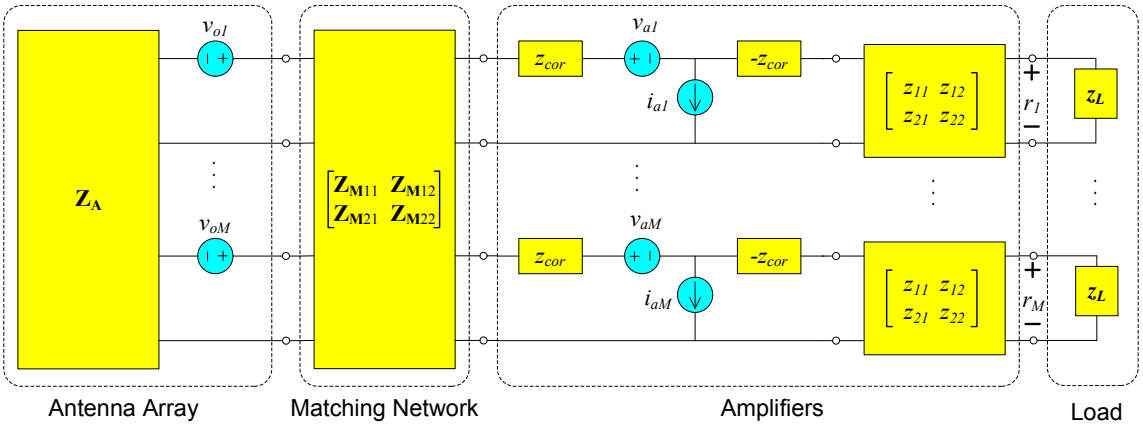


Figure 4.2: A simple front-end consisting of a matching network and M uncoupled amplifiers.

This completes our goal of developing low-noise design principles for multi-antenna receivers of the form in Fig. 4.1. The result is a *sufficient* condition for optimal MIMO front-end design. That is, suppose we have a set $\mathcal{F}$ of distinct noise factor matrices such that there exists an

$$\mathbf{F}^o = \inf \mathcal{F} \ , \tag{4.11}$$

where the infimum is with respect to the positive definite partial ordering of Hermitian matrices (cf., [40, Sec. 7.7]). From the above result we may conclude that the performance of a system built with $\mathbf{F}^o$ will be strictly better than one built with any other $\mathbf{F} \in \mathcal{F}$. For the general problem, however, we note that unlike a totally ordered set (e.g., the real numbers), an arbitrary partially ordered set may not have an infimum. For example, if $\mathbf{F}_2 - \mathbf{F}_1$ is indefinite, Theorem 2 does not apply and we cannot make any immediate conclusions about the relative performance of these systems. Fortunately, there are practical scenarios where one can find an infimum, as demonstrated in the next section.

4.3 Optimal Matching for Front-End Amplifiers

Consider a front-end consisting of a $2M$-port matching network and a bank of M uncoupled, identical amplifiers, as shown in Fig. 4.2. The matching network has impedance matrix

$$\begin{bmatrix} \mathbf{Z_{M}}_{11} & \mathbf{Z_{M}}_{12} \\ \mathbf{Z_{M}}_{21} & \mathbf{Z_{M}}_{22} \end{bmatrix}, \tag{4.12}$$

where each submatrix is $M \times M$. In addition to our usual assumption of nonsingularity, we shall restrict the matching network to be lossless so that it does not dissipate power or generate noise. Such a network satisfies [38, pg. 13] $\mathbf{Z_{M}}_{11} = -\mathbf{Z_{M}}_{11}^{\dagger}$, $\mathbf{Z_{M}}_{22} = -\mathbf{Z_{M}}_{22}^{\dagger}$, and $\mathbf{Z_{M}}_{21} = -\mathbf{Z_{M}}_{12}^{\dagger}$. The amplifiers are represented by Rothe-Dahlke equivalent networks (cf., Appendix B) with $\mathbf{v_a} \sim \mathcal{CN}(\mathbf{0}, 4kT_0Br_a\mathbf{I})$ and $\mathbf{i_a} \sim \mathcal{CN}(\mathbf{0}, 4kT_0Bg_a\mathbf{I})$ mutually independent. Our goal is to find the (possibly non-unique) optimal matching network for this system.

Before proceeding, we note that similar problems have been posed in the literature. For example, front-ends of this form were considered in [58] and [31], however, in the former antenna thermal noise was neglected and in the latter additional constraints were imposed upon the amplifiers. In these papers optimal matching schemes were suggested heuristically [58] or through numerical simulations [31]. More recently, in [94] a matrix Taylor series expansion was used to derive an optimal matching network for a phased array, and the authors conjectured that such a network may also be optimal from a channel capacity viewpoint. Here we use the low-noise design principles summarized in Theorem 2 to easily derive a matching condition that is optimal for a variety of communication schemes.

We first derive the infimum of the set of noise factor matrices with lossless matching, then seek matching networks that achieve this minimum. After some tedious but essentially straightforward circuit analysis, the input-referred front-end noise may be found as

$$\mathbf{z} = \mathbf{M}^{-1} \left[\mathbf{v_a} + (\mathbf{Z'_A} + z_{\rm cor}\mathbf{I})\, \mathbf{i_a} \right] \ , \tag{4.13}$$

where

$$\mathbf{M} = \mathbf{Z}_{\mathbf{M}21} \left(\mathbf{Z_A} + \mathbf{Z}_{\mathbf{M}11} \right)^{-1} \qquad \text{and} \qquad \mathbf{Z'_A} = -\mathbf{M}\mathbf{Z}_{\mathbf{M}12} + \mathbf{Z}_{\mathbf{M}22} \ . \tag{4.14}$$

Using this in (4.9), the noise factor matrix is

$$\mathbf{F} = \mathbf{I} + \mathbf{R_A}^{-1/2}\mathbf{M}^{-1} \left[r_a\mathbf{I} + g_a \left(\mathbf{Z'_A} + z_{\rm cor}\mathbf{I} \right) \left(\mathbf{Z'_A} + z_{\rm cor}\mathbf{I} \right)^\dagger \right] \mathbf{M}^{-\dagger}\mathbf{R_A}^{-1/2} \ , \tag{4.15}$$

which reduces to the classical noise factor (B.4) for $M = 1$. The noise factor of a single-antenna receiver achieves its minimum $F_{\rm min}$ (B.6) when the source impedance is $z_{\rm opt} = r_{\rm opt} + jx_{\rm opt}$ (B.7). We now derive a multi-antenna generalization of this result. Let us factor (4.15) as follows:

$$\begin{aligned}
\mathbf{F} &= \mathbf{I} + g_a\mathbf{R_A}^{-1/2}\mathbf{M}^{-1} \left[\mathbf{Z'_A}\mathbf{Z'^\dagger_A} + 2r_{\rm cor}\mathbf{R'_A} + 2x_{\rm cor}\mathbf{X'_A} + |z_{\rm opt}|^2\mathbf{I} \right] \mathbf{M}^{-\dagger}\mathbf{R_A}^{-1/2} \\
&= \mathbf{I} + g_a\mathbf{R_A}^{-1/2}\mathbf{M}^{-1} \left[(\mathbf{Z'_A} - z_{\rm opt}\mathbf{I})(\mathbf{Z'_A} - z_{\rm opt}\mathbf{I})^\dagger + 2(r_{\rm cor} + r_{\rm opt})\mathbf{R'_A} \right] \mathbf{M}^{-\dagger}\mathbf{R_A}^{-1/2} \\
&= F_{\rm min}\mathbf{I} + g_a\mathbf{R_A}^{-1/2}\mathbf{M}^{-1}(\mathbf{Z'_A} - z_{\rm opt}\mathbf{I})(\mathbf{Z'_A} - z_{\rm opt}\mathbf{I})^\dagger\mathbf{M}^{-\dagger}\mathbf{R_A}^{-1/2}
\end{aligned} \tag{4.16}$$

In the first step we used $|z_{\rm opt}|^2 = \frac{r_a}{g_a} + |z_{\rm cor}|^2$, in the second step we completed the square and observed that $x_{\rm opt} = -x_{\rm cor}$, and the third step made use of the identities $\mathbf{R'_A} = \mathbf{M}\mathbf{R_A}\mathbf{M}^\dagger$ (for lossless matching) and $F_{\rm min} = 1 + 2g_a(r_{\rm cor} + r_{\rm opt})$. The second term in (4.16) is nonnegative definite, so

$$\mathbf{F} \geq F_{\rm min}\mathbf{I} \ , \tag{4.17}$$

with equality if and only if $\mathbf{Z}'_{\mathbf{A}} = z_{\text{opt}}\mathbf{I}$. From Theorem 2 we conclude that the optimal matching network transforms the antenna array into a bank of uncoupled impedances z_{opt}, and any matching that does not perform this transformation is strictly suboptimal.

The problem of finding a lossless matching network that achieves the desired impedance transformation is heavily over-determined – we shall see there is no loss of optimality by imposing the additional restriction of reciprocity. This guarantees that the matching network can be constructed with passive, reactive elements (cf., [66, pg. 171]), and when combined with the lossless assumption restricts (4.12) to be of the form

$$j \begin{bmatrix} \mathbf{X}_{\mathbf{M}11} & \mathbf{X}_{\mathbf{M}12} \\ \mathbf{X}_{\mathbf{M}12}^T & \mathbf{X}_{\mathbf{M}22} \end{bmatrix}, \tag{4.18}$$

where each submatrix is real, $\mathbf{X}_{\mathbf{M}11}, \mathbf{X}_{\mathbf{M}22}$ are symmetric, and the T superscript denotes the matrix transpose. From (4.14) we obtain[2]

$$\mathbf{R}'_{\mathbf{A}} = \mathbf{X}_{\mathbf{M}12}^T \left[\mathbf{R}_{\mathbf{A}} + (\mathbf{X}_{\mathbf{A}} + \mathbf{X}_{\mathbf{M}11})\, \mathbf{R}_{\mathbf{A}}^{-1} (\mathbf{X}_{\mathbf{A}} + \mathbf{X}_{\mathbf{M}11})\right]^{-1} \mathbf{X}_{\mathbf{M}12} \tag{4.19}$$

$$\mathbf{X}'_{\mathbf{A}} = \mathbf{X}_{\mathbf{M}22} - \mathbf{X}_{\mathbf{M}12}^T \left[\mathbf{X}_{\mathbf{A}} + \mathbf{X}_{\mathbf{M}11} + \mathbf{R}_{\mathbf{A}} (\mathbf{X}_{\mathbf{A}} + \mathbf{X}_{\mathbf{M}11})^{-1} \mathbf{R}_{\mathbf{A}}\right]^{-1} \mathbf{X}_{\mathbf{M}12}\ . \tag{4.20}$$

Setting $\mathbf{Z}'_{\mathbf{A}} = z_{\text{opt}}\mathbf{I}$ and solving these equations for the matching impedances, we see that for any given $\mathbf{X}_{\mathbf{M}11}$ the solution is of the form

$$\mathbf{X}_{\mathbf{M}12} = \sqrt{r_{\text{opt}}} \left[\mathbf{R}_{\mathbf{A}} + (\mathbf{X}_{\mathbf{A}} + \mathbf{X}_{\mathbf{M}11})\, \mathbf{R}_{\mathbf{A}}^{-1} (\mathbf{X}_{\mathbf{A}} + \mathbf{X}_{\mathbf{M}11})\right]^{1/2} \mathbf{U} \tag{4.21}$$

$$\mathbf{X}_{\mathbf{M}22} = x_{\text{opt}}\mathbf{I} + \mathbf{X}_{\mathbf{M}12}^T \left[\mathbf{X}_{\mathbf{A}} + \mathbf{X}_{\mathbf{M}11} + \mathbf{R}_{\mathbf{A}} (\mathbf{X}_{\mathbf{A}} + \mathbf{X}_{\mathbf{M}11})^{-1} \mathbf{R}_{\mathbf{A}}\right]^{-1} \mathbf{X}_{\mathbf{M}12}\ , \tag{4.22}$$

where $\mathbf{U}$ is an arbitrary real-orthogonal matrix. Thus the optimal matching network is only unique up to the choice of $\mathbf{X}_{\mathbf{M}11}$ and $\mathbf{U}$. By picking $\mathbf{X}_{\mathbf{M}11} = -\mathbf{X}_{\mathbf{A}}$ (recall this

[2]We have made use of the identity $(\mathbf{R} + j\mathbf{X})^{-1} = (\mathbf{R} + \mathbf{X}\mathbf{R}^{-1}\mathbf{X})^{-1} - j(\mathbf{X} + \mathbf{R}\mathbf{X}^{-1}\mathbf{R})^{-1}$.

matrix is real and symmetric by reciprocity) and $\mathbf{U} = \mathbf{I}$, we obtain a multi-antenna extension of (B.8):

$$\begin{bmatrix} \mathbf{Z}_{\mathbf{M}11} & \mathbf{Z}_{\mathbf{M}12} \\ \mathbf{Z}_{\mathbf{M}21} & \mathbf{Z}_{\mathbf{M}22} \end{bmatrix} = j \begin{bmatrix} -\mathbf{X}_{\mathbf{A}} & (r_{\text{opt}}\mathbf{R}_{\mathbf{A}})^{1/2} \\ (r_{\text{opt}}\mathbf{R}_{\mathbf{A}})^{1/2} & x_{\text{opt}}\mathbf{I} \end{bmatrix} . \tag{4.23}$$

As a concluding remark, we note that the derivation of the optimal matching condition (4.17) relied heavily on the fact that the front-ends in Fig. 4.2 were uncoupled, uncorrelated, and identical. There are many practical scenarios that may violate any of these conditions, such as noise correlation between receiver chains from a common local oscillator or multi-band devices that use amplifiers tuned to different frequencies. In such cases one should not necessarily expect that decoupling the antenna array is optimal.

4.4 Numerical Results

With the analytical portion of the chapter complete, we present some numerical results to illustrate our theory. We first describe the antenna array and propagation model, then evaluate capacity for several front-ends of interest. Note that we may write the SNR matrix (3.22) as

$$\mathbf{\Sigma} = \sigma r_A F \mathbf{\Sigma}_{\mathbf{h}}^{1/2} \mathbf{R}_{\mathbf{A}}^{-1/2} \mathbf{F}^{-1} \mathbf{R}_{\mathbf{A}}^{-1/2} \mathbf{\Sigma}_{\mathbf{h}}^{1/2} , \tag{4.24}$$

where σ, r_A, and F are the SNR, antenna resistance, and noise factor of the corresponding SISO system in Appendix B, and $\mathbf{\Sigma_h}$ is normalized by the SISO channel attenuation $\mathrm{E}|h|^2$. In all our examples we take the SISO SNR as 10 dB. Results are given for $M = N = 1-4$ antennas with spacings of $0.01\lambda \leq d \leq \lambda$, where λ is the carrier wavelength.

All metrics except V-BLAST capacity were evaluated using numerical integration of closed-form expressions available in the literature, see our discussion at the end of Section 2.3. The V-BLAST ZF and MMSE capacities were evaluated using Monte Carlo simulations over 5000 realizations of $\mathbf{H}_w$, which were also used to verify the aforementioned numerical integrations.

4.4.1 Antenna Array and Propagation Model

The receive array is a ULA of half-wavelength dipoles. Expressions for the impedance matrix and radiation pattern of such an array are available for infinitesimally thin wire dipoles [4], however, we prefer to evaluate finite-thickness dipoles numerically. The Numerical Electromagnetics Code (NEC) [9], a well-known Method-of-Moments (MoM) based program, was used for this purpose. Each wire dipole is $10^{-3}\lambda$ thick and divided into 25 computational segments. Details of the NEC simulations and results are provided in Appendix A.

A sufficient condition for Rayleigh fading is that the received electric field be composed of a large number (in the sense of the central limit theorem) of plane waves modulated by independent and uniformly-distributed phases. We consider a two-dimensional model in which the plane waves are co-polarized with the dipole orientation and arrive with equal probability over the azimuthal range. The entries of the open-circuit correlation matrix $\mathbf{\Sigma_h}$ are then given by [17]

$$[\mathbf{\Sigma_h}]_{nm} = \frac{1}{2\pi}\int_0^{2\pi} g_n\left(\pi/2,\phi\right) g_m^*\left(\pi/2,\phi\right) e^{j2\pi\frac{d}{\lambda}(m-n)\cos\phi} d\phi \ , \tag{4.25}$$

where $g_n\left(\pi/2,\phi\right)$ is the voltage induced in the n^{th} antenna by a plane wave arriving with

azimuthal angle ϕ (cf., Appendix A).

The $g_n(\pi/2, \phi)$ functions were evaluated in NEC over 32 evenly spaced angles (for each inter-element spacing) and the results used in a numerical approximation of (4.25). This approach accounts for scattering between neighboring antennas – if we assumed the omnidirectionality of isolated dipoles, (4.25) would reduce to Clarke's [17] formula $J_0\left(2\pi\frac{d}{\lambda}(m-n)\right)$, where $J_0(x)$ is the zeroth-order Bessel function of the first kind.

4.4.2 Noise Factor Example

The noise factor matrix $\mathbf{F}$ introduced in this chapter provides a sufficient condition for low-noise design, whereas the two-port noise factor F (B.2) may be unable to predict performance in the presence of coupled antennas. We demonstrate the latter point in this example by evaluating the performance of two amplifiers with the same noise factor.

Referring to Fig. 4.1, consider two uncoupled front-ends (label them "A" and "B") with identical impedance matrices

$$\begin{bmatrix} \mathbf{Z}_{11} & \mathbf{Z}_{12} \\ \mathbf{Z}_{21} & \mathbf{Z}_{22} \end{bmatrix} = \begin{bmatrix} z_{11}\mathbf{I} & \mathbf{0} \\ z_{21}\mathbf{I} & z_{22}\mathbf{I} \end{bmatrix}, \tag{4.26}$$

where $z_{11} = z_{22} = 50\ \Omega$ and $z_{21} = 500\ \Omega$, and uncorrelated noise sources $\mathbf{\Sigma}_{12} = \mathbf{\Sigma}_{21} = \mathbf{0}$, $\mathbf{\Sigma}_{ii} = 4kT_0Br_i$, $i \in \{1, 2\}$. Let the noise resistances of each front-end be assigned as

$$\begin{array}{ll} \text{Front-end A}: & r_1 = 9r_A,\ r_2 = 0 \\ \text{Front-end B}: & r_1 = 0,\ r_2 = 9r_A|G|^2 \end{array}, \tag{4.27}$$

where $G = z_{21}/(z_A + z_{11})$ and $z_A = r_A + jx_A$ is the impedance of an isolated antenna. Note that the branches of both of these front-ends have a 10 dB noise figure, hence for uncoupled antennas they should perform identically.

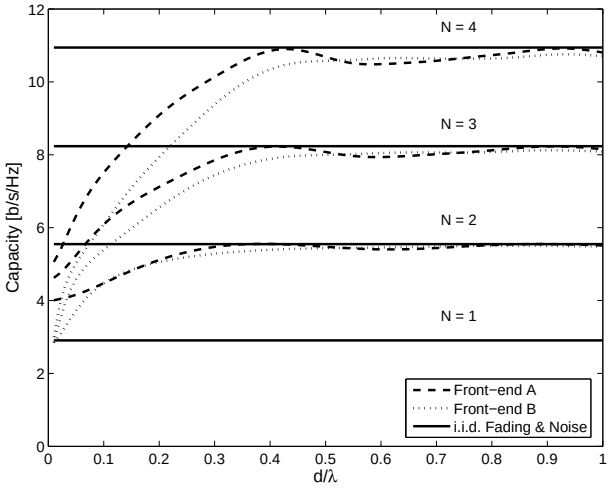


Figure 4.3: CSIR capacity versus inter-element spacing for two front-ends.

The CSIR capacity (2.20) of each of these systems for $N = 1 - 4$ is shown in Fig. 4.3, with the i.i.d. fading and noise case provided for reference. We immediately notice that for strongly coupled antennas (say, $d < 0.5\lambda$) these two front-ends behave quite differently, in stark contrast to the above statement for uncoupled antennas (which appears to be in agreement here for $d \geq 0.5\lambda$). In particular, a four-antenna receiver with $d = 0.1\lambda$ will benefit from a 25% increase in capacity by choosing front-end A over front-end B, even though they have the same noise figure. This somewhat peculiar result is understood by observing that the noise in front-end A correlates through the coupled antennas, while that of front-end B is isolated from the antennas and thus remains spatially white for all inter-element spacings. It is well-known that noise correlation can only increase capacity, thus front-end A should have a larger capacity at small spacings.

4.4.3 Matching Example

In Section 4.3 we derived an optimal matching network for the system of Fig. 4.2. While the network (4.23) achieves the minimum noise factor matrix, it may be difficult to realize in practice. A practical, suboptimal alternative is to apply to each antenna the (two-port) optimal matching network (B.8) for an isolated dipole. We shall follow the terminology introduced in [58] and refer to this as *self-matching*. Below we compare optimal and self matching for the six performance metrics presented in Sec. 2.3. In each plot the $N = 3$ case is suppressed for clarity.

The amplifier selected for this example is a SiGe low-noise amplifier [57] designed for use in the cellular band. In high-gain mode with $R_{\text{bias}} = 510\ \Omega$ and $f = 900$ MHz, its impedance matrix and Rothe-Dahlke noise parameters are:

$$\begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix} = \begin{bmatrix} 35.7\angle -82.0^{\text{o}} & 2.74\angle 91.8^{\text{o}} \\ 325\angle 119^{\text{o}} & 46.1\angle -23.3^{\text{o}} \end{bmatrix} \Omega$$

$$r_a = 9.45\ \Omega, \quad g_a = 3.24\ \text{mS}, \ z_{cor} = 35.3\angle -114^{\text{o}}\ \Omega$$

This amplifier is approximately unilateral ($z_{21} \gg z_{12}$) and the minimum noise figure is quite low (1.04 dB), ideal characteristics for front-end amplifiers.

The ergodic capacity for systems with CSIR (2.20) and full CSI (2.21) are shown in Fig. 4.4, with the i.i.d. fading and noise capacity provided for reference. At small inter-element spacings, the optimal match is superior to the self match, and is an upper bound at all spacings in accordance with the results derived in Section 4.3. For self-matching, the fading and noise correlations approach unity as the antennas are brought closer together,

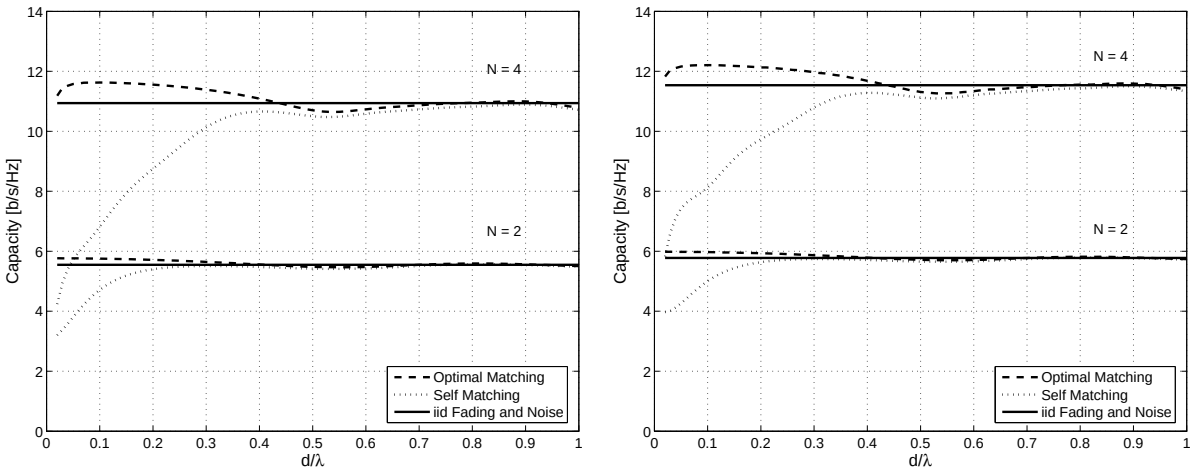


Figure 4.4: (a) CSIR and (b) full CSI capacity of $N \times N$ systems for optimal and self matching.

and the capacity reduces to that of a SISO system. In contrast, the optimal match is able to completely decouple the array at any distance, and the performance of an i.i.d. system is retained even at extremely close spacings. However, the results for optimal matching should be interpreted with caution since the array may exhibit supergain behavior [31]. Consequently, the system is highly sensitive to channel estimation errors, and, as shown in[3] [50], becomes extremely narrowband at small inter-element spacings.[4]

Capacity for V-BLAST systems with ZF-SC (2.26) and MMSE-SC (2.29) receivers are shown in Fig. 4.5. Similar observations about the performance of optimal and self matching can be made. Note that the MMSE-SC curve is identical to the CSIR capacity,

[3]In this study the match was optimized for power transfer instead of noise, however, the decoupling property responsible for these effects remains.

[4]In fact, this narrowband behavior may be used to explain the drop in the optimal match curve for $d < 0.1\lambda$ in Fig. 3 of [58] that is not observed in our results in Fig. 4.4(a). The authors of [58] use a time-limited sinusoid in their finite-difference time-domain (FDTD) program, while here the MoM-based NEC assumes time-harmonic fields.

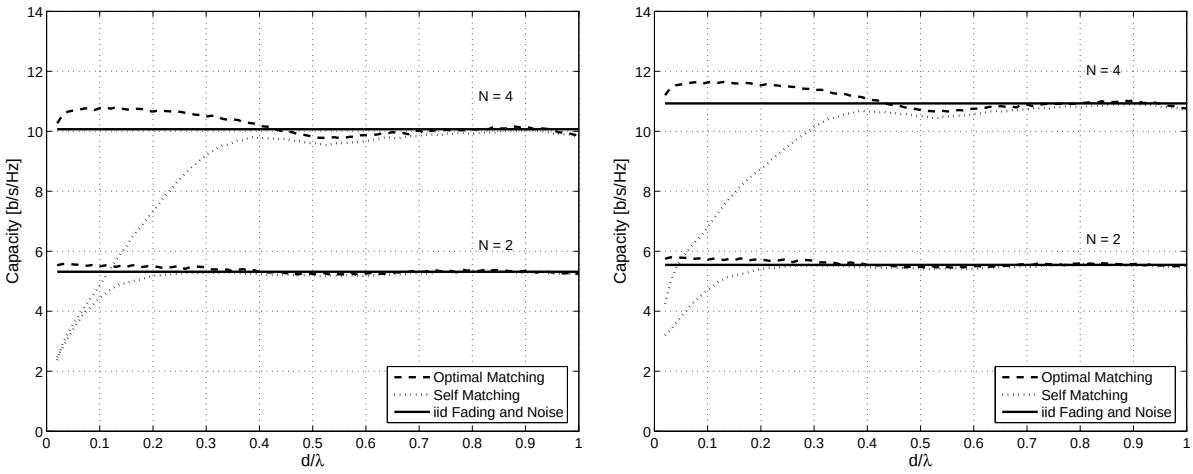
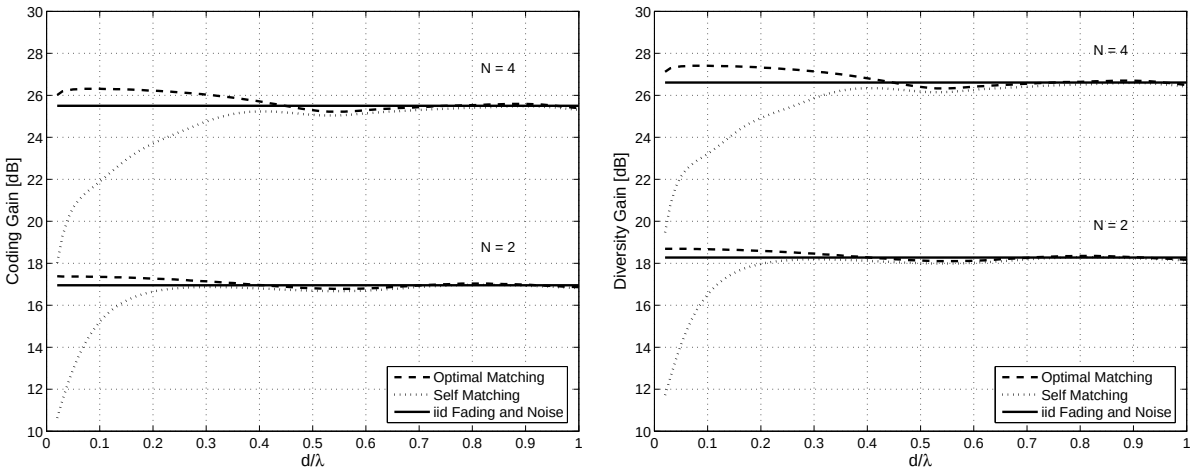


Figure 4.5: (a) ZF-SC and (b) MMSE-SC V-BLAST capacity of $N \times N$ systems for optimal and self matching.

confirming that MMSE with successive cancelation is still optimal for correlated noise. Finally, coding and diversity gains for orthogonal STBCs (2.34) and MIMO-MRC (2.37) are shown in Fig. 4.6. Here the coding gain is defined as the SNR difference between the PEP and $N = 1$ Rayleigh curves at an error probability of 10^{-3}, and diversity gain is similarly defined for the outage curves at a 1% probability of outage. In Fig. 4.6(a) we use a BPSK constellation with a rate-1 code (e.g., the Alamouti code [2]) for $N = 2$ and a rate-1/2 code for $N = 4$, so $NR = 2$ for each curve.

4.5 Conclusion

In this chapter we derived optimal front-end design principles from a communication theory perspective by examining the form of several MIMO communication schemes. Through this approach vector channel generalizations of SNR and noise factor were de-



(a) (b)

Figure 4.6: (a) Orthogonal STBC coding gain at 10^{-3} error probability and (b) MIMO-MRC diversity gain at 1% outage probability for $N \times N$ systems for optimal and self matching.

veloped, and a sufficient condition for low-noise design was derived in Theorem 2. The result applies to a larger class of front-ends than is typically considered in the communications literature, but may be readily applied to practical problems such as amplifier matching. In fact, with the developed theory we were able to derive optimal matching networks that were only conjectured to be optimal in previous studies. Through numerical simulations we verified this result, and also demonstrated the need for alternative noise measures such as the noise factor matrix by showing that the two-port noise factor may not be able to accurately predict performance in the presence of coupled antennas.

Chapter 5

Conclusion

In this concluding chapter we summarize the main results of this dissertation and provide some recommendations for future work.

5.1 Summary of Dissertation

We began our study in Chapter 2 by examining the impact of noise correlation on the performance of several MIMO communication schemes. In particular, we looked at capacity with CSIR and full CSI, spatial multiplexing with ZF-SC and MMSE-SC receivers, space-time block codes, and beamforming. In each case we observed that fading and noise correlation had very different effects on performance: fading correlation was detrimental and noise correlation was beneficial. This observation, and the lack of noise modeling in the MIMO literature, motivated the detailed study of noise modeling in subsequent chapters.

In Chapter 3 we introduced a circuit model for a compact multi-antenna receiver. The model allowed us to examine how noise from the antennas, front-end amplifiers, and downstream components may become spatially correlated. Moreover, it was shown that these noise sources affect the performance of an optimal diversity combiner in profoundly different ways. Each noise source is observed by the combiner only after interacting with the coupled antennas, which acts to correlate both the signal and noise components in each receiver branch. This effect is more significant for noise sources closer to the antennas, and is dependent on the unilaterality of the front-end amplifiers.

While Chapter 3 studied noise correlation for a specific receiver front-end, in Chapter 4 we considered a much more general front-end. By observing the form of the performance metrics introduced in Chapter 2, we were able to extend existing low-noise design principles for single-antenna receivers to MIMO systems. Specifically, we showed that the performance of each MIMO system was uniquely optimized by designing the front-end such that the noise factor matrix is minimized (with respect to the positive definite partial ordering). By applying these low-noise design principles to the amplifier matching problem, we were able to prove the optimality of a matching network that was only conjectured to be optimal in prior studies.

During the literature review in Chapter 1 we made the observation that the modeling and impact of fading correlation in MIMO has received much more attention than noise. While we believe this dissertation is a good first step towards closing the gap between our understanding of fading and noise correlation, there are still many issues we would like to see addressed. We conclude by reviewing some of these problems in the next section.

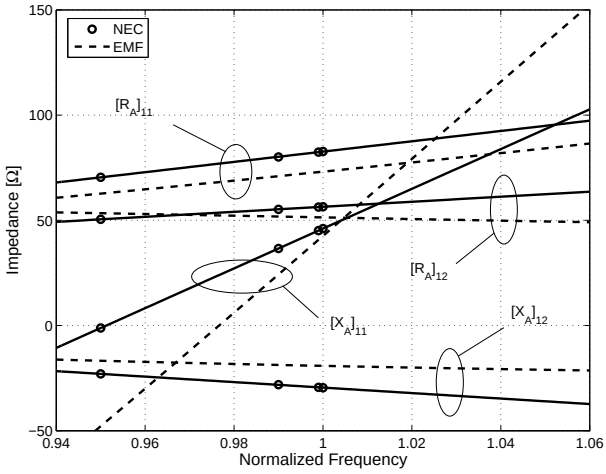


Figure 5.1: Frequency response of a two-element dipole array with inter-element spacing $d = 0.2\lambda$, as computed by the NEC and the induced EMF method.

5.2 Future Work

During the development of this dissertation several interesting research problems were encountered that still remain unsolved. In this section we briefly present some of these open problems and, in the hopes that future researchers consider them, comment on some potential solution approaches.

Optimal Broadband Matching. In Sec. 3.6.4 we found that the optimal matching network exhibited highly narrowband behavior at close inter-element spacings. An open problem is deriving the optimal matching network over a non-zero bandwidth (as opposed to the single-frequency derivation in Chapter 4). To begin, one would need to obtain the frequency response of the antenna impedances. In Fig. 5.1 we computed the frequency response of a two-element dipole array with inter-element spacing $d = 0.2\lambda$ using the two approaches presented in Appendix A: the Numerical Electromagnetics Code (NEC) and

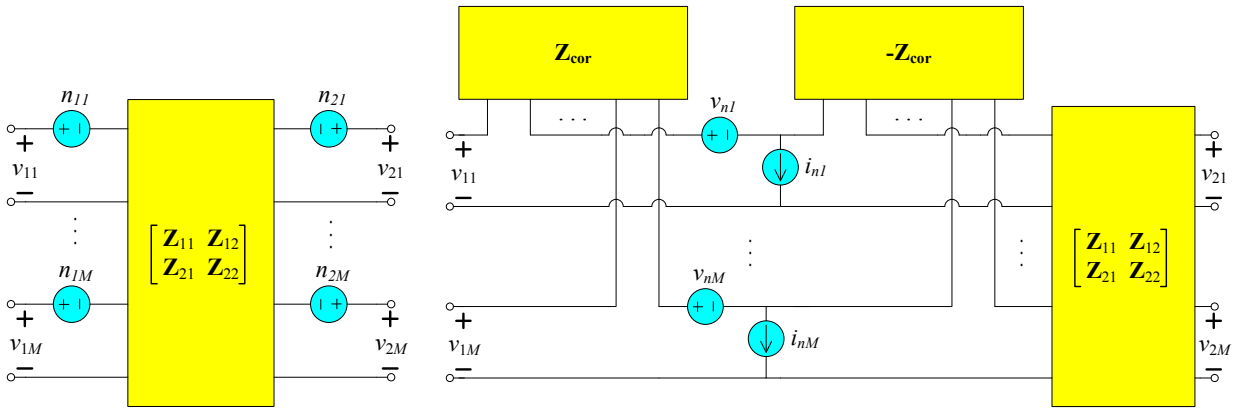
the induced EMF method. The NEC impedances were computed at the center frequency f_c and at $0.999f_c$, $0.99f_c$, and $0.95f_c$ (the value of f_c is immaterial since the antenna parameters may be normalized by the wavelength, e.g., we are using a half-wavelength dipole); a linear least-squares fit line is shown through the actual data points. The induced EMF curves were obtained by varying the wavelength λ in (A.18). From these results it appears that, over a bandwidth of 10% of the center frequency, a reasonably good approximation to the frequency response is given by

$$\mathbf{Z_A}(\omega) = \mathbf{Z}_0 + \omega\mathbf{Z}_1 \ , \tag{5.1}$$

where ω is the frequency in rad/s and $\mathbf{Z}_0, \mathbf{Z}_1 \in \mathcal{C}^{M\times M}$.

The research problem we now pose is to find the optimal matching network for the system of Fig. 4.2 over some bandwidth $B > 0$. From the following discussion, we anticipate that this problem may be considerably more difficult than its single-frequency counterpart. The pioneering studies of Bode [7] and Fano [28], conducted over 60 years ago, considered matching a purely resistive source to a frequency-dependent load. Subsequent studies have have expended considerable effort to generalize these results, first to a frequency-dependent source [15] and then to a bank of uncoupled, frequency-dependent sources [12]. After reviewing this literature and texts on broadband matching [11] and network synthesis [87], [95], we have not found any existing theory that appears to be immediately applicable to the coupled source impedances (5.1) with frequency-dependent real and imaginary parts.

Optimal Matching for Coupled Front Ends. In Sec. 4.3 we applied our noise factor matrix formulation to derive an optimal matching condition for a bank of uncoupled,



(a) (b)

Figure 5.2: (a) A $2M$-port noisy network and (b) its generalized Rothe-Dahlke equivalent network.

identical amplifiers. Also of interest is an optimal matching condition for coupled front-ends. One possible approach to this problem proceeds in a manner analogous to the Rothe-Dahlke [70] model for two-port noisy networks. Consider the $2M$-port network in Fig. 5.2(a), which was defined in the general front-end model of Sec. 4.1.2. By making the transformations

$$\mathbf{v_n} = \mathbf{n}_1 + (\mathbf{Z}_{\text{cor}} - \mathbf{Z}_{11})\mathbf{Z}_{21}^{-1}\mathbf{n}_2 \ , \quad \mathbf{i_n} = -\mathbf{Z}_{21}^{-1}\mathbf{n}_2 \ , \tag{5.2}$$

and $\mathbf{Z}_{\text{cor}} \triangleq \mathbf{Z}_{11} - \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{Z}_{21}$, where $\mathbf{v_n} \sim \mathcal{CN}(\mathbf{0}, 4kT_0B\mathbf{R_n})$ and $\mathbf{i_n} \sim \mathcal{CN}(\mathbf{0}, 4kT_0B\mathbf{G_n})$ are independent, we obtain the generalized version of the Rothe-Dahlke model illustrated in Fig. 5.2(b). By replacing the uncoupled front-end amplifiers in Fig. 4.2 with this generalized front-end, we see that the noise factor matrix (4.9) is

$$\mathbf{F} = \mathbf{I} + \mathbf{R}_{\mathbf{A}}'^{-1/2}\left[\mathbf{R_n} + (\mathbf{Z}_{\mathbf{A}}' + \mathbf{Z}_{\text{cor}})\mathbf{G_n}(\mathbf{Z}_{\mathbf{A}}' + \mathbf{Z}_{\text{cor}})^{\dagger}\right]\mathbf{R}_{\mathbf{A}}'^{-1/2} \ , \tag{5.3}$$

where $\mathbf{Z}_{\mathbf{A}}' = \mathbf{R}_{\mathbf{A}}' + j\mathbf{X}_{\mathbf{A}}'$ is the transformed antenna impedance matrix (3.31).

According to the low-noise design principles developed in Chapter 4, the optimal matching network produces the transformed impedance

$$\mathbf{Z}_{\text{opt}} = \arg \inf_{\mathbf{Z}'_{\mathbf{A}}} \mathbf{F} \ , \tag{5.4}$$

where the infimum is with respect to the positive definite partial ordering of Hermitian matrices (cf., [40, Sec. 7.7]). In the optimal matching derivation for uncoupled front-ends (4.16), finding an infimum was straightforward. In fact, if all the matrices in (5.3) were simultaneously diagonalizable, we could proceed in the same manner. However, for the more general case it is not clear how to proceed. Moreover, since the positive definite ordering is *partial*, an infimum is not even guaranteed to exist. Perhaps by using some alternative approach to the problem, i.e., not relying on noise factor matrices, one could find a solution; however, we have not explored this possibility. Motivation for considering this problem includes applications to receivers that are not well-modeled by uncoupled, identical front ends, e.g., noise correlation between receiver chains from a common local oscillator, or multi-band devices that use (non-identical) amplifiers tuned to different frequencies.

Additional Simulations and Experimental Work. The numerical examples of Chapters 3 and 4 considered an idealized model for thin-wire dipole antennas (cf., Appendix A). While this canonical model provides a good reference for initial investigations, we would also like to see how our results translate to more modern antennas, e.g., microstrip antennas [4, Ch. 14]. Alternate choices for other receiver components may also provide further research opportunities. For example, in Sec. 4.4.2 we observed that the location of noise sources within a front-end amplifier (i.e., the input and output ports) may have a

dramatic affect on performance. It would therefore be interesting to investigate different amplifier technologies, e.g., CMOS and HBT [66, Sec. 10.4], on the basis of how noise is correlated between input and output ports. We believe a detailed study of this type, combined with the insights developed in this dissertation, may have profound implications on the selection of front-end amplifiers. Finally, experimental work verifying the noise correlation predicted by the theoretical analysis of this dissertation, and its impact on the performance of real-world systems, would be a welcome addition to the literature.

Bibliography

[1] 4nec2, NEC based antenna modeler and optimizer. Available online: http://home.ict.nl/ arivoors/.

[2] S. M. Alamouti, "A simple transmitter diversity scheme for wireless communication," *IEEE J. Sel. Areas Commun.*, vol. 16, pp. 1451-1458, Oct. 1998.

[3] G. Alfano, A. M. Tulino, A. Lozano, and S. Verdu, "Capacity of MIMO channels with one-sided correlation," in *Proc. IEEE Int. Symp. Spread Spectrum Tech. Applications*, Aug. 2004.

[4] C. A. Balanis, *Antenna Theory: Analysis and Design*, 3rd ed. New York: Wiley, 2005.

[5] W. R. Bennett, *Electrical Noise*. New York: McGraw Hill, 1960.

[6] E. Biglieri, R. Caulderbank, A. Constantinides, A. Paulraj, and H. V. Poor, *MIMO Wireless Communications*. Cambridge: University Press, 2007.

[7] H. W. Bode, *Network Analysis and Feedback Amplifier Design*. Van Nostrand: New York, 1945.

[8] D. G. Brennan, "Linear diversity combining techniques," *Proc. IRE*, vol. 47, pp. 1075-1102, Jun. 1959.

[9] G. J. Burke and A. J. Poggio, "Numerical Electromagnetics Code (NEC) – method of moments," Tech. Doc. 11, Naval Ocean Systems Center, San Diego, CA, Jan. 1981.

[10] P. S. Carter, "Circuit relations in radiating systems and applications to antenna problems," *Proc. IRE*, vol. 20, pp. 1004-1041, Jun. 1932.

[11] W. Chen, *Broadband Matching: Theory and Implementations*. Teaneck, NJ: World Scientific, 1988.

[12] W. Chen, "Mathematical theory of broadband matching of multiport networks," *J. Franklin Institute*, vol. 326, pp. 737-747, May 1989.

[13] D. Chizhik, G. J. Foschini, M. J. Gans, and R. A. Valenzula, "Keyholes, correlations, and capacities of multielement transmit and receive antennas," *IEEE Trans. Wireless Commun.*, vol. 1, pp. 361-368, Apr. 2002.

[14] M. Chiani, M. Z. Win, and A. Zanella, "On the capacity of spatially correlated MIMO Rayleigh-fading channels," *IEEE Trans. Inform. Theory*, vol. 49, pp. 2363-2371, Oct. 2003.

[15] T. Chien, "A theory of broadband matching of a frequency-dependent generator and load – Part I: Theory," *J. Franklin Institute*, vol. 298, pp. 181-199, Sep. 1974.

[16] C. Chuah, D. N. C. Tse, J. M. Kahn, and R. A. Valenzula, "Capacity scaling in MIMO wireless systems under correlated fading," *IEEE Trans. Inform. Theory*, vol. 48, pp. 637-650, Mar. 2002.

[17] R. H. Clarke, "A statistical theory of mobile-radio reception," *Bell Syst. Tech. J.*, vol. 47, pg 957-1000, Jul. 1968.

[18] B. Clerckx, D. Vanhoenacker-Janvier, and C. Oestges, "Mutual coupling effects on the channel capacity and the space-time processing of MIMO communication systems," in *Proc. IEEE Inter. Conf. Commun.*, pp. 2638-2642, May 2003.

[19] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Hoboken, NJ: Wiley, 2006.

[20] W. C. Cox, J. A. Simpson, C. P. Domizioli, B. L. Hughes, and J. F. Muth, "An underwater optical communication system implementing Reed-Solomon channel coding," in *Proc. 2008 MTS/IEEE Oceans Conf.,* Quebec City, Canada.

[21] C. A. Desoer, "The maximum power transfer theorem for n-ports," *IEEE Trans. Circuit Theory*, vol. 20, pp. 328-330, May 1973.

[22] C. P. Domizioli, B. L. Hughes, K. G. Gard, and G. Lazzi, "Receive diversity revisited: correlation, coupling, and noise," in *Proc. 2007 IEEE Global Commun. Conf.,* Washington, D.C., pp. 3601-3606.

[23] C. P. Domizioli, B. L. Hughes, K. G. Gard, and G. Lazzi, "Optimal front-end design

for MIMO receivers," in *Proc. 2008 IEEE Global Commun. Conf.,* New Orleans, LA.

[24] C. P. Domizioli, B. L. Hughes, K. G. Gard, and G. Lazzi, "Noise correlation in compact diversity receivers," to appear in *IEEE Trans. Commun.*.

[25] C. P. Domizioli and B. L. Hughes, "Front-end design for compact MIMO receivers: a communication theory perspective," in preparation.

[26] Y. Dong, C. P. Domizioli, and B. L. Hughes, "Effects of mutual coupling and noise correlation on downlink coordinated beamforming with limited feedback," *EURASIP J. Advances Signal Process.*, Jul. 2009.

[27] R. S. Elliott, *Antenna Theory & Design*. Englewood-Cliffs, NJ: Prentice-Hall, 1981. Reprinted by Wiley – IEEE Press, New York, 2003.

[28] R. M. Fano, "Theoretical limitations on the broad-band matching of arbitrary impedances," *J. Franklin Institute*, vol. 249, pp. 57-83, Jan. 1950.

[29] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Commun.*, vol. 6, pp. 311-335, 1998.

[30] M. J. Gans, "Channel capacity between antenna arrays – Part I: Sky noise dominates," *IEEE Trans. Commun.*, vol. 54, pp. 1586-1592, Sep. 2006.

[31] M. J. Gans, "Channel capacity between antenna arrays – Part II: Amplifier noise dominates," *IEEE Trans. Commun.*, vol. 54, pp. 1983-1992, Nov. 2006.

[32] D. Gesbert, H. Bolcskei, D. A. Gore, and A. J. Paulraj, "Outdoor MIMO wireless channels: models and performance prediction," *IEEE Trans. Commun.*, vol. 50, pp. 1926-1934, Dec. 2002.

[33] A. Goldsmith, *Wireless Communications*. Cambridge: University Press, 2005.

[34] I. J. Gupta and A. A. Ksienski, "Effect of mutual coupling on the performance of adaptive arrays," *IEEE Trans. Antennas Propag.*, vol. AP-5, pp. 785-791, Sep. 1983.

[35] S. C. Hagness and A. Taflove, *Computational Electrodynamics: The Finite-Difference Time-Domain Method*, 2nd ed. Boston: Artec House, 2000.

[36] R. F. Harrington, *Time-Harmonic Electromagnetic Fields*. New York: McGraw-Hill, 1961. Reprinted by Wiley – IEEE Press, New York, 2001.

[37] R. F. Harrington, *Field Computation by Moment Methods*. New York: MacMillan, 1968. Reprinted by Wiley – IEEE Press, New York, 1993.

[38] H. A. Haus and R. B. Adler, *Circuit Theory of Linear Noisy Networks*. New York: Wiley, 1959.

[39] B. M. Hochwald, G. Caire, B. Hassabi, T. L. Marzetta, eds., *IEEE Trans. Inform. Theory*, vol. 49, Oct. 2003.

[40] R. A. Horn and C. R. Johnson, *Matrix Analysis*. New York: Cambridge, 1985.

[41] J. E. Hudson, *Adaptive Array Principles*. London: IEE Press, 1981.

[42] IEEE Std. 802.16e, *IEEE Standard for Local and Metropolitan Area Neworks: Part 16: Air Interface for Fixed Broadband Wireless Access Systems*, Dec. 2005.

[43] M. F. Iskander, M. Ando, R. G. Vaughan, and M. A. Jensen, eds., *IEEE Trans. Antennas Propag.*, vol. 54, Nov. 2006.

[44] W. C. Jakes, *Microwave Mobile Communications*. New York: Wiley, 1974.

[45] R. Janaswamy, "Effect of element mutual coupling on the capacity of fixed length linear arrays," *Antennas Wireless Propag. Lett.*, vol. 1, pp. 157-160, 2002.

[46] S. K. Jayaweera and H. V. Poor, "Capacity of multiple-antenna systems with both receiver and transmitter channel state information," *IEEE Trans. Inform. Theory*, vol. 49, pp. 2697-2709, Oct. 2003.

[47] R. W. P. King, "The linear antenna – eighty years of progress," *Proc. IEEE*, vol. 55, pp. 2-16, Jan. 1967.

[48] S. M. Krusevac, R. A. Kennedy, and P. B. Rapajic, "Effect of signal and noise mutual coupling on MIMO channel capacity," *Wireless Personal Commun.*, vol. 40, pp. 317-328, Feb. 2007.

[49] E. G. Larsson and P. Stoica, *Space-Time Block Coding for Wireless Communications*. Cambridge: University Press, 2003.

[50] B. K. Lau, J. B. Anderson, G. Kristensson, and A. F. Molish, "Impact of matching network on bandwidth of compact antenna arrays," *IEEE Trans. Antennas Propag.*, vol 54, no. 11, pp. 3225-3238, Nov. 2006.

[51] W. C. Y. Lee, "Mutual coupling effect on maximum-ratio diversity combiners and application to mobile radio," *IEEE Trans. Commun. Technol.*, vol. COM-18, pp. 779 - 791, Dec. 1970.

[52] M. C. Leifer, "Signal correlations in coupled cell MIMO antennas," in *Proc. IEEE Antennas Propag. Soc. Intern. Symp.*, pp. 194-197, Jun. 2002.

[53] S. Lin and D.J. Costello, *Error Control Coding*, 2nd ed. Upper Saddle River, NJ: Prentice Hall, 2004.

[54] T. K. Y. Lo, "Maximum ratio transmission," *IEEE Trans. Commun.*, vol. 47, pp. 1458-1461, Oct. 1999.

[55] S. L Loyka, "Channel capacity of MIMO architecture using the exponential correlation matrix," *IEEE Commun. Lett.*, vol. 5, pp. 396-371, Sep. 2001.

[56] A. Lozano, A. M. Tulino, and S. Verdu,"Multiple-antenna capacity in the low-power regime," *IEEE Trans. Inform. Theory*, vol. 49, pp. 2527-2544, Oct. 2003.

[57] Maxim 2642/2643 LNAs. Data sheet available online: http://datasheets.maxim-ic.com/en/ds/MAX2642-MAX2643.pdf

[58] M. L. Morris and M. A. Jensen, "Network model for MIMO systems with coupled antennas and noisy amplifiers," *IEEE Trans. Antennas Propag.*, vol. 53 pp. 545-552, Jan. 2005.

[59] M. L. Morris and M. A. Jensen, "Improved network analysis of coupled antenna

diversity performance," *IEEE. Trans. Wireless Commun.*, vol. 4, pp. 1928-1934, Jul. 2005.

[60] M. L. Morris, M. A. Jensen, and J. W. Wallace, "Superdirectivity in MIMO systems," *IEEE Trans. Antennas Propag.*, vol. 53, pp. 2850-2857, Sep. 2005.

[61] H. Nyquist, "Thermal agitation of electric charge in conductors," *Physical Review*, vol. 32, pp. 110-113, Jul. 1928.

[62] C. Oestges and B. Clerckx, *MIMO Wireless Communications: From Real-World Propagation to Space-Time Code Design*. Oxford: Academic Press, 2007.

[63] L. G. Ordonez, D. P. Palomar, and J. R. Fonollosa, "Ordered eigenvalues of a general class of Hermitian random matrices and performance analysis of MIMO systems," in *Proc. IEEE Intern. Conf. Commun. 2008*, pp. 3846-3852.

[64] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*. Cambridge: University Press, 2003.

[65] N. Pierce and S. Stein, "Multiple diversity with non-independent fading," *Proc. IRE*, vol. 48, pp. 89-104, Jan. 1960.

[66] D. M. Pozar, *Microwave Engineering*, 3rd ed. New York: Wiley, 2005.

[67] J. G. Proakis, *Digital Communications*, 4th ed. New York: McGraw-Hill, 2001.

[68] Python programming language, available online: www.python.org.

[69] T. S. Rappaport, *Wireless Communications: Principles and Practice*, 2nd ed. Saddle River, NJ: Prentice-Hall, 2001.

[70] H. Rothe and W. Dahlke, "Theory of noisy fourpoles," *Proc. IRE*, vol. 44, pp. 811-818, Jun. 1956.

[71] M. Shafi, D. Gesbert, D. Shiu, and P. J. Smith, ed., *IEEE J. Sel. Areas Commun.*, vol. 21, Apr. 2003.

[72] M. Shafi, D. Gesbert, D. Shiu, and P. J. Smith, ed., *IEEE J. Sel. Areas Commun.*, vol. 21, Jun. 2003.

[73] A. M. Sayeed, "Deconstructing multiantenna fading channels," *IEEE Trans. Signal Process.*, vol. 50, pp. 2563-2579, Oct. 2002.

[74] C. E. Shannon, "A mathematical theory of communication," *Bell Sys. Tech. J.*, vol. 27, pp. 379-423, 1948.

[75] D. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Trans. Commun.*, vol. 48, pp. 502-513, Mar. 2000.

[76] M. K. Simon and M. Alouini, *Digital Commnications over Fading Channels*, 2nd ed. Hoboken, NJ: Wiley, 2005.

[77] S. Stein, "On cross coupling in multiple-beam antennas," *IRE Trans. Antennas Propag.*, vol. 10, pp. 548-557, Sep. 1962.

[78] W. L. Stutzman and G. A. Thiele, *Antenna Theory and Design*, 2nd ed. New York: Wiley, 1998.

[79] T. S. Svantesson and A. Ranheim, "Mutual coupling effects on the capacity of multielement antenna systems," in *Proc. IEEE ICASSP*, 2001, pp. 2485-2488.

[80] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, pp. 744-764, Mar. 1998.

[81] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1456-1467, Jul. 1999.

[82] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. Telecommun.*, vol. 10, pp. 585-595, Nov. 1999.

[83] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge: University Press, 2005.

[84] A. M. Tulino and S. Verdu, "Random matrix theory and wireless communications," *Foundations and Trends in Communication and Information Theory*, vol. 1, no. 1, 2004.

[85] R. Q. Twiss, "Nyquist's and Thevenin's theorems generalized for nonreciprocal linear networks," *J. Applied Physics*, vol. 26, pp. 599-602, May 1955.

[86] H. L. Van Trees, *Detection, Estimation, Modulation Theory, Part I: Detection, Estimation, and Linear Modulation Theory*. New York: Wiley, 1968.

[87] M. E. Van Valkenburg, *Introduction to Modern Network Synthesis*. New York: Wiley, 1960.

[88] M. K. Varanasi and T. Guess, "Optimum decision feedback multiuser equalization and successive decoding acheives the total capacity of the Gaussian multiple-access channel," in *Proc. 1997 Asilomar Conf. Signals, Sys. Computers*.

[89] R. G. Vaughan and J. B. Anderson, "Antenna diversity in mobile communications," *IEEE Trans. Vehic. Technol.*, vol. 36, pp. 149-172, Nov. 1987.

[90] S. Verdu, *Multiuser Detection*. Cambridge: University Press, 1998.

[91] J. W. Wallace and M. A. Jensen, "Mutual coupling in MIMO wireless systems: a rigorous network theory analysis," *IEEE Trans. Wireless Commun.*, vol. 3, pp. 1317-1325, Jul. 2004.

[92] C. Waldschmidt, J. v. Hagen, and W. Wiesbeck,"Influence and modeling of mutual coupling in MIMO and diversity systems," in *Proc. IEEE Antennas Propag. Soc. Intern. Symp.*, pp. 190-193, Jun. 2002.

[93] C. Waldschmidt, S. Schulteis, and W. Wiesbeck, "Complete RF system model for analysis of compact MIMO arrays," *IEEE Trans. Vehic. Technol.*, vol. 53, pp. 579-586, May 2004.

[94] K. F. Warnick and M. A. Jensen, "Optimal noise matching for mutually coupled arrays," *IEEE Trans. Antennas Propag.*, vol. 55, pp. 1726-1731, Jun. 2007.

[95] L. Weinburg, *Network Analysis and Synthesis*. New York: R. E. Krieger, 1975.

[96] P. W. Wolniansky, "V-BLAST: an architecture for realizing very high data rates over the rich-scattering wireless channel," in *Proc. URSI Intern. Symp. on Signals, Systems, Electronics Conf.*, pp. 295-300, New York: 1998.

[97] A. Zanella, M. Chiani, and M. Z. Win, "On the marginal distribution of the eigenvalues of Wishart matrices," *IEEE Trans. Commun.*, vol. 57, pp. 1050-1060, Apr. 2009.

[98] Y. Zhang, K. Hirasawa, and K. Fujimoto, "Opened parasitic elements nearby a driven dipole," *IEEE Trans. Antennas Propag.*, vol. 34, pp. 711-713, May 1986.

Appendices

Appendix A

Computation of Antenna Impedances and Patterns

In this Appendix we provide an overview of the numerical technique used to compute the antenna impedances and patterns for the numerical examples in Chapters 3 and 4. The Numerical Electromagnetics Code (NEC) [9], a well-known Method-of-Moments (MoM) based program, is employed to compute the impedances and patterns of a uniform linear array (ULA) of half-wavelength dipole antennas. As a benchmark we also consider the induced EMF method, which is used to derive approximate formulas for the impedances and patterns. Detailed expositions on both the MoM and induced EMF method are available in various antenna texts, e.g., [4], [27], and [78], so we shall limit our coverage to discussing the similarities and differences between these two approaches. We begin with a description of the physical model and some simplifying assumptions frequently used in the analysis of thin-wire dipoles.

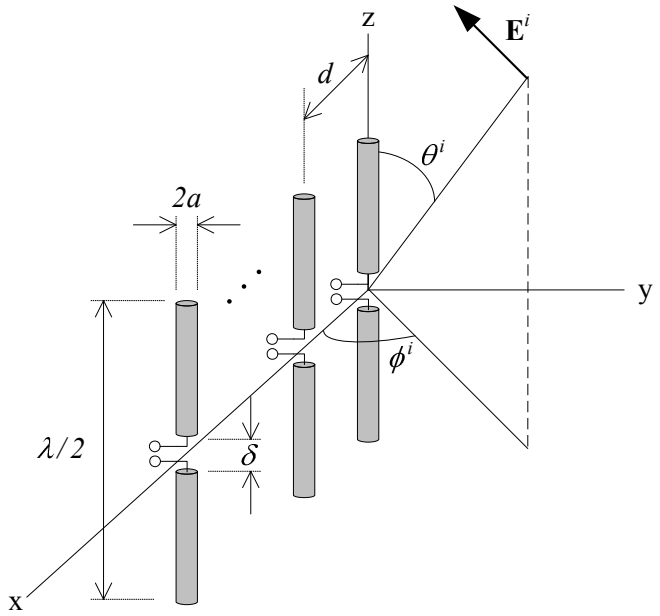


Figure A.1: Dipole array model.

A.1 Model and Solution Techniques

Consider the model for an M-element ULA of half-wavelength dipoles illustrated in Fig. A.1. Each dipole consists of two perfect electrically conducting (PEC) cylinders of radius a and length $\frac{\lambda}{4} - \frac{\delta}{2}$, where λ is the wavelength, separated by a *feed gap* of width δ. Typically both $a \ll \lambda$ and $\delta \ll \lambda$, in which case we have a *thin-wire dipole* (cf., [4, Ch. 4]). Wires connected to the feed gap form the antenna terminals, across which voltages $v_1, \ldots, v_M$ and currents $i_1, \ldots, i_M$ are defined. The dipoles are uniformly spaced a distance d apart along the x-axis in infinite free space, so the effect of the feed wires (or any other surrounding obstacle) is neglected. All sources and fields are sinusoidally-varying in time, in which case we are within the scope of standard time-harmonic field

analysis [36]. The complex baseband current density in the n^{th} dipole and the electric field it generates will be denoted by $\mathbf{J}_n$ and $\mathbf{E}_n$, respectively. Incident on the array is a complex baseband, uniform plane wave[1] (cf., [36, Sec. 2-11])

$$\mathbf{E}^i = \left(\hat{\mathbf{x}}E_x^i + \hat{\mathbf{y}}E_y^i + \hat{\mathbf{z}}E_z^i\right) e^{jk\left(x\sin\theta^i\cos\phi^i + y\sin\theta^i\sin\phi^i + z\cos\theta^i\right)} , \tag{A.1}$$

where $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, and $\hat{\mathbf{z}}$ are the standard orthonormal basis vectors of the rectangular coordinate system, the field components E_x^i, E_y^i, and E_z^i are constant, $k = \frac{2\pi}{\lambda}$ is the wave number, and θ^i and ϕ^i are the *zenith* and *azimuth* of the wave angle-of-arrival (AoA), respectively.

Let us consider the relationship between the terminal voltages and currents when $\mathbf{E}^i = \mathbf{0}$. The medium connected to the antenna terminals (i.e., the PECs and free space) is linear and isotropic, so the voltages and currents are related by an $M \times M$ *impedance matrix* $\mathbf{Z_A}$ (cf., [36, Sec. 3-8])

$$\mathbf{v} = \mathbf{Z_A}\mathbf{i} , \tag{A.2}$$

where $\mathbf{v} \triangleq [v_1 \cdots v_M]^T$ and $\mathbf{i} \triangleq [i_1 \cdots i_M]^T$. By definition, $[\mathbf{Z_A}]_{nm}$ is the ratio $\frac{v_n}{i^s}$ when the m^{th} dipole is driven by a current source $i^s \neq 0$ and every other dipole is open-circuited (i.e., $i_n = 0$ for all $n \neq m$). Now if $\mathbf{E}^i \neq \mathbf{0}$, by superposition (since the medium is linear)

$$\mathbf{v} = \mathbf{Z_A}\mathbf{i} + \mathbf{v_o} , \tag{A.3}$$

where $\mathbf{v_o} = [v_{o1} \cdots v_{oM}]^T$ is the voltage induced in the open-circuited array terminals by $\mathbf{E}^i$. When plotted as a function of AoA, the induced voltages $v_{on}(\theta, \phi)$ are often referred to as *receive patterns*. It is conventional to work with normalized receive patterns $g_n(\theta, \phi)$,

[1] For clarity, throughout the Appendix we consider a single, deterministic incident wave. The extension to the multiple, random waves used in the fading models of Chapters 3 and 4 is straightforward.

defined as the voltage induced in the n^{th} antenna by $\hat{\theta}$-polarized, unit-amplitude, zero-phase plane wave with AoA (θ, ϕ). Thus,

$$v_{on} = E^i_\theta e^{jkd(n-1)\sin\theta^i\cos\phi^i} \cdot g_n(\theta^i, \phi^i) \ , \tag{A.4}$$

where $E^i_\theta = E^i_x \cos\theta^i \cos\phi^i + E^i_y \cos\theta^i \sin\phi^i - E^i_z \sin\theta_i$ is the $\hat{\theta}$-component of the incident field, and the exponential term is simply the phase shift of the incident field between the dipole coordinates and the origin. In writing (A.4) we have made use of the fact that the thin-wire dipoles in Fig. A.1 only detect the $\hat{\theta}$-component of the incident field; for other antenna and array configurations we may also have to define patterns for the $\hat{\phi}$-component.

In summary, the terminal characteristics of the array are determined by two quantities: the impedance matrix and receive patterns. In principle, one could solve for the impedances and patterns by making the necessary terminal excitations, then solving Maxwell's equations subject to the boundary conditions[2] of zero tangential electric field over the surfaces of the PECs and a sphere of infinite radius. For example, to find $[\mathbf{Z_A}]_{12}$ we could drive dipole 2 with a 1 A current source while leaving all the other dipoles open-circuited and setting $\mathbf{E}^i = \mathbf{0}$, then solve for the voltage induced across the terminal of dipole 1. Unfortunately, closed-form solutions to Maxwell's equations are feasible only for problems in which the boundary conditions are coincident with constant surfaces of certain coordinate systems (e.g., [36, Ch. 4-6]). Since the boundary conditions specified above do not meet this requirement, we resort to approximate solution techniques.

Below we present two well-known techniques for thin-wire dipoles that invoke the

[2]By the uniqueness theorem (cf., [36, Sec. 3-3]), there is a unique solution to this problem.

following assumptions (cf., [4, Ch. 8], [78, Ch. 10]):

- The current density $\mathbf{J}_n$ in each dipole resides on a cylindrical shell of radius a about each dipole axis, and is axially-directed and rotationally-symmetric:

$$\mathbf{J}_n(x,y,z) = \begin{cases} \hat{\mathbf{z}}\frac{1}{2\pi a} I_n(z) \ , & [x-(n-1)d]^2 + y^2 = a^2 \text{ and } -\frac{\lambda}{4} \le z \le \frac{\lambda}{4} \\ \mathbf{0} \ , & \text{otherwise} \end{cases} \ , \quad \text{(A.5)}$$

  where $I_n(z)$ is the total current.

- The z-component of the electric field E_{nz} produced by $\mathbf{J}_n$ is equal to the field produced by a line current source with distribution equal to $I_n(z)$, positioned along the axis of the dipole in free space:

$$E_{nz}(x,y,z) = \frac{1}{j4\pi\omega\epsilon_0} \int_{-\lambda/4}^{\lambda/4} I_n(z') \left[k^2 + \frac{\partial^2}{\partial z^2} \right] \frac{e^{-jkr}}{r} dz' \ , \quad \text{(A.6)}$$

  where ω is the frequency in rad/s, $\epsilon_0 = 8.85 \times 10^{-12}$ F/m is the permittivity of free space, and

$$r = \sqrt{[x-(n-1)d]^2 + y^2 + (z-z')^2} \quad \text{(A.7)}$$

  is the distance between the source (x', y', z') and field (x, y, z) coordinates.

Since the validity of these approximations improves as $a \to 0$, they are often collectively referred to as the *thin-wire approximation*. In the first solution technique we shall consider, the *method of moments* (MoM), the current distributions $I_n(z)$ are evaluated numerically. The second technique, the *induced EMF method*, assumes a current distribution and derives formulas for the impedances and patterns by applying the reciprocity theorem. Some other possible approaches are briefly mentioned at the end of the section. The NEC, which we shall use for computing the impedances and patterns, is based on the MoM.

A.1.1 Method of Moments

The MoM for solving electromagnetic problems was introduced by Harrington in [37]. Among its many applications, the MoM is particularly suitable for computing fields and currents in the presence of thin wires. In this section we provide a brief overview of the MoM as it relates to our problem. The current distributions $\{I_1, \ldots, I_M\}$ are computed in the following manner:

1. Specify a source field $\mathbf{E}^s$. For impedance calculations $\mathbf{E}^s$ is produced by a voltage impressed across one of the dipole terminals; for pattern calculations $\mathbf{E}^s = \mathbf{E}^i$.

2. Enforce the PEC boundary condition $\sum_n E_{nz} = -E_z^s$ along the M lines specified by the coordinates $x \in \{a, a+d, \ldots, a+(M-1)d\}$, $y = 0$, and $-\frac{\lambda}{4} \leq z \leq \frac{\lambda}{4}$. This results in M integral equations in the unknown current distributions.

3. Approximate the integral equations with an appropriate linear equation. Solve the linear equations for the approximate current distributions.

A brief overview these steps is provided below; more detailed expositions may be found, for example, in [4, Ch. 8], [27, Ch. 7], and [78, Ch. 10]. Note that the MoM views the terminal voltages as known and computes terminal currents $i_n = I_n(0)$, which is opposite to the viewpoint of the Thevenin network model (A.3). It is therefore convenient to instead work with the Norton equivalent network

$$\mathbf{i} = \mathbf{Y_A}\mathbf{v} + \mathbf{i_s} \ , \tag{A.8}$$

where $\mathbf{Y_A}$ is the antenna admittance matrix and $\mathbf{i_s}$ is the short-circuit current induced by $\mathbf{E}^i$. The desired impedances and patterns can be computed from the circuit relations

$\mathbf{Z_A} = \mathbf{Y_A^{-1}}$ and $\mathbf{v_o} = -\mathbf{Z_A i_s}$.

Let us begin by calculating the first column of the admittance matrix, $[\mathbf{Y_A}]_{n1}$. We apply a voltage source v^s to dipole 1 and short-circuit the other dipoles, so $v_n = 0$ for all $n \neq 1$. The field of the voltage source is assumed to follow the *delta-gap model* (cf., [4, pg. 447]) in which $\mathbf{E}^s = \hat{\mathbf{z}}\frac{v^s}{\delta}$ within the gap of dipole 1 and $\mathbf{E}^s = \mathbf{0}$ elsewhere (i.e., there is no fringing). By enforcing the boundary conditions outlined in Step 2, and recalling the thin-wire approximation of the currents and fields, we obtain the following integral equations:

$$\int_{-\lambda/4}^{\lambda/4} \left[\sum_{n=1}^{M} G_{n1}(z,z')I_n(z')\right] dz' = \begin{cases} -j4\pi\omega\epsilon_0\frac{v^s}{\delta} \ , & -\frac{\delta}{2} \leq z \leq \frac{\delta}{2} \\ 0 \ , & \frac{\delta}{2} < |z| \leq \frac{\lambda}{4} \end{cases}$$

$$\int_{-\lambda/4}^{\lambda/4} \left[\sum_{n=1}^{M} G_{nm}(z,z')I_n(z')\right] dz' = 0 \ , \quad -\frac{\lambda}{4} \leq z \leq \frac{\lambda}{4} \ , \quad m = 2, \ldots, M \ ,$$

where

$$G_{nm}(z,z') \triangleq \begin{cases} \left[k^2 + \frac{\partial^2}{\partial z^2}\right] \frac{e^{-jk\sqrt{a^2+(z-z')^2}}}{\sqrt{a^2+(z-z')^2}} \ , & n = m = 1, 2, \ldots, M \\ \left[k^2 + \frac{\partial^2}{\partial z^2}\right] \frac{e^{-jk\sqrt{(n-m)^2d^2+(z-z')^2}}}{\sqrt{(n-m)^2d^2+(z-z')^2}} \ , & n \neq m \end{cases} . \tag{A.9}$$

In writing G_{nm} for $n \neq m$ we have assumed $d \gg a$ so that $d - a \approx d$, which is usually a valid approximation in practical systems.

The above equations may be expressed in a compact form by introducing the inner product

$$\langle f, g \rangle \triangleq \int_{-\lambda/4}^{\lambda/4} f(z)g(z)dz \ , \tag{A.10}$$

defining the M functions

$$\begin{aligned} e_1(z) &= \begin{cases} -j4\pi\omega\epsilon_0 \frac{v^s}{\delta} , & -\frac{\delta}{2} \le z \le \frac{\delta}{2} \\ 0 , & \frac{\delta}{2} < |z| \le \frac{\lambda}{4} \end{cases} \\ e_m(z) &= 0 , \quad m = 2, \ldots, M , \end{aligned}$$

and finally, the M^2 linear operators

$$L_{nm}(f) = \int_{-\lambda/4}^{\lambda/4} G_{nm}(z, z') f(z') dz' . \tag{A.11}$$

With these definitions the above integral equations take the following form:

$$e_m = \sum_{n=1}^{M} L_{nm}(I_n) , \quad m = 1, \ldots, M . \tag{A.12}$$

The MoM provides a numerical solution to these equations by considering solutions of the form

$$I_n(z) = \sum_{i=1}^{L} \alpha_{ni} f_i(z) , \tag{A.13}$$

where $\{f_1, \ldots, f_L\}$ are *basis functions* and the α_{ni} terms are unknown constants. For some *weighting functions* $\{w_1, \ldots, w_L\}$ defined on the space spanned by the functions $L_{nm}(f_i)$ and e_i, substituting (A.13) into (A.12) and taking inner products yields

$$\langle w_j, e_m \rangle = \sum_{i=1}^{L} \sum_{n=1}^{M} \alpha_{ni} \langle w_j, L_{nm}(f_i) \rangle , \qquad \begin{aligned} j &= 1, \ldots, L \\ m &= 1, \ldots, M \end{aligned} . \tag{A.14}$$

The inner products on both sides may be evaluated (either analytically or numerically), yielding LM linear equations in LM unknowns α_{ni}, which may be solved by standard techniques such as Gaussian elimination. Finally, we obtain the approximate current densities (A.13), from which we compute the admittances $[\mathbf{Y_A}]_{1n} = i_n / v^s = I_n(0)/v^s$.

To compute the short-circuit currents $\mathbf{i_s}$ we repeat the above steps, but with $\mathbf{E}^s = \mathbf{E}^i$. Clearly the accuracy of the MoM approach is heavily dependent on the choice of basis and weighting functions, a subject which has received much attention in the literature (cf., [4, Ch. 8], [27, Ch. 7], and [78, Ch. 10]).

A.1.2 Induced EMF

The induced EMF method assumes that the currents are sinusoidal,

$$I_n(z) = i_n \cos kz \ , \tag{A.15}$$

and derives the impedances from the following specialization of the reciprocity theorem (cf., [36, Sec. 3-8]):

$$[\mathbf{Z_A}]_{nm} = \frac{-1}{i_n i_m} \int \mathbf{E}_m \cdot \mathbf{J}_n d\tau \ , \tag{A.16}$$

where the volume integral is over all space. By applying the thin-wire approximation, we can reduce the integral to

$$[\mathbf{Z_A}]_{nm} = \frac{-1}{j4\pi\omega\epsilon_0} \int_{-\lambda/4}^{\lambda/4} \int_{-\lambda/4}^{\lambda/4} \cos kz \cos kz' G_{nm}(z,z') dz' dz \ , \tag{A.17}$$

where for $n \neq m$ we have again assumed $d \gg a$ so that the electric field E_{nz} is approximately independent of x and y over each circle $\{(x,y) : [x-(m-1)d]^2 + y^2 = a^2\}$. The above integral has been evaluated in closed form [10]; for the self impedances we have

$$\begin{aligned} [\mathbf{R_A}]_{nn} &= \frac{\eta_0}{4\pi} \left[2\mathrm{Ci}\,(ka) - \mathrm{Ci}\left(\sqrt{(ka)^2+\pi^2}+\pi\right) - \mathrm{Ci}\left(\sqrt{(ka)^2+\pi^2}-\pi\right)\right] \\ [\mathbf{X_A}]_{nn} &= \frac{-\eta_0}{4\pi} \left[2\mathrm{Si}\,(ka) - \mathrm{Si}\left(\sqrt{(ka)^2+\pi^2}+\pi\right) - \mathrm{Si}\left(\sqrt{(ka)^2+\pi^2}-\pi\right)\right] \ , \end{aligned} \tag{A.18}$$

where $\mathbf{Z_A} = \mathbf{R_A} + j\mathbf{X_A}$, $\eta_0 \approx 120\pi\ \Omega$ is the intrinsic impedance of free space, and

$$\mathrm{Si}(x) \triangleq \int_0^x \frac{\sin x}{x} dx \quad \text{and} \quad \mathrm{Ci}(x) \triangleq -\int_x^\infty \frac{\cos x}{x} dx \tag{A.19}$$

are the sine and cosine integrals (cf., [4, App. III]). As one example, for $a = 10^{-3}\lambda$ we obtain $[\mathbf{Z_A}]_{nn} = 73.1 + j42.5\ \Omega$. The mutual impedances $[\mathbf{Z_A}]_{nm}$, $n \neq m$, are given by (A.18) with a replaced by $(n-m)d$.

The receive patterns may also be derived using reciprocity by considering an element of current $\mathbf{i}^s = \hat{\theta} E_\theta^i i_n$ located far from the origin, from which one can show that [36, pg. 140]

$$\begin{aligned} v_{on} &= \frac{k}{2} \int_{-\lambda/4}^{\lambda/4} E_\theta^i e^{jk\left((n-1)d \sin\theta^i \cos\phi^i + z' \cos\theta^i\right)} \sin\theta^i I(z')\ dz' \\ &= E_\theta^i e^{jk(n-1)d \sin\theta^i \cos\phi^i} \frac{\cos\left(\frac{\pi}{2}\cos\theta^i\right)}{\sin\theta^i}\ , \end{aligned} \tag{A.20}$$

where we used (A.15) and evaluated the integral. From (A.4) we conclude that

$$g_n(\theta, \phi) = \frac{\cos\left(\frac{\pi}{2}\cos\theta\right)}{\sin\theta}\ . \tag{A.21}$$

A.1.3 Other Techniques

Dipole antennas have been extensively studied for well over a century, and a vast collection of other solution techniques have been developed. For example, before the advent of digital computers made numerical electromagnetics (e.g., the MoM) feasible, several authors proposed iterative solutions to the current integral equations in which successive approximations to the actual current distribution were derived (these techniques are well-chronicled in [27, Ch. 7] and [47]). More recently, the finite-difference

time-domain (FDTD) method [35], which discretizes Maxwell's equations in time and space, has gained widespread use in the antenna community, and is implemented in several commercial software packages. While the FDTD does not have to invoke the thin-wire approximation, other issues arise (e.g., absorbing boundary conditions) that are not present in the MoM approach. There are many other numerical techniques we could use to calculate the patterns and impedances; however, we shall proceed with the NEC as it is a widely-accepted approach to solving thin-wire problems [4, Ch. 8].

A.2 NEC Computation of Array Parameters

The Numerical Electromagnetics Code (NEC) [9] was developed at Lawrence Livermore National Laboratory as a user-oriented program for computing the fields and currents generated by sources operating in the presence of metallic structures. It is based on a MoM solution to generalized versions of the integral equations considered in the previous section: the electric field integral equation (EFIE) for wires and the magnetic field integral equation (MFIE) for surfaces. A detailed description of the code is available in [9]; here we provide a brief discussion on its connection to the prior section and application to our problem.

Recall that the primary consideration in developing a MoM solution (A.14) was the selection of the basis and weighting functions. The basis functions in NEC are

$$f_i(z) = \begin{cases} a_i + b_i \sin k(z - z_i) + c_i \cos k(z - z_i) \ , & |z - z_i| \leq \Delta \\ 0 \ , & \text{otherwise} \end{cases} \ , \tag{A.22}$$

where, for our problem, the points $\{z_1, \ldots, z_L\}$ are equally spaced over $[-\frac{\lambda}{4}, \frac{\lambda}{4}]$ in *segments*

of length Δ. For each i, two of the constants $\{a_i, b_i, c_i\}$ are computed by enforcing boundary conditions on the current and its derivative between segments, leaving one unknown constant for each segment. The weighting functions are

$$w_i(z) = \delta(z - p_i) \ , \tag{A.23}$$

where the *matching points* $\{p_1, \ldots, p_L\}$ are equally spaced over $[-\frac{\lambda}{4}, \frac{\lambda}{4}]$ along the axis of each segment.

The NEC accepts a text file – the *input card*[3] – which contains simulation information. After computing the MoM solution, it returns the requested computations in another text file – the *output card*. A variety of graphical-user interface (GUI) programs have been developed for the NEC, e.g., 4nec2 [1], so one does not have to use the cards directly. However, since we will require repeated executions of the NEC over various array configurations, the GUI approach is impractical. In order to automate the process, we created a Python [68] script that performs the following operations:

1. Create the NEC input card
2. Execute the NEC
3. Extract the desired simulation results from the NEC output card

Using this approach we were able to repeatedly run simulations for a variety of different array sizes M and inter-element spacings d. We set $a = 10^{-3}\lambda$ and $\lambda = 1$ m, and used 25 segments per dipole, so $\Delta = 0.02\lambda$. These values are consistent with the NEC

[3]This name dates back to the punched-card computers used during initial versions of the NEC.

```
CE
GW 1 25 0.0  0 -0.25 0.0  0 0.25 0.001
GW 2 25 0.25 0 -0.25 0.25 0 0.25 0.001
GE
FR 0 0 0 0 299.8 0
PT 2
GN -1
EX  0 1 13  0  1  0
XQ
EX  0 2 13  0  1  0
XQ
EN
```

Figure A.2: NEC input card for admittance calcuations.

recommendations [9] of selecting $10^{-3}\lambda < \Delta < 10^{-1}\lambda$, $2\pi a \ll \lambda$, and, for less than 1 % error in the computed fields, $\Delta > 8a$.

An NEC input card that performs an admittance measurement for $M = 2$ and $d = 0.25\lambda$ is shown in Fig. A.2. A detailed description of the various NEC commands is provided in [9]. Briefly, the `GW` lines specify the coordinates of each dipole, the `EX` commands drive the center segment of each antenna with a 1 V voltage source, and the `XQ` commands run the simulation, generating the output cards. In the first output card we find the currents induced in the center segment of each dipole by a 1 V source applied at the first antenna – these are equal to the first row of $\mathbf{Y_A}$. Similarly, the second row of $\mathbf{Y_A}$ is extracted from the second output card. To compute the short-circuit currents $\mathbf{i_s}$, uniform plane waves were propagated towards the array by driving $\hat{\theta}$-oriented "source" dipoles located $10,000\lambda$ from the origin with 1 V sources. The source dipoles were uniformly spaced over 11.25° increments of azimuth and zenith.

The NEC impedances computed for a 2-element array with $0.01\lambda \leq d \leq \lambda$ are shown in Fig. A.3. Also included are NEC results for $a = 10^{-5}\lambda$ and the induced EMF

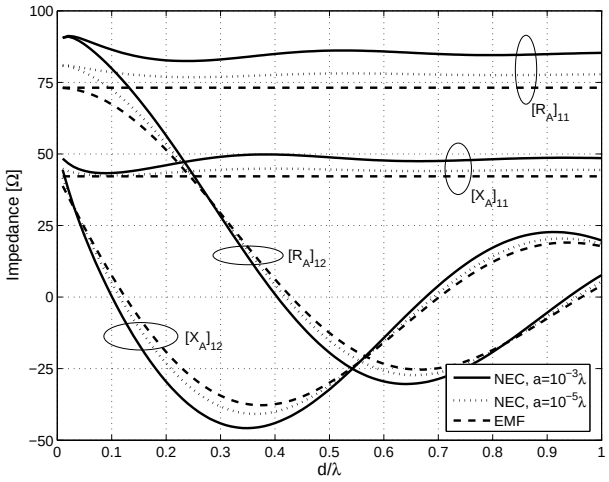


Figure A.3: Impedances computed with NEC and EMF.

impedances (A.18) (the difference between the EMF curves for $a = 10^{-3}\lambda$ and $a = 10^{-5}\lambda$ is negligible). All three curves are in reasonably good agreement, with the $a = 10^{-5}$ NEC curve closer to the EMF curve than the $a = 10^{-3}$ curve. This is expected since the sinusoidal current distribution (A.15) assumed by the induced EMF method becomes a better approximation as $a \to 0$ (cf., [4, Ch. 8]). Note that the self-impedance $[\mathbf{Z_A}]_{11}$ varies with d when computed with NEC, but is independent of d under the induced EMF model. From (A.15), we see that for the induced EMF method, when a dipole is open-circuited its current density is zero, and therefore its effect on neighboring dipoles is neglected. On the other hand, NEC computes the current in open-circuited dipoles, so the presence of neighboring dipoles does affect the self-impedance (again, more for $a = 10^{-3}\lambda$ than for $a = 10^{-5}\lambda$). Results similar to Fig. A.3 were obtained in [4, Ch. 8].

To compute the NEC receive patterns we first found the voltage v_{iso} induced in an isolated ($M = 1$) dipole by a plane wave with AoA $(\frac{\pi}{2}, \frac{\pi}{2})$, then normalized the induced

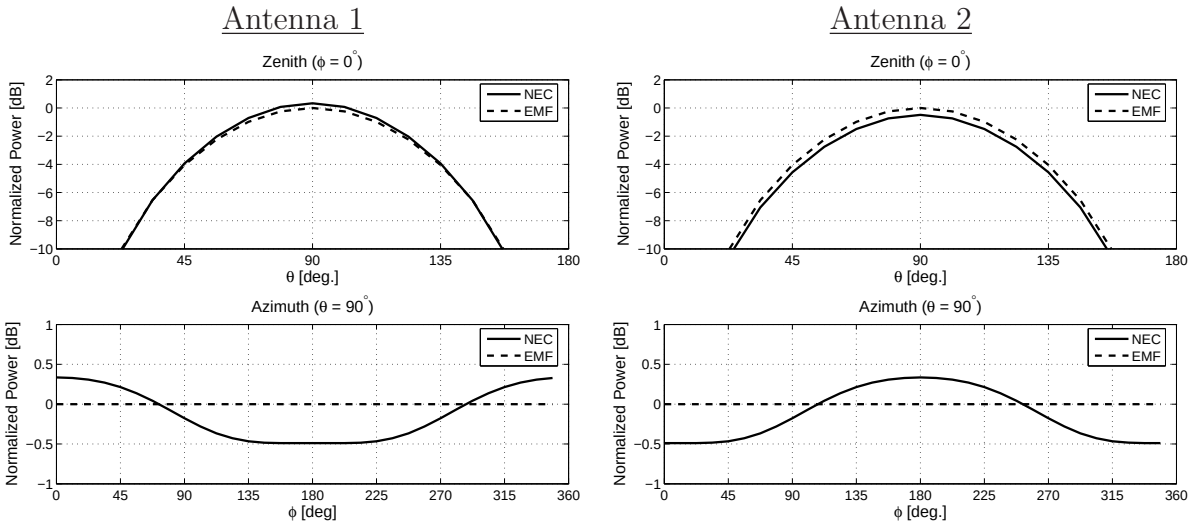


Figure A.4: Power patterns computed with NEC and EMF.

voltages $\mathbf{v_o}$ by v_{iso}. This approach removed the dependance of $\mathbf{v_o}$ on the unknown E_θ^i term in (A.4). The resulting power $|v_{on}/v_{\text{iso}}|^2$ and phase $\angle v_{on} - \angle v_{\text{iso}}$ patterns for $d = 0.2\lambda$ (results for other inter-element spacings are similar) are shown for $a = 10^{-3}\lambda$ in Figs. A.4 and A.5, respectively, with zenith patterns ($0 \leq \theta \leq \pi$, $\phi = 0$) in the upper plots and azimuth patterns ($0 \leq \phi \leq 2\pi$, $\theta = \frac{\pi}{2}$) in the lower plots. Again we include the EMF patterns (A.20) for comparison. In contrast to the impedance calculations, there is very little difference between the NEC and EMF patterns. The NEC power patterns deviate at most by 0.5 dB from the EMF pattern, and the phase patterns are indistinguishable. It is well known (cf., [36, pg. 81]) that far-field quantities such as patterns are relatively insensitive to minor changes in the current distribution, so we should not expect much difference in the two approaches (this result is also confirmed using NEC in [98]).

Finally, we compute the open-circuit correlation matrix $\mathbf{\Sigma_h}$ for the 2D scattering

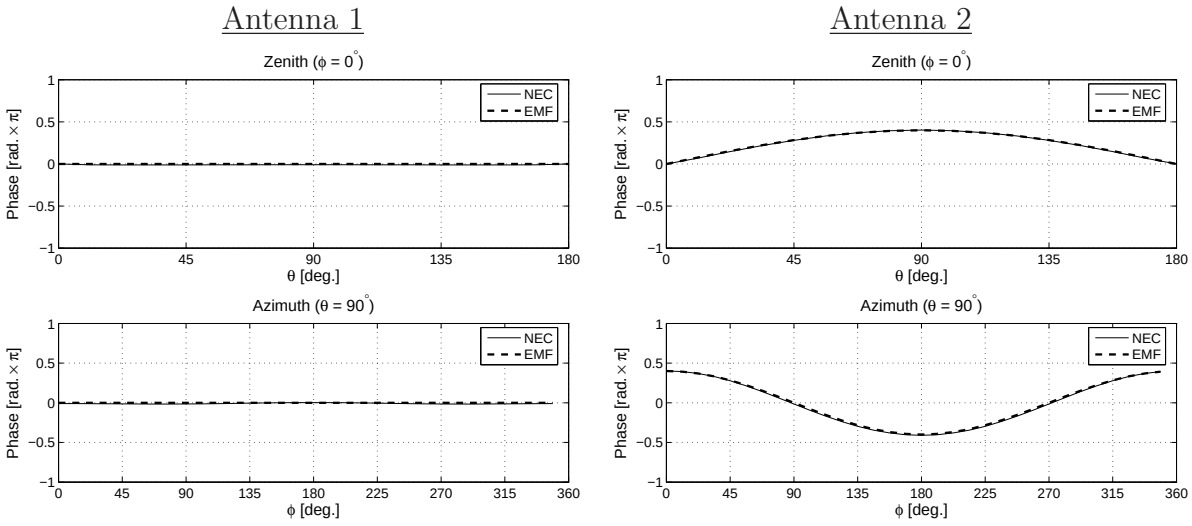


Figure A.5: Phase patterns computed with NEC and EMF.

model used in Chapters 3 and 4. The incident field consists of a large number of $\hat{\mathbf{z}}$-polarized plane waves with independent and uniformly-distributed AoAs and phases. Recall from (3.33) that

$$[\mathbf{\Sigma_h}]_{nm} = \frac{1}{2\pi}\int_0^{2\pi} g_n\left(\pi/2,\phi\right) g_m^*\left(\pi/2,\phi\right) e^{jkd(n-m)\cos\phi} d\phi \ . \tag{A.24}$$

The normalized power $\mathcal{P} = [\mathbf{\Sigma_h}]_{11}/P_{\text{iso}}$ and correlation coefficient $\rho = [\mathbf{\Sigma_h}]_{12}/[\mathbf{\Sigma_h}]_{11}$ are shown in Fig. A.6 for NEC and EMF. As expected from our previous results on patterns, there is little difference between the two curves. The NEC accounts for the slight fluctuations in the received field due to scattering from open-circuited dipoles, while the EMF approach assumes each antenna receives the incident field unperturbed. In fact, the induced EMF patterns (A.21) are *omnidirectional* (i.e., $g_n\left(\pi/2,\phi\right) = 1$), in

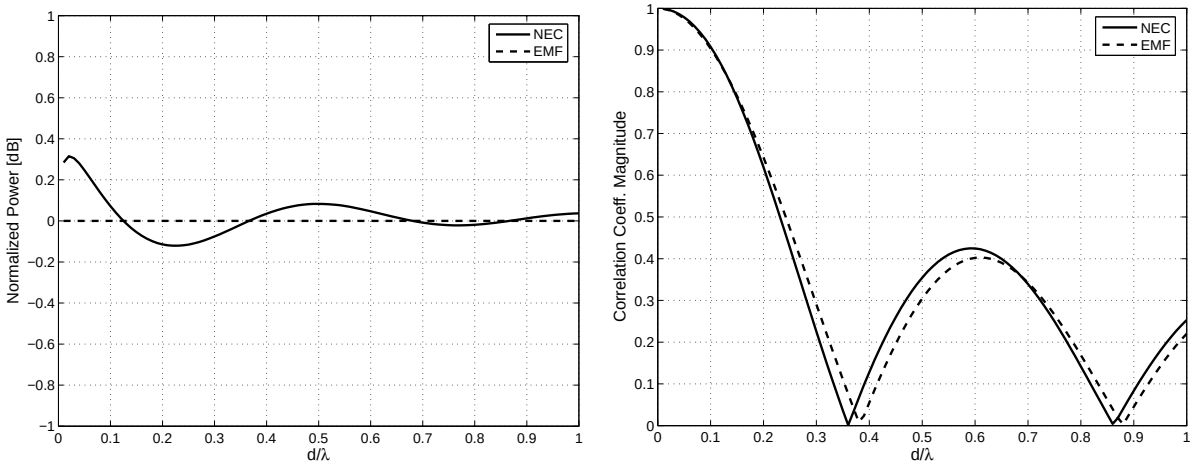


Figure A.6: Open-circuit voltage (a) power and (b) correlation for 2D scattering.

which case we recover Clarke's [17] model

$$\begin{aligned} [\mathbf{\Sigma_h}]_{nm} &= \frac{1}{2\pi}\int_0^{2\pi} e^{jkd(n-m)\cos\phi} d\phi \\ &= J_0\left(kd\left(n-m\right)\right) \ , \end{aligned} \tag{A.25}$$

where $J_0(x)$ is the zeroth-order Bessel function of the first kind.

Appendix B

SISO Low-Noise Design

We briefly review the standard noise model for a single-antenna receiver, shown in Fig. B.1. The antenna is modeled by a Thevenin equivalent circuit with impedance $z_A = r_A + jx_A$ and open-circuit voltage $v_o = hx + n_o$, where $h \sim \mathcal{CN}(0,1)$ is Rayleigh fading, x is the transmitted signal with $\mathrm{E}[|x|^2] = P$, and $n_o \sim \mathcal{CN}(0, 4kT_0Br_A)$ is antenna thermal noise.

The Rothe-Dahlke model [70], a commonly accepted standard for representing noisy,

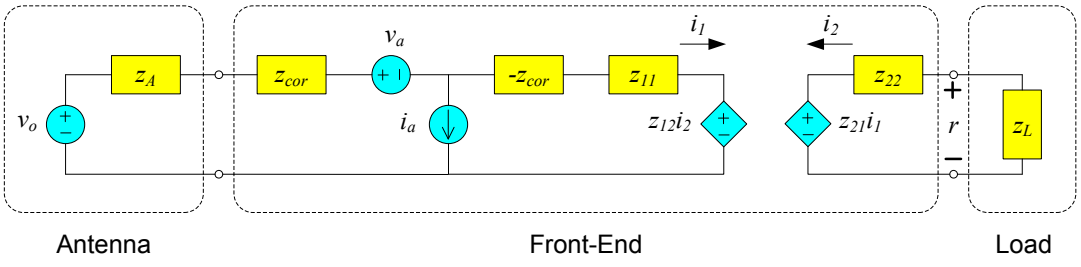


Figure B.1: Circuit model for a SISO receiver front-end.

linear two-ports, is used to model the front-end. The noise sources $v_a \sim \mathcal{CN}(0, 4kT_0Br_a)$ and $i_a \sim \mathcal{CN}(0, 4kT_0Bg_a)$ are mutually independent, where r_a is the *equivalent noise resistance* and g_a is the *equivalent noise conductance*. Noise correlation between the input and output ports is accounted for by the *correlation impedance* z_{cor}.

The voltage across the load is

$$r = \frac{Gz_L}{z_{22} + z_L - Gz_{12}} \left(hx + n_o - z\right) \ , \tag{B.1}$$

where $G = z_{21}/(z_A + z_{11})$ and $z = v_a + (z_A + z_{\text{cor}})i_a$ are the front-end open-circuit voltage gain and noise, respectively. A useful metric associated with a two-port network is the *noise factor*,[1] F, defined as the ratio of the output noise power to the (thermal) noise power contributed by the source alone. Identifying the noise component of r as $n \propto n_o - z$, we have

$$F = \frac{\mathrm{E}[|n|^2]}{\mathrm{E}[|n|^2]|_{z=0}} \ = \ 1 + \frac{\mathrm{E}[|z|^2]}{4kT_0Br_A} \tag{B.2}$$

Now consider a lossless (and therefore noise-free) matching network with impedance matrix

$$\begin{bmatrix} z_{M11} & z_{M12} \\ z_{M21} & z_{M22} \end{bmatrix} \tag{B.3}$$

inserted between the antenna and front-end. The noise factor of the composite (matching plus front-end) network is

$$F = 1 + \frac{r_a + g_a|z'_A + z_{\text{cor}}|^2}{r'_A} \ , \tag{B.4}$$

where the transformed antenna impedance $z'_A = r'_A + jx'_A$ may be chosen arbitrarily with the appropriate matching network.

[1] The *noise figure* of a two-port network is $10 \log_{10} F$.

The SNR (2.15) at the load is

$$\sigma = \frac{P}{4kT_0Br_AF} \ . \tag{B.5}$$

Thus we can maximize σ by minimizing (B.4). The minimum $F_{\min}$ occurs at $z'_A = z_{\text{opt}} = r_{\text{opt}} + jx_{\text{opt}}$, where [70]

$$F_{\min} = 1 + 2\left(g_a r_{\text{cor}} + \sqrt{g_a r_a + (g_a r_{\text{cor}})^2}\right) \tag{B.6}$$

$$z_{\text{opt}} = \sqrt{r_a/g_a + r_{\text{cor}}^2} - jx_{\text{cor}}. \tag{B.7}$$

One particular matching network that achieves $z'_A = z_{\text{opt}}$ is

$$\begin{bmatrix} z_{M11} & z_{M12} \\ z_{M21} & z_{M22} \end{bmatrix} = j \begin{bmatrix} -x_A & \sqrt{r_A r_{\text{opt}}} \\ \sqrt{r_A r_{\text{opt}}} & x_{\text{opt}} \end{bmatrix} \ . \tag{B.8}$$

Appendix C

Proofs for Chapter 4

Here we show that the metrics presented in Section 2.3 satisfy the two properties specified in (4.6). In each case we identify a simplified version of the functional $g(\mathbf{A})$. The proof for ergodic capacity with CSIR also applies to the full CSI (2.19) and MMSE V-BLAST (2.30) capacities.

Ergodic capacity with CSIR (2.20), $g(\mathbf{A}) = \log\det\mathbf{A}$. If $\mathbf{A} \geq \mathbf{B} > \mathbf{0}$ then $\lambda_i(\mathbf{A}) \geq \lambda_i(\mathbf{B}) > 0$ and Property 1 follows immediately from writing the determinant as the product of eigenvalues and using the monotonicity of $\log(\cdot)$. The forward part of Property 2 is trivial, so we verify the converse. If $\log\det\mathbf{C}^\dagger\mathbf{A}\mathbf{C} = \log\det\mathbf{C}^\dagger\mathbf{B}\mathbf{C}$, then the argument of each logarithm must be equal, and by using basic properties of determinants and factoring the Hermitian square root of the positive definite matrix (since $\mathbf{C}$ is full-rank) $\mathbf{C}^\dagger\mathbf{B}\mathbf{C}$ we obtain

$$\prod_{i=1}^{N} \lambda_i(\mathbf{D}) = 1 \ , \tag{C.1}$$

where $\mathbf{D} \triangleq (\mathbf{C}^\dagger\mathbf{B}\mathbf{C})^{-1/2}\mathbf{C}^\dagger\mathbf{A}\mathbf{C}(\mathbf{C}^\dagger\mathbf{B}\mathbf{C})^{-1/2}$. But $\mathbf{C}^\dagger\mathbf{A}\mathbf{C} \geq \mathbf{C}^\dagger\mathbf{B}\mathbf{C}$, so $\mathbf{D} \geq \mathbf{I}$ and each

$\lambda_i(\mathbf{D}) \geq 1$. Together we conclude that $\lambda_i(\mathbf{D}) = 1$, so $\mathbf{D} = \mathbf{I}$ and $\mathbf{C}^\dagger(\mathbf{A} - \mathbf{B})\mathbf{C} = \mathbf{0}$ for all full-rank $\mathbf{C}$. For any nonzero $\mathbf{x} \in \mathcal{C}^N$ it is clear that $\{\mathbf{Cx} \mid \mathbf{C} \in \mathcal{C}^{M\times N} \text{ is full-rank}\} = \mathcal{C}^M$, so the previous condition implies $\mathbf{y}^\dagger(\mathbf{A} - \mathbf{B})\mathbf{y} = 0 \; \forall \; \mathbf{y} \in \mathcal{C}^M$. But $\mathbf{A} - \mathbf{B}$ is nonnegative definite, so we must have $\mathbf{A} = \mathbf{B}$.

STBC pairwise error probability (2.33), $g(\mathbf{A}) = Q(tr\ \mathbf{A})$. The proof is similar to the previous case since $\mathrm{Q}(\cdot)$ is also monotonic and the trace of a matrix is the sum of its eigenvalues, so we only show the converse of Property 2. Let $\mathrm{Q}(\mathrm{tr}\ \mathbf{C}^\dagger\mathbf{AC}) = \mathrm{Q}(\mathrm{tr}\ \mathbf{C}^\dagger\mathbf{BC})$ so that the arguments are equal and we may obtain

$$\sum_{i=1}^{N} \lambda_i \left[\mathbf{C}^\dagger(\mathbf{A} - \mathbf{B})\mathbf{C}\right] = 0 \; . \tag{C.2}$$

But $\mathbf{A} \geq \mathbf{B}$, so $\lambda_i \left[\mathbf{C}^\dagger(\mathbf{A} - \mathbf{B})\mathbf{C}\right] \geq 0$, and together we conclude each eigenvalue is zero, and finally that $\mathbf{A} = \mathbf{B}$.

MIMO-MRC outage probability (2.37), $g(\mathbf{A}) = \lambda_1(\mathbf{A})$. Once again the only non-trivial part of the proof is the converse of Property 2. Suppose that $\lambda_1(\mathbf{C}^\dagger\mathbf{AC}) = \lambda_1(\mathbf{C}^\dagger\mathbf{BC})$ for all full-rank $\mathbf{C} \in \mathcal{C}^{M\times N}$. In particular, let us write $\mathbf{B}^{-1/2}\mathbf{AB}^{-1/2} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^\dagger$, where $\mathbf{U}$ is unitary and $\mathbf{\Lambda}$ is diagonal, and consider $\mathbf{C}$ such that its columns $\{\mathbf{c}_1, \ldots, \mathbf{c}_N\}$ satisfy

$$\mathbf{c}_i = \begin{cases} \mathbf{B}^{-1/2}\mathbf{U}\mathbf{e}_i \; , & i \in \mathcal{I} \\ \mathbf{0} \; , & \text{otherwise} \end{cases} \; , \tag{C.3}$$

where $\mathbf{e}_i$ is the i^{th} canonical basis vector (1 in the i^{th} component, 0 elsewhere) for $\mathcal{C}^M$ and $\mathcal{I} \subset \{1, \ldots, M\}$. For $M < N$ we take $|\mathcal{I}| = M$, while for $M \geq N$ we shall take $|\mathcal{I}| = N$ (note that only the former is uniquely defined). With this choice of $\mathbf{C}$ we obtain

the $N \times N$ matrices

$$\mathbf{C}^\dagger \mathbf{A} \mathbf{C} = \begin{cases} \begin{bmatrix} \mathbf{\Lambda} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, & M < N \\ \mathbf{\Lambda}(\mathcal{I}), & M \geq N \end{cases} \quad \text{and} \quad \mathbf{C}^\dagger \mathbf{B} \mathbf{C} = \begin{cases} \begin{bmatrix} \mathbf{I}_M & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, & M < N \\ \mathbf{I}_N, & M \geq N \end{cases},$$

where $\mathbf{\Lambda}(\mathcal{I})$ denotes the principal submatrix of $\mathbf{\Lambda}$ indexed by $\mathcal{I}$.

For any M, N we see that $\lambda_1(\mathbf{C}^\dagger \mathbf{B} \mathbf{C}) = 1$, and we can always pick our index set such that $1 \in \mathcal{I}$, in which case $\lambda_1(\mathbf{C}^\dagger \mathbf{A} \mathbf{C}) = \lambda_1(\mathbf{B}^{-1/2} \mathbf{A} \mathbf{B}^{-1/2})$. Applying our initial assumption yields $\lambda_1(\mathbf{B}^{-1/2} \mathbf{A} \mathbf{B}^{-1/2}) = 1$. But $\mathbf{A} \geq \mathbf{B}$, so $\lambda_1(\mathbf{B}^{-1/2} \mathbf{A} \mathbf{B}^{-1/2}) \geq \cdots \geq \lambda_N(\mathbf{B}^{-1/2} \mathbf{A} \mathbf{B}^{-1/2}) \geq 1$. Together we conclude that each $\lambda_i(\mathbf{B}^{-1/2} \mathbf{A} \mathbf{B}^{-1/2}) = 1$, so $\mathbf{A} = \mathbf{B}$.

V-BLAST capacity with ZF-SC receiver (2.25), $g(\mathbf{A}) = \sum_i \log \frac{1}{[\mathbf{A}^{-1}]_{ii}}$. Property 1 follows from observing $\mathbf{A} \geq \mathbf{B} \Rightarrow \mathbf{A}^{-1} \leq \mathbf{B}^{-1}$ and the diagonal elements of a nonnegative definite matrix are nonnegative. The converse of Property 2 follows by utilizing the convexity of $\sum_k \log(\cdot)$ and choosing $\mathbf{C}$ as in the prior proof so that the initial assumption implies each $\lambda_i(\mathbf{B}^{-1/2} \mathbf{A} \mathbf{B}^{-1/2}) = 1$.