UNIVERSITY OF CALIFORNIA, SAN DIEGO

Analysis of Multiple Antenna Ultra-Wideband and Millimeter Wave Communication Systems

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy
in
Electrical Engineering
(Communications Theory and Systems)

by

Preeti Nagvanshi

Committee in charge:

Professor Elias Masry, Co-Chair
Professor Laurence B. Milstein, Co-Chair
Professor William S. Hodgkiss
Professor Bhaskar D. Rao
Professor Ruth Williams

2007

The dissertation of Preeti Nagvanshi is approved, and it is acceptable in quality and form for publication on microfilm:

_____

_____

_____

_____
Co-Chair

_____
Co-Chair

University of California, San Diego

2007

To my husband Mahim,

my parents,

and my brother Manpreet.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to thank my advisors, Professor Elias Masry and Professor Laurence Milstein for their support and guidance throughout the duration of my Ph.D. studies. Their insight and technical knowledge was invaluable and their dedication to research has been a constant source of inspiration and encouragement. It has been a pleasure working under their supervision.

I would also like to express my gratitude to my committee members, Professors Bhaskar Rao, William Hodgkiss and Ruth Williams.

I am grateful to Professor Bhaskar Rao for giving me the opportunity to participate in his group meetings and for the course on MIMO Wireless Communications, which were very helpful and informative. Thanks to Professors William Coles and Gabriel Rebeiz for the immensely fruitful discussions on antennas. Thanks also to Professor Alon Orlitsky who's course on Information Theory was inspirational.

I would like to thank my friends and colleagues who have made my stay at UCSD a memorable and enjoyable experience. Thanks in particular to Aditya Jagannatham, Ramesh Annavajjala, Thomas Svantesson, Patrick Amihood, Yoav Nebat, Narayana Santhanam, Chandra Murthy, Qi Qu, June Chul Roh, Yogananda Isukapalli, Farzad Parvaresh, Claudio da Silva, Yeliz Tokgoz and Joseph Soriaga.

Above all, I owe my deepest gratitude to my husband Mahim, my parents and my brother Manpreet. I dedicate this work to them for their belief in me and for their unwavering support and encouragement.

The text of chapter II, III and IV is, in part, a reprint of the following works:

1. P. Nagvanshi, E. Masry and L.B. Milstein, "Optimum transmit-receive beamforming with noisy channel estimates for correlated MIMO Rayleigh channels," accepted for publication in the *IEEE Transactions on Communications.*

2. P. Nagvanshi, E. Masry and L.B. Milstein, "Performance of Multiple Antenna DS-CDMA UWB System With Noisy Channel Estimates and Narrow-Band Interference ," *IEEE Journal of Selected Topics in Signal Processing,* vol. 1, no. 3,

pp. 470 - 482, Oct. 2007.

3. P. Nagvanshi, E. Masry and L.B. Milstein, "The effect of channel estimation errors on transceiver optimization for correlated MIMO Rayleigh channels," *Proc. 39th Annual Conference on Information Sciences and Systems (CISS)*, March 2005.

4. P. Nagvanshi, E. Masry and L.B. Milstein, "Performance Analysis of Multiple Antenna DS-CDMA UWB System with Noisy Channel Estimates and Narrow-Band Interference ," *IEEE Military Communications Conference (MILCOM)*, pp. 1-7, Oct. 2006.

5. P. Nagvanshi, E. Masry and L.B. Milstein, "Performance Analysis of QAM Multiple Antenna Systems at 60GHz ," *IEEE Transactions on Communications,* to be submitted.

The dissertation author was the primary researcher and author, and the co-authors listed in these publications directed and supervised the research which forms the basis for chapters II-IV.

VITA

1999	Integrated M.Tech., (Integrated Bachelor's and Master's degrees) Electrical Engineering, Indian Institute of Technology, Bombay
1999–2001	Systems Engineer, Cradle Technologies, Pune, India, and Fremont, CA
2001–2007	Research Assistant, Dept. of Electrical and Computer Engineering, University of California, San Diego
2005	M.S., Electrical and Computer Engineering, University of California, San Diego
2007	Ph.D., Electrical and Computer Engineering, University of California, San Diego

PUBLICATIONS

P. Nagvanshi, E. Masry and L.B. Milstein, "Optimum transmit-receive beamforming with noisy channel estimates for correlated MIMO Rayleigh channels," accepted for publication in the *IEEE Transactions on Communications.*

P. Nagvanshi, E. Masry and L.B. Milstein, "Performance of Multiple Antenna DS-CDMA UWB System With Noisy Channel Estimates and Narrow-Band Interference ," *IEEE Journal of Selected Topics in Signal Processing,* vol. 1, no. 3, pp. 470 - 482, Oct. 2007.

P. Nagvanshi, E. Masry and L.B. Milstein, "The effect of channel estimation errors on transceiver optimization for correlated MIMO Rayleigh channels," *Proc. 39th Annual Conference on Information Sciences and Systems (CISS)*, March 2005.

P. Nagvanshi, E. Masry and L.B. Milstein, "Performance Analysis of Multiple Antenna DS-CDMA UWB System with Noisy Channel Estimates and Narrow-Band Interference ," *IEEE Military Communications Conference (MILCOM)*, pp. 1-7, Oct. 2006.

P. Nagvanshi, E. Masry and L.B. Milstein, "Performance Analysis of QAM Multiple Antenna Systems at 60GHz ," *IEEE Transactions on Communications,* to be submitted.

ABSTRACT OF THE DISSERTATION

Analysis of Multiple Antenna Ultra-Wideband and Millimeter Wave Communication Systems

by

Preeti Nagvanshi
Doctor of Philosophy in Electrical Engineering (Communications Theory and Systems)
University of California, San Diego, 2007
Professor Elias Masry, Co-Chair
Professor Laurence B. Milstein, Co-Chair

The home and office environments are experiencing an ever increasing penetration of consumer electronic devices, often requiring data rates well in excess of tens of megabits per second. Communication to and from such devices has mostly relied on wireline technologies such as USB, DVI and IEEE1394.

Ultra-wideband (UWB) and millimeter-wave (mmW) systems have been proposed to replace these wireline communication systems with short range high speed wireless networks. The significantly higher occupied bandwidth of UWB and mmW systems provides immense advantages in terms of higher data rates, while at the same time presenting new challenges such as greater susceptibility to interferers and possibly complex transceiver design. This dissertation addresses several technical challenges in the design of UWB and mmW systems. Multiple antenna techniques to improve the interference suppression capabilities and reliability of the UWB and mmW systems are employed. First, a MIMO beamforming system is analyzed. In the presence of antenna correlation and noisy channel estimates, an optimal MIMO beamforming scheme is proposed. The performance of this scheme is analyzed through a closed-form expression for the probability of error, and the combined effects of channel estimation errors and diversity on the system performance are studied. At the receiver, a fixed length antenna

array is considered due to spatial constraints. For such an array, it is shown that there exists an optimal number of receive antenna elements for a given array length.

The performance of a DS-CDMA-based UWB system with multiple antennas at the receiver is then analyzed. An optimal spatio-temporal receiver is proposed and its performance evaluated in the presence of narrowband interference, multiple access interference, antenna correlation and channel estimation errors. For a fixed array length and fixed maximum diversity level, the tradeoff between the number of antennas and the number of temporal taps in order to achieve the best performance is investigated.

A 60 GHz mmW system is considered next. Multiple antenna equalization scheme to suppress both the intersymbol interference and multiple access interference is employed. A spectrally efficient multilevel quadrature amplitude modulation and a realistic IEEE channel model are used for analysis. The combined effect of interference suppression and spatial correlation on the system performance is studied through an analytically derived expression for bit error rate. It is shown that joint spatial and temporal processing can significantly improve the system performance.

I

Introduction

I.A Consumer Electronics and Home Networking

The recent surge of digital consumer electronic devices within the home, such as DVD players, cable and satellite televisions, digital cameras and portable music players, has created a need for high-speed, short-range home entertainment networks. The increase in number of devices increases the complexity and number of wires required for connecting these devices together.

A typical living room entertainment setup could consist of a DVD player, a cable set-top box and a digital video recorder, amongst other devices, such as the one shown in Figure I.1. These devices may then connect to a television and to an audio receiver via cables. The complexity and sheer mess of cables surrounding such a home entertainment setup is calling for high speed, short-range wireless networks which could make an average user's entertainment setup aesthetically more pleasing, and easier to connect.

A home or office desk is another instance which could benefit significantly from such high speed networks. A typical desk may have a computer, a monitor, a printer, a high speed external storage device and a personal digital assistant, among other devices. All these consumer electronics and personal computing peripherals frequently require data rates in excess of a few tens of megabits per second for reliable connectiv-

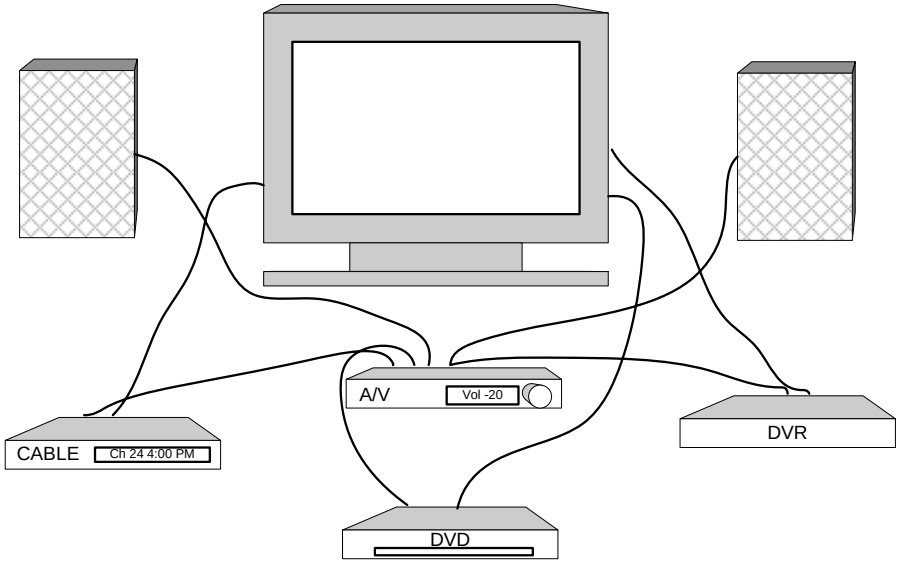

Figure I.1: A typical living room entertainment setup.

ity. The next section enumerates the wireless networking technologies available to the average home or office user.

I.B Home Wireless Networking

Various wireless data networking technologies are available to the home or office user. Most of them can be classified in one of the following two categories:

1. Wireless Personal Area Networks (WPAN): These networks are designed for relatively short-range communication, within distances of a few feet. Bluetooth [1] and infra-red are prime examples of this technology. To maintain low cost and low power consumption, Bluetooth was designed for very low data rates. The latest rendition of Bluetooth supports a maximum raw data rate of 3.0 Mbps, which is too low to support the above mentioned consumer electronic devices. Due to the limitation of it's data-rate, Bluetooth has been relegated to low data-rate applica-

tions such as wireless headsets.

2. Wireless Local Area Networks (WLAN): WLAN networks encompass technologies such as the ubiquitous WiFi (802.11 a/b/g) networks and home RF. These networks are designed to operate across an average user's home or office, and can cover a few hundred feet in distance. WiFi offers raw data rates up to 54 Mbps [2]. WiFi also consumes a significant amount of power, and sustained usage of WiFi in mobile devices can drastically reduce battery life.

I.C Short Range High Data Rate Wireless Networking

The Federal Communications Commission (FCC) has recently allocated two large chunks of spectrum (3.1 - 10.6 GHz and 57 - 64 GHz) for unlicensed wireless communication usage. These spectral bands have opened up new possibilities for realizing short range, high data rate wireless communication systems and set off a flurry of activity in both industry and academia. Several proposals have been presented to efficiently utilize these newly available spectra.

I.C.1 Ultra Wideband

The spectrum from 3.1 - 10.6 GHz is designated for Ultra Wideband (UWB) communication usage. The FCC defines UWB as any transmission system that occupies either a minimum bandwidth of 500 MHz or more than 20% of the center frequency [3] [4]. Since this spectrum overlaps with various other bands such as WiFi, Bluetooth, radar and cellular, the transmit power of the UWB system has to be low enough so as to have minimal impact on these other systems. To this end, the FCC has mandated that the power spectral density (PSD) of a UWB transmitter should be less than -41.3 dBm/MHz. The allowed spectral mask for UWB systems is shown in Figure I.2. Due to its low transmission power levels and large spectral occupancy, UWB makes it possible to realize a large number of low power applications requiring high speed access over short distances. The fine time resolution of UWB signals also makes it a highly effective

solution for high precision position location in dense environments. In order to realize the many attractive features of UWB, a system designer faces several formidable challenges. Since a UWB system is envisioned to operate in a dense wireless environment, it has to cope with the interference it receives from other UWB-enabled devices operating in close proximity, as well as strong inband interference from the overlaying narrowband systems such as WiFi. In addition, they need to comply with the FCC imposed power limitations and be of sufficiently low complexity to reduce cost and increase ease of implementation.

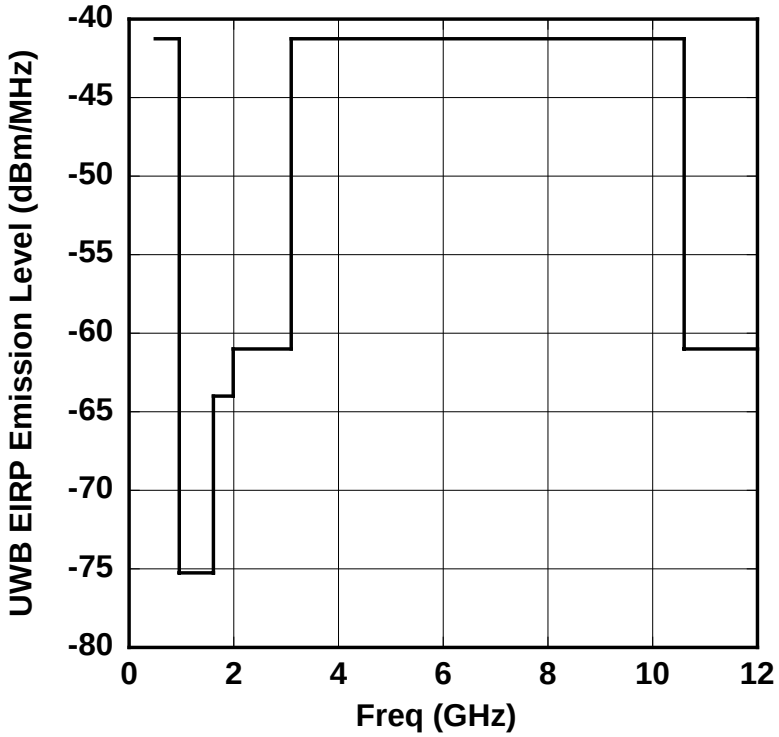

Figure I.2: FCC designated spectral mask for Ultra Wideband systems. Allowed emission level is in dBm/MHz [3].

I.C.2 Millimeter Wave Band

In the 57 - 64 GHz band, the FCC mandates a system bandwidth greater than 100 MHz, and limits the transmit power to less than 10 W [5]. This frequency band is

of great interest because of the attenuation characteristics of millimeter waves (mmW). Since free space path loss increases quadratically with frequency, at 60GHz, there is a large attenuation of 68 dB at a distance of 1m [6], [7]. Considering even larger distances (several km), there is an additional attenuation due to atmospheric oxygen of 15.1 dB/km [8], thus, making the band ideally suited for short range communication and a dense deployment scenario, such as a home or office. Multiple independent links can be created in a house, as the interference level experienced from similar mmW devices placed in adjacent rooms would be low. This is because the signal would be highly attenuated when passing through a wall. Due to the relaxed emission specifications (when compared to the spectral mask for UWB), the use of directional antennas is permitted for mmW devices. This is advantageous, as directional antennas can help improve the range of mmW devices. Also, due to its shorter wavelength, the mmW band offers the possibility of smaller components and antennas, thus reducing the overall size of these devices.

Systems in both these bands (UWB and mmW) can offer very high data rates, due to the significantly higher available bandwidth. Due to large bandwidth, both of these systems experience fine multipath resolution and lack of significant fading. The received energy is distributed thinly over a large number of these multipath components, consequently, the energy-per-path is very low. This, thus, presents great difficulties in accomplishing tasks such as channel estimation, synchronization and multipath combining. Also, the presence of co-channel interference when multiple such systems are operating in close proximity in a dense wireless environment further complicates these tasks.

I.D Dissertation Focus and Outline

This dissertation addresses several technical challenges in the design of UWB and mmW systems. One of the major issues is the ability to coexist in an environment with multiple interferers. Both of these systems would benefit significantly from the use

of efficient interference-suppression techniques. One way to mitigate interference is to use multiple antenna beamforming, which exploits the spatial signature of an interferer in order to suppress it. In this thesis, we consider such spatial processing techniques in conjunction with temporal processing to enhance the interference suppression capabilities and reliability of these systems.

MIMO systems, with multiple antennas at the transmitter and the receiver, have been shown to improve robustness and increase throughput in wireless communication channels [9], [10]. Effective utilization of spatial domain through techniques such as spatial multiplexing and space-time coding have been proposed to achieve higher data rates, diversity and coding gains [11], [12], [13]. These multiple antenna techniques, therefore, subsume great significance as powerful tools for improving the performance of a wireless communication system. Therefore, in Chapter II, we first consider a MIMO system. In the presence of both imperfect channel estimation and antenna correlation, an optimum MIMO beamforming scheme is proposed, and an exact closed-form expression for the probability of error of the system is derived. The impact of channel estimation accuracy and diversity on the performance of the system is studied.

In Chapter III, we consider a DS-CDMA UWB system and analyze its error rate performance under the conditions of imperfect channel estimation and the presence of both narrowband interference (NBI) and multiple access interference (MAI). An optimal spatio-temporal receiver is designed that first forms an estimate of the channel in the presence of NBI and MAI, and then uses it to optimally combine the multipath components such that the output signal-to-interference-plus-noise-ratio is maximized. An exact closed-form expression for the probability of error is established when the temporal correlation between the NBI during the channel-estimation phase and the data-detection phase is negligible.

Chapter IV focuses on the performance analysis of a millimeter wave communication system at 60 GHz. A multiple-antenna equalization scheme is employed to combat both intersymbol interference and multiple-access interference. In order to increase the spectral efficiency, a higher order modulation format is used. The perfor-

mance of this system through an analytically derived expression for the bit error rate is investigated. Finally, Chapter V concludes the thesis with an overview of this research.

II

Optimum Transmit-Receive Beamforming with Noisy Channel Estimates for Correlated MIMO Rayleigh Channels

II.A Introduction

MIMO systems can significantly enhance link reliability by realizing transmit and receive diversity to mitigate the effects of channel fading [14, 15]. Full spatial diversity can be achieved using space-time coding when no channel knowledge is available at the transmitter, whereas when the transmitter and the receiver have access to perfect channel state information (CSI), a spatial diversity scheme known as maximum ratio transmission (also known as beamforming) [14] was proposed. In this scheme, the transmitter utilizes CSI to select the strongest eigenmode for transmission so that the output signal-to-noise ratio is maximized, and the independently fading MIMO channel enables multiple independent reception of the same signal to realize the full diversity order. The performance of this scheme in terms of exact probability of error expression was studied by Dighe et.al [15]. Central to [14, 15] is the assumption of perfect channel

knowledge and independent fading.

Practical channel estimation schemes are, however, prone to errors, and fades at the antenna elements could be significantly correlated due to proximity of the elements. Typically, known pilot symbols are exploited for channel estimation. The accuracy of the estimates depends on the particular channel estimation scheme, the channel conditions and the pilot and data multiplexing scheme [16]. Also, in a practical scenario, spatial constraints limit the size of an antenna array. In such a situation, an increase in the number of antenna elements is brought about by reduction in inter-element spacing, and a subsequent increase in correlation between the fades. In this chapter, we extend the work in [14], [15] and design a jointly optimum transmitter receiver scheme for MIMO channels that takes into account the combined effects of correlated fading and imperfect channel estimation, with the aim of minimizing the average probability of error.

The effect of antenna correlation on different aspects of MIMO system design and performance has been addressed previously. Specifically, the impact of antenna correlation on the capacity of MIMO channels has been studied in [17]-[19]. In [20], an optimum transmitter scheme and a necessary and sufficient condition under which beamforming is capacity-achieving is determined. However, it is assumed that perfect CSI is available at the receiver, and only partial CSI at the transmitter. In the context of channel estimation for MIMO systems, issues related to design of an optimal training sequence for uncorrelated and correlated MIMO channels, respectively, have been studied in [21] and [22]. For uncorrelated MIMO channels, the effect of channel estimation on the precoder and decoder design for a multimode MIMO system with MSE as the performance criterion was considered in [23]. Also addressed in [23] was the power allocation tradeoff between channel estimation and data transmission. This work was later extended in [24], [25] to include the effects of receiver antenna correlation. However, note that work in [23]-[25] considers MSE as the performance metric and does not present any probability of error analysis.

Our work, both here and in [26], differs significantly from all of the above works in the following respect. We design an optimal MIMO beamforming scheme in

the presence of both imperfect channel estimation and antenna correlation, and analyze its performance through a closed-form expression for the probability of error [27]. In our system, the receiver forms an estimate of the MIMO channel using the pilot symbols and makes it available to the transmitter through an error-free feedback channel. We assume that spatially correlated fading exists at one end of the communication channel, and independent fading at the other. To model the correlated fading, the physical MIMO channel correlation model of [28] has been adopted.

Based on knowledge of the channel estimate, we design the optimum transmitter and receiver beamforming structures that maximizes the output signal-to-noise ratio (SNR) and therefore minimizes the average probability of error. We establish an exact closed-form expression for the probability of error of the system and, based on this expression, study the combined effect of channel estimation errors and diversity on the performance. At the receiver, we consider a physically constrained linear array, where the effective aperture of the array is *fixed*. For such an array, we show that there exists an optimal number of receive antenna elements, and that the optimum value depends both on the available transmit and receive diversity and the quality of the estimates. Note that higher fade correlation leads to improved channel estimation on one hand, but lower diversity on the other. We next show that for a given number of transmit and receive antennas, as the correlation between the channel coefficients increases due to poor scattering, the system performance degrades. This is because the benefits of improved channel estimation accuracy are offset by the loss in diversity order.

The rest of the chapter is organized as follows. In Section II.B, the system model and channel estimation scheme are described. The optimal transceiver is derived in Section II.C. In Section II.D, we obtain an exact closed-form expression for the probability of error. This is followed by numerical results and discussion in Section II.E. Finally, we state the conclusions in Section II.F.

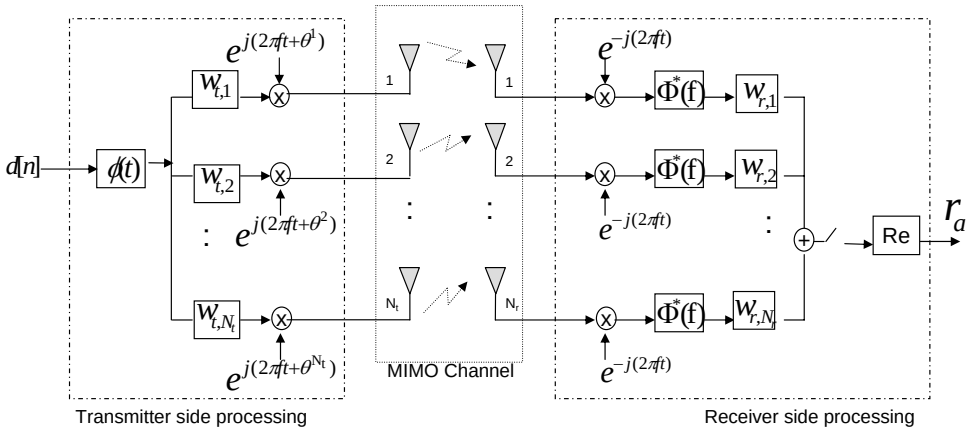

Figure II.1: System block diagram - Data transmission

II.B System Model

II.B.1 Signal and Channel Model

Consider a single-user MIMO system with binary signaling, operating over a frequency flat slowly fading channel with N_t transmitter antennas and N_r receiver antennas. We employ the transmitter/receiver beamforming technique where the single binary data stream at the input, denoted as $d[n]$, is transmitted from the N_t transmit antennas with antenna weights $\{w_{t,i}\}_{i=1}^{N_t}$, as shown in the Figure II.1. In the figure, $\phi(t)$ is the impulse response of a wave shaping filter satisfying the Nyquist criterion. At the receiver, after demodulation, matched filtering and sampling, the signal from each of the receiver antennas is combined with weights ($\{w_{r,i}\}_{i=1}^{N_r}$) across the array. Assuming that perfect carrier and bit synchronization have been established, the output test statistic for the a^{th} transmitted bit can be written as

$$r_a = \text{Re}(\mathbf{w}_r^H \mathbf{H} \mathbf{w}_t) d[a] + \text{Re}(\mathbf{w}_r^H \mathbf{n}_a) \tag{II.1}$$

where $(\cdot)^H$ denotes the Hermitian operator, $\mathbf{w}_r$ is an $N_r \times 1$ weight vector at the receiver, $\mathbf{w}_t$ is an $N_t \times 1$ weight vector at the transmitter, $d[a]$ is the binary data at the ath sampling

instant, and $\mathbf{n}_a$ is the iid complex Gaussian noise vector with zero mean and covariance $E[\mathbf{n}_a\mathbf{n}_a^H] = N_o\mathbf{I}$. We assume that the total transmit power is constant, irrespective of the number of transmit antennas, i.e., $\mathbf{w}_t^H\mathbf{w}_t = P_o$.

The channel is assumed to be Rayleigh fading, therefore, the channel fade between each transmit and receive antenna pair is complex Gaussian. The $N_r \times N_t$ MIMO channel matrix $\mathbf{H}$, formed by collecting all $N_r\,N_t$ channel fades, is multivariate circularly symmetric complex Gaussian distributed. We denote this distribution as $\sim CN(\mathbf{0}, \mathbf{P})$, where $CN(\mathbf{A}, \mathbf{B})$ represents a complex normal distribution with mean matrix $\mathbf{A}$ and covariance matrix $\mathbf{B}$. The covariance matrix $\mathbf{P}$ captures the spatially correlated fading. We will adopt the well-known Kronecker product model for channel covariance matrix [29] and denote $\mathbf{P} = E[\mathrm{vec}(\mathbf{H}^T)\mathrm{vec}(\mathbf{H}^T)^H] = \mathbf{\Sigma} \otimes \mathbf{\Psi}$, where $\otimes$ is the Kronecker product [31], and the $\mathrm{vec}(\cdot)$ operator stacks up all the columns of its argument to form one long column vector [31]. This implies that the covariance of the ij^{th} and lm^{th} entry of $\mathbf{H}$ is modelled as $E[(\mathbf{H})_{i,j}(\mathbf{H})^*_{l,m}] = (\mathbf{\Sigma})_{i,l}(\mathbf{\Psi})_{j,m}$, where the elements of the $N_r \times N_r$ matrix $\mathbf{\Sigma}$ denote the covariance of any column of $\mathbf{H}$, and is the fade covariance at the receiver side, and the $N_t \times N_t$ matrix $\mathbf{\Psi}$ is the covariance of any row of $\mathbf{H}$ and is the fade covariance at the transmitter. Both $\mathbf{\Sigma}$ and $\mathbf{\Psi}$ are positive definite Hermitian matrices.

In our setup, it is assumed that independent fading exists at one end of the communication link and correlated fading at the other. This implies that either $\mathbf{\Sigma}$ or $\mathbf{\Psi}$ is an identity matrix, while the other one is an arbitrary covariance matrix. This is a reasonable assumption in an uplink or downlink of a communications system where a *fixed length* antenna array is deployed at the mobile station. The closely spaced antenna elements would lead to correlated fading at the mobile end, whereas the space considerations, and consequently, the correlatedness assumption, can be relaxed at the base station. To model correlated fading, we adopt the physical MIMO channel correlation model of [28], the specifics of which will be discussed in Section II.E.

II.B.2 Channel Estimation

The channel matrix $\mathbf{H}$ is unknown at the receiver. An estimate, $\hat{\mathbf{H}}$, of $\mathbf{H}$ is obtained based on a received pilot. For simplicity, orthogonal pilot symbols are used on each transmitter antenna. In [21], it was shown that an orthogonal pilot is optimal for estimating uncorrelated MIMO channels. The transmitted pilot vector from N_t transmitters is denoted by $\sqrt{\frac{P_p}{N_t}}\,\mathbf{c}[n]^T \triangleq \sqrt{\frac{P_p}{N_t}}\,\{c_1[n], c_2[n], \ldots, c_{N_t}[n]\}$, where $c_i[n]$ is the pilot symbol transmitted from the i^{th} antenna at the n^{th} time instant, and P_p is the pilot power-per-symbol. The pilot transmission takes place for a duration of q, for integer q, symbol intervals. The output pilot from all the antennas, after matched filtering and sampling, can be written in matrix form as

$$\mathbf{Y} = \sqrt{\frac{P_p}{N_t}}\,\mathbf{H}\mathbf{C}^H + \widetilde{\mathbf{N}} \tag{II.2}$$

where $\mathbf{C}$ is a $q \times N_t$ matrix of known pilot symbols with the i^{th} row of the matrix given by $\mathbf{c}[i]^T$, for $i = 0, \ldots, q-1$, and $\widetilde{\mathbf{N}}$ is an $N_r \times q$ spatio - temporal uncorrelated complex Gaussian noise matrix, i.e., $\widetilde{\mathbf{N}} \sim CN(0, N_o(\mathbf{I} \otimes \mathbf{I}))$. Since the transmitted pilot is orthogonal in space, we have $\mathbf{C}^H\mathbf{C} = q\mathbf{I}$. Multiplying the pilot matrix $\mathbf{Y}$ in (II.2) by $\mathbf{C}$, we get

$$\mathbf{X} \triangleq \mathbf{Y}\mathbf{C} = \sqrt{\frac{P_p\, q^2}{N_t}}\,\mathbf{H} + \widetilde{\mathbf{N}}\mathbf{C} \qquad (\text{let } \mathbf{G} \triangleq \widetilde{\mathbf{N}}\mathbf{C}). \tag{II.3}$$

An example of the pilot transmission is shown in Figure II.2. As $\mathbf{C}$ is a deterministic matrix, $\widetilde{\mathbf{N}}\mathbf{C}$ is complex Gaussian with zero mean, and its covariance matrix is given by $N_o(\mathbf{I} \otimes \mathbf{C}^H\mathbf{C}) = qN_o(\mathbf{I} \otimes \mathbf{I})$. Thus, $\mathbf{G}$ in (II.3) can be viewed as the new Gaussian noise matrix with the aforementioned properties.

Next, we use the received pilot $\mathbf{X}$ to form a linear minimum mean-square error (LMMSE) estimate, $\hat{\mathbf{H}}$. Since the statistics are Gaussian, $\hat{\mathbf{H}}$ is given by the mean of the posterior pdf [30], i.e., $\hat{\mathbf{H}} = E[\mathbf{H}|\mathbf{X}]$. Note that an expression for $\hat{\mathbf{H}}$ can easily be obtained from an equivalent expression for $\hat{\mathbf{h}} \triangleq \text{vec}(\hat{\mathbf{H}}^T)$. Thus, we define $\mathbf{h} \triangleq \text{vec}(\mathbf{H}^T)$ and $\mathbf{g} \triangleq \text{vec}(\mathbf{G}^T)$, so that $\mathbf{x} \triangleq \text{vec}(\mathbf{X}^T) = \sqrt{\frac{P_p\, q^2}{N_t}}\mathbf{h} + \mathbf{g}$. From these definitions

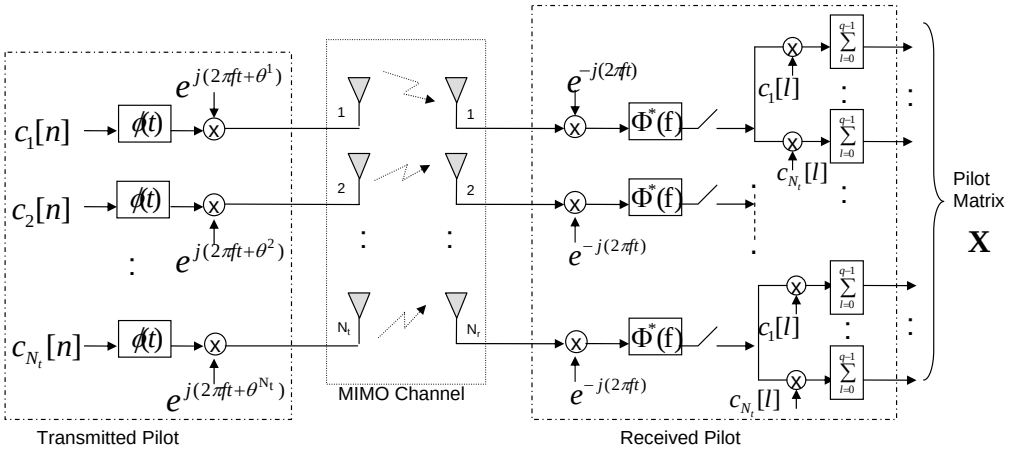

Figure II.2: System block diagram - Pilot transmission

and [30, eq. 15.61], we have

$$\begin{aligned}\hat{\mathbf{h}} = E[\mathbf{h}|\mathbf{x}] &= E[\mathbf{h}] + \mathbf{C_{hx}}\mathbf{C_{xx}^{-1}}(\mathbf{x} - E[\mathbf{x}]) \\ &= \sqrt{\gamma}\mathbf{C_{hh}}(\gamma\mathbf{C_{hh}} + \mathbf{C_{gg}})^{-1}\mathbf{x} \\ &= \sqrt{\gamma}(\mathbf{\Sigma}\otimes\mathbf{\Psi})(\gamma\mathbf{\Sigma}\otimes\mathbf{\Psi} + qN_o(\mathbf{I}\otimes\mathbf{I}))^{-1}\mathbf{x} \qquad \text{(II.4)}\end{aligned}$$

where $\gamma \triangleq \frac{P_p\, q^2}{N_t}$. In the second equality above, $E[\mathbf{h}] = 0$, $E[\mathbf{x}] = 0$, the covariance matrix of the received pilot $\mathbf{C_{xx}}$ is given by $\mathbf{C_{xx}} = E[(\sqrt{\gamma}\mathbf{h} + \mathbf{g})(\sqrt{\gamma}\mathbf{h} + \mathbf{g})^H] = \gamma\mathbf{C_{hh}} + \mathbf{C_{gg}}$, where $\mathbf{C_{hh}} = \mathbf{P} = (\mathbf{\Sigma}\otimes\mathbf{\Psi})$ is the channel covariance matrix and $\mathbf{C_{gg}} = qN_o(\mathbf{I}\otimes\mathbf{I})$ is the covariance matrix of the noise during the channel estimation phase. The covariance between the channel and the received pilot is given by $\mathbf{C_{hx}} = E[\mathbf{h}(\sqrt{\gamma}\mathbf{h}+\mathbf{g})^H] = \sqrt{\gamma}\mathbf{C_{hh}}$. Now, letting $\mathbf{\Psi} = \mathbf{I}$ in (II.4) yields

$$\begin{aligned}\hat{\mathbf{h}} &= \sqrt{\gamma}(\mathbf{\Sigma}\otimes\mathbf{I})(\gamma\mathbf{\Sigma}\otimes\mathbf{I} + qN_o\mathbf{I}\otimes\mathbf{I})^{-1}\mathbf{x} \\ &= \text{vec}(\{\sqrt{\gamma}\mathbf{\Sigma}(\gamma\mathbf{\Sigma} + qN_o\mathbf{I})^{-1}\mathbf{X}\}^T). \qquad \text{(II.5)}\end{aligned}$$

The properties [31, (see eqs. (4), (11a), (8), (5) and (14) of Section 2.4)] of the Kronecker product were used to simplify the second equality in (II.5). Similarly, if we set

$\mathbf{\Sigma} = \mathbf{I}$ in (II.4), we can show that $\hat{\mathbf{h}} = \text{vec}(\{\mathbf{X}(\sqrt{\gamma}\mathbf{\Psi}(\gamma\mathbf{\Psi} + qN_o\mathbf{I})^{-1})^T\}^T)$. Since $\hat{\mathbf{h}} = \text{vec}(\hat{\mathbf{H}}^T)$, an expression for $\hat{\mathbf{H}}$, under these special cases, can be shown to be given by

$$\hat{\mathbf{H}} = \begin{cases} \sqrt{\frac{P_p\, q^2}{N_t}}\, \mathbf{\Sigma}(qN_o\mathbf{I} + \frac{P_p\, q^2}{N_t}\mathbf{\Sigma})^{-1}\mathbf{X}, & \text{if } \mathbf{\Psi} = \mathbf{I} \\ \mathbf{X}(\sqrt{\frac{P_p\, q^2}{N_t}}\, \mathbf{\Psi}(qN_o\mathbf{I} + \frac{P_p\, q^2}{N_t}\mathbf{\Psi})^{-1})^T, & \text{if } \mathbf{\Sigma} = \mathbf{I}. \end{cases} \tag{II.6}$$

Notice that $\hat{\mathbf{H}}$ in (II.6) depends on the covariance matrices $\mathbf{\Sigma}$ or $\mathbf{\Psi}$, contingent on whether correlation exists at the receiver or at the transmitter. Thus, the optimal linear estimator in (II.6) effectively utilizes the antenna correlation information while forming the estimate. The extent of correlation between the fades impacts the channel estimation accuracy, which in turn effects the system performance. Later, in the numerical results, we will study this effect by examining the behavior of MMSE and the probability of error as the correlation is varied.

Also, $\hat{\mathbf{H}}$ is a statistical quantity and is characterized by a Gaussian distribution with zero mean and covariance $\mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}} = \mathbf{C}_{\mathbf{hx}}\mathbf{C}_{\mathbf{xx}}^{-1}\mathbf{C}_{\mathbf{xh}} = \sqrt{\gamma}\mathbf{C}_{\mathbf{hh}}(\gamma\mathbf{C}_{\mathbf{hh}} + \mathbf{C}_{\mathbf{gg}})^{-1}\sqrt{\gamma}\mathbf{C}_{\mathbf{hh}} = (\sqrt{\gamma}\mathbf{\Sigma} \otimes \mathbf{\Psi})(\gamma\mathbf{\Sigma} \otimes \mathbf{\Psi} + qN_o\mathbf{I} \otimes \mathbf{I})^{-1}(\sqrt{\gamma}\mathbf{\Sigma} \otimes \mathbf{\Psi})$.

As before, we simplify $\mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}$ for the case when $\mathbf{\Psi} = \mathbf{I}$ or $\mathbf{\Sigma} = \mathbf{I}$. Upon simplification, the estimator can be shown to have the distribution $\hat{\mathbf{H}} \sim CN(0, \bar{\mathbf{\Sigma}} \otimes \bar{\mathbf{\Psi}})$, where

$$\begin{aligned} \bar{\mathbf{\Sigma}} &= \begin{cases} \frac{P_p q}{N_t N_o}\, \mathbf{\Sigma}(\mathbf{I} + \frac{P_p q}{N_t N_o}\mathbf{\Sigma})^{-1}\mathbf{\Sigma}, & \text{if } \mathbf{\Psi} = \mathbf{I} \\ \mathbf{I}, & \text{if } \mathbf{\Sigma} = \mathbf{I} \end{cases} \\ \text{and } \bar{\mathbf{\Psi}} &= \begin{cases} \mathbf{I}, & \text{if } \mathbf{\Psi} = \mathbf{I} \\ \frac{P_p q}{N_t N_o}\, \mathbf{\Psi}(\mathbf{I} + \frac{P_p q}{N_t N_o}\mathbf{\Psi})^{-1}\mathbf{\Psi}, & \text{if } \mathbf{\Sigma} = \mathbf{I}. \end{cases} \end{aligned} \tag{II.7}$$

II.C Optimum Transceiver

With knowledge of the channel estimate available, we wish to find the optimal transmitter and receiver weights, $\mathbf{w}_t$ and $\mathbf{w}_r$, that maximize the conditional output SNR (conditioned on $\hat{\mathbf{H}}$) under a fixed transmit power constraint. For the LMMSE estimate, we can express the unknown channel matrix $\mathbf{H}$ as $\mathbf{H} = \hat{\mathbf{H}} + \widetilde{\mathbf{H}}$, where $\widetilde{\mathbf{H}}$ is the estimation error matrix. From the properties of LMMSE [30], we know that $\hat{\mathbf{H}}$ and $\widetilde{\mathbf{H}}$ are

uncorrelated and $\widetilde{\mathbf{H}}$ is Gaussian distributed $\widetilde{\mathbf{H}} \sim CN(0, \mathbf{\Sigma}_e \otimes \mathbf{\Psi}_e)$ (further details are provided below). Thus, (II.1) can be written as

$$r_a = \mathrm{Re}(\mathbf{w}_r^H \hat{\mathbf{H}} \mathbf{w}_t) d[a] + \mathrm{Re}(\mathbf{w}_r^H \widetilde{\mathbf{H}} \mathbf{w}_t) d[a] + \mathrm{Re}(\mathbf{w}_r^H \mathbf{n}_a). \tag{II.8}$$

Conditioned on $\hat{\mathbf{H}}$, the three terms in (II.8) are uncorrelated. We then state the following lemma:

Lemma 1 *The conditional output SNR as a function of* $\mathbf{w}_r$ *and* $\mathbf{w}_t$ *is given by*

$$\mathit{SNR}(\mathbf{w}_r, \mathbf{w}_t | \hat{\mathbf{H}}) = \frac{2(\mathrm{Re}(\mathbf{w}_r^H \hat{\mathbf{H}} \mathbf{w}_t))^2}{((\mathbf{w}_r^H \mathbf{\Sigma}_e \mathbf{w}_r)(\mathbf{w}_t^H \mathbf{\Psi}_e \mathbf{w}_t)) + N_o \mathbf{w}_r^H \mathbf{w}_r} \tag{II.9}$$

where the parameters $\mathbf{\Sigma}_e$ *and* $\mathbf{\Psi}_e$ *in* (II.9) *are given by*

$$\mathbf{\Sigma}_e = \begin{cases} \mathbf{\Sigma} - \frac{P_p q}{N_t N_o} \mathbf{\Sigma}(\mathbf{I} + \frac{P_p q}{N_t N_o} \mathbf{\Sigma})^{-1} \mathbf{\Sigma}, & \text{if } \mathbf{\Psi} = \mathbf{I} \\ \mathbf{I}, & \text{if } \mathbf{\Sigma} = \mathbf{I} \end{cases}$$

$$\text{and} \quad \mathbf{\Psi}_e = \begin{cases} \mathbf{I}, & \text{if } \mathbf{\Psi} = \mathbf{I} \\ \mathbf{\Psi} - \frac{P_p q}{N_t N_o} \mathbf{\Psi}(\mathbf{I} + \frac{P_p q}{N_t N_o} \mathbf{\Psi})^{-1} \mathbf{\Psi}, & \text{if } \mathbf{\Sigma} = \mathbf{I}. \end{cases} \tag{II.10}$$

Proof. The proof is presented in Appendix II.H.1. ■

It is interesting to contrast expression (II.9) with the corresponding expression when the channel is perfectly known. In the latter case, $\mathrm{SNR}(\mathbf{w}_r, \mathbf{w}_t | \mathbf{H}) = \frac{2(\mathrm{Re}(\mathbf{w}_r^H \mathbf{H} \mathbf{w}_t))^2}{N_o \mathbf{w}_r^H \mathbf{w}_r}$. Thus, there is an additional term in the denominator of (II.9) due to the channel estimation errors. $\mathbf{\Sigma}_e$ and $\mathbf{\Psi}_e$ in the first term in the denominator of (II.9) define the covariance of the channel estimation error matrix $\widetilde{\mathbf{H}}$. Our goal is to maximize (II.9) with respect to $\mathbf{w}_r$ and $\mathbf{w}_t$, such that $\mathbf{w}_t^H \mathbf{w}_t = P_o$. Thus, we consider the following optimization problem:

$$\max_{\mathbf{w}_r,\, \mathbf{w}_t} \quad \mathrm{SNR}(\mathbf{w}_r, \mathbf{w}_t | \hat{\mathbf{H}}) \qquad \text{such that } \ \mathbf{w}_t^H \mathbf{w}_t = P_o. \tag{II.11}$$

The above constrained optimization can be reduced to an equivalent unconstrained optimization problem in (II.12), where $\mathbf{w}_t$ is set equal to $\mathbf{w}_t = \sqrt{P_o}\, \mathbf{z} / \|\mathbf{z}\|$

(which defines $\mathbf{z}$) in the SNR expression, and the resulting SNR is now maximized with respect to $\mathbf{w}_r$ and $\mathbf{z}$.

$$\begin{aligned}\max_{\mathbf{w}_r,\,\mathbf{z}\neq 0} \quad & \mathrm{SNR}\left(\mathbf{w}_r, \tfrac{\sqrt{P_o}\mathbf{z}}{\|\mathbf{z}\|}|\hat{\mathbf{H}}\right) \\ &= \left(2\frac{P_o}{N_o}\right) \max_{\mathbf{w}_r,\,\mathbf{z}} \left(\frac{(\mathrm{Re}(\mathbf{w}_r^H\hat{\mathbf{H}}\,\mathbf{z}))^2}{((\mathbf{w}_r^H\boldsymbol{\Sigma}_e\mathbf{w}_r)(\mathbf{z}^H\boldsymbol{\Psi}_e\mathbf{z})\frac{P_o}{N_o}) + \mathbf{w}_r^H\mathbf{w}_r\,\|\mathbf{z}\|^2}\right). \qquad \text{(II.12)}\end{aligned}$$

We want to solve (II.12) for the special cases: $\boldsymbol{\Psi} = \mathbf{I}$ or $\boldsymbol{\Sigma} = \mathbf{I}$. Next we show that, under these special cases, (II.12) reduces to the maximization problem in Lemma 2.

Lemma 2 *Suppose we wish to obtain*

$$\max_{\mathbf{x},\,\mathbf{y}} \left(\frac{(\mathrm{Re}(\mathbf{x}^H\mathbf{A}\mathbf{y}))^2}{(\mathbf{x}^H\mathbf{B}\mathbf{x})\;(\mathbf{y}^H\mathbf{C}\mathbf{y})}\right) \qquad \text{(II.13)}$$

where $\mathbf{B}$*,* $p \times p$*, and* $\mathbf{C}$*,* $q \times q$*, are arbitrary Hermitian positive definite matrices and* $\mathbf{A}$ *is a* $p \times q$ *matrix. Then the maximum value in* (II.13) *is* $\lambda_{\max}(\mathbf{C}^{-1}\mathbf{A}^H\mathbf{B}^{-1}\mathbf{A})$ *and is obtained when* $\mathbf{x}^* = \alpha\mathbf{B}^{-1}\mathbf{A}\mathbf{y}^*$ *and* $\mathbf{y}^* = \mathbf{v}_{\max}(\mathbf{C}^{-1}\mathbf{A}^H\mathbf{B}^{-1}\mathbf{A})$*. Here,* $\mathbf{x}^*$ *and* $\mathbf{y}^*$ *stand for the optimum* $\mathbf{x}$ *and* $\mathbf{y}$ *respectively, and* α *is a real-valued scalar.* $\lambda_{\max}(\mathbf{D})$ *denotes the maximum eigenvalue of the matrix* $\mathbf{D}$*, and* $\mathbf{v}_{\max}(\mathbf{D})$ *denotes the eigenvector of* $\mathbf{D}$ *corresponding to* $\lambda_{\max}(\mathbf{D})$.

Proof. The proof is presented in Appendix II.H.2. ∎

When $\boldsymbol{\Psi} = \mathbf{I}$, (II.12) reduces to $(2\frac{P_o}{N_o}) \max_{\mathbf{w}_r,\,\mathbf{z}} \left(\frac{(\mathrm{Re}(\mathbf{w}_r^H\hat{\mathbf{H}}\,\mathbf{z}))^2}{(\mathbf{w}_r^H(\frac{P_o}{N_o}\boldsymbol{\Sigma}_e+\mathbf{I})\mathbf{w}_r)\;(\mathbf{z}^H\mathbf{z})}\right)$, which has the same form as (II.13) with the following substitution: $\mathbf{x} = \mathbf{w}_r$, $\mathbf{y} = \mathbf{z}$, $\mathbf{A} = \hat{\mathbf{H}}$, $\mathbf{B} = (\frac{P_o}{N_o}\boldsymbol{\Sigma}_e + \mathbf{I})$ and $\mathbf{C} = \mathbf{I}$. Thus, the optimum weights and the corresponding value for the maximum SNR are given by

$$\begin{aligned}\mathbf{w}_t &= \sqrt{P_o}\,\mathbf{z}/\|\mathbf{z}\| = \sqrt{P_o}\,\mathbf{v}_{\max}\left(\hat{\mathbf{H}}^H\left[\tfrac{P_o}{N_o}\boldsymbol{\Sigma}_e + \mathbf{I}\right]^{-1}\hat{\mathbf{H}}\right) \\ \mathbf{w}_r &= \alpha\mathbf{B}^{-1}\mathbf{A}\,\mathbf{z} = \left[\tfrac{P_o}{N_o}\boldsymbol{\Sigma}_e + \mathbf{I}\right]^{-1}\hat{\mathbf{H}}\,\mathbf{v}_{\max}\left(\hat{\mathbf{H}}^H\left[\tfrac{P_o}{N_o}\boldsymbol{\Sigma}_e + \mathbf{I}\right]^{-1}\hat{\mathbf{H}}\right) \qquad \text{(II.14)}\end{aligned}$$

$$\max \quad \mathrm{SNR}(\hat{\mathbf{H}}) = \left(2\frac{P_o}{N_o}\right)\lambda_{\max}\left(\hat{\mathbf{H}}^H\left[\tfrac{P_o}{N_o}\boldsymbol{\Sigma}_e + \mathbf{I}\right]^{-1}\hat{\mathbf{H}}\right). \qquad \text{(II.15)}$$

Since the matrix $\hat{\mathbf{H}}^H[\frac{P_o}{N_o}\boldsymbol{\Sigma}_e + \mathbf{I}]^{-1}\hat{\mathbf{H}}$ is Hermitian, the norm of the vector $\mathbf{z}$ is unity in the above.

Similarly, when $\mathbf{\Sigma} = \mathbf{I}$, (II.12) simplifies to $(2\frac{P_o}{N_o})\ \max\limits_{\mathbf{w}_r,\,\mathbf{z}} \left(\frac{(\mathrm{Re}(\mathbf{w}_r^H \hat{\mathbf{H}}\,\mathbf{z}))^2}{(\mathbf{w}_r^H \mathbf{w}_r)(\mathbf{z}^H[\frac{P_o}{N_o}\mathbf{\Psi}_e + \mathbf{I}]\mathbf{z})} \right)$. Substituting $\mathbf{x} = \mathbf{w}_r$, $\mathbf{y} = \mathbf{z}$, $\mathbf{A} = \hat{\mathbf{H}}$, $\mathbf{B} = \mathbf{I}$ and $\mathbf{C} = (\frac{P_o}{N_o}\mathbf{\Psi}_e + \mathbf{I})$, the optimum $\mathbf{w}_t$ and $\mathbf{w}_r$ and the maximum SNR for this case are

$$\begin{aligned} \mathbf{w}_t &= \sqrt{P_o}\,\mathbf{z}/\|\mathbf{z}\| = \sqrt{P_o}\,\frac{\mathbf{v}_{\max}\left(\left[\frac{P_o}{N_o}\mathbf{\Psi}_e + \mathbf{I}\right]^{-1}\hat{\mathbf{H}}^H\hat{\mathbf{H}}\right)}{\|\mathbf{v}_{\max}(\cdot)\|} \\ \mathbf{w}_r &= \alpha\mathbf{B}^{-1}\mathbf{A}\,\mathbf{z} = \hat{\mathbf{H}}\,\mathbf{v}_{\max}\left(\left[\frac{P_o}{N_o}\mathbf{\Psi}_e + \mathbf{I}\right]^{-1}\hat{\mathbf{H}}^H\hat{\mathbf{H}}\right) \end{aligned} \qquad \text{(II.16)}$$

$$\max \quad \mathrm{SNR}(\hat{\mathbf{H}}) = \left(2\frac{P_o}{N_o}\right)\ \lambda_{\max}\left(\left[\frac{P_o}{N_o}\mathbf{\Psi}_e + \mathbf{I}\right]^{-1}\hat{\mathbf{H}}^H\hat{\mathbf{H}}\right). \qquad \text{(II.17)}$$

The scalar α in the expressions for $\mathbf{w}_r$, (II.14) and (II.16), has been dropped, as it does not affect the output SNR. Upon examining the set of equations (II.14), (II.15), (II.16) and (II.17), we note their dependence on $\mathbf{\Psi}_e$ or $\mathbf{\Sigma}_e$, the channel estimation error covariance matrices. Also, recall that the channel estimate, $\hat{\mathbf{H}}$, is a Gaussian random matrix where its covariance, as given by (II.7), depends on the covariance structure at the transmitter or at the receiver. Dighe et.al [15] analyzed a special case when perfect knowledge of the channel is available and the entries of the MIMO channel matrix, $\mathbf{H}$, are iid. As can be readily verified, the above results can be reduced to the corresponding results in [15].

Mathematically, the two expressions, (II.15) and (II.17), for maximum SNR, are equivalent. This follows because the maximum SNR in each of the two cases is expressed in terms of the maximum eigenvalue of a matrix quadratic form, which is either correlated Wishart or pseudo-Wishart distributed [36], depending on whether the quadratic form is full rank or reduced rank, respectively. Thus, the maximum eigenvalue in the two cases can be shown to have the same distribution with the same set of parameters. The distribution parameters depend on the matrix $\mathbf{\Psi}_e$ in one case and on $\mathbf{\Sigma}_e$ in the other. Therefore, if we appropriately replace $\mathbf{\Psi}_e$ by $\mathbf{\Sigma}_e$ and vice versa, the two distributions would be the same. This also implies that the statistical properties of the maximum SNR, and therefore the average probability of error, in the two cases, are

alike. From a systems point-of-view, this means that with imperfect channel estimation the performance is identical if there is independent fading at, say L, transmitters (i.e., $\mathbf{\Psi} = \mathbf{I}$) and correlated fading at, say M, receivers, compared to the case when there is correlated fading at M transmitters and independent fading at L receivers (i.e., $\mathbf{\Sigma} = \mathbf{I}$). The equivalence of these two systems is a consequence of the reciprocity of the wireless channel. Thus, from now on, we only consider the case where there is independent fading at the transmitter and correlated fading at the receiver ($\mathbf{\Sigma}$ is an arbitrary covariance matrix).

In the following section, we present a closed-form expression for the probability of error and evaluate the transceiver's performance using this expression.

II.D Probability of Error

In the BPSK system with Gaussian statistics, the average probability of error is given by

$$P_e = \int_0^\infty Q(\sqrt{x}) f_{\lambda_{\max}}(x)\, \mathrm{d}x \tag{II.18}$$

where $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{y^2}{2}} \mathrm{d}y$ and $f_{\lambda_{\max}}(x)$ is the probability density function (pdf) for the maximum output SNR (see (II.15), (II.17)). In order to evaluate the integral in (II.18), we need to find an expression for $f_{\lambda_{\max}}(x)$. Considering the maximum SNR in (II.15), the exact closed-form expression for its pdf, $f_{\lambda_{\max}}(x)$, is given by Theorem 1 below.

We define an $N_r \times N_r$ matrix $\mathbf{\Phi}$ by

$$\mathbf{\Phi} \triangleq 2\frac{P_o}{N_o} [\frac{P_o}{N_o}\mathbf{\Sigma}_e + \mathbf{I}]^{-\frac{1}{2}} \bar{\mathbf{\Sigma}} [\frac{P_o}{N_o}\mathbf{\Sigma}_e + \mathbf{I}]^{-\frac{1}{2}}, \tag{II.19}$$

and denote its eigenvalues by $\{\frac{1}{\lambda_i}\}_{i=1}^{N_r}$. Throughout the rest of this chapter, we assume that the eigenvalues of the matrix $\mathbf{\Phi}$ are distinct and satisfy $0 < \frac{1}{\lambda_{N_r}} < \frac{1}{\lambda_{N_r-1}} < \cdots < \frac{1}{\lambda_1}$.

Theorem 1 *The expression for $f_{\lambda_{\max}}(x)$ is given by the following:*

$$\mathbf{A} : (N_t \geq N_r) \qquad f_{\lambda_{\max}}(x) = \tilde{a}\, x^{N_t - N_r} |\mathbf{G}(x)| \tag{II.20}$$

where $|\cdot|$ *denotes the determinant, and the* ij^{th} *element of the* $N_r \times N_r$, $\mathbf{G}(x)$ *is given by*

$$\begin{aligned}(\mathbf{G}(x))_{1,j} &= e^{-\lambda_j x} \\ (\mathbf{G}(x))_{i,j} &= x\,\lambda_j^{-N_t+i-1}\,\Gamma(N_t-i+1, x\,\lambda_j) \\ &\quad -\lambda_j^{-N_t+i-2}\,\Gamma(N_t-i+2, x\,\lambda_j)\end{aligned}$$

for $i = 2, \ldots, N_r,\ j = 1, \ldots, N_r$ *in the above, and*

$$\tilde{a} = \frac{(|\mathbf{\Phi}|)^{-N_t}}{\left[\prod_{i=1}^{N_r}\prod_{j=i+1}^{N_r}(\lambda_j - \lambda_i)\right]\left[\prod_{i=1}^{min(N_r,N_t)}\Gamma(N_t-i+1)\right]}. \tag{II.21}$$

$$\mathbf{B} : (N_t < N_r) \qquad f_{\lambda_{\max}}(x) = \tilde{a}\ |\mathbf{F}(x)| \tag{II.22}$$

where the ij^{th} *element of the* $N_r \times N_r$ *matrix* $\mathbf{F}(x)$ *is given by*

$$\begin{aligned}(\mathbf{F}(x))_{1,j} &= e^{-\lambda_j x} \\ (\mathbf{F}(x))_{i,j} &= x\,\lambda_j^{-N_t+i-1}\,\Gamma(N_t-i+1, x\,\lambda_j) \\ &\quad -\lambda_j^{-N_t+i-2}\,\Gamma(N_t-i+2, x\,\lambda_j) \quad \textit{for } i = 2, \ldots, N_t \\ (\mathbf{F}(x))_{i,j} &= \lambda_j^{i-N_t-1} \qquad \textit{for } i = N_t+1, \ldots, N_r\end{aligned}$$

$j = 1, \ldots, N_r$ *in all of the above and the constant* $\tilde{a}$ *is defined in* (II.21)*. The function,* $\Gamma(j+1, xa) = \int_0^{xa} e^{-y} y^j \,\mathrm{d}y$ *is the incomplete gamma function and* $\Gamma(j+1) = j!$ *is the gamma function [33].*

Proof. The proof is given in Appendix II.H.3. ∎

Remark. In [17], the authors provided without proof an expression for the cumulative distribution function (cdf) of the maximum eigenvalue of the correlated Wishart and pseudo-Wishart matrices. The expression for the density function presented in Theorem 1 can be obtained by direct differentiation of the cdf in [17]. This approach is adopted in Appendix II.H.3. In particular, in the proof of Theorem 1 in Appendix II.H.3, we first exploit the specific structure of the matrices $\mathbf{L}(x)$ and $\mathbf{M}(x)$ that appears in the corresponding cdf expressions, and are given in (II.33) and (II.34) of Appendix II.H.3, respectively, and then apply the simplified formula, (II.36) (also in Appendix II.H.3), for

the derivative of a determinant to obtain the desired densities. Alternatively, Theorem 1 can be proved by obtaining the marginal density of the largest eigenvalue, starting with the joint density of all the eigenvalues given in [39] and [36]. This latter approach is detailed in Appendix II.H.6.

Note that our approach in Theorem 1 is different from that of [34], [35], and results in much simpler pdf expressions that render themselves easily to further manipulations and, in particular, allow us to obtain a closed-form expression for the probability of error in Theorem 3.

Figure II.3 is a plot of $f_{\lambda_{\max}}(x)$, in Theorem 1, for the two cases of $N_t = 3, N_r = 7$ and $N_t = 7, N_r = 3$. Note that the two curves are different due to different covariance structures at the transmitter and the receiver. Here, we have assumed that independent fading exists at the transmitter and correlated fading at the receiver. For the case when both transmitter and the receiver experience independent fading, the two curves in Figure II.3 would coincide.

Next we proceed to evaluate the probability of error, P_e, given by (II.18), i.e., carry out the integration in (II.18). To this end, $f_{\lambda_{\max}}(x)$ from (II.20) and (II.22) must be simplified in a manner that the functional dependence on the variable x becomes explicit, as presented in Theorem 2.

Theorem 2 *The density $f_{\lambda_{\max}}(x)$ is given by*

$$f_{\lambda_{\max}}(x) \;=\; \sum_{j=1}^{min(N_r,N_t)} \binom{N_r}{j} \sum_{m=1} \sum_{l=(N_t-N_r)^+}^{jN_t-j^2} e^{-\eta_{j,m}(\lambda)\,x}\, x^l\, d_{j,m,l} \tag{II.23}$$

where $\eta_{j,m}(\lambda) \;\triangleq\; (\lambda_{u_{1,m}} + \lambda_{u_{2,m}} + \cdots + \lambda_{u_{j,m}})$ is a function of the eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_{N_r}$ and $(x)^+ = max(x, 0)$. The integer set $S_{1,j,m} = \{u_{1,m}, u_{2,m}, \cdots, u_{j,m}\}$ represents the m^{th} way of picking j distinct integers from the set $S_1 = \{1, 2, \cdots, N_r\}$,

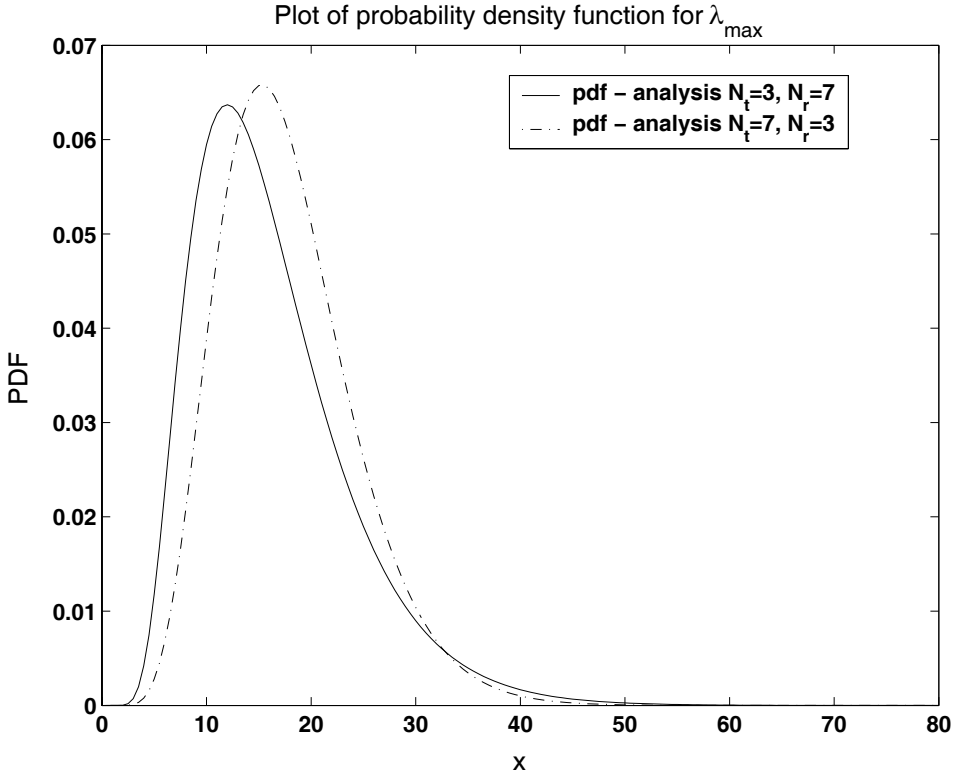

Figure II.3: Probability density function for $\lambda_{\max}$

with the total number of possible ways being $\binom{N_r}{j}$*. The constant* $d_{j,m,l}$ *is given by*

$$
\begin{aligned}
d_{j,m,l} = & \\
& \tilde{a} \sum_{i=1}^{\binom{Q-1}{j-1}} \sum\nolimits_1 \sum\nolimits_2 \operatorname{sign}(T) \\
& \times \overbrace{\sum_{k_1=0}^{1} \cdots \sum_{k_{Q-j}=0}^{1}}^{Q-j \text{ summations}} \overbrace{\sum_{p_1=0}^{N_t-v_{1,i}} \cdots \sum_{p_{j-1}=0}^{N_t-v_{j-1,i}}}^{j-1 \text{ summations}} \Big((a_{f_{1,i},t_{f_{1,i}},k_1}) \cdots \\
& \qquad {\scriptstyle k_1+\cdots+k_{Q-j}+p_1+\cdots+p_{j-1}=l-(N_t-N_r)^+} \\
& \times \cdots (a_{f_{Q-j,i},t_{f_{Q-j,i}},k_{Q-j}}).(\zeta_{v_{1,i},t_{v_{1,i}},p_1}) \cdots (\zeta_{v_{j-1,i},t_{v_{j-1,i}},p_{j-1}}) \Big) \left(\prod_{p=0}^{(N_r-N_t-1)^+} \lambda^p_{t_{p+N_t+1}} \right)
\end{aligned}
\tag{II.24}
$$

where $Q = min(N_r, N_t)$*,* $(v_{1,i}, \cdots, v_{j-1,i}, f_{1,i}, \cdots, f_{Q-j,i})$ *are* $Q-1$ *distinct integers from the set* $S_2 = \{2, 3, \cdots, Q\}$*. Index* i *denotes the* i*th way of partitioning the set* S_2

into two sets of size $(j-1)$ *and* $(Q-j)$.

$\sum_1$ *is the sum over all permutations* $(t_1, t_{v_{1,i}}, \cdots, t_{v_{j-1,i}})$ *of* $S_{1,j,m}$ *and* $\sum_2$ *is the sum over all permutations* $(t_{f_{1,i}}, \cdots, t_{f_{Q-j,i}})$ *of* $S_1 - S_{1,j,m}$*, if* $N_t \geq N_r$*, and sum over all permutations* $(t_{f_{1,i}}, \cdots, t_{f_{Q-j,i}}, t_{N_t+1}, \cdots, t_{N_r})$ *of* $S_1 - S_{1,j,m}$*, if* $N_t < N_r$*. Thus,* $T \triangleq (t_1, t_{v_{1,i}}, \cdots, t_{v_{j-1,i}}, t_{f_{1,i}}, \cdots \cdots, t_{f_{Q-j,i}}, t_{N_t+1}, \cdots, t_{N_r})$*, arranged with increasing values of the subscripts, are permutations of the set* S_1*. The function* $\mathrm{sign}(T)$ *is positive if this permutation is even and negative if it is odd [31]. Also,* $a_{i,j,0} \triangleq \xi_{i,j}$*,* $a_{i,j,1} \triangleq \theta_{i,j}$*,* $\zeta_{i,j,l}$ *are defined in* $(C.3)$*, and* $\tilde{a}$ *is given in* (II.21).

Using $f_{\lambda_{\max}}(x)$ from Theorem 2, we now present a closed-form expression for P_e in Theorem 3.

Theorem 3 *The probability of error is given by*

$$P_e = \sum_{j=1}^{min(N_r,N_t)} \sum_{m=1}^{\binom{N_r}{j}} \sum_{l=max(N_t-N_r,0)}^{jN_t-j^2} \frac{d_{j,m,l}\, l!}{2\, \eta_{j,m}(\lambda)^{l+1}} \times \left[1 - \frac{1}{\sqrt{1+2\eta_{j,m}(\lambda)}} \sum_{k=0}^{l} \binom{2k}{k} \left[\frac{\eta_{j,m}(\lambda)}{2+4\eta_{j,m}(\lambda)}\right]^k\right] \tag{II.25}$$

where $d_{j,m,l}$ *and* $\eta_{j,m}(\lambda)$ *are defined in Theorem 2.*

Theorem 2 and Theorem 3 are proved in Appendix II.H.4 and Appendix II.H.5, respectively.

The expression for $f_{\lambda_{\max}}(x)$ in Theorem 1 and 2, and the P_e expression in Theorem 3, are valid under the assumption that all the eigenvalues of the matrix $\mathbf{\Phi}$ are distinct. In the case when some of the eigenvalues, $\lambda_1, \lambda_2, \cdots, \lambda_{N_r}$, are equal, the approach in [37] can be adopted to provide corresponding *limiting* expressions. Specifically, it can be shown that when all the eigenvalues are equal to unity, our results reduce to the corresponding $f_{\lambda_{\max}}(x)$ and P_e expressions in [15].

II.E Numerical Results

We consider independent fading at the transmitter ($\mathbf{\Psi} = \mathbf{I}$) and correlated fading at the receiver. Using the correlation model of [28], the lm^{th} element of the receiver covariance matrix $\mathbf{\Sigma}$ is given by

$$(\mathbf{\Sigma})_{l,m} = \frac{I_o\left(\sqrt{\kappa^2 - \frac{4\pi^2\Delta_{l,m}^2}{\lambda^2} + \frac{j4\pi\kappa\sin(\mu)\Delta_{l,m}}{\lambda}}\right)}{I_o(\kappa)} \tag{II.26}$$

where $l, m = 1, 2, \ldots, N_r$, $I_o(.)$ is the modified Bessel function of the first kind and zeroth order, $\kappa\,(\geq 0)$ controls the width of the angle of arrival (AOA) (or, in other words, the angular spread) and ranges from 0 (isotropic scattering) to infinity (non-isotropic scattering), the parameter $\mu(\in [-\pi, \pi))$ accounts for the mean direction of the AOA, and $\Delta_{l,m}$ is the spacing between the l^{th} and m^{th} antenna elements. In (II.26), we notice that there are three parameters controlling the correlation: κ, $\Delta_{l,m}$ and μ. In all our results, we will concentrate on a broadside array with $\mu = 0$. Correlation increases as κ is increased, and/or $\Delta_{l,m}$ is decreased.

At the receiver, we consider a physically constrained linear array where the effective aperture of the array is fixed. For such an antenna array, increasing the number of receiver antenna elements would lead to reduced inter-element spacing. This, in turn, would cause higher correlation between the channel coefficients at each element. Also, for a fixed aperture array, the total power collected by the array is constant, irrespective of the number of elements [75]. For such an antenna array, an increase in the number of the elements is brought about by decreasing the size of each element [78]. In turn, the power captured by each element decreases with the decrease in its effective aperture. Examples of such an array are a collinear array of dipoles and an array of microstrip antennas [42]. In our results, the total length of the antenna array is constrained to λ (the carrier wavelength). To capture the effect of fixed received power in our analysis, we introduce an additional factor, say β, in the signal term in Equations (II.1) and (II.2). The factor β can be conveniently absorbed in the signal power P_o and the pilot power P_p, and can be thought of as a normalizing constant, i.e., βP_o and βP_p, respectively. With

$\beta = 1/N_r$, the total received power remains fixed. Lastly, for simplicity, we ignore the effects of mutual coupling between the antenna elements [75], [42].

Thus, for a fixed aperture antenna array, the power received by each antenna element reduces as more antenna elements are added to the array. This effectively means that the SNR-per-element decreases. Therefore, during the channel estimation phase, with a larger number of antenna elements, there are more parameters to estimate and less power available to make each estimate, resulting in higher channel estimation noise. Also, the non-coherent combining loss increases when a larger number of these noisy estimates are combined.

The quality of the channel estimate is determined by both the fade correlation and the received pilot power-per-antenna element. We evaluate the impact of these factors on the channel estimation with the help of an example. Consider the MMSE for the estimator, i.e., $e_{\text{MMSE}} = tr(\mathbf{C}_{\tilde{\mathbf{h}}\tilde{\mathbf{h}}}) = tr(\mathbf{\Sigma}_e \otimes \mathbf{\Psi}_e) = tr(\mathbf{\Sigma}_e)tr(\mathbf{\Psi}_e)$, where e_{MMSE} denotes the MMSE and $\mathbf{C}_{\tilde{\mathbf{h}}\tilde{\mathbf{h}}}$ is given in (II.29). For the case $\mathbf{\Psi} = \mathbf{I}$, the MMSE is given by $e_{\text{MMSE}} = N_t\, tr(\mathbf{\Sigma}_e)$. Substituting for $\mathbf{\Sigma}_e$, and denoting $\mathbf{U\Omega U}^H$ as the eigenvalue decomposition of $\mathbf{\Sigma}$, where $\mathbf{\Omega} \triangleq \text{diag}(\omega_1, \cdots, \omega_{N_r})$ and $\{\omega_i\}_{i=1}^{N_r}$ are the eigenvalues, e_{MMSE} can be simplified as

$$\begin{aligned} e_{\text{MMSE}} &= N_t\, tr\left(\mathbf{\Omega} - \gamma\mathbf{\Omega}\left(\gamma\mathbf{\Omega} + qN_o\mathbf{I}\right)^{-1}\mathbf{\Omega}\right) \\ &= N_t \sum_{i=1}^{N_r}\left(\omega_i - \frac{\omega_i^2\rho}{1+\rho\,\omega_i}\right) \\ &= N_t \sum_{i=1}^{N_r}\left(\frac{\omega_i}{1+\rho\,\omega_i}\right) \end{aligned} \tag{II.27}$$

where $\rho \triangleq \frac{\gamma}{qN_o}$. Substituting for γ, and replacing P_p by βP_p, we obtain $\rho = \frac{\beta P_p q}{N_t N_o} = \frac{SNR_p}{N_r N_t}$, where $SNR_p \triangleq \frac{P_p q}{N_o}$ is the pilot SNR, and β is set to $1/N_r$. Now consider the two extreme cases, when receive fades are uncorrelated ($\mathbf{\Sigma} = \mathbf{I}$) i.e., $\{\omega_i\}_{i=1}^{N_r}$ are all unity, and the case when receive fades are perfectly correlated i.e., w.l.o.g, $\omega_1 = N_r$ and $\{\omega_i\}_{i=2}^{N_r} = 0$ (note that $\sum_{i=1}^{N_r}\omega_i = tr(\mathbf{\Sigma}) = N_r$). For the former case, (II.27) reduces to $e_{\text{MMSE}}^{\text{uncorr}} = (\frac{N_t N_r}{1+\rho}) = (\frac{N_t N_r}{1+\frac{SNR_p}{N_r N_t}})$, whereas for the latter case it becomes

$e^{\text{corr}}_{\text{MMSE}} = (\frac{N_t N_r}{1+N_r\rho}) = (\frac{N_t N_r}{1+\frac{SNR_p}{N_t}})$. Note that the MMSE is lower for the perfect correlation case. Thus, for a given N_r, as correlation is increased by reducing the inter-element spacing and/or reducing the scattering, the MMSE reduces. This implies that an increase in correlation between the channel fades makes channel estimation easier. However, note that both these expressions are increasing in N_r and N_t. It can be shown that $e^{\text{uncorr}}_{\text{MMSE}}$ and $e^{\text{corr}}_{\text{MMSE}}$, respectively, upper and lower bound the MMSE (i.e., e_{MMSE}). Furthermore, notice that the term SNR_p/N_r, which is the pilot SNR-per-element, appears in the denominator of $e^{\text{uncorr}}_{\text{MMSE}}$. Thus, for the uncorrelated fades, the estimator loses performance because of the decreasing pilot SNR-per-element as N_r is increased.

The performance of the system is governed by the net effect of the available diversity and the channel estimation errors. First, we investigate the performance of the channel estimator. Figure II.4 is a plot of the MMSE normalized by the number of transmit antennas (e_{MMSE}/N_t), as a function of κ, and for different N_r. The number of transmitters is fixed to $N_t = 2$, the number of receivers is varied from $N_r = 1$ to 6, $\mu = 0$, and the pilot SNR is 20dB. Also plotted in the figure is $e^{\text{uncorr}}_{\text{MMSE}}/N_t$, and is shown by the curves without any markers. There is one such curve in each of the six groups for the six N_r values. Note that the curves for $e^{\text{uncorr}}_{\text{MMSE}}$ are not a function of κ, because, by assumption, they always correspond to uncorrelated fading. Except for the curves $N_r = 1$ and $N_r = 2$, the MMSE (i.e., e_{MMSE}/N_t) decreases as κ is increased. This is because a higher value of κ (poor scattering) leads to higher correlation, which aids in channel estimation. For the case $N_r = 2$, the receiver antenna elements are spaced far enough apart so that the correlation does not change much for the observed values of κ. The case $N_r = 1$ corresponds to a single receiver antenna, thus, κ does not affect its performance according to the model in (II.26). Note that both MMSE and $e^{\text{uncorr}}_{\text{MMSE}}/N_t$ increase with N_r, but the MMSE is lower than the corresponding $e^{\text{uncorr}}_{\text{MMSE}}/N_t$ for higher values of N_r ($N_r \geq 3$). For $N_r < 3$, the curves corresponding to the MMSE and $e^{\text{uncorr}}_{\text{MMSE}}/N_t$ are almost indistinguishable because the fade correlation is sufficiently low. Thus, as correlation is increased by reducing the inter-element spacing and/or reducing scattering, the MMSE value decreases when compared to $e^{\text{uncorr}}_{\text{MMSE}}/N_t$.

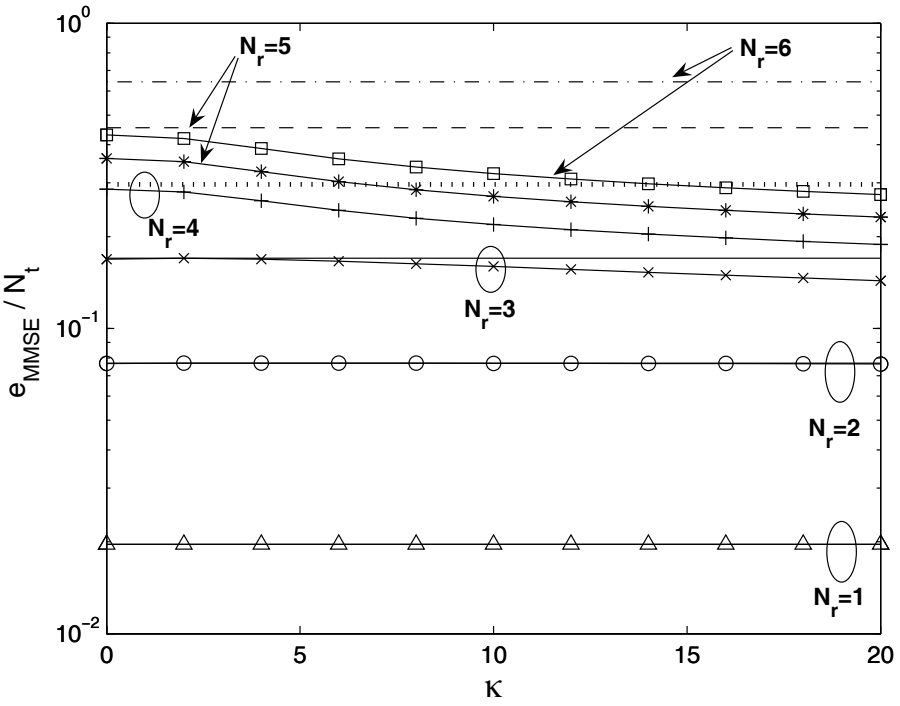

Figure II.4: Minimum mean-square error normalized by the number of transmitters vs. κ for different number of receivers (N_r), $N_t = 2$, pilot SNR is 20dB, and $\mu = 0$. The MMSE (i.e., e_{MMSE}/N_t) is given by the curves with markers. Also plotted in the figure is $e_{\text{MMSE}}^{\text{uncorr}}/N_t$, and is shown by the curves without any markers. There is one such curve in each of the six groups for the six N_r values.

Figure II.5 is a plot of probability of error, (II.25), versus the number of receiver antennas N_r, when the data SNR and the pilot SNR are both 20dB, N_t is varied from 1 to 2, and $\kappa = 2$. Observe the convexity of the curves and the existence of an optimal value for N_r. For the given parameters, $N_r = 4$ is optimal for the two curves. Note that the performance of the system is determined by the net effect of the available diversity and the channel estimation errors, as mentioned earlier. When N_r is increased from 1 to 4 in Figure II.5, the diversity order of the system increases initially, whereas the performance of the channel estimator degrades (e_{MMSE} increases). Clearly, the effect of diversity is winning over the effect of channel estimation, as the overall performance improves. Any further increase in N_r provides only a marginal increase in diversity,

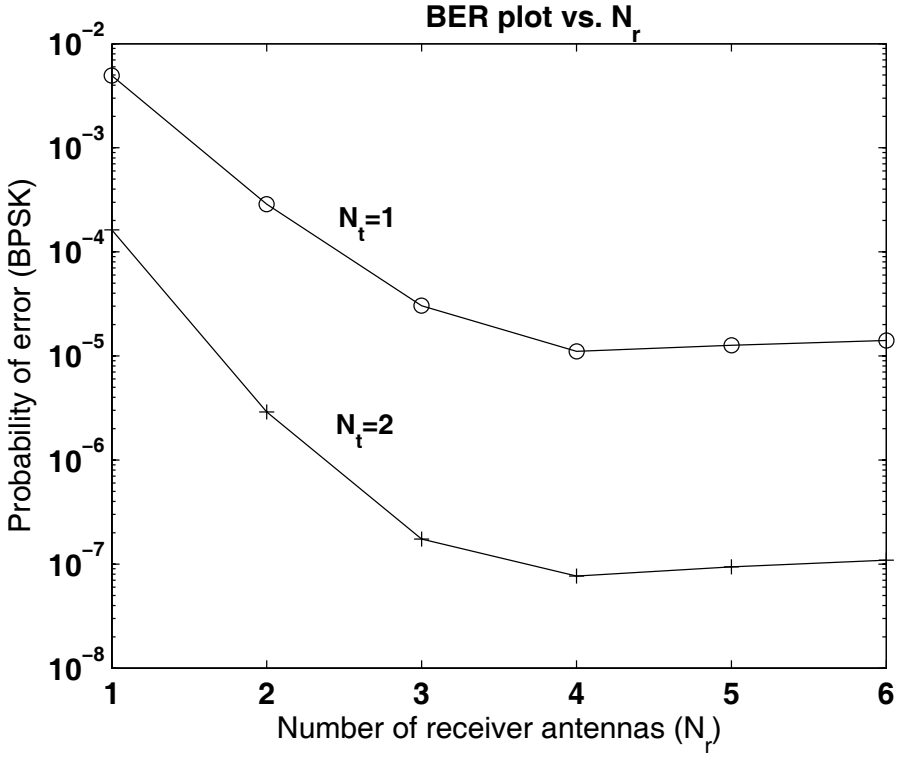

Figure II.5: BER vs number of receivers (N_r) for fixed $\kappa(= 2)$, pilot and data SNR are 20dB, and $\mu = 0$

and a continued increase in the estimation noise (as the overall MMSE increases), resulting in an increase in the bit error rate. Note also, from Figure II.5, that the system with $N_t = 2$ performs better than the one with $N_t = 1$ due to higher transmit diversity, although the channel estimates are noisier for $N_t = 2$ than for $N_t = 1$. This is because the total transmit power in the two systems is the same and there are more parameters to estimate when $N_t = 2$ than when $N_t = 1$, leading to higher estimation noise in the former than the latter.

Figure II.6 is a bit error rate plot for the same set of parameters as in Figure II.5, except that the pilot and data SNR are now lowered to 9dB. Note that the optimal N_r for each of the two N_t values is different, i.e., $N_r = 2$ is optimal for $N_t = 2$, whereas $N_r = 3$ is optimal for $N_t = 1$. A lower value of N_r is optimal for $N_t = 2$ due to higher estimation noise in this case when compared to $N_t = 1$.

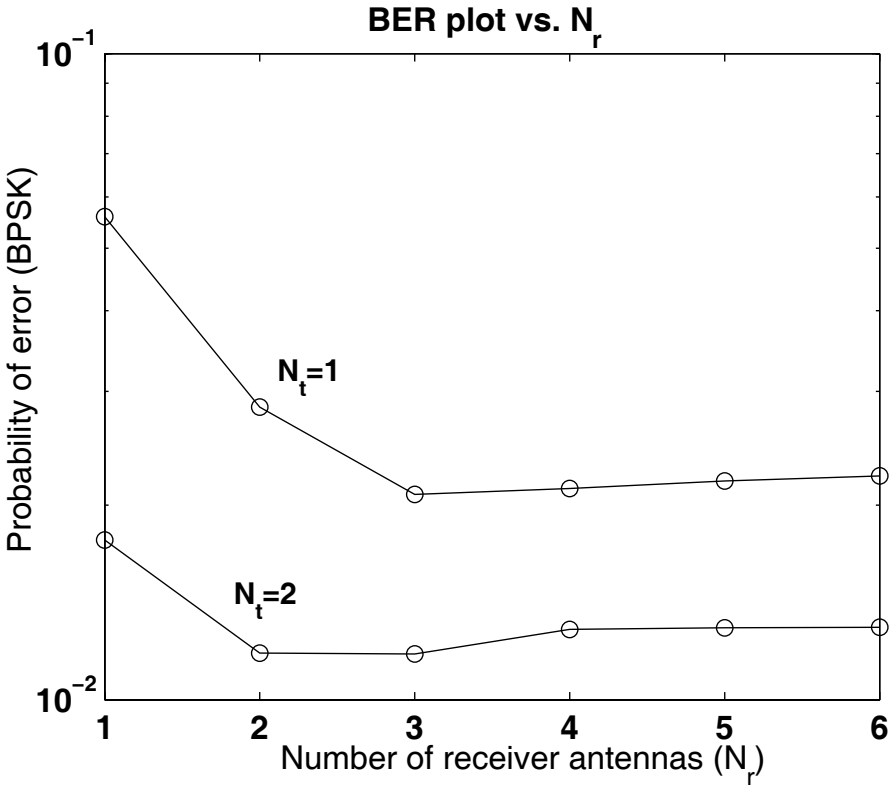

Figure II.6: BER vs number of receivers (N_r) for $\kappa = 2$, $\mu = 0$, and pilot and data SNR are 9dB

Figure II.7 is a plot of the probability of bit error for various κ (the extent of scattering). We fix $N_t = 2$ and vary N_r from 1 to 6, while maintaining the data and pilot SNRs at 20dB. Higher values of κ (reduced scattering) result in lower diversity and higher channel estimation accuracy due to increased correlation between the receive fades. Note from Figure II.7 that the performance of the system degrades as κ increases. This implies that the loss in diversity order outweighs the gain in channel estimation accuracy as κ is increased. The plot for $N_t = 2$ and $N_r = 2$ is unaffected by varying κ, as the receiver antenna elements are spaced far apart.

Observing the curves in Figure II.7 along any vertical line, corresponding to a fixed value of κ, reveals the same trends as in Figure II.5, i.e., the existence of an optimal N_r. As an example, the curve corresponding to $N_t = 2$ in Figure II.5 can be obtained from Figure II.7 for $\kappa = 2$. Also, it is interesting to note the crossovers by the curve

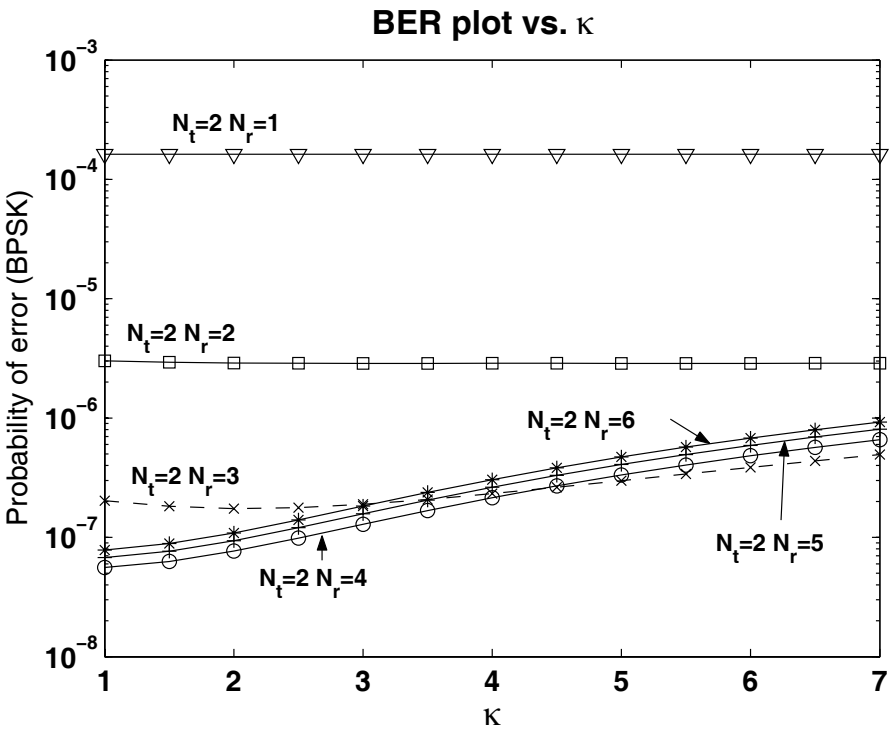

Figure II.7: BER vs. κ for different number of receivers (N_r), $N_t = 2$, pilot and data SNR are 20dB, and $\mu = 0$

corresponding to $N_t = 2$, $N_r = 3$, as κ is increased. For lower κ, the configuration $N_t = 2$, $N_r = 4$ performs better than $N_t = 2$, $N_r = 3$ because the increase in diversity order more than compensates for the increase in the non-coherent combining loss in going from $N_t = 2$, $N_r = 3$ to $N_t = 2$, $N_r = 4$. Thus, the optimum N_r depends on the value of κ and the data and the pilot SNRs.

Finally, in Figure II.8, bit error rate is plotted versus κ, where the product of the number of transmitters and receivers is kept constant. This fixes the maximum diversity of the system in the two cases, where the product is either 12 or 16 in the figure. Since the transmit fades are assumed uncorrelated, it is evident that the system with more transmitters leads to better performance when the product $N_t N_r$ is fixed, due to higher diversity.

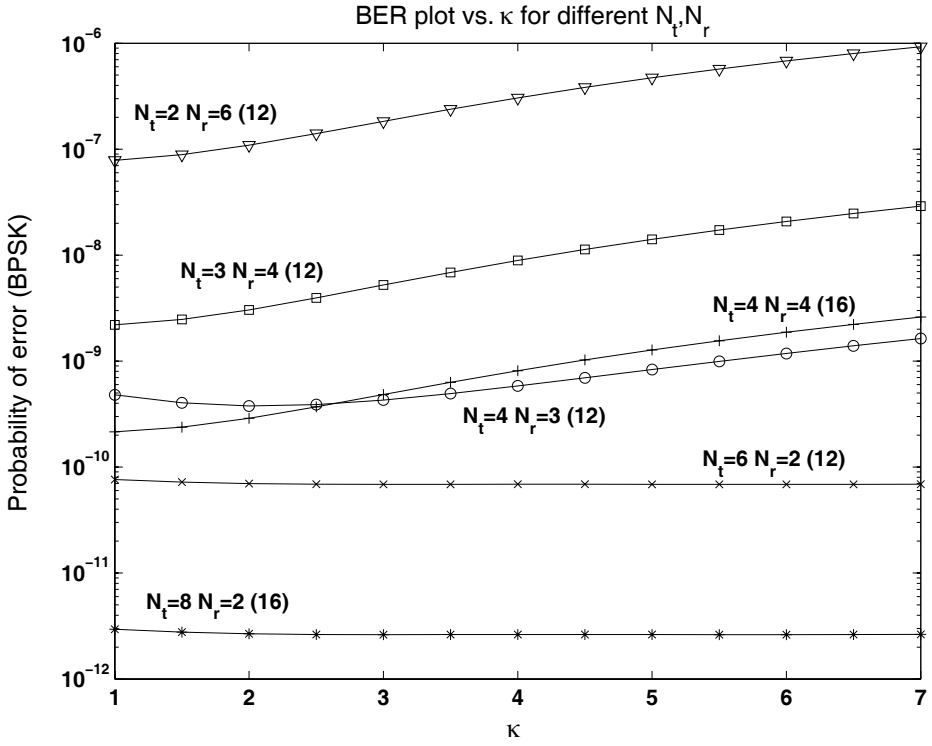

Figure II.8: BER vs. κ when the product of number of transmitters and receivers, $N_t N_r$, is fixed, pilot and data SNR are 20dB, and $\mu = 0$

II.F Conclusion

In this work, the error rate performance of a MIMO beamforming system in correlated Rayleigh fading was analyzed under the conditions of imperfect channel estimation. An optimum beamforming structure and an exact closed-form expression for the probability of error of the system were derived. The impact of channel estimation accuracy and diversity on the performance of the system was studied. It was shown that performance degrades as the extent of scattering decreases (i.e., the width of the angle-of-arrival decreases). The increased fade correlation due to poor scattering leads to improved channel estimation on one hand, but lower diversity on the other. The performance degrades because the benefits of improved channel estimation accuracy are offset by the loss in diversity order. For a fixed length antenna array where the effective aperture of the array is fixed, it was also shown that there is an optimum number of re-

ceiver antennas that depends on the available transmit and receive diversity, the quality of the estimates, and the length of the array.

II.G Acknowledgement

This chapter has been published, in part, in the following publications:

1. P. Nagvanshi, E. Masry and L.B. Milstein, "The effect of channel estimation errors on transceiver optimization for correlated MIMO Rayleigh channels," *Proc. 39th Annual Conference on Information Sciences and Systems*, March 2005.

2. P. Nagvanshi, E. Masry and L.B. Milstein, "Optimum transmit-receive beamforming with noisy channel estimates for correlated MIMO Rayleigh channels," accepted for publication in the *IEEE Transactions on Communications.*

The dissertation author was the primary researcher and author, and the co-authors listed in these publications directed and supervised the research which forms the basis for this chapter.

II.H Appendices for Chapter II

II.H.1 Proof of Lemma 1

From (II.8), the conditional output SNR is defined as

$$\text{SNR}(\mathbf{w}_r, \mathbf{w}_t|\hat{\mathbf{H}}) \triangleq \frac{(\text{Re}(\mathbf{w}_r^H \hat{\mathbf{H}} \mathbf{w}_t))^2}{E[(\text{Re}(\mathbf{w}_r^H \widetilde{\mathbf{H}} \mathbf{w}_t))^2|\hat{\mathbf{H}}] + E[(\text{Re}(\mathbf{w}_r^H \mathbf{n}_a))^2|\hat{\mathbf{H}}]}. \quad \text{(II.28)}$$

We denote the two terms in the denominator of (II.28) as Var1 $\triangleq E[(\text{Re}(\mathbf{w}_r^H \widetilde{\mathbf{H}} \mathbf{w}_t))^2|\hat{\mathbf{H}}]$ and Var2 $\triangleq E[(\text{Re}(\mathbf{w}_r^H \mathbf{n}_a))^2|\hat{\mathbf{H}}]$. In order to obtain a simplified expression for Var1 and Var2, we first note that $\tilde{\mathbf{h}} \sim CN(0, \mathbf{\Sigma}_e \otimes \mathbf{\Psi}_e)$. This follows because the estimation error vector $\tilde{\mathbf{h}}(= \mathbf{h} - \hat{\mathbf{h}})$ is Gaussian with zero mean and covariance given by $\mathbf{C}_{\tilde{\mathbf{h}}\tilde{\mathbf{h}}} = \mathbf{C}_{\mathbf{hh}} - \mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}$. Using the expression for $\mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}$ in Section II.B, $\mathbf{C}_{\tilde{\mathbf{h}}\tilde{\mathbf{h}}}$ simplifies to

$$\begin{aligned} \mathbf{C}_{\tilde{\mathbf{h}}\tilde{\mathbf{h}}} &= \mathbf{\Sigma} \otimes \mathbf{\Psi} - (\sqrt{\gamma}\mathbf{\Sigma} \otimes \mathbf{\Psi})(\gamma \mathbf{\Sigma} \otimes \mathbf{\Psi} + qN_o \mathbf{I} \otimes \mathbf{I})^{-1}(\sqrt{\gamma}\mathbf{\Sigma} \otimes \mathbf{\Psi}) \\ &= \begin{cases} (\mathbf{\Sigma} - \gamma \mathbf{\Sigma}(\gamma \mathbf{\Sigma} + qN_o \mathbf{I})^{-1}\mathbf{\Sigma}) \otimes \mathbf{I}, & \text{if } \mathbf{\Psi} = \mathbf{I} \\ \mathbf{I} \otimes (\mathbf{\Psi} - \gamma \mathbf{\Psi}(\gamma \mathbf{\Psi} + qN_o \mathbf{I})^{-1}\mathbf{\Psi}), & \text{if } \mathbf{\Sigma} = \mathbf{I}. \end{cases} \end{aligned} \quad \text{(II.29)}$$

For compactness, we can express $\mathbf{C}_{\tilde{\mathbf{h}}\tilde{\mathbf{h}}} = E[\text{vec}(\widetilde{\mathbf{H}}^T)\text{vec}(\widetilde{\mathbf{H}}^T)^H] = \mathbf{\Sigma}_e \otimes \mathbf{\Psi}_e$, where $\mathbf{\Sigma}_e$ and $\mathbf{\Psi}_e$ are defined in Lemma 1. The variance of the linear transform $\mathbf{w}_r^H \widetilde{\mathbf{H}} \mathbf{w}_t$ of $\widetilde{\mathbf{H}}$ can now be obtained as

$$\begin{aligned} C_{\mathbf{w}_r^H \tilde{\mathbf{H}} \mathbf{w}_t} &\triangleq E[\mathbf{w}_r^H \widetilde{\mathbf{H}} \mathbf{w}_t \mathbf{w}_t^H \widetilde{\mathbf{H}}^H \mathbf{w}_r|\hat{\mathbf{H}}] \\ &= \mathbf{w}_r^H (\mathbf{w}_t^T \otimes \mathbf{I}) \mathbf{K}_{N_t N_r} E[\text{vec}(\widetilde{\mathbf{H}}^T)\text{vec}(\widetilde{\mathbf{H}}^T)^H|\hat{\mathbf{H}}] \mathbf{K}_{N_r N_t} (\mathbf{w}_t^* \otimes \mathbf{I}) \mathbf{w}_r \\ &= \mathbf{w}_r^H (\mathbf{w}_t^T \otimes \mathbf{I})(\mathbf{\Psi}_e \otimes \mathbf{\Sigma}_e)(\mathbf{w}_t^* \otimes \mathbf{I}) \mathbf{w}_r \\ &= (\mathbf{w}_t^H \mathbf{\Psi}_e \mathbf{w}_t)(\mathbf{w}_r^H \mathbf{\Sigma}_e \mathbf{w}_r). \end{aligned} \quad \text{(II.30)}$$

The second equality in (II.30) follows since $\widetilde{\mathbf{H}} \mathbf{w}_t = \text{vec}(\widetilde{\mathbf{H}} \mathbf{w}_t) = (\mathbf{w}_t^T \otimes \mathbf{I})\text{vec}(\widetilde{\mathbf{H}}) = (\mathbf{w}_t^T \otimes \mathbf{I}) \mathbf{K}_{N_t N_r} \text{vec}(\widetilde{\mathbf{H}}^T)$ and $\mathbf{K}_{N_r N_t} = \mathbf{K}_{N_t N_r}^T$, where $\mathbf{K}_{N_t N_r}$ is a commutation matrix [31]. In the third equality in (II.30) we have used $\mathbf{K}_{N_t N_r}(\mathbf{\Sigma}_e \otimes \mathbf{\Psi}_e)\mathbf{K}_{N_r N_t} = (\mathbf{\Psi}_e \otimes \mathbf{\Sigma}_e)$. From [30], we have Var1 $= C_{\text{Re}(\mathbf{w}_r^H \tilde{\mathbf{H}} \mathbf{w}_t)} = \frac{1}{2} C_{\mathbf{w}_r^H \tilde{\mathbf{H}} \mathbf{w}_t} = \frac{1}{2}(\mathbf{w}_t^H \mathbf{\Psi}_e \mathbf{w}_t)(\mathbf{w}_r^H \mathbf{\Sigma}_e \mathbf{w}_r)$. Finally, Var2 can be simplified as Var2 $= C_{\text{Re}(\mathbf{w}_r^H \mathbf{n}_a)} = \frac{1}{2} C_{\mathbf{w}_r^H \mathbf{n}_a} = \frac{1}{2} N_o \mathbf{w}_r^H \mathbf{w}_r$.

II.H.2 Proof of Lemma 2

Without loss of generality, we define $\tilde{\mathbf{x}} \stackrel{\Delta}{=} \mathbf{B}^{1/2}\mathbf{x}$ and $\tilde{\mathbf{y}} \stackrel{\Delta}{=} \mathbf{C}^{1/2}\mathbf{y}$ and simplify (II.13) as

$$\max_{\tilde{\mathbf{x}},\,\tilde{\mathbf{y}}} \left(\frac{(\mathrm{Re}(\tilde{\mathbf{x}}^H \mathbf{B}^{-1/2}\mathbf{A}\mathbf{C}^{-1/2}\tilde{\mathbf{y}}))^2}{(\tilde{\mathbf{x}}^H\tilde{\mathbf{x}})\,(\tilde{\mathbf{y}}^H\tilde{\mathbf{y}})} \right). \tag{II.31}$$

The function of $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{y}}$ that we are trying to maximize in (II.31) can be upper bounded as shown below:

$$\begin{aligned} \tfrac{(\mathrm{Re}(\tilde{\mathbf{x}}^H \mathbf{B}^{-1/2}\mathbf{A}\mathbf{C}^{-1/2}\tilde{\mathbf{y}}))^2}{(\tilde{\mathbf{x}}^H\tilde{\mathbf{x}})\,(\tilde{\mathbf{y}}^H\tilde{\mathbf{y}})} &= \tfrac{(\mathrm{Re}(\tilde{\mathbf{x}}^H\mathbf{a}))^2}{\tilde{\mathbf{x}}^H\tilde{\mathbf{x}}} \stackrel{(i)}{\leqslant} \mathbf{a}^H\mathbf{a} \\ \mathbf{a}^H\mathbf{a} = \tfrac{\tilde{\mathbf{y}}^H\mathbf{C}^{-1/2}\mathbf{A}^H\mathbf{B}^{-1}\mathbf{A}\mathbf{C}^{-1/2}\tilde{\mathbf{y}}}{\|\tilde{\mathbf{y}}\|^2} &\stackrel{(ii)}{\leqslant} \lambda_{\max}(\mathbf{C}^{-1/2}\mathbf{A}^H\mathbf{B}^{-1}\mathbf{A}\mathbf{C}^{-1/2}) \end{aligned} \tag{II.32}$$

where $\mathbf{a} \stackrel{\Delta}{=} \mathbf{B}^{-1/2}\mathbf{A}\mathbf{C}^{-1/2}\frac{\tilde{\mathbf{y}}}{\|\tilde{\mathbf{y}}\|}$. Inequality (i), above, follows from the Cauchy-Schwarz inequality [32, eq. 5.1.5] and inequality (ii) follows from the Rayleigh-Ritz Theorem [32, page 176]. From (II.32), it is evident that the maximum value of the function in (II.31) is $\lambda_{\max}(\mathbf{C}^{-1/2}\mathbf{A}^H\mathbf{B}^{-1}\mathbf{A}\mathbf{C}^{-1/2})$, or equivalently $\lambda_{\max}(\mathbf{C}^{-1}\mathbf{A}^H\mathbf{B}^{-1}\mathbf{A})$, and is obtained when the weights $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{y}}$ are chosen such that the inequalities in (i) and (ii), above, are satisfied with an equality.

The equality in (i) holds iff $\tilde{\mathbf{x}} = \alpha\,\mathbf{a}$, [32, page 261]. This implies that, $\tilde{\mathbf{x}} = \alpha\mathbf{B}^{-1/2}\mathbf{A}\mathbf{C}^{-1/2}\frac{\tilde{\mathbf{y}}}{\|\tilde{\mathbf{y}}\|}$. Thus, the optimum $\mathbf{x}$ is $\mathbf{x} = \alpha\mathbf{B}^{-1}\mathbf{A}\mathbf{C}^{-1/2}\frac{\tilde{\mathbf{y}}}{\|\tilde{\mathbf{y}}\|} \Rightarrow \mathbf{x} = \alpha\mathbf{B}^{-1}\mathbf{A}\frac{\mathbf{y}}{\|\tilde{\mathbf{y}}\|}$. On the other hand, the equality in (ii) is satisfied if $\tilde{\mathbf{y}}$ is the eigenvector corresponding to the eigenvalue $\lambda_{\max}(\mathbf{C}^{-1/2}\mathbf{A}^H\mathbf{B}^{-1}\mathbf{A}\mathbf{C}^{-1/2})$, i.e., $\tilde{\mathbf{y}} = \mathbf{v}_{\max}(\mathbf{C}^{-1/2}\mathbf{A}^H\mathbf{B}^{-1}\mathbf{A}\mathbf{C}^{-1/2})$.

Since the matrix $\mathbf{C}^{-1/2}\mathbf{A}^H\mathbf{B}^{-1}\mathbf{A}\mathbf{C}^{-1/2}$ is Hermitian, the norm of the vector $\mathbf{v}_{\max}(\mathbf{C}^{-1/2}\mathbf{A}^H\mathbf{B}^{-1}\mathbf{A}\mathbf{C}^{-1/2})$ is unity, i.e., $\|\tilde{\mathbf{y}}\| = 1$. This implies that optimum $\mathbf{x}$ in terms of optimum $\mathbf{y}$ is given by $\mathbf{x}^* = \alpha\mathbf{B}^{-1}\mathbf{A}\mathbf{y}^*$. The optimum $\mathbf{y}$ is $\mathbf{y}^* = \mathbf{C}^{-1/2}\mathbf{v}_{\max}(\mathbf{C}^{-1/2}\mathbf{A}^H\mathbf{B}^{-1}\mathbf{A}\mathbf{C}^{-1/2})$, simplifying using matrix algebra [31], $\mathbf{y}^* = \mathbf{v}_{\max}(\mathbf{C}^{-1}\mathbf{A}^H\mathbf{B}^{-1}\mathbf{A})$.

II.H.3 Proof of Theorem 1

Using the notation developed so far, it can be shown that the cdf in [17] can be written as

$$1)\ \ (N_t \geq N_r) \qquad F_{\lambda_{\max}}(x) = \tilde{a}\,|\mathbf{L}(x)| \tag{II.33}$$

where the ij^{th} $(i,j = 1,2,\ldots,N_r)$ element of the $N_r \times N_r$ matrix $\mathbf{L}(x)$ is given by $(\mathbf{L}(x))_{i,j} = \lambda_j^{-N_t+i-1}\,\Gamma(N_t - i + 1, x\,\lambda_j)$.

$$2)\ \ (N_t < N_r) \qquad F_{\lambda_{\max}}(x) = \tilde{a}\ |\mathbf{M}(x)| \tag{II.34}$$

where the ij^{th} $(i,j = 1,2,\ldots,N_r)$ element of the $N_r \times N_r$ matrix $\mathbf{M}(x)$ is given by $(\mathbf{M}(x))_{i,j} = \lambda_j^{-N_t+i-1}\,\Gamma(N_t - i + 1, x\,\lambda_j)$ for $i = 1,2,\ldots,N_t$ and $(\mathbf{M}(x))_{i,j} = \lambda_j^{i-N_t-1}$ for $i = N_t+1,\ldots,N_r$.

The constant $\tilde{a}$ in the above is the same as in Theorem 1. In order to differentiate (II.33) and (II.34), we use the following formula for the derivative of a determinant [38]:

$$\begin{aligned}\frac{\mathrm{d}\,|\mathbf{D}(x)|}{\mathrm{d}\,x} &= |\mathbf{D}(x)|\ \left(\mathbf{D}(x)^{-1}\ \frac{\mathrm{d}\,\mathbf{D}(x)}{\mathrm{d}\,x}\right)\\ &= trace\left(\mathbf{D}^{\text{Adj}}(x)\ \frac{\mathrm{d}\,\mathbf{D}(x)}{\mathrm{d}\,x}\right)\end{aligned} \tag{II.35}$$

since $|\mathbf{D}(x)|\mathbf{D}(x)^{-1} = \mathbf{D}^{\text{Adj}}(x)$, where $\mathbf{D}^{\text{Adj}}(x)$ is the adjoint of an $N_r \times N_r$ matrix $\mathbf{D}(x)$ [31]. From [38], (II.35) can be simplified to

$$\frac{\mathrm{d}\,|\mathbf{D}(x)|}{\mathrm{d}\,x} = \sum_{i=1}^{N_r}\left|\left\{(\mathbf{D}(x))_1, (\mathbf{D}(x))_2, \cdots, \frac{\mathrm{d}\,(\mathbf{D}(x))_i}{\mathrm{d}\,x}, \cdots, (\mathbf{D}(x))_{N_r}\right\}\right| \tag{II.36}$$

where $(\mathbf{D}(x))_j$ denotes the j^{th} row of $\mathbf{D}(x)$, thus $\mathbf{D}(x) = \{(\mathbf{D}(x))_1, (\mathbf{D}(x))_2, \cdots, (\mathbf{D}(x))_{N_r}\}$. Note that the matrix $\mathbf{L}(x)$ has a specific structure that we will exploit in arriving at the density

expressions. To do this, we first define $\gamma_{i,j}(x)$ in terms of entries of $\mathbf{L}(x)$ as shown below:

$$\gamma_{i,j}(x) \triangleq x\ (\mathbf{L}(x))_{i,j} - (\mathbf{L}(x))_{i-1,j}.$$

Next we establish that the derivative of $\gamma_{i,j}(x)$ (see (II.37) below) equals the ij^{th} element of $\mathbf{L}(x)$. This is an important property that leads to much simplification in carrying out the differentiation of (II.33) using (II.36).

$$\begin{aligned}\frac{\mathrm{d}\,\gamma_{i,j}(x)}{\mathrm{d}\,x} &= x\,\frac{\mathrm{d}\,(\mathbf{L}(x))_{i,j}}{\mathrm{d}\,x} + (\mathbf{L}(x))_{i,j} - \frac{\mathrm{d}\,(\mathbf{L}(x))_{i-1,j}}{\mathrm{d}\,x} \\ &= (\mathbf{L}(x))_{i,j}. \end{aligned} \tag{II.37}$$

The above follows, since $\frac{\mathrm{d}\,(\mathbf{L}(x))_{i,j}}{\mathrm{d}\,x} = \lambda_j^{-N_t+i-1}\,\frac{\mathrm{d}\,\Gamma(N_t-i+1,x\,\lambda_j)}{\mathrm{d}\,x} = x^{N_t-i}\;e^{-\lambda_j\,x}$, where we have used $\frac{\mathrm{d}\,\Gamma(n,x\,a)}{\mathrm{d}\,x} = a^n\,x^{n-1}\,e^{-ax}$. From this, it follows that $x\;\frac{\mathrm{d}\,(\mathbf{L}(x))_{i,j}}{\mathrm{d}\,x} = \frac{\mathrm{d}\,(\mathbf{L}(x))_{i-1,j}}{\mathrm{d}\,x}$.

We first provide proof of Part A of Theorem 1, and then outline a similar method to prove Part B. We define a new matrix $\widetilde{\mathbf{L}}(x)$ and use the basic properties of a determinant to write

$$|\mathbf{L}(x)| = x^{-N_r+1}\,|\widetilde{\mathbf{L}}(x)| \tag{II.38}$$

$$\begin{aligned}\text{where}\quad (\widetilde{\mathbf{L}}(x))_{1,j} &\triangleq (\mathbf{L}(x))_{1,j} \qquad \text{for } j = 1,2,\ldots,N_r \\ (\widetilde{\mathbf{L}}(x))_{i,j} &\triangleq \gamma_{i,j}(x) = x\,(\mathbf{L}(x))_{i,j} - (\mathbf{L}(x))_{i-1,j} \qquad \text{for } i = 2,\ldots,N_r.\end{aligned} \tag{II.39}$$

In the above, $\widetilde{\mathbf{L}}(x)$ is an $N_r \times N_r$ matrix whose i^{th} ($i = 2,\cdots,N_r$) row is obtained by multiplying the i^{th} row of $\mathbf{L}(x)$ by x and then subtracting the $(i-1)^{\text{th}}$ row of $\mathbf{L}(x)$ from it. The determinant of $\widetilde{\mathbf{L}}(x)$ would thus differ from that of $\mathbf{L}(x)$ by a multiplicative factor of x^{-N_r+1}. Note that the matrix $\widetilde{\mathbf{L}}(x)$ is designed so as to make use of (II.37). Now we are ready to apply (II.36) to (II.33) to obtain the density in (II.20):

$$\begin{aligned} f_{\lambda_{\max}}(x) &= \frac{\mathrm{d}\,F_{\lambda_{\max}}(x)}{\mathrm{d}\,x} = \tilde{a}\,\frac{\mathrm{d}\,|\mathbf{L}(x)|}{\mathrm{d}\,x} \\ &= \tilde{a}\,\frac{\mathrm{d}\,(x^{-N_r+1}\,|\widetilde{\mathbf{L}}(x)|)}{\mathrm{d}\,x} \\ &= \tilde{a}\,x^{-N_r+1}\sum_{i=1}^{N_r}\left|\left\{(\widetilde{\mathbf{L}}(x))_1,\cdots,\frac{\mathrm{d}\,(\widetilde{\mathbf{L}}(x))_i}{\mathrm{d}\,x},\cdots,(\widetilde{\mathbf{L}}(x))_{N_r}\right\}\right| \\ &\qquad -\tilde{a}\,(N_r-1)x^{-N_r}|\widetilde{\mathbf{L}}(x)| \end{aligned}$$

$$
\begin{aligned}
&= \tilde{a}\, x^{-N_r+1} \left| \left\{ \frac{\mathrm{d}\,(\widetilde{\mathbf{L}}(x))_1}{\mathrm{d}\,x},\ (\widetilde{\mathbf{L}}(x))_2, \cdots, (\widetilde{\mathbf{L}}(x))_{N_r} \right\} \right| \\
&\quad + \tilde{a}\, x^{-N_r+1} \sum_{i=2}^{N_r} \left| \left\{ (\widetilde{\mathbf{L}}(x))_1, \cdots, \frac{\mathrm{d}\,(\widetilde{\mathbf{L}}(x))_i}{\mathrm{d}\,x}, \cdots, (\widetilde{\mathbf{L}}(x))_{N_r} \right\} \right| \\
&\quad - \tilde{a}\,(N_r-1) x^{-N_r} x^{N_r-1} |\mathbf{L}(x)| \\
&\overset{(a)}{=} \tilde{a}\, x^{-N_r+1} \left| \left\{ (\{x^{N_t-1}\, e^{-\lambda_j\, x}\}_{j=1}^{N_r}),\ (\widetilde{\mathbf{L}}(x))_2, \cdots, (\widetilde{\mathbf{L}}(x))_{N_r} \right\} \right| \\
&\quad + \tilde{a}\, x^{-N_r+1} \sum_{i=2}^{N_r} \left| \left\{ (\widetilde{\mathbf{L}}(x))_1, \cdots, (\mathbf{L}(x))_i, \cdots, (\widetilde{\mathbf{L}}(x))_{N_r} \right\} \right| \\
&\quad - \tilde{a}\,(N_r-1)\ x^{-1}\, |\mathbf{L}(x)| \\
&= \tilde{a}\, x^{N_t-N_r} \left| \left\{ (\{x^{N_t-1}\, e^{-\lambda_j\, x}\}_{j=1}^{N_r}),\ (\widetilde{\mathbf{L}}(x))_2, \cdots, (\widetilde{\mathbf{L}}(x))_{N_r} \right\} \right| \\
&\quad + \tilde{a}\, x^{-N_r+1} \sum_{i=2}^{N_r} \left(x^{N_r-2}\, |\mathbf{L}(x)| \right) - \tilde{a}\,(N_r-1) x^{-1}\ |\mathbf{L}(x)| \\
&= \tilde{a}\, x^{N_t-N_r} \left| \left\{ \left(\{x^{N_t-1}\, e^{-\lambda_j\, x}\}_{j=1}^{N_r} \right),\ (\widetilde{\mathbf{L}}(x))_2, \cdots, (\widetilde{\mathbf{L}}(x))_{N_r} \right\} \right|.
\end{aligned}
\tag{II.40}
$$

In the equality (a) in (II.40), we have used $\frac{\mathrm{d}\,(\widetilde{\mathbf{L}}(x))_1}{\mathrm{d}\,x} = \frac{\mathrm{d}\,(\mathbf{L}(x))_1}{\mathrm{d}\,x} = \{x^{N_t-1} \quad e^{-\lambda_j\, x}\}_{j=1}^{N_r}$ and $\frac{\mathrm{d}\,(\widetilde{\mathbf{L}}(x))_i}{\mathrm{d}\,x} = \frac{\mathrm{d}\,(x\ (\mathbf{L}(x))_i - (\mathbf{L}(x))_{i-1})}{\mathrm{d}\,x} = (\mathbf{L}(x))_i$, for $i = 2, \cdots, N_r$ (from (II.39) and (II.37)). A close examination of the determinant inside the summation in the second term of the equality (a) in (II.40) reveals that $|\{(\widetilde{\mathbf{L}}(x))_1, \cdots, (\mathbf{L}(x))_i, \cdots, (\widetilde{\mathbf{L}}(x))_{N_r}\}| = x^{N_r-2}\, |\mathbf{L}(x)|$. This follows from the definition in (II.39) and the basic properties of determinants. Finally, note that the last equality in (II.40) is the desired expression in (II.20).

The proof of part B parallels that of part A. In order to prove part B, we define a new matrix $\widetilde{\mathbf{M}}(x)$ and write a similar expression for $\mathbf{M}(x)$ as in (B.6):

$$
|\mathbf{M}(x)| = x^{-N_t+1}\, |\widetilde{\mathbf{M}}(x)| \tag{II.41}
$$

where

$$
\begin{aligned}
(\widetilde{\mathbf{M}}(x))_{1,j} &\triangleq (\mathbf{M}(x))_{1,j} \\
(\widetilde{\mathbf{M}}(x))_{i,j} &\triangleq \rho_{i,j}(x) \triangleq x\ (\mathbf{M}(x))_{i,j} - (\mathbf{M}(x))_{i-1,j} \qquad \text{for } i = 2, \ldots, N_t \\
(\widetilde{\mathbf{M}}(x))_{i,j} &\triangleq (\mathbf{M}(x))_{i,j} \qquad \text{for } i = N_t+1, \ldots, N_r.
\end{aligned}
\tag{II.42}
$$

In the above, $\widetilde{\mathbf{M}}(x)$ is an $N_r \times N_r$ matrix and $\frac{\mathrm{d}\,\rho_{i,j}(x)}{\mathrm{d}\,x} = (\mathbf{M}(x))_{i,j}$. To arrive at the

density in (II.22), we now apply (II.36) to (II.34) and follow the same steps as in Part A.

II.H.4 Proof of Theorem 2.

From (II.20), we have

$$f_{\lambda_{\max}}(x) = \tilde{a}\, x^{N_t - N_r} |\mathbf{G}(x)|. \tag{II.43}$$

Using the fact that $\Gamma(j+1, x\,a) = \Gamma(j+1)(1 - \sum_{k=0}^{j} \frac{(ax)^k}{k!} e^{-ax})$ and $\Gamma(j+1) = j!$ [33], the ij^{th} ($i = 2, \ldots, Q$ and $j = 1, 2, \ldots, N_r$) element of $\mathbf{G}(x)$ can be simplified to

$$(\mathbf{G}(x))_{i,j} = x\ \theta_{i,j} + \xi_{i,j} + e^{-\lambda_j x}\left(\sum_{l=0}^{N_t - i} x^l\ \zeta_{i,j,l}\right) \tag{II.44}$$

$$\begin{aligned} \text{where} \quad \theta_{i,j} &= \lambda_j^{-N_t+i-1}\ \Gamma(N_t - i + 1); \\ \xi_{i,j} &= -\lambda_j^{-N_t+i-2}\ \Gamma(N_t - i + 2) \\ \text{and}\ \ \zeta_{i,j,l} &= \lambda_j^{-N_t+i+l-2}\ \left[\frac{(N_t - i)!}{l!}\right]\ (N_t - i - l + 1) \\ &\qquad\qquad \text{for}\ \ 0 \le l \le (N_t - i), \end{aligned} \tag{II.45}$$

$$\text{while}\ \ (\mathbf{G}(x))_{1,j} = e^{-\lambda_j x} \qquad \text{for } j = 1, 2, \ldots, N_r.$$

Next, we note the following well-known formula for determinant of an $m \times m$ matrix $\mathbf{A}$ [31]:

$$|\mathbf{A}| = \sum{}' \text{sign}(t_1, t_2, \cdots, t_m)\ (\mathbf{A})_{1,t_1}\, (\mathbf{A})_{2,t_2} \cdots (\mathbf{A})_{m,t_m}. \tag{II.46}$$

The sum in (II.71) is over all the permutations $(t_1, t_2, \cdots, t_m)$ of $(1, 2, \cdots, m)$. Using (II.71), $f_{\lambda_{\max}}(x)$ in (II.43) can then be written as

$$\begin{aligned} &f_{\lambda_{\max}}(x) \\ &= \tilde{a}\, x^{N_t - N_r} \sum{}' \text{sign}(t_1, t_2, \cdots, t_{N_r})(\mathbf{G}(x))_{1,t_1}\ (\mathbf{G}(x))_{2,t_2}\ \cdots\ (\mathbf{G}(x))_{N_r,t_{N_r}}. \end{aligned} \tag{II.47}$$

We now want to find a simplified expression for $\sum{}' \text{sign}(t_1, t_2, \cdots, t_{N_r})(\mathbf{G}(x))_{1,t_1}\ (\mathbf{G}(x))_{2,t_2}\ \cdots\ (\mathbf{G}(x))_{N_r,t_{N_r}}$. For the time being,

we neglect sign$(t_1, t_2, \cdots, t_{N_r})$ and simplify the expression $\sum'(\mathbf{G}(x))_{1,t_1}\ (\mathbf{G}(x))_{2,t_2} \cdots\ (\mathbf{G}(x))_{N_r,t_{N_r}}$. The above sum consists of terms that have a common factor of $e^{-(\lambda_{u_{1,m}}+\lambda_{u_{2,m}}+\cdots+\lambda_{u_{j,m}})x}$, where the set $S_{1,j,m} = \{u_{1,m}, u_{2,m}, \cdots u_{j,m}\}$ is comprised of j distinct integers, as has been defined in the statement of Theorem 2. We collect all such terms, and after some algebraic manipulations, write it in a general form as

$$\sum_{j=1}^{N_r}\sum_{m=1}^{\binom{N_r}{j}} e^{-(\lambda_{u_{1,m}}+\lambda_{u_{2,m}}+\cdots+\lambda_{u_{j,m}})x} \sum_{k=0}^{jN_t-N_t+N_r-\frac{1}{2}j^2-\frac{3}{2}j+1} x^k \varrho_{\,j,m,k} \tag{II.48}$$

where $\varrho_{\,j,m,k}$ is the coefficient of $e^{-(\lambda_{u_{1,m}}+\lambda_{u_{2,m}}+\cdots+\lambda_{u_{j,m}})x}\ x^k$ and is given by $\varrho_{\,j,m,k} = d_{\,j,m,k+N_t-N_r}$ when $N_t \geq N_r$; $d_{j,m,l}$ is given in (II.24).

Now, the presence of sign$(t_1, t_2, \cdots, t_{N_r})$ simplifies the sum in (II.47), since various terms can be shown to cancel. Specifically, it can be shown that $(\frac{1}{2}j^2 - \frac{3}{2}j + 1)$ terms are cancelled from the third summation in (II.48). This gives us the final expression for $f_{\lambda_{\max}}(x)$ as

$$\begin{aligned}
&\text{Case:}(N_t \geq N_r)\\
&\quad f_{\lambda_{\max}}(x) = \tilde{a}\, x^{N_t-N_r} \sum_{j=1}^{N_r}\sum_{m=1}^{\binom{N_r}{j}} e^{-\eta_{j,m}(\lambda)\,x} \sum_{k=0}^{jN_t-N_t+N_r-j^2} x^k \varrho_{\,j,m,k}\\
&\quad\quad = \tilde{a} \sum_{j=1}^{N_r}\sum_{m=1}^{\binom{N_r}{j}} e^{-\eta_{j,m}(\lambda)\,x} \sum_{l=N_t-N_r}^{jN_t-j^2} x^l \varrho_{\,j,m,l-N_t+N_r}.
\end{aligned} \tag{II.49}$$

Starting from (II.22), and following the same steps as above, it can be shown that

$$\begin{aligned}
&\text{Case:}(N_t < N_r)\\
&\quad f_{\lambda_{\max}}(x) = \tilde{a} \sum_{j=1}^{N_t}\sum_{m=1}^{\binom{N_r}{j}} e^{-\eta_{j,m}(\lambda)\,x} \sum_{l=0}^{jN_t-j^2} x^l \varsigma_{j,m,l}\\
&\qquad\qquad \text{where } \varsigma_{j,m,l} = d_{\,j,m,l}\\
&\qquad\qquad\quad \text{if } N_t < N_r.
\end{aligned} \tag{II.50}$$

The two expressions in (II.49) and (II.50) can be combined to yield (II.23).

II.H.5 Proof of Theorem 3.

Using the expression for $f_{\lambda_{\max}}(x)$ from Theorem 2, we simplify (II.18). In order to do this, we need to compute the integral $\int_0^\infty Q(\sqrt{x})\, e^{-\rho\, x} x^l \, \mathrm{d}x$:

$$\begin{aligned}\int_0^\infty Q(\sqrt{x})\, e^{-\rho\, x} x^l \, \mathrm{d}x &= \frac{1}{\pi}\int_0^{\frac{\pi}{2}} \int_0^\infty e^{\frac{-x}{2\sin^2\theta}}\, e^{-\rho\, x} x^l \, \mathrm{d}x \, \mathrm{d}\theta \\ &= \frac{l!}{\rho^{l+1}} \frac{1}{\pi} \int_0^{\frac{\pi}{2}} \left(\frac{\sin^2\theta}{\frac{1}{2\rho} + \sin^2\theta} \right)^{l+1} \mathrm{d}\theta \\ &= \frac{l!}{\rho^{l+1}} \frac{1}{2} \left[1 - \frac{1}{\sqrt{1+2\rho}} \sum_{k=0}^{l} \binom{2k}{k} \left[\frac{\rho}{2+4\rho} \right]^k \right]. \end{aligned} \tag{II.51}$$

The last equality in (II.51) follows from [43]. From (II.18), (II.23) and (II.51), we obtain the desired expression for the probability of error in (II.25).

II.H.6 Alternative Proof of Theorem 1

In order to prove Theorem 1, we first note the following lemmas and corollaries.

Lemma 3 *Let* $\mathbf{Z}$ *be* $m \times n$ *and* $\mathbf{Z} \sim CN(0, \mathbf{\Theta} \otimes \mathbf{\Omega})$*, where* $\mathbf{\Theta}$*,* $m \times m$*, and* $\mathbf{\Omega}$*,* $n \times n$*, are positive definite matrices and* $m \leq n$*. Then,* $\mathbf{S} \triangleq \mathbf{Z}\mathbf{Z}^H$ *is Hermitian positive definite and its density is given by*

$$f_{\mathbf{S}}(\mathbf{X}) = \frac{|\mathbf{\Theta}|^{-n} |\mathbf{\Omega}|^{-m}}{\widetilde{\Gamma}_m(n)} \, |\mathbf{X}|^{n-m} \, \widetilde{F}_{00}^{(n)}(\mathbf{\Omega}^{-1}, -\mathbf{\Theta}^{-1}\mathbf{X}) \tag{II.52}$$

where $\widetilde{F}_{00}^{(n)}(\mathbf{A}, \mathbf{B}) \triangleq \sum_{k=0}^{\infty} \sum_{\kappa} \frac{\widetilde{C}_\kappa(\mathbf{A}) \widetilde{C}_\kappa(\mathbf{B})}{k!\, \widetilde{C}_\kappa(\mathbf{I}_n)}$ *is the hypergeometric function of matrix argument and* $\widetilde{\Gamma}_m(n) \triangleq \pi^{\frac{1}{2}m(m-1)} (\prod_{i=1}^m \Gamma(n-i+1))$*.* $\mathbf{A}$ *and* $\mathbf{B}$ *are* $n \times n$ *matrices and* $\widetilde{C}_\kappa(\mathbf{A})$ *is the zonal polynomial of* $\mathbf{A}$ *(for the definitions see [36]).*

Proof. The pdf of $\mathbf{S}$ when $\mathbf{Z}$ is a real matrix is given in [44]. Following the same recipe for the case when $\mathbf{Z}$ is complex, the pdf can be shown to be given by (II.52). The above expression in terms of an arbitrary constant q appears in [45]. ∎

Lemma 4 *Given the density of* $\mathbf{S}$ *in Lemma* 1*, the joint pdf of its eigenvalues,* $\widetilde{w}_1 > \widetilde{w}_2 > \cdots > \widetilde{w}_m > 0$*, is given by*

$$f_{\widetilde{w}_1,\cdots,\widetilde{w}_m}(w_1,\cdots,w_m) \;=\; g(\{w_l\}_{l=1}^m)\; c\; |\mathbf{W}|^{n-m} \sum_{k=0}^{\infty} \sum_{\kappa} \frac{\widetilde{C}_\kappa(\mathbf{\Omega}^{-1})}{k!\;\widetilde{C}_\kappa(\mathbf{I}_n)} \frac{\widetilde{C}_\kappa(\mathbf{\Theta}^{-1})\widetilde{C}_\kappa(-\mathbf{X})}{\widetilde{C}_\kappa(\mathbf{I}_m)} \tag{II.53}$$

where $\mathbf{W} = diag(w_1, w_2, \cdots, w_m)$, $c \triangleq \frac{|\mathbf{\Theta}|^{-n}|\mathbf{\Omega}|^{-m}}{\widetilde{\Gamma}_m(n)}$ *and* $g(\{w_l\}_{l=1}^m) \triangleq \frac{\pi^{m(m-1)}}{\widetilde{\Gamma}_m(m)} \prod_{i=1}^m \prod_{j=i+1}^m (w_i - w_j)^2$.

Proof. In order to find the joint pdf of the eigenvalues of $\mathbf{S}$, we use the following result in [36]. If $f_{\mathbf{A}}(\mathbf{Y})$ is the density function of an $m \times m$ Hermitian positive definite matrix $\mathbf{A}$, then the density of its latent roots (eigenvalues), $\widetilde{\mu}_1, \widetilde{\mu}_2, \cdots, \widetilde{\mu}_m$, is given by

$$f_{\widetilde{\mu}_1,\cdots,\widetilde{\mu}_m}(\mu_1,\cdots,\mu_m) = \tfrac{\pi^{m(m-1)}}{\widetilde{\Gamma}_m(m)} \prod_{i<j}^{m} (\mu_i - \mu_j)^2 \int_{\mathbf{U}(m)} f_{\mathbf{A}}(\mathbf{U}\mathbf{Y}\mathbf{U}^H)\, \mathrm{d}\mathbf{U} \tag{II.54}$$

where, $\widetilde{\mu}_1 > \widetilde{\mu}_2 > \cdots > \widetilde{\mu}_m$, are the m distinct roots, $\mathbf{U}(m)$ is $m \times m$ unitary matrix and $\mu_1, \cdots, \mu_m$ are m eigenvalues of $\mathbf{Y}$. Applying the above result to the pdf of $\mathbf{S}$ in (II.52), the joint pdf of all the eigenvalues, $(\widetilde{w}_1 > \widetilde{w}_2 > \cdots > \widetilde{w}_m)$, of $\mathbf{S}$ can be written as

$$\begin{aligned} f_{\widetilde{w}_1,\cdots,\widetilde{w}_m}(w_1,\cdots,w_m) &= g(\{w_l\}_{l=1}^m) \int_{\mathbf{U}(m)} f_{\mathbf{S}}(\mathbf{U}\mathbf{X}\mathbf{U}^H)\, \mathrm{d}\mathbf{U} \\ &= g(\{w_l\}_{l=1}^m)\, c \\ &\quad \times \int_{\mathbf{U}(m)} |\mathbf{U}\mathbf{X}\mathbf{U}^H|^{n-m}\; \widetilde{F}_{00}^{(n)}(\mathbf{\Omega}^{-1}, -\mathbf{\Theta}^{-1}\mathbf{U}\mathbf{X}\mathbf{U}^H)\, \mathrm{d}\mathbf{U} \end{aligned} \tag{II.55}$$

Simplifying $|\mathbf{U}\mathbf{X}\mathbf{U}^H|^{n-m} = (|\mathbf{U}\mathbf{U}^H||\mathbf{X}|)^{n-m} = (|\mathbf{I}||\mathbf{X}|)^{n-m} = (|\mathbf{W}|)^{n-m}$, where $\mathbf{W}$ is the diagonal matrix of the eigenvalues, $(w_1, \cdots, w_m)$, of $\mathbf{X}$ and using the definition of $\widetilde{F}_{00}^{(\cdot)}(\cdot,\cdot)$, we have

$$= \; g(\{w_l\}_{l=1}^m)\; c\; |\mathbf{W}|^{n-m} \sum_{k=0}^{\infty} \sum_{\kappa} \frac{\widetilde{C}_\kappa(\mathbf{\Omega}^{-1})}{k!\;\widetilde{C}_\kappa(\mathbf{I}_n)} \int_{\mathbf{U}(m)} \widetilde{C}_\kappa(-\mathbf{\Theta}^{-1}\mathbf{U}\mathbf{X}\mathbf{U}^H)\, \mathrm{d}\mathbf{U}. \tag{II.56}$$

Finally, applying the following property of the zonal polynomials [45]

$$\int_{\mathbf{U}(m)} \widetilde{C}_\kappa(\mathbf{A}\mathbf{U}\mathbf{B}\mathbf{U}^H)\, \mathrm{d}\mathbf{U} = \frac{\widetilde{C}_\kappa(\mathbf{A})\widetilde{C}_\kappa(\mathbf{B})}{\widetilde{C}_\kappa(\mathbf{I}_m)} \tag{II.57}$$

in (II.56), we obtain (II.53). ■

We now specialize the result of Lemma 4 when $\mathbf{\Omega} = \mathbf{I}$ or $\mathbf{\Theta} = \mathbf{I}$

Corollary 1 *a) If $\mathbf{\Omega} = \mathbf{I}$, the result of Lemma 4 can be simplified to*

$$\begin{aligned} f_{\widetilde{w}_1,\cdots,\widetilde{w}_m}(w_1,\cdots,w_m) &= g(\{w_l\}_{l=1}^m)\; c\; |\mathbf{W}|^{n-m} \sum_{k=0}^{\infty} \sum_{\kappa} \frac{\widetilde{C}_\kappa(\mathbf{I}_n)\widetilde{C}_\kappa(\mathbf{\Theta}^{-1})\widetilde{C}_\kappa(-\mathbf{X})}{k!\; \widetilde{C}_\kappa(\mathbf{I}_n)\widetilde{C}_\kappa(\mathbf{I}_m)} \\ &= \frac{\pi^{m(m-1)}}{\widetilde{\Gamma}_m(m)} \prod_{i<j}^{m}(w_i - w_j)^2\; \frac{|\mathbf{\Theta}|^{-n}}{\widetilde{\Gamma}_m(n)}\; |\mathbf{W}|^{n-m} \widetilde{F}_{00}^{(m)}(-\mathbf{\Theta}^{-1}, \mathbf{X}). \end{aligned} \tag{II.58}$$

b) If $\mathbf{\Theta} = \mathbf{I}$, the result of Lemma 4 reduces to

$$\begin{aligned} f_{\widetilde{w}_1,\cdots,\widetilde{w}_m}(w_1,\cdots,w_m) &= g(\{w_l\}_{l=1}^m)\; c\; |\mathbf{W}|^{n-m} \sum_{k=0}^{\infty} \sum_{\kappa} \frac{\widetilde{C}_\kappa(\mathbf{\Omega}^{-1})\widetilde{C}_\kappa(\mathbf{I}_m)\widetilde{C}_\kappa(-\mathbf{X})}{k!\; \widetilde{C}_\kappa(\mathbf{I}_n)\widetilde{C}_\kappa(\mathbf{I}_m)} \\ &= \frac{\pi^{m(m-1)}}{\widetilde{\Gamma}_m(m)} \prod_{i<j}^{m}(w_i - w_j)^2\; \frac{|\mathbf{\Omega}|^{-m}}{\widetilde{\Gamma}_m(n)}\; |\mathbf{W}|^{n-m} \widetilde{F}_{00}^{(n)}(-\mathbf{\Omega}^{-1}, \mathbf{X}). \end{aligned} \tag{II.59}$$

The equalities in (II.58) and (II.59) follow from the definition of $\widetilde{F}_{00}^{(\cdot)}(\cdot,\cdot)$. The forms of the expressions in (II.58) and (II.59) appear identical, but are not. The matrix arguments of the function $\widetilde{F}_{00}^{(\cdot)}(\cdot,\cdot)$ in (II.58) have the same dimension, namely $m \times m$, whereas in (II.59) they have different dimensions, i.e., $n \times n$ and $m \times m$ for the first and the second arguments, respectively. Also, note that $\widetilde{F}_{00}^{(\cdot)}(\cdot,\cdot)$ depends on its matrix arguments only through its eigenvalues [36]. Using this fact it is shown later that different expressions for $\widetilde{F}_{00}^{(\cdot)}(\cdot,\cdot)$ result for the two cases. To reach this end, we note the following result in [37], where a simplified expression for the hypergeometric function is given

$$\widetilde{F}_{00}^{(k)}(\mathbf{U},\mathbf{V}) = |(e^{u_i v_j})|\; \Gamma_{(k)}(k)/[\alpha_k(\mathbf{U})\alpha_k(\mathbf{V})] \tag{II.60}$$

where $\alpha_k(\mathbf{U}) = \prod_{i<j}^{k}(u_i - u_j)$, $\Gamma_{(k)}(k) = \widetilde{\Gamma}_k(k)/\pi^{\frac{1}{2}k(k-1)}$, $\mathbf{U} = \text{diag}(u_1,\cdots,u_k)$ such that $u_1 > \cdots > u_k$ and $\mathbf{V} = \text{diag}(v_1,\cdots,v_k)$ such that $v_1 > \cdots > v_k$. $|(e^{u_i v_j})|$ denotes determinant of the matrix whose ij^{th} entry is $e^{u_i v_j}$, where $i,j = 1,2,\cdots,k$. Upon examination, we note that the above expression holds only when all the u_i's and v_i's are distinct. This is true because the determinant $|(e^{u_i v_j})|$ and $\alpha_k(\mathbf{U})$ become zero

when some of the u_i are equal, and $|(e^{u_i v_j})|$ and $\alpha_k(\mathbf{V})$ become zero when some of the v_i are equal, leading to an indeterminate form of $0/0$ for (II.60). In this case a limiting form for (II.60) must be obtained, as has been suggested in [37]. For our purpose we need to obtain an expression for (II.60) when some of u_i's are equal while all v_i's are distinct. Note that (II.60) is symmetric in the arguments, $\mathbf{U}$ and $\mathbf{V}$. Hence, the limiting expression for the case when $\mathbf{U}$ has some of the eigenvalues equal whereas $\mathbf{V}$ has all the eigenvalues distinct, is the same (after an appropriate interchange of arguments $\mathbf{U}$ and $\mathbf{V}$) when $\mathbf{U}$ has all distinct whereas $\mathbf{V}$ has some of the eigenvalues equal. Therefore, only one of these cases is considered and the result is summarized in Lemma 5.

Lemma 5 *Assume that* $\mathbf{U}$ *has* r *distinct eigenvalues, namely,* $\widetilde{u}_1, \widetilde{u}_2, \cdots, \widetilde{u}_r$ *with multiplicities* $h_1, h_2, \cdots, h_r$ *such that* $h_1 + h_2 + \cdots + h_r = k$. *Without loss of generality we can assume that* $\widetilde{u}_1 > \widetilde{u}_2 > \cdots > \widetilde{u}_r$. *Then*

$$\widetilde{F}^{(k)}_{00\ lim}(\mathbf{U}, \mathbf{V}) \;=\; \frac{\Gamma_{(k)}(k)}{\alpha_k(\mathbf{V})}\ \frac{N(\mathbf{U}, \mathbf{V})}{D(\mathbf{U})} \tag{II.61}$$

where

$$D(\mathbf{U}) \;=\; \left(\prod_{i=1}^{r-1}\prod_{j=i+1}^{r}(\widetilde{u}_i - \widetilde{u}_j)^{h_i h_j}\right)\left(\prod_{j=1}^{r}(\prod_{l=1}^{h_j-1} l!)\right) \tag{II.62}$$

and

$$N(\mathbf{U}, \mathbf{V}) \;=\; \begin{vmatrix} e^{\widetilde{u}_1 v_1} & \cdots & e^{\widetilde{u}_1 v_k} \\ -v_1 e^{\widetilde{u}_1 v_1} & \cdots & -v_k e^{\widetilde{u}_1 v_k} \\ \vdots & \ddots & \vdots \\ (-1)^{h_1-1} v_1^{h_1-1} e^{\widetilde{u}_1 v_1} & \cdots & (-1)^{h_1-1} v_k^{h_1-1} e^{\widetilde{u}_1 v_k} \\ \vdots & \ddots & \vdots \\ \vdots & \ddots & \vdots \\ e^{\widetilde{u}_r v_1} & \cdots & e^{\widetilde{u}_r v_k} \\ -v_1 e^{\widetilde{u}_r v_1} & \cdots & -v_k e^{\widetilde{u}_r v_k} \\ \vdots & \ddots & \vdots \\ (-1)^{h_r-1} v_1^{h_r-1} e^{\widetilde{u}_r v_1} & \cdots & (-1)^{h_r-1} v_k^{h_r-1} e^{\widetilde{u}_r v_k} \end{vmatrix}. \tag{II.63}$$

Proof. When some of the u_i's are equal the expression in (II.60) becomes indeterminate of the form $0/0$. Hence, we apply the Ĺ Hospital rule to obtain the limiting expression in (II.61). ■

Corollary 2 *The expressions in* (II.62) *and* (II.63) *can be simplified for the following two cases:*

1. *Let* $r = k$ *and* $h_1 = h_2 = \cdots = h_k = 1$*: This corresponds to the case when all the eigenvalues are distinct. Thus* (II.62) *reduces to* $D(\mathbf{U}) = \left(\prod_{i=1}^{k-1} \prod_{j=i+1}^{k} (\widetilde{u}_i - \widetilde{u}_j) \right) = \alpha_k(\mathbf{U})$ *and* (II.63) *becomes* $N(\mathbf{U}, \mathbf{V}) = |(e^{\widetilde{u}_i v_j})|$*. Substitution of this in* (II.61) *results in* (II.60)*, as expected.*

2. *Let* $r = 1$*,* $h_1 = k$*: This corresponds to the case when all the eigenvalues are equal. Thus* (II.62) *and* (II.63) *reduce to* $D(\mathbf{U}) = (\prod_{l=1}^{k-1} l!)$ *and* $N(\mathbf{U}, \mathbf{V}) = e^{\widetilde{u}_1(v_1+v_2+\cdots+v_k)} \alpha_k(\mathbf{V})$*, respectively. Substituting these in* (II.61) *and assigning* $\widetilde{u}_1 = 1$*, results in the expected form for* $\widetilde{F}_{00}^{(k)}{}_{lim}(\mathbf{I}, \mathbf{V}) = e^{tr(\mathbf{V})}$.

Using the results in Lemma 3, 4 and 5, we now give the proof for Theorem 1.

Proof of Theorem 1. We want to find an expression for the pdf, $f_{\lambda_{\max}}(x)$, of the maximum output SNR in , or in other words, we want to find the pdf of the largest eigenvalue of the random matrix $\mathbf{A} \triangleq (\frac{2P_o}{\sigma^2} \hat{\mathbf{H}}^H [\frac{P_o}{\sigma^2} \mathbf{\Sigma}_e + \mathbf{I}]^{-1} \hat{\mathbf{H}})$. $\mathbf{A}$ can be written as $\mathbf{A} = (\hat{\mathbf{H}}^H \mathbf{F}^{-1/2} \mathbf{F}^{-1/2} \hat{\mathbf{H}}) = \mathbf{Y}^H \mathbf{Y}$, where $\mathbf{F} \triangleq [\frac{P_o}{\sigma^2} \mathbf{\Sigma}_e + \mathbf{I}](\frac{\sigma^2}{2P_o})$ and $\mathbf{Y} \triangleq \mathbf{F}^{-1/2} \hat{\mathbf{H}}$. Recall that $\hat{\mathbf{H}} \sim CN(0, \bar{\mathbf{\Sigma}} \otimes \bar{\mathbf{\Psi}})$. It is easy to verify that $\mathbf{Y} \sim CN(0, \mathbf{F}^{-1/2} \bar{\mathbf{\Sigma}} \mathbf{F}^{-1/2} \otimes \bar{\mathbf{\Psi}})$. For the case, $\mathbf{\Psi} = \mathbf{I}$, we have $\mathbf{Y} \sim CN(0, \mathbf{F}^{-1/2} \bar{\mathbf{\Sigma}} \mathbf{F}^{-1/2} \otimes \mathbf{I})$. Note also that, if $N_t < N_r$, then $\mathbf{Y}^H \mathbf{Y}$ is positive definite with probability 1 and its density exists. If $N_t \geq N_r$, then $\mathbf{Y}^H \mathbf{Y}$ is positive semidefinite, whereas $\mathbf{Y} \mathbf{Y}^H$ is positive definite, with probability 1 . The non-zero eigenvalues of $\mathbf{Y} \mathbf{Y}^H$ and $\mathbf{Y}^H \mathbf{Y}$ are identical. In short,

1. If $N_t \geq N_r$, then $\mathbf{Y} \mathbf{Y}^H > 0$ (positive definite) and $\mathbf{Y} \sim CN(0, \mathbf{\Phi} \otimes \mathbf{I})$, where $\mathbf{\Phi} = \mathbf{F}^{-1/2} \bar{\mathbf{\Sigma}} \mathbf{F}^{-1/2}$.

2. If $N_t < N_r$, then $\mathbf{Y}^H \mathbf{Y} > 0$ (positive definite). Let $\mathbf{B} \triangleq \mathbf{Y}^H$ then $\mathbf{B} \sim CN(0, \mathbf{I} \otimes \mathbf{\Phi})$ and $\mathbf{Y}^H \mathbf{Y} = \mathbf{B} \mathbf{B}^H > 0$.

As a generalization of the above two cases, consider

$$\mathbf{Z} \mathbf{Z}^H, \text{ where } \mathbf{Z} \text{ is a } m \times n,\ m \leq n \text{ and } \mathbf{Z} \sim CN(0, \mathbf{\Theta} \otimes \mathbf{\Omega}). \tag{II.64}$$

$\mathbf{Z}\mathbf{Z}^H$ is positive definite with probability 1 and its density is given in Lemma 3. Also, cases 1 and 2 above follow from (II.64) due to the following assignment: $m = N_r$, $n = N_t$, $\mathbf{Z} = \mathbf{Y}$, $\mathbf{\Theta} = \mathbf{\Phi}$ and $\mathbf{\Omega} = \mathbf{I}$ for case 1 and $m = N_t$, $n = N_r$, $\mathbf{Z} = \mathbf{B} = \mathbf{Y}^H$, $\mathbf{\Theta} = \mathbf{I}$ and $\mathbf{\Omega} = \mathbf{\Phi}$ for case 2.

The joint pdf of all the eigenvalues of $\mathbf{S} = \mathbf{Z}\mathbf{Z}^H$ is given in Lemma 4. The corresponding expressions for the special cases 1 and 2, above, are dealt with in part a) and b), respectively, of Corollary 1, after performing the appropriate substitutions for m, n, $\mathbf{Z}$, $\mathbf{\Theta}$ and $\mathbf{\Omega}$, as indicated above. Using (II.58) and (II.59), we wish to obtain the desired expression for the density of the largest eigenvalue. We present two separate proofs for Part A and Part B of Theorem 1. We prove Part A, below, starting with (II.58) and then outline a similar method for proving Part B, starting with (II.59).

PART A: Starting with (II.58), the joint density of the eigenvalues $(\widetilde{w}_1, \widetilde{w}_2, \cdots, \widetilde{w}_m)$ of $\mathbf{S}$, or $\mathbf{Y}\mathbf{Y}^H$ when $N_t \geq N_r$, is given by

$$f_{\widetilde{w}_1,\cdots,\widetilde{w}_m}(w_1, \cdots, w_m) \;=\; a\,(\alpha_m(\mathbf{W}))^2 |\mathbf{W}|^{n-m} \widetilde{F}_{00}^{(m)}(-\mathbf{\Theta}^{-1}, \mathbf{X}) \qquad \text{(II.65)}$$

In (II.65), the notation developed in the text after Equation (II.60) for $\alpha_m(\mathbf{W})$ is used and $a \stackrel{\Delta}{=} \frac{\pi^{m(m-1)}}{\widetilde{\Gamma}_m(m)} \frac{|\mathbf{\Theta}|^{-n}}{\widetilde{\Gamma}_m(n)}$. Using (II.60) in (II.65),

$$\begin{aligned} f_{\widetilde{w}_1,\cdots,\widetilde{w}_m}(w_1, \cdots, w_m) &= a\,(\alpha_m(\mathbf{W}))^2 |\mathbf{W}|^{n-m} \frac{|(e^{-\lambda_i w_j})|\;\Gamma_{(m)}(m)}{[\alpha_m(-\mathbf{\Lambda})\alpha_m(\mathbf{W})]} \\ &= \widetilde{a} \quad \prod_{i=1}^{m} \prod_{j=i+1}^{m} (w_i - w_j)\;\; (\prod_{i=1}^{m} w_i^{n-m})\;\; |(e^{-\lambda_i w_j})|. \qquad \text{(II.66)} \end{aligned}$$

$\widetilde{a} \stackrel{\Delta}{=} \frac{a\,\Gamma_{(m)}(m)}{\alpha_m(-\mathbf{\Lambda})}$, $(-\lambda_1, \cdots, -\lambda_m)$ are eigenvalues of $-\mathbf{\Theta}^{-1}$ and $\mathbf{\Lambda} \stackrel{\Delta}{=} \text{diag}(\lambda_1, \cdots, \lambda_m)$ where $\lambda_1 < \cdots < \lambda_m$. Notice that $\prod_{i=1}^{m} \prod_{j=i+1}^{m} (w_i - w_j)$ is the determinant of a Vandermonde matrix [31], i.e.,

$$\prod_{i=1}^{m} \prod_{j=i+1}^{m} (w_i - w_j) = \begin{vmatrix} 1 & 1 & \cdots & 1 \\ w_m & w_{m-1} & \cdots & w_1 \\ w_m^2 & w_{m-1}^2 & \cdots & w_1^2 \\ \vdots & \vdots & \ddots & \vdots \\ w_m^{m-1} & w_{m-1}^{m-1} & \cdots & w_1^{m-1} \end{vmatrix} \qquad \text{(II.67)}$$

where $w_1 > w_2 > \cdots > w_m$. Clearly, the density of the largest eigenvalue $\widetilde{w}_1$ is given by the marginal density

$$\begin{aligned} f_{\widetilde{w}_1}(x) &= \int\limits_{0<w_m<\cdots<w_2<x} f_{\widetilde{w}_1,\widetilde{w}_2,\cdots,\widetilde{w}_m}(x, w_2, \cdots, w_m) \prod_{i=2}^{m} \mathrm{d}\, w_i \\ &= \widetilde{a} \int_R \prod_{i=2}^{m} \prod_{j=i+1}^{m} (w_i - w_j) \prod_{j=2}^{m} (x - w_j) \\ &\qquad \times (\prod_{i=2}^{m} w_i^{n-m}) \; x^{n-m} \; |(e^{-\lambda_i w_j})| \prod_{i=2}^{m} \mathrm{d}\, w_i \\ &= \widetilde{a}\, x^{n-m} \int_R \prod_{i=2}^{m} \prod_{j=i+1}^{m} (w_i - w_j) \prod_{i=2}^{m} a_i(x) \; |(e^{-\lambda_i w_j})| \prod_{i=2}^{m} \mathrm{d}\, w_i \end{aligned} \tag{II.68}$$

where $R \stackrel{\Delta}{=} 0 < w_m < \cdots < w_2 < x$ and $a_i(x) \stackrel{\Delta}{=} (x - w_i) w_i^{n-m}$. The second equality in the above is obtained by substituting (II.66) and simplifying such that the dependence on variable x is explicit. In (II.68) note that

$$\prod_{i=2}^{m} \prod_{j=i+1}^{m} (w_i - w_j) = \begin{vmatrix} 1 & 1 & \cdots & 1 & 0 \\ w_m & w_{m-1} & \cdots & w_2 & 0 \\ w_m^2 & w_{m-1}^2 & \cdots & w_2^2 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ w_m^{m-2} & w_{m-1}^{m-2} & \cdots & w_2^{m-2} & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{vmatrix}. \tag{II.69}$$

Through matrix manipulations, the integrand in (II.68) can be expressed as the determinant shown below.

$$\prod_{i=2}^{m} \prod_{j=i+1}^{m} (w_i - w_j) \;\; \prod_{i=2}^{m} a_i(x) \;\; |(e^{-\lambda_i w_j})|$$

$$= \begin{vmatrix} \sum_{i=2}^{m} e^{-\lambda_m w_i} a_i(x) & \sum_{i=2}^{m} e^{-\lambda_{m-1} w_i} a_i(x) & \cdots & \sum_{i=2}^{m} e^{-\lambda_1 w_i} a_i(x) \\ \sum_{i=2}^{m} e^{-\lambda_m w_i} w_i a_i(x) & \sum_{i=2}^{m} e^{-\lambda_{m-1} w_i} w_i a_i(x) & \cdots & \sum_{i=2}^{m} e^{-\lambda_1 w_i} w_i a_i(x) \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=2}^{m} e^{-\lambda_m w_i} w_i^{m-2} a_i(x) & \sum_{i=2}^{m} e^{-\lambda_{m-1} w_i} w_i^{m-2} a_i(x) & \cdots & \sum_{i=2}^{m} e^{-\lambda_1 w_i} w_i^{m-2} a_i(x) \\ e^{-\lambda_m x} & e^{-\lambda_{m-1} x} & \cdots & e^{-\lambda_1 x} \end{vmatrix}.$$

Using the compact notation for determinant, i.e., $|\cdot|$, the above can be simplified to

$$= \sum_{(l_2, l_3, \cdots, l_m)} |\mathbf{V}|. \tag{II.70}$$

The ij^{th} entry of the matrix $\mathbf{V}$, denoted as $(\mathbf{V})_{i,j}$, is given by $(\mathbf{V})_{i,j} = (e^{-\lambda_{m+1-j}\, w_{l_{i+1}}} \;\; w_{l_{i+1}}^{i-1} \;\; a_{l_{i+1}}(x))$ and $(\mathbf{V})_{m,j} = e^{-\lambda_{m+1-j}x}$ for $i = 1, 2, \cdots, (m-1)$ and $j = 1, 2, \cdots, m$. In (II.70), $(l_2, l_3, \cdots, l_m)$ are all the permutations of $(2, 3, \cdots, m)$. We also note the following formula for a determinant of, $m \times m$, $\mathbf{A}$:

$$|\mathbf{A}| = \sum\nolimits' \text{sign}(t_1, t_2, \cdots, t_m)\ (\mathbf{A})_{1,t_1}\,(\mathbf{A})_{2,t_2}\cdots(\mathbf{A})_{m,t_m}. \tag{II.71}$$

The sum in (II.71) is over all the permutations $(t_1, t_2, \cdots, t_m)$ of $(1, 2, \cdots, m)$, and sign$(t_1, t_2, \cdots, t_m)$ is positive if the permutation is even and negative if the permutation is odd [31]. Thus, (II.68) can be written as

$$f_{\widetilde{w}_1}(x) = \widetilde{a}\; x^{n-m} \int_R \sum_{(l_2, l_3, \cdots, l_m)} |\mathbf{V}| \;\; \prod_{i=2}^{m} \mathrm{d}\, w_i. \tag{II.72}$$

Using (II.71) in (II.72)

$$\begin{aligned} f_{\widetilde{w}_1}(x) \;\; &= \;\; \widetilde{a}\; x^{n-m} \sum_{(l_2, l_3, \cdots, l_m)} \int_R \;\; \sum\nolimits' \text{sign}(t_1, t_2, \cdots, t_m) \\ &\qquad \times (\mathbf{V})_{1,t_1}\,(\mathbf{V})_{2,t_2}\cdots(\mathbf{V})_{m,t_m} \;\; \prod_{i=2}^{m} \mathrm{d}\, w_i \\ &= \;\; \widetilde{a}\; x^{n-m} \sum_{(l_2, l_3, \cdots, l_m)} \;\; \sum\nolimits' \text{sign}(t_1, t_2, \cdots, t_m) \;\; (e^{-\lambda_{m+1-t_m} x}) \\ &\qquad \times \;\; \int_R \;\; \prod_{k=2}^{m} (e^{-\lambda_{m+1-t_{k-1}}\, w_{l_k}} \;\; w_{l_k}^{k-2} \;\; a_{l_k}(x)) \;\; \prod_{i=2}^{m} \mathrm{d}\, w_i. \end{aligned} \tag{II.73}$$

Using Lemma 1 of [46] in (II.73) and substituting for $a_k(x) = (x - w_k) w_k^{n-m}$

$$\begin{aligned} f_{\widetilde{w}_1}(x) \;\; &= \;\; \widetilde{a}\; x^{n-m} \;\; \sum\nolimits' \text{sign}(t_1, t_2, \cdots, t_m) \;\; (e^{-\lambda_{m+1-t_m} x}) \\ &\qquad \times \;\; \prod_{k=2}^{m} \;\; \int_0^x (e^{-\lambda_{m+1-t_{k-1}}\, w_k} \;\; w_k^{k-2} \;\; (x - w_k) w_k^{n-m}) \;\; \mathrm{d}\, w_k \end{aligned} \tag{II.74}$$

In order to simplify the integral in (II.74), consider $\int_0^x e^{-aw} w^j \, \mathrm{d}\, w$. With the substitution $w = y/a \;\; \Rightarrow \;\; \mathrm{d}\, w = \mathrm{d}\, y/a$, we get $\int_0^{xa} e^{-y} (y/a)^j \;(\mathrm{d}\, y/a) = a^{-j-1} \int_0^{xa} e^{-y} y^j \, \mathrm{d}\, y = a^{-j-1}\, \Gamma(j+1, xa)$. Using this in (II.74), we obtain

$$f_{\widetilde{w}_1}(x) = \widetilde{a}\; x^{n-m} \;\; \sum\nolimits' \text{sign}(t_1, t_2, \cdots, t_m) \;\; (e^{-\lambda_{m+1-t_m} x}) (\prod_{k=2}^{m} \beta_{k-1, t_{k-1}}(x)) \tag{II.75}$$

where

$$\begin{aligned}\beta_{i,j}(x) \;&\stackrel{\Delta}{=}\; \big(x[\lambda_{m+1-j}^{-n+m-i}\,\Gamma(n-m+i, x\lambda_{m+1-j})] \\ &-\; [\lambda_{m+1-j}^{-n+m-i-1}\,\Gamma(n-m+i+1, x\lambda_{m+1-j})]\big)\,. \end{aligned} \tag{II.76}$$

From (II.71), it is immediately seen that (II.75) can be expressed in terms of the determinant shown below:

$$f_{\widetilde{w}_1}(x) = \widetilde{a}\; x^{n-m} \begin{vmatrix} \beta_{1,1}(x) & \beta_{1,2}(x) & \cdots & \beta_{1,m}(x) \\ \beta_{2,1}(x) & \beta_{2,2}(x) & \cdots & \beta_{2,m}(x) \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{m-1,1}(x) & \beta_{m-1,2}(x) & \cdots & \beta_{m-1,m}(x) \\ e^{-\lambda_m x} & e^{-\lambda_{m-1}x} & \cdots & e^{-\lambda_1 x} \end{vmatrix} \tag{II.77}$$

where $\widetilde{a}$ is defined earlier and can be simplified to

$$\widetilde{a} = \frac{|\boldsymbol{\Theta}|^{-n}}{[\prod_{i=1}^{m}\prod_{j=i+1}^{m}(\lambda_j-\lambda_i)]\prod_{i=1}^{m}\Gamma(n-i+1)}. \tag{II.78}$$

Substituting $m = N_r$, $n = N_t$ and $\boldsymbol{\Theta} = \boldsymbol{\Phi}$ in (II.77) and (II.78), we obtain (II.20) and (II.21), respectively. This proves Part A of Theorem 1. Next, we give the outline for proving Part B.

PART B: We intend to find the pdf of the largest eigenvalue using the joint pdf of all the eigenvalues in (II.59), i.e., the joint pdf of $\mathbf{Y}^H\mathbf{Y}$ for the case $N_t < N_r$. Equation (II.59) can be rewritten as

$$f_{\widetilde{w}_1,\cdots,\widetilde{w}_m}(w_1,\cdots,w_m) \;=\; b\;\alpha_m(\mathbf{W})^2\;\;|\mathbf{W}|^{n-m}\widetilde{F}_{00}^{(n)}(-\boldsymbol{\Omega}^{-1},\mathbf{X}) \tag{II.79}$$

where $b \stackrel{\Delta}{=} \frac{\pi^{m(m-1)}}{\widetilde{\Gamma}_m(m)}\frac{|\boldsymbol{\Omega}|^{-m}}{\widetilde{\Gamma}_m(n)}$. As in the proof for the Part A, we want to simplify (II.79) by finding a expression for $\widetilde{F}_{00}^{(n)}(-\boldsymbol{\Omega}^{-1},\mathbf{X})$. Since $\widetilde{F}_{00}^{(n)}(-\boldsymbol{\Omega}^{-1},\mathbf{X})$ depends on $-\boldsymbol{\Omega}^{-1}$ and $\mathbf{X}$ only through its eigenvalues, we can write $\widetilde{F}_{00}^{(n)}(-\boldsymbol{\Omega}^{-1},\mathbf{X}) = \widetilde{F}_{00}^{(n)}(-\boldsymbol{\Lambda},\mathbf{W})$. Recall that $-\boldsymbol{\Lambda}$ and $\mathbf{W}$ are the diagonal matrices of the eigenvalues of $-\boldsymbol{\Omega}^{-1}$ and $\mathbf{X}$, respectively.

The expression in (II.60) assumes that $\mathbf{U}$ and $\mathbf{V}$ are of the same dimension and all their eigenvalues are distinct, whereas in Lemma 5, $\mathbf{U}$ and $\mathbf{V}$ are still of the

same dimension but all the eigenvalues of $\mathbf{U}$ are *not* necessarily distinct. Also, notice that some of the eigenvalues can be zero in Lemma 5. An expression for $\widetilde{F}_{00}^{(n)}(-\mathbf{\Lambda}, \mathbf{W})$, where $-\mathbf{\Lambda}$ and $\mathbf{W}$ are of *different* dimensions, i.e., $n \times n$ and $m \times m$, respectively, and $m \leq n$, can be written. Since $\mathbf{W}$ is of lower dimension, a positive semidefinite matrix, say $\mathbf{U}$, of dimension $n \times n$, which has m non zero distinct eigenvalues of $\mathbf{W}$ and $n-m$ zero eigenvalues can be constructed. Next, we apply Lemma 5, when the arguments to the function $\widetilde{F}_{00\ lim}^{(n)}(\cdot,\cdot)$ are $\mathbf{U}$ and $-\mathbf{\Lambda}$, and with the substitution $k = n$, $r = m+1$, $h_1 = h_2 = \cdots = h_m = 1$, $h_{m+1} = n-m$ and $\widetilde{u}_1 = w_1, \cdots, \widetilde{u}_m = w_m, \widetilde{u}_{m+1} = 0$. The following limiting expression is thus obtained:

$$\widetilde{F}_{00\ lim}^{(n)}(\mathbf{U}, -\mathbf{\Lambda}) \;=\; \frac{\Gamma_{(n)}(n)}{\alpha_n(-\mathbf{\Lambda})} \frac{N(\mathbf{U}, -\mathbf{\Lambda})}{D(\mathbf{U})} \tag{II.80}$$

where

$$\begin{aligned} D(\mathbf{U}) &= \left(\prod_{i=1}^{m}\prod_{j=i+1}^{m}(\widetilde{u}_i - \widetilde{u}_j)\right)\left(\prod_{i=1}^{m}\widetilde{u}_i^{n-m}\right)\left(\prod_{i=1}^{n-m-1} i!\right) \\ &= \alpha_m(\mathbf{W})\,|\mathbf{W}|^{n-m}\left(\prod_{l=1}^{n-m-1} l!\right) \end{aligned} \tag{II.81}$$

$$N(\mathbf{U}, -\mathbf{\Lambda}) \;=\; \begin{vmatrix} e^{-\lambda_1 w_1} & \cdots & e^{-\lambda_1 w_m} & 1 & \lambda_1 & \lambda_1^2 & \cdots & \lambda_1^{n-m-1} \\ e^{-\lambda_2 w_1} & \cdots & e^{-\lambda_2 w_m} & 1 & \lambda_2 & \lambda_2^2 & \cdots & \lambda_2^{n-m-1} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ e^{-\lambda_n w_1} & \cdots & e^{-\lambda_n w_m} & 1 & \lambda_n & \lambda_n^2 & \cdots & \lambda_n^{n-m-1} \end{vmatrix} \tag{II.82}$$

Substituting (II.80) in the joint pdf of all the eigenvalues of $\mathbf{S}$ in (II.79), we get

$$\begin{aligned} f_{\widetilde{w}_1,\cdots,\widetilde{w}_m}(w_1,\cdots,w_m) &= b\;\alpha_m(\mathbf{W})^2\;|\mathbf{W}|^{n-m} \\ &\quad \times \left[\frac{\Gamma_{(n)}(n)\;N(\mathbf{U},-\mathbf{\Lambda})}{\alpha_n(-\mathbf{\Lambda})\alpha_m(\mathbf{W})\;|\mathbf{W}|^{n-m}\left(\prod_{l=1}^{n-m-1} l!\right)}\right] \\ &= \tilde{b}\;\alpha_m(\mathbf{W})\;N(\mathbf{U},-\mathbf{\Lambda}). \end{aligned} \tag{II.83}$$

In the above, $\tilde{b} \triangleq b\;\Gamma_{(n)}(n)/\big(\alpha_n(-\mathbf{\Lambda})\;\prod_{l=1}^{n-m-1} l!\big)$. From this point onwards, proceeding in the same manner as in Part A, the following expression for the largest eigenvalue

is obtained:

$$f_{\widetilde{w}_1}(x) = \tilde{b} \begin{vmatrix} e^{-\lambda_1 x} & e^{-\lambda_2 x} & \cdots & e^{-\lambda_n x} \\ \gamma_{2,1}(x) & \gamma_{2,2}(x) & \cdots & \gamma_{2,n}(x) \\ \gamma_{3,1}(x) & \gamma_{3,2}(x) & \cdots & \gamma_{3,n}(x) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{m,1}(x) & \gamma_{m,2}(x) & \cdots & \gamma_{m,n}(x) \\ 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_n \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1^{n-m-1} & \lambda_2^{n-m-1} & \cdots & \lambda_n^{n-m-1} \end{vmatrix} \tag{II.84}$$

where $\gamma_{i,j}(x) \stackrel{\Delta}{=} x\ \lambda_j^{-m+i-1}\ \Gamma(m-i+1, x\,\lambda_j) - \lambda_j^{-m+i-2}\ \Gamma(m-i+2, x\,\lambda_j)$ for $i = 2, \cdots, m$ and $j = 1, \cdots, n$; $\tilde{b}$ can be simplified to

$$\tilde{b} = \left[\frac{|\mathbf{\Omega}|^{-m}}{\left(\prod_{i=1}^{m} \Gamma(m-i+1)\right)\left(\prod_{i=1}^{n} \prod_{j=i+1}^{n} (\lambda_j - \lambda_i)\right)} \right]. \tag{II.85}$$

Substituting $m = N_t$, $n = N_r$ and $\mathbf{\Omega} = \mathbf{\Phi}$ in (II.84) and (II.85), we obtain (II.22). ■

III

Performance of Multiple Antenna DS-CDMA UWB System With Noisy Channel Estimates and Narrow-Band Interference

III.A Introduction

Due to the large spectral occupancy and low transmission power levels (-41.3 dBm/MHz for indoor and hand-held peer-to-peer systems [3]), UWB systems have been proposed to overlay licensed narrowband services that occupy various portions of the radio spectrum. Interference arising from the coexisting narrowband systems, however, limits the performance of a UWB system (for example, the strong inband interference from the IEEE 802.11a WLAN services that operate with output power in the range 40–800 mW when compared to maximum average EIRP of 0.56 mW for UWB systems [2], [55]). This mandates that an adequate interference suppression measure be employed at the UWB receiver, which in addition has to estimate a large number of low energy multipath components to leverage multipath diversity.

For UWB systems, temporal processing techniques involving the use of an

MMSE receiver and tunable notch filters have been proposed to enhance the system performance [47]–[50]. Our work, here, considers the use of spatial processing techniques in addition to temporal processing to enhance the interference suppression capabilities and reliability of a UWB system [51], [52]. A DS-CDMA UWB system is considered, as it is one of the candidate techniques for multiple access UWB communications [53]. Note that the DS-CDMA system proposed in [53] is *carrier based*. Initially, UWB communications systems were based on carrier-less impulse radio (IR) technology [54]. However, the modulation formats considered for the commercial UWB applications are carrier-based [55], [56], such as the single carrier DS-CDMA [53] and the multicarrier frequency-hopping orthogonal frequency division multiplexing (FH-OFDM) [57] systems. The carrier-based solution was preferred because the severe restrictions placed by the FCC on the out-of-band emission for a UWB-enabled device makes the use of the baseband pulse, such as the one employed in the IR technology, very difficult [56].

There has been a growing research interest in applying multiple-input-multiple-output (MIMO) techniques to UWB systems [62]–[68]. MIMO systems, with multiple antennas at the transmitter and the receiver, have been shown to increase throughput and improve robustness in wireless communication channels [9], [58], [59]. The use of multiple antennas to improve the interference suppression capability of a CDMA system, over and above what is achievable through their processing gain, has been addressed in [60], [61]. In the context of UWB, employing multiple antennas to achieve higher data rates through spatial multiplexing has been considered in [62], [63]. In these latter works, an interference-free UWB link is assumed, where the channel between each transmit and receive antenna pair is independently faded. In [64], beamforming to combat interference in a time hopping impulse radio UWB system is examined. However, both [63] and [64] assume that perfect knowledge of the channel is available at the receiver. Issues related to the design of a space-time coding scheme for a single-band UWB system and space-time-frequency coding for a multi-band UWB system are studied in [65], [66], respectively. Characterization of a multiple antenna UWB channel and the study of antenna effects, such as coupling on the performance, is car-

ried out in [67], [68] through measurements. In particular, [68] considers the design of a UWB antenna array operating in the 3.1–4.5 GHz range, and presents the performance of a space-time RAKE receiver based upon measured data.

In this work, the performance of a single-input-multiple-output UWB link with multiple users and NBI is analytically investigated. A spatio-temporal receiver is employed, where each antenna element at the receiver array utilizes an FIR filter. To benefit from the large multipath diversity inherent in a UWB system, the receiver first forms an estimate of the channel in the presence of multi-access interference (MAI) and NBI, and then uses it to optimally determine the spatio-temporal tap weights such that the output signal-to-interference-plus-noise-ratio (SINR) is maximized. We investigate the performance of this system, in particular, studying the effects of both quality of channel estimates and antenna correlation. We show that for a fixed total diversity, there is an optimal combination of the number of antennas and the number of temporal taps in order to achieve the best performance, and that the optimal combination depends both on the signal-to-interference ratio and the quality of the estimates.

The rest of the chapter is organized as follows. In Section III.B, the system model along with the data detection and channel estimation scheme are described. The optimal receiver is derived in Section III.C. In Section III.D, we obtain an exact closed-form expression for the probability of error for a special case when the temporal correlation between the NBI during the estimation and the data detection phases is negligible. This is followed by numerical results and discussion in Section III.E. Finally, we state the conclusions in Section III.F.

III.B System Model

III.B.1 Transmitted Signal and Channel Model

Consider a conventional DS-CDMA system with binary signaling operating over a frequency selective slowly fading channel with a single transmitter antenna and M receiver antennas. A block diagram for the system is shown in Figure III.1. With K

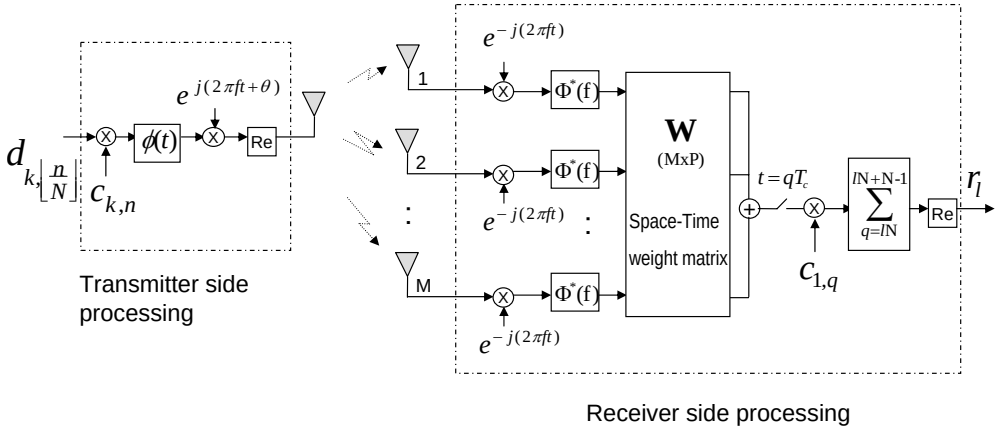

Figure III.1: System block diagram

users in the system, the transmitted signal of the kth user is given by

$$y_k(t) = \mathrm{Re}\left(\sqrt{\widetilde{P}_k}\sum_{n=-\infty}^{\infty} d_{k,\lfloor\frac{n}{N}\rfloor}\, c_{k,n}\, \phi(t-nT_c)\, e^{j2\pi ft}\right) \tag{III.1}$$

where $d_{k,\lfloor\frac{n}{N}\rfloor}$ and $c_{k,n}$ are the transmitted binary data sequence and the spreading sequence, respectively, $\phi(t)$ is the impulse response of the chip wave shaping filter and conforms to FCC regulations [3], and $\widetilde{P}_k$ is the transmitted power. We assume that the chip sequence of the desired user ($k = 1$) is deterministic, whereas for each of the $K-1$ interfering users it is an iid random binary sequence. The processing gain, N, of the system is $N = \frac{T}{T_c}$, where T is the bit duration and T_c is the chip duration. Denoting $\Phi(f)$ as the Fourier transform of $\phi(t)$, the wave-shaping filter has the following characteristics:

$$\Phi(f) = \begin{cases} \sqrt{T_c}, & \text{if } f \in [-1/2T_c, 1/2T_c] \\ 0, & \text{else.} \end{cases} \tag{III.2}$$

We will adopt the tapped delay line UWB channel model proposed in [69]. Using this model, the complex low-pass equivalent impulse response of the kth user's channel as seen at the mth receive antenna can be written as

$$h_{k,m}(t) = \sum_{i=0}^{J-1} \beta_{k,m,i}\, \delta(t - \widetilde{\tau}_{k,m,i}) \tag{III.3}$$

where J is the number of multipath components, $\beta_{k,m,i}$ is the complex channel fade for the ith path,[1] and the $\widetilde{\tau}_{k,m,i}$ are the multipath delays given by $\widetilde{\tau}_{k,m,i} = iT_c$, where T_c is the time resolution of the channel. It has been shown through measurements that the temporal correlation between various multipath components is negligible [69]. Thus, $\beta_{k,m,i}$ for all i and k are independent random variables. However, for different values of m, the $\beta_{k,m,i}$ are spatially correlated. According to the UWB channel model proposed in [69], the amplitude of the channel fade is modeled as a Nakagami distributed random variable where the parameters of the Nakagami distribution, i.e., fading figure (μ) and the second moment, are modeled as random variables. We will adopt an approximation to this model by first fixing the random distribution parameters to their mean values, as in [71], and then approximating the Nakagami density with a Ricean. Such an approximation holds when the Nakagami fading figure μ is greater than unity. In this case, the Ricean factor, K_r, can be related to μ as [70]

$$K_r = \frac{\sqrt{\mu^2 - \mu}}{\left(\mu - \sqrt{\mu^2 - \mu}\right)}. \tag{III.4}$$

Since μ is greater than 2.5 in our results, we model the complex fades $\beta_{k,m,i}$ as Gaussian. Expressing these fades in vector form, we denote the $(MJ \times 1)$ channel vector for user k by $\boldsymbol{\beta}_k$, which comprises the multipath coefficients at each of the receive antennas, i.e.,

$$\begin{aligned} \boldsymbol{\beta}_k^T &\triangleq [\boldsymbol{\beta}_{k,1}^T, \boldsymbol{\beta}_{k,2}^T, \cdots, \boldsymbol{\beta}_{k,M}^T], \\ \text{where } \boldsymbol{\beta}_{k,m}^T &\triangleq [\beta_{k,m,0}, \cdots, \beta_{k,m,J-1}], \quad \text{for } m = 1, \cdots, M. \end{aligned} \tag{III.5}$$

The distribution of $\boldsymbol{\beta}_k$ can now be written as $\boldsymbol{\beta}_k \sim CN(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$, where $CN(\mathbf{y}, \mathbf{X})$ represents a complex normal distribution with mean vector $\mathbf{y}$ and covariance matrix $\mathbf{X}$. Since the multipaths are assumed temporally uncorrelated, the covariance matrix, $\boldsymbol{\Sigma}_k$, has the following structure: $\boldsymbol{\Sigma}_k = S_{MJ}\ diag\{\mathbf{B}_{k,0}, \mathbf{B}_{k,1}, \cdots\ \cdots, \mathbf{B}_{k,J-1}\}\ S_{JM}$, where the $M \times M$ matrix $\mathbf{B}_{k,i}$ denotes the spatial correlation matrix of the ith path, and S_{MJ} and S_{JM} are the commutation matrices [32].

[1] Since in our work we consider a carrier-based system, it is common to include a phase shift. Thus, the path gain is complex.

III.B.2 Data Detection

The lowpass equivalent of the signal after passing through the channel in the presence of NBI and received at the mth antenna is given by

$$\begin{aligned} r_m(t) &= \sum_{k=1}^{K}\sum_{i=0}^{J-1}\beta_{k,m,i}\sum_{n=-\infty}^{\infty}\sqrt{P_k}\,d_{k,\lfloor\frac{n}{N}\rfloor}\,c_{k,n} \\ &\quad\times\phi(t-\tau_k-iT_c-nT_c)+I_m^{NBI}(t)+I_m^{Therm}(t) \end{aligned} \tag{III.6}$$

where P_k is the received power and τ_k is the time delay of user k. Note that the P_k and $\widetilde{P}_k$ (see (III.1)) are related by the path loss model of [69]. We assume that the signal appears to be narrowband to the antenna due to the fact that the fractional bandwidth of the signal is much less than unity [77]. The time delay of the desired user ($k = 1$) is assumed to be tracked, i.e., $\tau_1 = 0$, but for the interfering users, $k = 2, \cdots, K$, τ_k, is a random variable uniformly distributed over $[0, T]$. The thermal noise component, $I_m^{Therm}(t)$, is modeled as a zero-mean complex Gaussian process with two-sided spectral density N_o, and is assumed spatially uncorrelated. $I_m^{NBI}(t)$ denotes the lowpass equivalent of the narrow-band interference, which is modeled as a zero-mean WSS complex Gaussian process with power spectral density given by

$$P_{NBI}(f) = \begin{cases} J_o, & \text{if } f \in [-W/2, W/2] \\ 0, & \text{elsewhere.} \end{cases} \tag{III.7}$$

The space-time autocorrelation function of the NBI can be written as $\theta_{m_1,m_2}^{NBI}(\tau) = (\mathbf{B})_{m_1,m_2}\left(\frac{sin\pi W\tau}{\pi\tau}\right)\ J_o$, where m_1 and m_2 denote the antenna indices, and $\mathbf{B}$ is the $M \times M$ spatial correlation matrix of the NBI. The spatial correlation depends on the physical characteristics of the propagation environment, which in the case of an indoor operating environment can be assumed to be slowly changing due to low mobility. In such an environment, it is reasonable to assume that the spatial correlation is time-invariant. As a result, the joint space-time correlation function of the NBI can be written in a factored form as shown above. The specific model for the spatial correlation is presented in Section III.E.

The signal at each receive antenna is passed through a P-tap transversal filter with tap weights $\{w_{m,p}\}_{p=1}^{P}$ at the mth antenna, and tap separation T_c. Despreading the

chip matched filter output samples and then combining the weighted signal across the antenna array yields the output test statistic for the lth transmitted bit

$$r_l \;=\; \mathrm{Re}(\mathbf{w}^H\,\mathbf{z}_l) \tag{III.8}$$

where $\mathbf{w}$ is an $MP \times 1$ vector containing the spatio-temporal weights, i.e., $\mathbf{w}^T \triangleq [\mathbf{w}_1^T, \mathbf{w}_2^T, \cdots, \mathbf{w}_M^T]$ and $\mathbf{w}_m^T \triangleq [w_{m,1}, w_{m,2}, \cdots, w_{m,P}]$, $m = 1, \cdots, M$, and the vector $\mathbf{z}_l$ is given by

$$\mathbf{z}_l = \sqrt{P_1}\,\mathbf{D}_{1,l}\,\boldsymbol{\beta}_1 d_{1,l} + \sum_{k=2}^{K} \mathbf{D}_{k,l}\,\boldsymbol{\beta}_k + \boldsymbol{\gamma}_l^{NBI} + \boldsymbol{\gamma}_l^{Therm}. \tag{III.9}$$

The first term in (III.9) is due to the desired user's signal and its multipaths, but the inter-symbol interference component is ignored both because the multipath spread is assumed to be less than the symbol duration and the multipath intensity profile is exponentially decaying. The second term is the contribution from the multiple interfering users, and the third and the fourth terms are due to NBI and thermal noise, respectively. The matrices $\{\mathbf{D}_{k,l}\}_{k=2}^{K}$, of dimension $MP \times MJ$, are expressed in terms of the cross-correlation of the desired user's spreading sequence with those of the interfering users over a bit duration of N chips, and the matrix $\mathbf{D}_{1,l}$ is defined by the partial correlation of the desired user's spreading sequence with itself. These matrices are given by

$$\mathbf{D}_{k,l} \;\triangleq\; \mathbf{I}_M \otimes \mathbf{G}_{k,l} \qquad (k = 1, \cdots, K) \tag{III.10}$$

where

$$(\mathbf{G}_{1,l})_{i,j} \;\triangleq\; \sum_{q=lN+(j-i)^+}^{lN+N-1+(j-i)^-} c_{1,q+i-j}\; c_{1,q}$$

and

$$(\mathbf{G}_{k,l})_{i,j} \;\triangleq\; \sqrt{P_k} \sum_{q=lN}^{lN+N-1} \sum_{n=-\infty}^{\infty} c_{1,q}\; d_{k,\lfloor \frac{n}{N} \rfloor}\; c_{k,n}$$

$$\times\; x((q+i-j-n)T_c - \tau_k) \qquad (k = 2, \cdots, K). \tag{III.11}$$

In the above, $\otimes$ stands for the Kronecker product, $(z)^+ \triangleq \max(z, 0)$, $(z)^- \triangleq \min(z, 0)$, $(\mathbf{G}_{k,l})_{i,j}$, denotes the ijth entry of the $(P \times J)$ matrix $\mathbf{G}_{k,l}$, and $x(t)$ is defined by

$x(t) = \frac{sin(\pi t/T_c)}{\pi t/T_c}$. The vectors $\boldsymbol{\gamma}_l^{NBI}$ and $\boldsymbol{\gamma}_l^{Therm}$ can each be written as

$$\boldsymbol{\gamma}_l^s \stackrel{\Delta}{=} \sum_{q=lN}^{lN+N-1} c_{1,q} \left\{ \mathbf{u}^s(t) \odot \phi(-t) \right\}_{|t=qT_c} \tag{III.12}$$

where $\odot$ denotes convolution operation, the superscript 's' stands for either 'NBI' or 'Therm', $\mathbf{u}^s(t)^T \stackrel{\Delta}{=} [\mathbf{u}_1^s(t)^T, \mathbf{u}_2^s(t)^T, \cdots, \mathbf{u}_M^s(t)^T]$ and $\mathbf{u}_m^s(t)^T \stackrel{\Delta}{=} [I_m^s(t), I_m^s(t + T_c), \cdots, I_m^s(t + PT_c - T_c)]$, for $m = 1, \cdots, M$.

Note that the $\mathbf{D}_{k,l}$, $k = 2, \cdots, K$, is a random matrix due to τ_k and $c_{k,n}$. In our analysis, the MAI term, namely, $\sum_{k=2}^{K} \mathbf{D}_{k,l}\, \boldsymbol{\beta}_k$ in (III.9), is approximated as being Gaussian due to a large number of interfering users. With this approximation, the vector $\mathbf{z}_l$, conditioned on the data $d_{1,l}$, is Gaussian.

III.B.3 Channel Estimation

The channel vector $\boldsymbol{\beta}_1$ is unknown at the receiver. An estimate of the channel is obtained based on a transmitted pilot which is time-multiplexed with the data. Since the receiver combines the first P ($\leq J$) multipaths, an estimate of P multipath coefficients at each receive antenna is formed. A spread tone is used for the pilot signal. At the receiver, the pilot symbols are demodulated, chip-matched filtered, despread and accumulated over LN chip intervals, where L is a positive integer. The output pilot from all the antennas can be written in an analogous form to (III.9) as

$$\mathbf{y} \;=\; \sqrt{P_p}\, \widetilde{\mathbf{D}}_{1,\widetilde{l}}\, \boldsymbol{\beta}_1 + \sum_{k=2}^{K} \widetilde{\mathbf{D}}_{k,\widetilde{l}} \boldsymbol{\beta}_k + \boldsymbol{\gamma}_{\widetilde{l}}^{NBI} + \boldsymbol{\gamma}_{\widetilde{l}}^{Therm} \tag{III.13}$$

where $\widetilde{l}$ denotes the start of the bit boundary for the estimation period of LN chips, P_p is the received pilot power and $\boldsymbol{\gamma}_{\widetilde{l}}^{NBI}$, $\boldsymbol{\gamma}_{\widetilde{l}}^{Therm}$ are $MP \times 1$ spatio - temporal complex Gaussian vectors due to NBI and thermal noise, respectively. Equation (III.13) is similar to (III.9), except that the data $d_{1,l}$, in (III.9), is set to unity, the sampled time instant, l in (III.9), has been changed to $\widetilde{l}$ in (III.13) for pilot reception/channel estimation, and the pilot is accumulated over LN chips, as opposed to N chips in (III.9). The matrix $\widetilde{\mathbf{D}}_{k,\widetilde{l}}$ is defined as $\widetilde{\mathbf{D}}_{k,\widetilde{l}} \stackrel{\Delta}{=} \mathbf{I}_M \otimes \widetilde{\mathbf{G}}_{k,\widetilde{l}}$, where $(\widetilde{\mathbf{G}}_{1,\widetilde{l}})_{i,j} \stackrel{\Delta}{=} \sum_{q=\widetilde{l}N}^{\widetilde{l}N+LN-1} c_{1,q+i-j}\, c_{1,q}$, and $\widetilde{\mathbf{G}}_{k,\widetilde{l}}$, for

$k = 2, \cdots, K$, is the same as (III.11) except the first summation in (III.11) is replaced by $\sum_{q=\widetilde{l}N}^{\widetilde{l}N+LN-1}$. A similar change of limits in the summation in (III.12) would lead to the definitions for $\boldsymbol{\gamma}_{\widetilde{l}}^{NBI}$, $\boldsymbol{\gamma}_{\widetilde{l}}^{Therm}$ in (III.13).

Using vector notation, the MP multipath coefficients to be estimated can be represented by the vector $\boldsymbol{\beta}_1^{(1)}$. The remaining $(J-P)M$ coefficients are represented as $\boldsymbol{\beta}_1^{(2)}$. Note that taken together, $\boldsymbol{\beta}_1^{(1)}$ and $\boldsymbol{\beta}_1^{(2)}$ constitute all the elements of $\boldsymbol{\beta}_1$. The mean and covariance matrices associated with $\boldsymbol{\beta}_1^{(1)}$ and $\boldsymbol{\beta}_1^{(2)}$ are denoted as $\{\boldsymbol{\mu}_1^{(1)}, \boldsymbol{\Sigma}_1^{(1)}\}$, and $\{\boldsymbol{\mu}_1^{(2)}, \boldsymbol{\Sigma}_1^{(2)}\}$, respectively, and are obtained from $\{\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1\}$ of $\boldsymbol{\beta}_1$. With these definitions, the first term in (III.9) can be further simplified to separate out the contribution due to self-interference (SI), i.e., we have

$$\begin{aligned}&\sqrt{P_1}\,\mathbf{D}_{1,l}\,\boldsymbol{\beta}_1 d_{1,l}\\ &\quad= \sqrt{P_1}\,(N\boldsymbol{\beta}_1^{(1)} + \mathbf{D}_{1,l}^{(1)}\,\boldsymbol{\beta}_1^{(1)} + \mathbf{D}_{1,l}^{(2)}\,\boldsymbol{\beta}_1^{(2)})\,d_{1,l}.\end{aligned} \tag{III.14}$$

Note that the second and the third terms in (III.14) are the SI terms. The off-main-diagonal entries of the matrix $\mathbf{G}_{1,l}$, constituting $\mathbf{D}_{1,l}$, result from the non-zero autocorrelation of the spreading sequence for non-zero lag. This, in turn, leads to SI which, by an appropriate partition of the matrix $\mathbf{G}_{1,l}$, is identified in (III.14). Thus, the following definitions are in order: $\mathbf{D}_{1,l}^{(1)} \triangleq \mathbf{I}_M \otimes \mathbf{G}_{1,l}^{(1)}$ and $\mathbf{D}_{1,l}^{(2)} \triangleq \mathbf{I}_M \otimes \mathbf{G}_{1,l}^{(2)}$, where the $(P \times P)$ matrix $\mathbf{G}_{1,l}^{(1)}$ and the $(P \times (J-P))$ matrix $\mathbf{G}_{1,l}^{(2)}$ are formed by partitioning $\mathbf{G}_{1,l}$. If the main diagonal entries of $\mathbf{G}_{1,l}$ are first set to zero, then the first P columns of $\mathbf{G}_{1,l}$ is the matrix $\mathbf{G}_{1,l}^{(1)}$, whereas the last $(J-P)$ columns is $\mathbf{G}_{1,l}^{(2)}$. Similarly, the first term of (III.13) can be simplified and quantities $\widetilde{\mathbf{D}}_{1,\widetilde{l}}^{(1)}$ and $\widetilde{\mathbf{D}}_{1,\widetilde{l}}^{(2)}$ can be defined.

Next we use the pilot $\mathbf{y}$ to form a linear minimum mean-square error (LMMSE) estimate, $\widehat{\boldsymbol{\beta}_1}^{(1)}$. From [30], $\widehat{\boldsymbol{\beta}_1}^{(1)}$ is given by the mean of the posterior pdf,[2] i.e., $\widehat{\boldsymbol{\beta}_1}^{(1)} = E[\boldsymbol{\beta}_1^{(1)}|\mathbf{y}]$. Thus, we have

$$\widehat{\boldsymbol{\beta}_1}^{(1)} = \boldsymbol{\mu}_1^{(1)} + \mathbf{G}^H\,(\mathbf{y} - E[\mathbf{y}]), \tag{III.15}$$

where $\mathbf{G} = C_{\mathbf{yy}}^{-1} C_{\mathbf{y}\boldsymbol{\beta}_1^{(1)}}$. The channel estimate, $\widehat{\boldsymbol{\beta}_1}^{(1)}$, is Gaussian-distributed with mean

[2] Since the LMMSE channel estimator is employed, the second order statistics of the channel and interference are assumed known.

$\boldsymbol{\mu}_1^{(1)}$ and covariance given by

$$C_{\widehat{\boldsymbol{\beta}}_1^{(1)}\widehat{\boldsymbol{\beta}}_1^{(1)}} = C_{\boldsymbol{\beta}_1^{(1)}\mathbf{y}} C_{\mathbf{yy}}^{-1} C_{\mathbf{y}\boldsymbol{\beta}_1^{(1)}}. \tag{III.16}$$

It can be shown that the mean-square error in estimation tends to zero at a rate $1/L$ as L tends to infinity. Note that the same processes, NBI and MAI, interfere during channel estimation and data detection. Since the NBI during channel estimation is correlated to itself during data detection, there is temporal correlation between the received pilot in (III.13), and the received data in (III.9). Later, in the numerical results, we will study the impact of this correlation on the system performance.

III.C Receiver Design

With knowledge of the channel estimate available, we wish to design the optimal receiver, i.e., determine the weights, $\mathbf{w}$, such that the conditional output SINR (conditioned on $\widehat{\boldsymbol{\beta}}_1^{(1)}$) is maximized. The optimal weights thus determined depend on the channel estimates. Before designing the receiver, we first establish an expression for the conditional probability of error.

III.C.1 Conditional Probability of Error

The test statistic in (III.8) can be rewritten in the form shown below:

$$r_l \;=\; x_1\, d_{1,l} + x_2 + x_3 + x_4 \tag{III.17}$$

where $x_1 \triangleq \mathrm{Re}(\sqrt{P_1}\mathbf{w}^H \mathbf{D}_{1,l}\, \boldsymbol{\beta}_1)$, $x_2 \triangleq \mathrm{Re}(\mathbf{w}^H \sum_{k=2}^{K} \mathbf{D}_{k,l}\, \boldsymbol{\beta}_k)$, $x_3 \triangleq \mathrm{Re}(\mathbf{w}^H \boldsymbol{\gamma}_l^{NBI})$, and $x_4 \triangleq \mathrm{Re}(\mathbf{w}^H \boldsymbol{\gamma}_l^{Therm})$. The optimum weight vector, $\mathbf{w}$, depends on the channel estimate, $\widehat{\boldsymbol{\beta}}_1^{(1)}$ (see part (a) of Lemma 6 in Section III.C.2), thus, conditioned on $\widehat{\boldsymbol{\beta}}_1^{(1)}$ and $d_{1,l}$, the test statistic in (III.17) is Gaussian. Note that, conditioned on the estimate $\widehat{\boldsymbol{\beta}}_1^{(1)}$, the terms in (III.17) are correlated as well. An expression for the conditional

probability of error, as a function of $\mathbf{w}$, can be written as

$$\begin{aligned} & P_e(\widehat{\boldsymbol{\beta}_1}^{(1)}) \\ &= \tfrac{1}{2}P[\nu_1 < 0 | d_{1,l} = 1, \widehat{\boldsymbol{\beta}_1}^{(1)}] \\ &\quad + \tfrac{1}{2}P[\nu_2 > 0 | d_{1,l} = -1, \widehat{\boldsymbol{\beta}_1}^{(1)}] \\ &= \tfrac{1}{2}Q\left(\frac{E[\nu_1|\widehat{\boldsymbol{\beta}_1}^{(1)}]}{\sigma_{\nu_1|\widehat{\boldsymbol{\beta}_1}^{(1)}}}\right) + \tfrac{1}{2}Q\left(\frac{-E[\nu_2|\widehat{\boldsymbol{\beta}_1}^{(1)}]}{\sigma_{\nu_2|\widehat{\boldsymbol{\beta}_1}^{(1)}}}\right) \end{aligned} \tag{III.18}$$

where $\nu_1 \triangleq x_1+x_2+x_3+x_4$ and $\nu_2 \triangleq -x_1+x_2+x_3+x_4$. Simplifying the conditional mean and variances that appear in (III.18), the conditional probability of error can be written as shown in Theorem 4.

Theorem 4 *The conditional probability of error is given by*

$$P_e(\widehat{\boldsymbol{\beta}_1}^{(1)}) = \tfrac{1}{2}Q\left(\frac{m_1+m_2}{\sigma_1}\right) + \tfrac{1}{2}Q\left(\frac{m_1-m_2}{\sigma_2}\right) \tag{III.19}$$

where $Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^\infty e^{-y^2/2}\mathrm{d}y$. *Expressions for the parameters appearing inside the* $Q(\cdot)$ *function in* (III.19) *are given below:*

$$\begin{aligned} m_1 &\triangleq \mathrm{Re}\left(\mathbf{w}^H\left(\mathbf{h} + \mathbf{Y}\,\widehat{\boldsymbol{\beta}_1}^{(1)}\right)\right) \\ m_2 &\triangleq \mathrm{Re}\left(\mathbf{w}^H\left(\boldsymbol{\Lambda}\, C^{-1}_{\widehat{\boldsymbol{\beta}_1}^{(1)}\widehat{\boldsymbol{\beta}_1}^{(1)}}\left[\widehat{\boldsymbol{\beta}_1}^{(1)} - \boldsymbol{\mu}_1^{(1)}\right]\right)\right) \end{aligned} \tag{III.20}$$

where $\mathbf{h} = \sqrt{P_1}\,\mathbf{D}^{(2)}_{1,l}(\,\boldsymbol{\mu}_1^{(2)} - \mathbf{Z}\,\boldsymbol{\mu}_1^{(1)})$, $\mathbf{Y} = \sqrt{P_1}\,(\mathbf{D}^{(2)}_{1,l}\,\mathbf{Z} + \hat{\mathbf{D}})$, $\mathbf{Z} = \sqrt{P_p}\,\boldsymbol{\Sigma}_1^{(2)}\,(\widetilde{\mathbf{D}}^{(2)}_{1,\tilde{l}})^T\,\mathbf{G}\,C^{-1}_{\widehat{\boldsymbol{\beta}_1}^{(1)}\widehat{\boldsymbol{\beta}_1}^{(1)}}$, $\hat{\mathbf{D}} = (N\mathbf{I}_{MP} + \mathbf{D}^{(1)}_{1,l})$, *and* $\boldsymbol{\Lambda} = (\boldsymbol{\Omega}_{l,\tilde{l}} + \boldsymbol{\Xi}_{l,\tilde{l}})^H\,\mathbf{G}$.

The variances in (III.19) *are given by*

$$\begin{aligned} \sigma_1^2 &\triangleq \tfrac{1}{2}\left(\mathbf{w}^H\left(C_1 - C_2\,C^{-1}_{\widehat{\boldsymbol{\beta}_1}^{(1)}\widehat{\boldsymbol{\beta}_1}^{(1)}}C_2^H\right)\mathbf{w}\right) \\ \textit{and}\quad \sigma_2^2 &\triangleq \tfrac{1}{2}\left(\mathbf{w}^H\left(C_3 - C_4\,C^{-1}_{\widehat{\boldsymbol{\beta}_1}^{(1)}\widehat{\boldsymbol{\beta}_1}^{(1)}}C_4^H\right)\mathbf{w}\right). \end{aligned} \tag{III.21}$$

Proof. The proof is given in Appendix III.H.1. ∎

The expressions m_1 and m_2, in (III.20), are obtained by simplifying the conditional means in (III.18). Specifically, the mean m_2 results because of the temporal correlation between the pilot and the data. This can be seen through its dependence on matrices $\mathbf{\Omega}_{l,\widetilde{l}}$ and $\mathbf{\Xi}_{l,\widetilde{l}}$ constituting $\mathbf{\Lambda}$ above. The matrix $\mathbf{\Omega}_{l,\widetilde{l}} = E[(\boldsymbol{\gamma}_{\widetilde{l}}^{NBI})(\boldsymbol{\gamma}_{l}^{NBI})^H]$, appearing in the definition of $\mathbf{\Lambda}$, is given by the cross-correlation of the NBI component during estimation and data detection phases, and $\mathbf{\Xi}_{l,\widetilde{l}} = E[(\sum_{k=2}^{K} \widetilde{\mathbf{D}}_{k,\widetilde{l}}\boldsymbol{\beta}_k) \times (\sum_{\widetilde{k}=2}^{K} \mathbf{D}_{\widetilde{k},l}\boldsymbol{\beta}_{\widetilde{k}})^H]$ is given by the cross-correlation of the MAI component during the estimation phase and the MAI during the data detection phase. Similarly, the two separate expression for the variances in (III.21) are obtained by simplifying the conditional variances that appear in (III.18). Expressions for the covariance matrices C_1, C_2, C_3 and C_4, in (III.21), and matrices $\mathbf{\Omega}_{l,\widetilde{l}}$ and $\mathbf{\Xi}_{l,\widetilde{l}}$, are given in Appendix III.H.2.

Consider now a special case when the correlation between the NBI during the channel estimation phase and the data detection phase is weak and can be neglected, i.e., mathematically, $W[(l-\widetilde{l}-L)N]T_c >> 1$. For this case, the expression in (III.19) simplifies to

$$P_e(\widehat{\boldsymbol{\beta}_1}^{(1)}) = Q\left(\frac{m_1}{\sigma}\right) \tag{III.22}$$

where $\sigma_1 = \sigma_2 = \sigma$. Note that the expression in (III.22) is exact when the cross-correlation of the NBI during estimation and data detection phases is zero, and the corresponding cross-correlation of the MAI components is zero. In this case, the matrices $\mathbf{\Omega}_{l,\widetilde{l}}$ and $\mathbf{\Xi}_{l,\widetilde{l}}$ are each zero, and the two expressions for σ_1 and σ_2, in (III.21), are equal. This is denoted as $\sigma_1^2 = \sigma_2^2 = \sigma^2 = \frac{1}{2}(\mathbf{w}^H\, C\, \mathbf{w})$. The mean m_2 depends on $\mathbf{\Omega}_{l,\widetilde{l}}$ and $\mathbf{\Xi}_{l,\widetilde{l}}$ and is equal to zero in (III.22).

III.C.2 SINR Maximization

To determine the tap weights, $\mathbf{w}$, we maximize the output SINR. The conditional SINR (conditioned on $\widehat{\boldsymbol{\beta}_1}^{(1)}$) is defined as

$$\text{SINR}(\mathbf{w}|\widehat{\boldsymbol{\beta}_1}^{(1)}) = \frac{\left(\text{Re}\,\left(\mathbf{w}^H\left(\mathbf{h} + \mathbf{Y}\,\widehat{\boldsymbol{\beta}_1}^{(1)}\right)\right)\right)^2}{\frac{1}{2}\left(\mathbf{w}^H\, C\mathbf{w}\right)}. \tag{III.23}$$

The SINR expression in (III.23) corresponds to the case when the cross correlation of the NBI during the channel estimation and the data detection phases is zero. Thus, the SINR in (III.23) relates to the argument of the $Q(.)$ function appearing in the error probability in (III.22). Specifically, SINR$(\mathbf{w}|\widehat{\boldsymbol{\beta}_1}^{(1)}) = m_1^2/\sigma^2$.

The function in (III.23) can be maximized using the Cauchy-Schwarz (CS) inequality. The optimum weights resulting from such a maximization (see Appendix III.H.3) are given below in part (a) of Lemma 6.

Lemma 6 *: (a) The optimum weights are:* $\mathbf{w} = C^{-1}\,(\mathbf{h} + \mathbf{Y}\,\widehat{\boldsymbol{\beta}_1}^{(1)})$.

By substituting the optimal weights in (III.23), the maximum SINR can be written as

$$\begin{aligned}\text{max SINR}(\widehat{\boldsymbol{\beta}_1}^{(1)}) \;\; = \;\; 2\;\Big(&\widehat{\boldsymbol{\beta}_1}^{(1)\,H}\;\mathbf{Y}^H C^{-1}\mathbf{Y}\;\widehat{\boldsymbol{\beta}_1}^{(1)} \\ &+ \widehat{\boldsymbol{\beta}_1}^{(1)\,H}\mathbf{Y}^H C^{-1}\,\mathbf{h} \\ &+ \mathbf{h}^H C^{-1}\mathbf{Y}\,\widehat{\boldsymbol{\beta}_1}^{(1)} + \mathbf{h}^H C^{-1}\,\mathbf{h}\Big)\,.\end{aligned}$$

Let us define $\mathbf{a}^T \triangleq [\text{Re}(\widehat{\boldsymbol{\beta}_1}^{(1)})^T \;\; \text{Img}(\widehat{\boldsymbol{\beta}_1}^{(1)})^T]$, and

$$\mathbf{F}(\mathbf{X}) \triangleq \begin{bmatrix} \text{Re}(\mathbf{X}) & -\text{Img}(\mathbf{X}) \\ \text{Img}(\mathbf{X}) & \text{Re}(\mathbf{X}) \end{bmatrix}, \tag{III.24}$$

which is a matrix-valued function of a matrix. Since the complex Gaussian vector $\widehat{\boldsymbol{\beta}_1}^{(1)}$ is of length PM, the dimension of $\mathbf{a}$ is $(2PM \times 1)$. Recall that $\widehat{\boldsymbol{\beta}_1}^{(1)} \sim CN(\boldsymbol{\mu}_1^{(1)}, C_{\widehat{\boldsymbol{\beta}_1}^{(1)}\widehat{\boldsymbol{\beta}_1}^{(1)}})$; this implies that $\mathbf{a} \sim N(\boldsymbol{\mu}_{\mathbf{a}}, C_{\mathbf{aa}})$, where $\boldsymbol{\mu}_{\mathbf{a}}^T = [\text{Re}(\boldsymbol{\mu}_1^{(1)})^T \,\text{Img}(\boldsymbol{\mu}_1^{(1)})^T]$ and $C_{\mathbf{aa}} = \frac{1}{2}\,\mathbf{F}(C_{\widehat{\boldsymbol{\beta}_1}^{(1)}\widehat{\boldsymbol{\beta}_1}^{(1)}})$ (see [30]).

The maximum output SINR can be further simplified to (III.25) in part (b) of Lemma 6 shown below.

Lemma 6: (b) The maximum output SINR is given by

$$\textit{max SINR}(\widehat{\boldsymbol{\beta}_1}^{(1)}) = \;\left(\mathbf{a}^T\,\mathbf{A}\,\mathbf{a} + \mathbf{a}^T\mathbf{g} + e\right) \tag{III.25}$$

$$\begin{aligned} \textit{where } \mathbf{A} &= 2\,\mathbf{F}(\mathbf{Y}^H C^{-1}\mathbf{Y}), \;\; e = 2\,\mathbf{h}^H C^{-1}\,\mathbf{h}, \\ \textit{and } \mathbf{g}^T &= 4\,\left[\mathrm{Re}(\mathbf{Y}^H C^{-1}\,\mathbf{h})^T,\; \mathrm{Img}(\mathbf{Y}^H C^{-1}\,\mathbf{h})^T\right]. \end{aligned}$$

Note that (III.25) is a quadratic form in real Gaussian random variables. The maximum SINR in (III.25) also minimizes the probability of error for the special case in (III.22).

With this choice of the optimal $\mathbf{w}$, the conditional probability of error in (III.19) and (III.22) can be simplified, and the unconditional probability of error can be obtained by numerical averaging using (III.19), and in a closed-form using (III.22) as shown in the next section.

III.D Probability of Error

For the special case when the cross correlation between the received pilot and the received data can be neglected, and using the optimal weights in Lemma 6 (a), we establish a closed-form expression for the unconditional probability of error given in Theorem 5.

Theorem 5 *For the case when the correlation between the NBI during the channel estimation phase and the data detection phase is zero, the probability of error is given by*

$$P_e = \frac{B(\frac{1}{2}, \frac{\nu+1}{2})}{2\,\pi\;\varrho^{\nu/2}} \left[\sum_{k=0}^{\infty} \xi_k \;{}_2F_1\left(k + \frac{\nu}{2}, \frac{\nu+1}{2}; \frac{\nu}{2}+1; \frac{-1}{\varrho}\right)\right] \qquad \text{(III.26)}$$

where $\nu = 2PM$, $B(x,y)$ is the beta function, and ${}_2F_1(\alpha_1, \alpha_2; b; z)$ is the hypergeometric function [74]. The coefficients ξ_k in (III.26) *are obtained recursively as*

$$\begin{aligned} \xi_o &= 1 \\ \xi_k &= \frac{1}{k}\sum_{r=0}^{k-1} \chi_{k-r}\;\; \xi_r, \qquad k \geq 1 \end{aligned}$$

$$\chi_k = \frac{1}{2}\sum_{j=1}^{\nu}\left(1 - \frac{\lambda_j}{\varrho}\right)^k - \frac{k}{8\;\varrho}\sum_{j=1}^{\nu}\frac{b_j^2}{\lambda_j}\left(1 - \frac{\lambda_j}{\varrho}\right)^{k-1}, \qquad k \geq 1. \qquad \text{(III.27)}$$

Let $\mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$ *be the eigenvalue decomposition (evd) of* $C_{\mathbf{aa}}^{1/2}\mathbf{A}C_{\mathbf{aa}}^{1/2}$*, and* $\mathbf{b} \triangleq \mathbf{U}^T C_{\mathbf{aa}}^{1/2}(\mathbf{g}+2\mathbf{A}\boldsymbol{\mu}_{\mathbf{a}})$*. Then in* (III.27)*,* b_j *is the* j*th component of the vector* $\mathbf{b}$*,* λ_j *is the* j*th diagonal element of the matrix* $\mathbf{\Lambda}$*, and* ϱ *is a positive constant such that* $\varrho > \lambda_{\max} > 0$*, where* $\lambda_{\max}$ *denotes the maximum eigenvalue.*

Proof. The proof is given in Appendix III.H.4. ■

III.E Numerical Results

The system parameter values considered for the numerical results are as follows: the signal and NBI bandwidths are 500 MHz and 5 MHz, respectively, the processing gain, N, is 50, the number of users (with equal received power), K, is 20, the number of multipath components, J, is 20, the signal-to-pilot ratio is 0 dB, the channel estimation period is over 50 chips, and signal-to-NBI (SNBI) ratio is fixed to -10 dB at $E_b/N_o = 0$ dB unless specified otherwise. Note that we fix the NBI power independent of the signal power, thus the signal-to-NBI ratio improves as signal power is increased. At the receiver, we consider a physically constrained linear array where the effective aperture of the array is fixed. For such an antenna array, increasing the number of receiver antenna elements leads to reduced inter-element spacing. This, in turn, causes higher correlation between the channel coefficients at each element. Also, for a fixed aperture array, the total power collected by the array is constant, irrespective of the number of elements [75]. For such an antenna array, an increase in the number of the elements is brought about by decreasing the size of each element [78]. In turn, the power captured by each element decreases with the decrease in its effective aperture. In our results, the total length of the antenna array is constrained to 2λ, where λ is the carrier wavelength.

Thus, for a fixed aperture antenna array, the power received by each antenna element reduces as more antenna elements are added to the array.[3] This effectively means that the SNR-per-element decreases. Therefore, during the channel estimation

[3]Note that in our results the total received power is kept fixed through normalization.

phase, with a larger number of antenna elements, there are more parameters to estimate and less power available to make each estimate. The normalized minimum mean square error (NMMSE) in estimation (normalized by MP) depends on the combination of two factors; pilot power-per-antenna element and fade correlation. When the number of antenna elements M is initially increased, the fade correlation is low, and the NMMSE increases due to reduced pilot power-per-antenna element. However, when the inter-element spacing is sufficiently reduced as M continues to increase, the NMMSE value remains largely the same. This is because the advantage of increased correlation is balanced out by a corresponding reduction in the pilot SNR. However, note also that the non-coherent combining loss keeps increasing as a larger number of these noisy estimates are combined.

Also, for a fixed length array, we know, from antenna theory [77], that the beamwidth or the spatial resolution of an antenna array depends on the length of the array. Consequently, the ability to discriminate and suppress interference through spatial selectivity, for a fixed length array, is potentially limited.

We assume an indoor operating environment. It has been shown through measurements that for an indoor channel (2–8 GHz band) the Non Line Of Sight (NLOS) component experiences wide-angle scattering, which is, in general, frequency dependent [72]. The data measurement results in [72] focused on a system bandwidth of 6 GHz (2–8 GHz band), and the wide angle-of-arrival for the NLOS link was shown to be particularly true for the 2.5–3 GHz and 6–8 GHz frequency intervals. For these frequency intervals of operation and a signal bandwidth of 0.5 GHz, as in our case, an assumption of wide angle scattering for the NLOS is justified. To capture this effect through a mathematically tractable model for spatial correlation, we employ the model of [28], according to which the $m_1\, m_2$th element of the spatial correlation matrix (say $\mathbf{Q}$) of each multipath component can be written as

$$(\mathbf{Q})_{m_1,m_2} = I_o\left(\sqrt{\kappa^2 - \frac{4\pi^2\Delta^2_{m_1,m_2}}{\lambda^2} + \frac{j4\pi\kappa\sin(\widetilde{\mu})\Delta_{m_1,m_2}}{\lambda}}\right) / \; I_o(\kappa) \qquad \text{(III.28)}$$

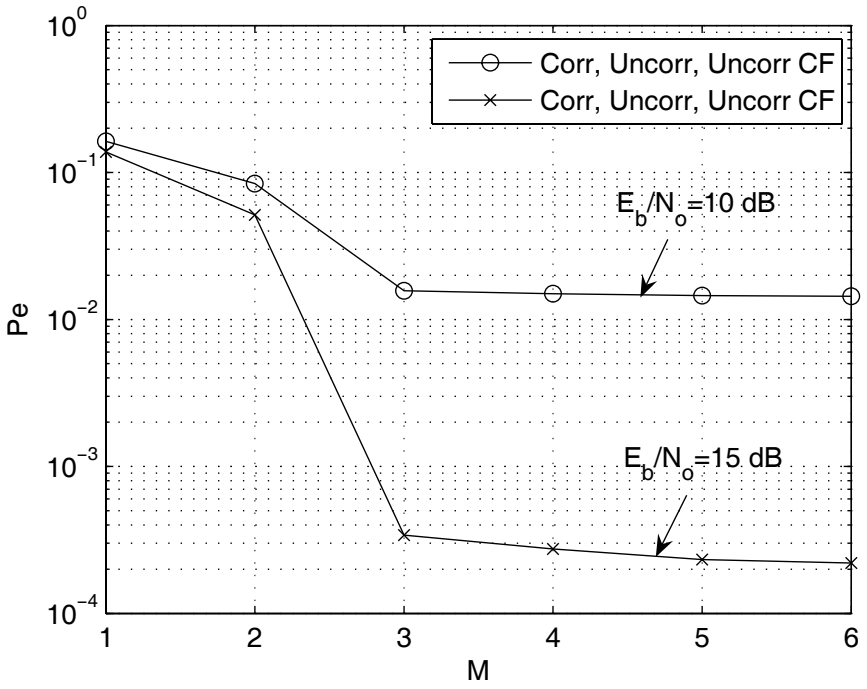

Figure III.2: BER plot for varying number of antenna elements M for fixed P and E_b/N_o.

where $m_1, m_2 = 1, 2, \ldots, M$, $I_o(.)$ is the modified Bessel function of the first kind and zeroth order, $\kappa\,(\geq 0)$ controls the width of the angle-of-arrival (AOA) and ranges from zero (isotropic scattering) to infinity (non-isotropic scattering), the parameter $\widetilde{\mu}(\in [-\pi, \pi))$ accounts for the mean direction of the AOA, and Δ_{m_1,m_2} is the spacing between the m_1th and m_2th antenna elements. A rich scattering environment is simulated by setting the parameter κ to small values (close to zero).

For the optimal combining scheme in Section III.C.2, the system performance for a varying number of antenna elements M and a fixed value of number of temporal taps $(P = 6)$ with E_b/N_o being a parameter is shown in Figure III.2. For a fixed P, as the number of antennas is initially increased, the system is able to suppress interference better through spatial filtering. However, as seen from Figure III.2, for the fixed length array, as M is continued to increase the improvement in performance is marginal. This is because, firstly, for fixed length array, interference suppression capability is limited,

and secondly, the continued increase in correlation as more antennas are added provides only small improvement in diversity. On the other hand, since the channel is estimated, the non-coherent combining loss increases as a larger number of the noisy estimates are combined, as mentioned earlier. Figure III.2 shows the net effect of these factors on the performance.

The labels 'Corr' and 'Uncorr', in Figure III.2, correspond to the average probability of error obtained through numerical averaging in (III.19) and (III.22), respectively, and 'Uncorr CF' corresponds to the closed-form expression in (III.26). Note that the performance for all these three cases is almost indistinguishable. This is shown in Figure III.2 by the curve with a circle that corresponds to $E_b/N_o = 10$ dB. Similarly, when $E_b/N_o = 15$ dB, the curve with a cross in Figure III.2 shows essentially the same performance for the three average BER expressions. The curves labeled 'Corr' and 'Uncorr' are very close because the optimum weights suppress the NBI component of the output test statistic, which in turn lowers the impact of temporal correlation between the NBI during channel estimation and during data detection on the performance. Also, as expected, curves 'Uncorr' and 'Uncorr CF' are very close, i.e., the results obtained through numerical averaging in (III.22) agree with the theoretical results predicted from (III.26).

Figure III.3 plots the performance when P is fixed at 2, and M takes on values $\{1, 4\}$. The SNBI is lowered to -25 dB. With a smaller number of spatial and temporal taps, as when $M = 1$ and $P = 2$ in Figure III.3, the system is unable to suppress the NBI. For this case, accounting for the temporal correlation between the NBI during estimation and data detection phases has a significant effect on the performance, in particular, the performance improves. This is evident from Figure III.3 where, for $M = 1$, the curve 'Corr' achieves a lower BER when compared to both 'Uncorr' and 'Uncorr CF'. However, as the number of antennas is increased to 4 and P is kept fixed at 2, the NBI is suppressed sufficiently because of which performance for the three cases is almost the same (shown by the curve with a cross in Figure III.3). Thus, the average BER predicted by the two expressions, namely (III.26) and numerical averaging of the condi-

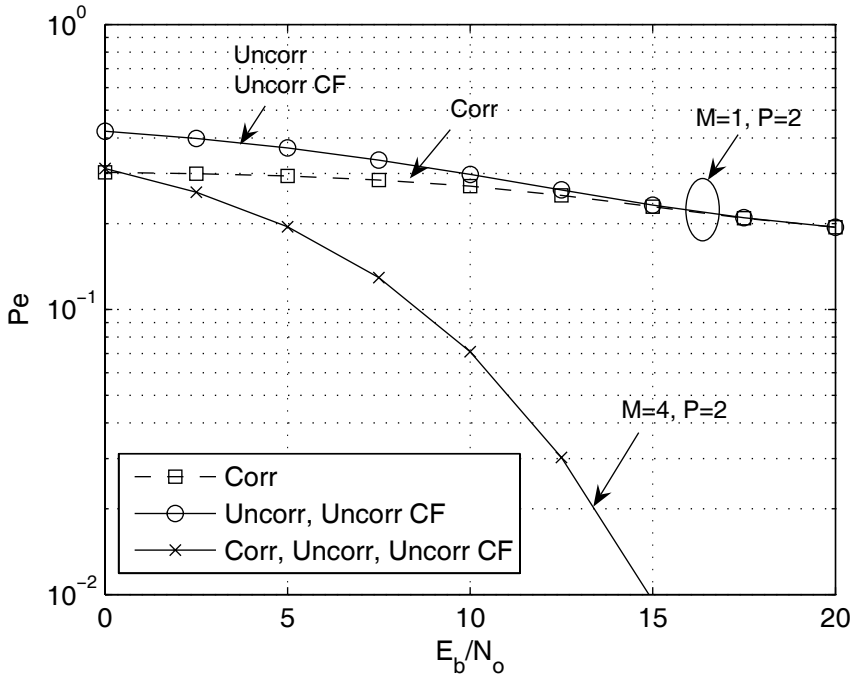

Figure III.3: BER plot for $P = 2$ and M takes on values $\{1, 4\}$. The SNBI is -25 dB.

tional probability of (III.19), would differ when there is insufficient NBI suppression.

For all the subsequent plots, the BER obtained from numerical averaging in (III.19) and (III.22), and in closed-form in (III.26), resulted in nearly the same performance. Thus, for clarity, we only present BER results from (III.26). Figure III.4 shows the probability of error when M is fixed ($M = 2$) while P is varied from 2 to 20. By increasing P, the system is able to capture additional multipath energy and suppress the interference better; however, the estimation quality degrades. This is because the lower SNR associated with the multipath coefficients with larger delays due to the exponentially decaying intensity profile results in higher estimation noise. For the system parameter values considered in Figure III.4, we observe that the performance improves as P is increased, indicating that the enhanced interference suppression with larger P dominates the performance.

Figure III.5 plots the performance when the product of the number of antennas and temporal taps is fixed at 12. This also fixes the maximum diversity of the system.

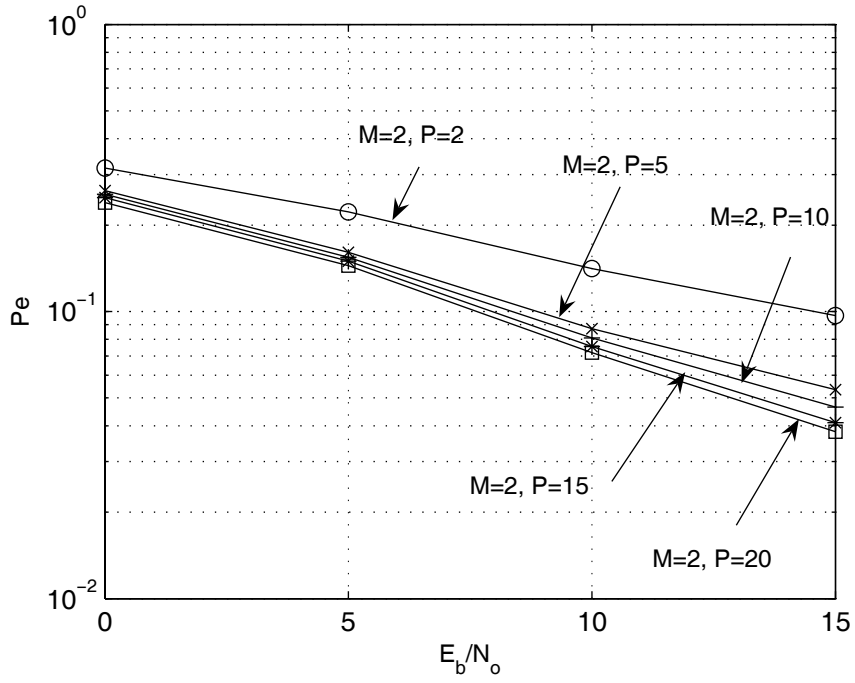

Figure III.4: BER plot for varying number of temporal taps P for fixed M.

Note from Figure III.5, that the configuration $M = 3, P = 4$ performs the best. Thus, there is an optimal combination of the number of antennas and the temporal taps that achieves the best performance, and this optimal combination depends on the signal-to-interference ratio and the estimation quality. Also note in Figure III.5 the crossover by the curves $M = 6, P = 2$ and $M = 2, P = 6$. As mentioned earlier, signal-to-NBI ratio improves as signal power is increased. This implies that NBI is more relevant in the lower SNR regime. For the higher SNR values, the performance is mainly limited by the MAI. Therefore, as SNR is increased, the configuration $M = 6, P = 2$ has an advantage over $M = 2, P = 6$ mainly because the larger number of antenna elements leads to enhanced MAI suppression.

Figure III.6 plots the performance for the same set of parameters as in Figure III.5, except that now ISI is taken into account. The dashed curves in Figure III.6 are for the case when ISI is considered, whereas the solid curves corresponds to when ISI is neglected. Since we have an interference-limited UWB system, where the presence

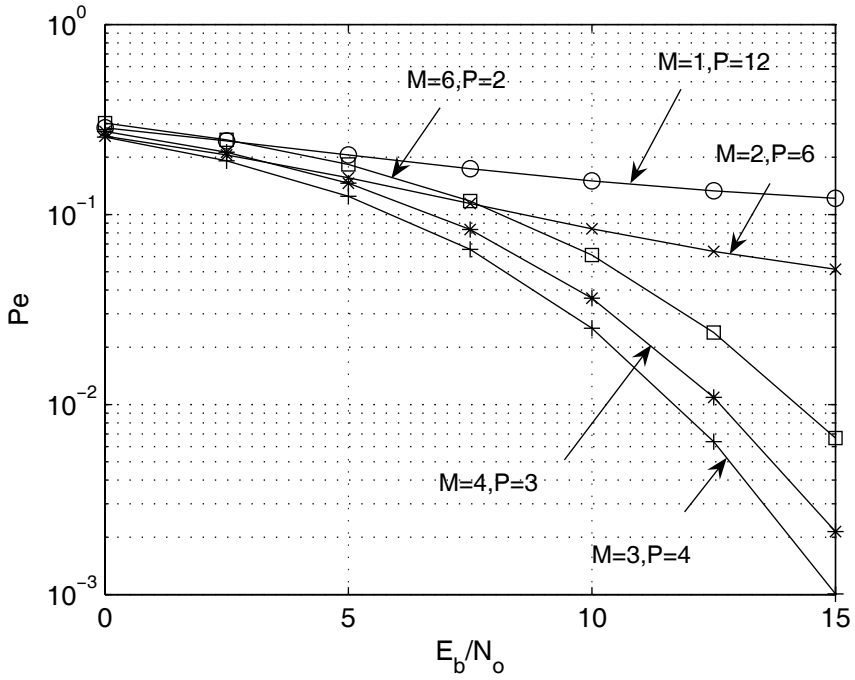

Figure III.5: BER plot for the case when the product of M times P is fixed to 12.

of both narrowband interference and multiple access interference, in addition to channel estimation errors, limit the performance, the impact of ISI is not significant. This is evident from Figure III.6.

III.F Conclusion

In this work, the error rate performance of a multiple antenna DS-CDMA UWB system was analyzed under the conditions of imperfect channel estimation and the presence of NBI and MAI. An optimal spatio-temporal receiver was designed that first forms an estimate of the channel in the presence of NBI and MAI, and then uses it to optimally combine the received signal. An exact closed-form expression for the probability of error for the special case when the correlation between the NBI during the channel estimation phase and the data detection phase is negligible was derived. Spatial processing, in addition to temporal processing, was shown to improve the system performance. The effects of channel estimation accuracy, interference suppression, and

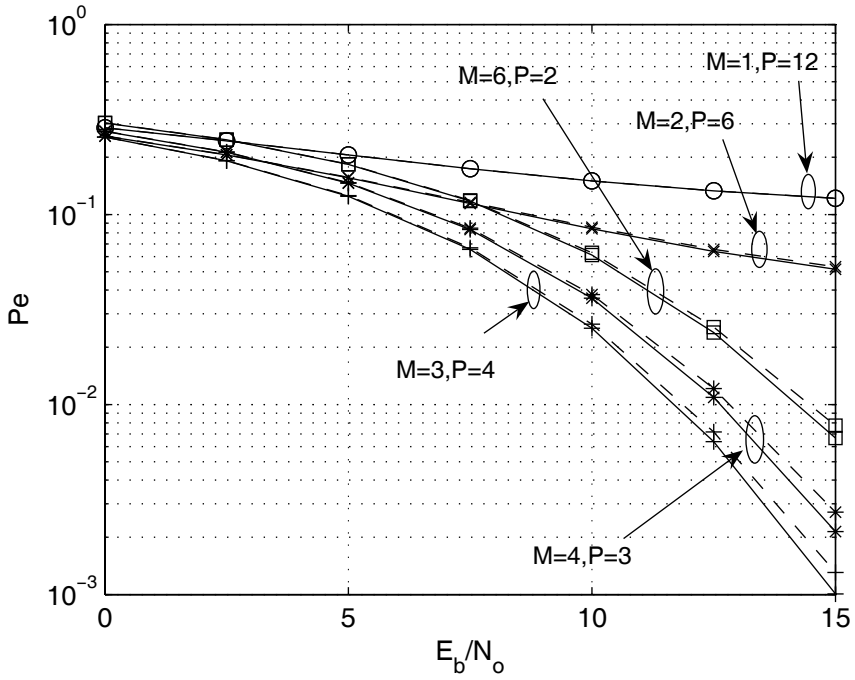

Figure III.6: BER plot when the product of M times P is fixed to 12 for both the cases when ISI is neglected (solid lines) and when it is taken into account (dashed lines).

spatial correlation on the performance were studied. It was shown that, for a fixed total diversity, there is an optimal combination of the number of antennas and the number of temporal taps in order to achieve the best performance, and that the optimal combination depends on the signal-to-interference ratio and the quality of the estimates.

III.G Acknowledgement

This chapter has been published, in part, in the following publications:

1. P. Nagvanshi, E. Masry and L.B. Milstein, "Performance Analysis of Multiple Antenna DS-CDMA UWB System with Noisy Channel Estimates and Narrow-Band Interference ," *IEEE Military Communications Conference (MILCOM)*, pp. 1-7, Oct. 2006.

2. P. Nagvanshi, E. Masry and L.B. Milstein, “Performance of Multiple Antenna DS-CDMA UWB System With Noisy Channel Estimates and Narrow-Band Interference ,” *IEEE Journal of Selected Topics in Signal Processing,* vol. 1, no. 3, pp. 470 - 482, Oct. 2007.

The dissertation author was the primary researcher and author, and the co-authors listed in these publications directed and supervised the research which forms the basis for this chapter.

III.H Appendices for Chapter III

III.H.1 Proof of Theorem 4

The probability of error in (III.19) is obtained from (III.18) by simplifying the conditional means and variances that appear in the expression in (III.18). From the definition of ν_1 (see after (III.18)), we have $E[\nu_1|\widehat{\boldsymbol{\beta}_1}^{(1)}] = E[x_1|\widehat{\boldsymbol{\beta}_1}^{(1)}]+E[x_2+x_3|\widehat{\boldsymbol{\beta}_1}^{(1)}]$, where $E[x_4|\widehat{\boldsymbol{\beta}_1}^{(1)}] = 0$, since the thermal noise component x_4 is independent of $\widehat{\boldsymbol{\beta}_1}^{(1)}$ and has zero mean. Consider

$$\begin{aligned} & E\Big[x_1\Big|\widehat{\boldsymbol{\beta}_1}^{(1)}\Big] \\ &= E\left[\mathrm{Re}\left\{\sqrt{P_1}\mathbf{w}^H\,\mathbf{D}_{1,l}\,\boldsymbol{\beta}_1\right\}\Big|\widehat{\boldsymbol{\beta}_1}^{(1)}\right] \\ &= E\left[\mathrm{Re}\left\{\sqrt{P_1}\mathbf{w}^H\,\Big(\,\hat{\mathbf{D}}\,\boldsymbol{\beta}_1^{(1)} + \,\mathbf{D}_{1,l}^{(2)}\,\boldsymbol{\beta}_1^{(2)}\Big)\right\}\Big|\widehat{\boldsymbol{\beta}_1}^{(1)}\right] \\ &= \mathrm{Re}\Bigg\{\sqrt{P_1}\mathbf{w}^H\,\Big(\,\hat{\mathbf{D}}\,E\big[\boldsymbol{\beta}_1^{(1)}\big|\widehat{\boldsymbol{\beta}_1}^{(1)}\big] \\ &\qquad\qquad + \,\mathbf{D}_{1,l}^{(2)}\,E\big[\boldsymbol{\beta}_1^{(2)}\big|\widehat{\boldsymbol{\beta}_1}^{(1)}\big]\,\Big)\Bigg\}, \end{aligned} \tag{III.29}$$

where the first equality in (III.29) is obtained by substituting for x_1 defined after (III.17), and the second equality is obtained by using the decomposition of $\mathbf{D}_{1,l}\,\boldsymbol{\beta}_1$ in (III.14). Since $\boldsymbol{\beta}_1^{(1)}, \boldsymbol{\beta}_1^{(2)}$ and $\widehat{\boldsymbol{\beta}_1}^{(1)}$ are jointly Gaussian, the conditional means in the last equality in (III.29) can be further simplified using the relation $E[\mathbf{x}|\mathbf{y}] = E[\mathbf{x}] + C_{\mathbf{xy}}C_{\mathbf{yy}}^{-1}(\mathbf{y} - E[\mathbf{y}])$ for jointly Gaussian $\mathbf{x}$ and $\mathbf{y}$, from [30]. The expression obtained after carrying out this simplification is the expression for the mean m_1 in (III.20). Now consider $\mathbf{q}_1 \triangleq (\sum_{k=2}^{K}\mathbf{D}_{k,l}\,\boldsymbol{\beta}_k + \boldsymbol{\gamma}_l^{NBI})$ and the definition of x_2 and x_3 after (III.17). We have

$$E\big[(x_2+x_3)\big|\widehat{\boldsymbol{\beta}_1}^{(1)}\big] = \mathrm{Re}\big(\mathbf{w}^H\,E\big[\mathbf{q}_1\big|\widehat{\boldsymbol{\beta}_1}^{(1)}\big]\big). \tag{III.30}$$

The mean m_2 in (III.20) is obtained from (III.30) after simplifying the conditional mean in (III.30) using the joint Gaussianity approach mentioned above.

To obtain a simplified expression for the conditional variances, we first define the following quantities: $\mathbf{q}_2 \triangleq [\sqrt{P_1}\,\hat{\mathbf{D}}\,(\boldsymbol{\beta}_1^{(1)} - \widehat{\boldsymbol{\beta}_1}^{(1)}) + \mathbf{D}_{1,l}^{(2)}\,\boldsymbol{\beta}_1^{(2)}]$, $\widetilde{\mathbf{p}} \triangleq \mathbf{q}_1 + \mathbf{q}_2 +$

$\boldsymbol{\gamma}_l^{Therm}$, and $\hat{\mathbf{p}} \triangleq \mathbf{q}_1 - \mathbf{q}_2 + \boldsymbol{\gamma}_l^{Therm}$. Then we have

$$\begin{aligned} \sigma_1^2 &= \sigma^2_{\nu_1|\widehat{\boldsymbol{\beta}}_1^{(1)}} &&= E\Big[\big(\nu_1 - E[\nu_1|\widehat{\boldsymbol{\beta}}_1^{(1)}]\big)^2\Big|\widehat{\boldsymbol{\beta}}_1^{(1)}\Big] \\ &&&= \tfrac{1}{2}(\mathbf{w}^H\, C_{\widetilde{\mathbf{p}}|\widehat{\boldsymbol{\beta}}_1^{(1)}}\mathbf{w}), \\ \text{and}\quad \sigma_2^2 &= \sigma^2_{\nu_2|\widehat{\boldsymbol{\beta}}_1^{(1)}} &&= E\Big[\big(\nu_2 - E[\nu_2|\widehat{\boldsymbol{\beta}}_1^{(1)}]\big)^2\Big|\widehat{\boldsymbol{\beta}}_1^{(1)}\Big] \\ &&&= \tfrac{1}{2}(\mathbf{w}^H\, C_{\hat{\mathbf{p}}|\widehat{\boldsymbol{\beta}}_1^{(1)}}\mathbf{w}). \end{aligned} \tag{III.31}$$

As $\hat{\mathbf{p}}$, $\widetilde{\mathbf{p}}$ and $\widehat{\boldsymbol{\beta}}_1^{(1)}$ are jointly Gaussian, the property $C_{\mathbf{x}|\mathbf{y}} = C_{\mathbf{xx}} - C_{\mathbf{xy}}C_{\mathbf{yy}}^{-1}C_{\mathbf{yx}}$ from [30] can be used to further simplify the conditional covariances $C_{\widetilde{\mathbf{p}}|\widehat{\boldsymbol{\beta}}_1^{(1)}}$ and $C_{\hat{\mathbf{p}}|\widehat{\boldsymbol{\beta}}_1^{(1)}}$ in (III.31). Upon doing so, the conditional variances σ_1^2 and σ_2^2 in (III.31) can be expressed as in (III.21), and the covariance matrices C_1, C_2, C_3 and C_4, in (III.21), can now be identified as: $C_1 = C_{\widetilde{\mathbf{p}}\widetilde{\mathbf{p}}}$, $C_2 = C_{\widetilde{\mathbf{p}}\widehat{\boldsymbol{\beta}}_1^{(1)}}$, $C_3 = C_{\hat{\mathbf{p}}\hat{\mathbf{p}}}$ and $C_4 = C_{\hat{\mathbf{p}}\widehat{\boldsymbol{\beta}}_1^{(1)}}$. An expression for these matrices is given in Appendix III.H.2.

III.H.2 Covariance Matrices

Matrices C_1, C_2, C_3 and C_4 are given by: $C_1 = \boldsymbol{\Psi} - \boldsymbol{\Upsilon}$, $C_2 = \boldsymbol{\Lambda} + \boldsymbol{\Gamma}$, $C_3 = \boldsymbol{\Psi} + \boldsymbol{\Upsilon}$, and $C_4 = \boldsymbol{\Lambda} - \boldsymbol{\Gamma}$, where $\boldsymbol{\Upsilon} = \sqrt{P_1}\,(\hat{\mathbf{D}}\boldsymbol{\Lambda}^H + \boldsymbol{\Lambda}\hat{\mathbf{D}}^T)$, and $\boldsymbol{\Gamma}$ and $\boldsymbol{\Psi}$ are defined by

$$\begin{aligned} \boldsymbol{\Gamma} &\triangleq \sqrt{P_1 P_p}\,\mathbf{D}_{1,l}^{(2)}\,\boldsymbol{\Sigma}_1^{(2)}\,(\widetilde{\mathbf{D}}_{1,\widetilde{l}}^{(2)})^T\,\mathbf{G}, \\ \boldsymbol{\Psi} &\triangleq \mathbf{A}_1 + \mathbf{A}_2 + \mathbf{A}_3 + \mathbf{A}_4 - \mathbf{A}_5 + \mathbf{A}_6, \end{aligned} \tag{III.32}$$

respectively. In the above, $\mathbf{A}_5 = \sqrt{P_1}\,(\hat{\mathbf{D}}\,\boldsymbol{\Gamma}^H + \boldsymbol{\Gamma}\,\hat{\mathbf{D}}^T)$, and the matrices appearing in the sum in (III.32) are defined as follows:

$$\begin{aligned} \mathbf{A}_1 &\triangleq P_1\ \hat{\mathbf{D}}\,C_{\widetilde{\boldsymbol{\beta}}_1^{(1)}\widetilde{\boldsymbol{\beta}}_1^{(1)}}\,\hat{\mathbf{D}}^T, \\ \mathbf{A}_2 &\triangleq \sum_{q=lN}^{lN+N-1}\ \sum_{\hat{q}=lN}^{lN+N-1} c_{1,q}\,c_{1,\hat{q}}\,\mathbf{R}_{q,\hat{q}}, \\ \mathbf{A}_3 &\triangleq \big(\mathbf{B}\otimes(\mathbf{I}_P\otimes\mathbf{c}_{1,l}^T)\boldsymbol{\Theta}_{NBI}(\mathbf{I}_P\otimes\mathbf{c}_{1,l})\big) \\ \mathbf{A}_4 &\triangleq P_1\ \mathbf{D}_{1,l}^{(2)}\,\boldsymbol{\Sigma}_1^{(2)}\,(\mathbf{D}_{1,l}^{(2)})^T \\ \boldsymbol{\Omega}_{l,\widetilde{l}} &\triangleq (\mathbf{I}_{MP}\otimes\widetilde{\mathbf{c}}_{1,\widetilde{l}}^T)\,\widetilde{\boldsymbol{\Omega}}_{l,\widetilde{l}}\,(\mathbf{I}_{MP}\otimes\mathbf{c}_{1,l}) \end{aligned} \tag{III.33}$$

where $\widetilde{\boldsymbol{\beta}_1^{(1)}} = \boldsymbol{\beta}_1{}^{(1)} - \widehat{\boldsymbol{\beta}_1^{(1)}}$ is the channel estimation error vector, and $C_{\widetilde{\boldsymbol{\beta}_1^{(1)}}\widetilde{\boldsymbol{\beta}_1^{(1)}}}$ denotes the estimation error covariance matrix. Also, $\mathbf{B}$ is the $M \times M$ spatial correlation matrix of the NBI. The vectors $\mathbf{c}_{1,l}$ and $\widetilde{\mathbf{c}}_{1,\widetilde{l}}$, in (III.33), are of dimensions $(N \times 1)$ and $(LN \times 1)$, respectively, and are defined as $\mathbf{c}_{1,l}^T \triangleq [c_{1,lN}, c_{1,lN+1}, \cdots, c_{1,lN+N-1}]$ and $\widetilde{\mathbf{c}}_{1,\widetilde{l}}^T \triangleq [c_{1,\widetilde{l}N}, c_{1,\widetilde{l}N+1}, \cdots, c_{1,\widetilde{l}N+LN-1}]$.

Matrix $\boldsymbol{\Xi}_{l,\widetilde{l}}$ has the same definition as of $\mathbf{A}_2$, except the first sum is changed to $\sum_{q=\widetilde{l}N}^{\widetilde{l}N+LN-1}$, and $\mathbf{A}_6$, the covariance of the thermal noise component, is obtained from $\mathbf{A}_3$ by setting $\mathbf{B}$ to $\mathbf{I}_M$, and $\boldsymbol{\Theta}_{NBI}$ to $\boldsymbol{\Theta}_{Therm}$.

Matrices $\mathbf{R}_{q,\hat{q}}$ and $\widetilde{\boldsymbol{\Omega}}_{l,\widetilde{l}}$ are of dimension $MP \times MP$ and $PLNM \times PNM$, respectively, whereas $\boldsymbol{\Theta}_s$, where the subscript 's' stands for either 'NBI or 'Therm', is of dimension $PN \times PN$. These are given by

$$\begin{aligned}
(\boldsymbol{\Theta}_s)_{i,j} &= \widetilde{\theta}_s((g(i) - g(j))T_c) \\
(\mathbf{R}_{q,\hat{q}})_{i,j} &= \sum_{\hat{k}=(g_1(i)-1)J+1}^{g_1(i)J} \sum_{\widetilde{k}=(g_1(j)-1)J+1}^{g_1(j)J} (\mathbf{H})_{\widetilde{k},\hat{k}} \\
&\quad \times \ \delta_{q-\hat{q}-u(j,P)+u(i,P)+u(\hat{k},J)-u(\widetilde{k},J)} \\
(\widetilde{\boldsymbol{\Omega}}_{l,\widetilde{l}})_{i,j} &= (\mathbf{B})_{v(i,PNL),v(j,PN)} \\
&\quad \times \ \widetilde{\theta}_{NBI}\Big(\big((\widetilde{l}-l)N + g(i,LN) - g(j,N)\big)T_c\Big)
\end{aligned}$$

respectively, where $\mathbf{H} = \sum_{k=2}^{K} P_k(\boldsymbol{\Sigma}_k + \boldsymbol{\mu}_k\boldsymbol{\mu}_k^H)$. The autocorrelation functions are given by

$$\widetilde{\theta}_{NBI}(\alpha T_c) = J_o\frac{Sin\pi W\alpha T_c}{\pi\alpha} \quad \text{and} \quad \widetilde{\theta}_{Therm}(\alpha T_c) = N_o\delta_\alpha.$$

Also, $u(i,A) = (i-1)_{\text{mod}\,A}$, $v(i,A) = \lceil \frac{i}{A} \rceil$, $g(i) = u(i,N) + v(i,N)$, $g_1(i) = v(i,P)$, and $g(i,A) = u(i,A) + u(v(i,A),P)$.

III.H.3 Optimum Weights

Without loss of generality, we define $\tilde{\mathbf{x}} \triangleq C^{1/2}\mathbf{w}$ and simplify the conditional SINR in (III.23) to the quantity on the left side of the inequality below:

$$2\left(\text{Re}\,(\tilde{\mathbf{x}}^H C^{-1/2}\widetilde{\mathbf{b}})\right)^2/(\tilde{\mathbf{x}}^H\tilde{\mathbf{x}}) \ \leqslant \ 2\,\widetilde{\mathbf{b}}^H C^{-1}\widetilde{\mathbf{b}} \tag{III.34}$$

where $\widetilde{\mathbf{b}} \triangleq (\mathbf{h} + \mathbf{Y}\,\widehat{\boldsymbol{\beta}_1}^{(1)})$. The upper bound in (III.34) follows from the CS inequality [32, eq. (5.1.5)]. The equality in (III.34) holds iff $\tilde{\mathbf{x}} = \alpha\, C^{-1/2}\widetilde{\mathbf{b}}$, [32, page 261]. Thus, the optimum weight $\tilde{\mathbf{x}}$, or equivalently $\mathbf{w}$ that achieves the bound in (III.34) is $\mathbf{w} = \alpha\, C^{-1}\widetilde{\mathbf{b}}$. The real scalar α in the expression for $\mathbf{w}$ can be dropped as it does not affect the output SINR.

III.H.4 Proof of Theorem 5

The conditional probability of error expression for the special case is given in (III.22). Using the optimal weights in Lemma $1\,(a)$, the average probability of error for this case can be written as

$$P_e = \int_0^\infty Q(\sqrt{x}) f_{\max \mathrm{SINR}}(x)\,\mathrm{d}x \tag{III.35}$$

where $f_{\max \mathrm{SINR}}(x)$ is the probability density function (pdf) for the maximum output SINR (see (III.25)). Using the Craig's formula for $Q(.)$, the P_e can be simplified to

$$P_e = \frac{1}{\pi}\int_0^\infty \left(\int_0^{\frac{\pi}{2}} e^{\frac{-x}{2\sin^2\theta}}\ \mathrm{d}\theta \right) f_{\max \mathrm{SINR}}(x)\,\mathrm{d}x. \tag{III.36}$$

Since the integrand is positive in (III.36), the order of integration can be swapped and P_e can be written as

$$\begin{aligned} P_e &= \frac{1}{\pi}\int_0^{\frac{\pi}{2}} \int_0^\infty e^{\frac{-x}{2\sin^2\theta}}\ f_{\max \mathrm{SINR}}(x)\,\mathrm{d}x\,\mathrm{d}\theta \\ &= \frac{1}{\pi}\int_0^{\frac{\pi}{2}} \varphi_{\max \mathrm{SINR}}\left(\frac{-1}{2\sin^2\theta}\right)\,\mathrm{d}\theta \end{aligned} \tag{III.37}$$

where $\varphi_{\max \mathrm{SINR}}(x)$ is the moment generating function (mgf) of the max SINR. As seen in (III.25), the max SINR is quadratic in the Gaussian random variables. From [73], the mgf of a *general* quadratic form denoted as $Q(\widetilde{\mathbf{a}}) = \left(\widetilde{\mathbf{a}}^T\ \widetilde{\mathbf{A}}\widetilde{\mathbf{a}} + \widetilde{\mathbf{a}}^T\widetilde{\mathbf{g}} + \widetilde{e}\right)$, is given by

$$\begin{aligned} \varphi_{Q(\widetilde{\mathbf{a}})}(s) &= exp\Bigg\{ s\left(\boldsymbol{\mu}_{\widetilde{\mathbf{a}}}^T\widetilde{\mathbf{A}}\boldsymbol{\mu}_{\widetilde{\mathbf{a}}} + \boldsymbol{\mu}_{\widetilde{\mathbf{a}}}^T\widetilde{\mathbf{g}} + \widetilde{e}\right) \\ &\quad + \frac{s^2}{2}\sum_{j=1}^{\nu} b_j^2\left(1-2s\lambda_j\right)^{-1} \Bigg\} \prod_{j=1}^{\nu}\left(1-2s\lambda_j\right)^{-1/2} \end{aligned}$$

for all real s such that for λ_j with odd multiplicity

$$s < \frac{1}{2\lambda_j}. \tag{III.38}$$

In (III.38) $\widetilde{\mathbf{a}}$ is defined as $\widetilde{\mathbf{a}} \sim N(\boldsymbol{\mu}_{\widetilde{\mathbf{a}}}, C_{\widetilde{\mathbf{a}}\widetilde{\mathbf{a}}})$ such that $C_{\widetilde{\mathbf{a}}\widetilde{\mathbf{a}}} > 0$ and the constant matrix $\widetilde{\mathbf{A}}$ is symmetric, i.e., $\widetilde{\mathbf{A}} = \widetilde{\mathbf{A}}^T$. Let $\mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^T$ be the evd of $C_{\widetilde{\mathbf{a}}\widetilde{\mathbf{a}}}^{1/2}\widetilde{\mathbf{A}}C_{\widetilde{\mathbf{a}}\widetilde{\mathbf{a}}}^{1/2}$, and $\mathbf{b} \stackrel{\Delta}{=} \mathbf{U}^T C_{\widetilde{\mathbf{a}}\widetilde{\mathbf{a}}}^{1/2}(\widetilde{\mathbf{g}} + 2\widetilde{\mathbf{A}}\boldsymbol{\mu}_{\widetilde{\mathbf{a}}})$. Then in (III.38), b_j is the jth component of the vector $\mathbf{b}$, and λ_j is the jth diagonal element of the matrix $\boldsymbol{\Lambda}$. The constant ν is the rank of the matrix $C_{\widetilde{\mathbf{a}}\widetilde{\mathbf{a}}}^{1/2}\widetilde{\mathbf{A}}C_{\widetilde{\mathbf{a}}\widetilde{\mathbf{a}}}^{1/2}$.

Note that the max SINR has the same form as the general quadratic $Q(\widetilde{\mathbf{a}})$ after the following assignments: $\widetilde{\mathbf{A}} = \mathbf{A}$, $\widetilde{\mathbf{g}} = \mathbf{g}$, $\widetilde{e} = e$, $\boldsymbol{\mu}_{\widetilde{\mathbf{a}}} = \boldsymbol{\mu}_{\mathbf{a}}$, and $C_{\widetilde{\mathbf{a}}\widetilde{\mathbf{a}}} = C_{\mathbf{aa}}$, where the covariance matrix, namely $C_{\mathbf{aa}}$, of the Gaussian vector $\mathbf{a}$, and the matrix $\mathbf{A}$ appearing in the SINR expression are both positive definite.

Based on the limits of the integration in (III.37), the mgf is evaluated over the range $s \in (-\infty, -1/2]$. Notice that for this range of s, the expression for the mgf in (III.38) is valid and can be used to simplify the integral in (III.37). Thus, we have the probability of error (Pe) expression in (III.37), which is expressed as a single integral over the mgf obtained from (III.38). Notice that the Pe expression as given in (III.37) is not in closed-form and requires numerical integration. We are thus motivated to obtain a closed-form expression for Pe.

Using the following series expansion

$$ln(1+x) = \sum_{k=1}^{\infty}(-1)^{k+1}\frac{x^k}{k} \quad \text{and} \quad (1-x)^{-1} = \sum_{k=0}^{\infty} x^k \quad \text{valid for } |x| < 1 \text{(III.39)}$$

and adopting the approach in [73], we will show that mgf in (III.38) can be expressed as a series. This particular form for the mgf is useful because when applied to the integral equation in (III.37), and after swapping the order of integration and summation (justified later), integration over each term of the series can be performed. This, thus, allows us to write a closed-form expression for Pe as a series. We next present the series result for the mgf as given in the following lemma.

Lemma 7 *By defining the constant* $\eta \stackrel{\Delta}{=} (\boldsymbol{\mu}_{\widetilde{\mathbf{a}}}^T\widetilde{\mathbf{A}}\boldsymbol{\mu}_{\widetilde{\mathbf{a}}} + \boldsymbol{\mu}_{\widetilde{\mathbf{a}}}^T\widetilde{\mathbf{g}} + \widetilde{e} - \frac{1}{4}\mathbf{b}^T\boldsymbol{\Lambda}^{-1}\mathbf{b})$ *and assuming that* $\lambda_j > 0$, $\forall j$, *the mgf of the random variable* $Q(\widetilde{\mathbf{a}}) - \eta$ *can be written as*

$$\varphi_{Q(\widetilde{\mathbf{a}})-\eta}(s) \;=\; \sum_{k=0}^{\infty} \xi_k \left(\frac{(-2s\varrho)^k}{(1-2s\varrho)^{k+\nu/2}}\right) \qquad \text{(III.40)}$$

where the coefficients ξ_k *are given in* (III.27) *and* ϱ *is a positive constant satisfying* $\varrho > \lambda_{\max} > 0$. *The expression in* (III.40) *is valid for any real* s, *such that*

$$s < \frac{1}{2\varrho\,(1+\vartheta)} \tag{III.41}$$

where $\vartheta \triangleq \max_j \left(\left|1 - \frac{\lambda_j}{\varrho}\right|\right)$. *With* ϱ *satisfying* $\varrho > \lambda_{\max} > 0$, *we have* $0 < \vartheta < 1$.

Proof. The proof is given in Appendix III.H.5. ∎

Based on the limits of the integral in (III.37), we have the mgf of the max SINR evaluated over the interval $s \in (-\infty, -1/2]$. Clearly, this interval satisfies the condition in (III.41), thus, the mgf in (III.40) can be applied to the integral in (III.37). With the substitutions $\widetilde{\mathbf{A}} = \mathbf{A}$, $\widetilde{\mathbf{g}} = \mathbf{g}$, $\widetilde{e} = e$, $\boldsymbol{\mu}_{\widetilde{\mathbf{a}}} = \boldsymbol{\mu}_{\mathbf{a}}$, and $C_{\widetilde{\mathbf{a}}\widetilde{\mathbf{a}}} = C_{\mathbf{aa}}$ in (III.40), the mgf of the max SINR can be written as

$$\varphi_{\max \mathrm{SINR}}(s) \;=\; \sum_{k=0}^{\infty} \xi_k \left(\frac{(-2s\varrho)^k}{(1-2s\varrho)^{k+\nu/2}}\right) e^{s\eta}. \tag{III.42}$$

Due to the structures of the matrix $\mathbf{A}$, vector $\mathbf{g}$, and scalar e (see after (III.25)) we show later in this Appendix that η is zero in our case. Thus, we can write the mgf for the max SINR as

$$\varphi_{\max \mathrm{SINR}}(s) \;=\; \sum_{k=0}^{\infty} \xi_k \left(\frac{(-2s\varrho)^k}{(1-2s\varrho)^{k+\nu/2}}\right). \tag{III.43}$$

Using (III.43), we simplify the P_e integral in (III.37).

$$\begin{aligned}
P_e &= \frac{1}{\pi}\int_0^{\frac{\pi}{2}} \sum_{k=0}^{\infty} \xi_k \left(\varrho/\sin^2\theta\right)^k \left(1+\left(\varrho/\sin^2\theta\right)\right)^{-k-\nu/2} \mathrm{d}\theta \\
&= \frac{1}{\pi}\int_0^{\frac{\pi}{2}} \sum_{k=0}^{\infty} \xi_k\, \varrho^k \left(\frac{(\sin^2\theta)^{\nu/2}}{(\sin^2\theta+\varrho)^{k+\nu/2}}\right) \mathrm{d}\theta \\
&= \frac{1}{\pi}\sum_{k=0}^{\infty} \xi_k\, \varrho^k \left[\int_0^{\frac{\pi}{2}} \left(\frac{(\sin^2\theta)^{\nu/2}}{(\sin^2\theta+\varrho)^{k+\nu/2}}\right) \mathrm{d}\theta\right] \\
&= \frac{1}{\pi}\sum_{k=0}^{\infty} \xi_k\, \varrho^k \; [I_{1,k}]
\end{aligned} \tag{III.44}$$

where $I_{1,k}$ is defined as the integral inside the square brackets in the third equality in the set of equations in (III.44) and can be further simplified by substituting $(\sin^2\theta) = z$.

The integral then reduces to

$$\begin{aligned} I_{1,k} &= \frac{1}{2}\int_0^1 (z+\varrho)^{-(k+\nu/2)}\ (1-z)^{-1/2}\, z^{(\nu-1)/2}\ \mathrm{d}z \\ &= \frac{1}{2}B\left(\frac{1}{2}, \frac{\nu+1}{2}\right)\ \varrho^{-(k+\nu/2)} \\ &\quad \times\ {}_2F_1\left(k+\frac{\nu}{2}, \frac{\nu+1}{2}; \frac{\nu}{2}+1; \frac{-1}{\varrho}\right). \end{aligned} \tag{III.45}$$

The expression in (III.45) is obtained by using the formula [74, (3.383, eq. (8), page 315)]. Upon substituting (III.45) in (III.44), we obtain the desired expression for P_e in (III.26).

We next show that the interchange of the order of summation and integration in the third equality in (III.44) is justified. In that equation, since the range of integration is finite, it suffices to show that the series converges uniformly in θ, i.e., we need to show that

$$\begin{aligned} I_2 &\triangleq \frac{1}{\pi}\sum_{k=0}^{\infty} |\,\xi_k|\ (\varrho/\sin^2\theta)^k \left(1+(\varrho/\sin^2\theta)\right)^{-k-\nu/2} \\ &\leq \tilde{L} < \infty \\ &\text{where}\ \ \tilde{L}\ \ \text{is independent of}\ \theta \in \left[0, \frac{\pi}{2}\right]. \end{aligned} \tag{III.46}$$

In order to establish a bound on I_2, we first bound the coefficients ξ_k. In [76], the authors gave a bound using Cauchy's inequality for the coefficients of a general power series. We use this bound for the coefficients ξ_k corresponding to the power series in (III.43), as given in the following lemma.

Lemma 8 *As suggested in [76], the coefficients ξ_k can be bounded using Cauchy's inequality, i.e.,*

$$|\xi_k| \ \leq\ \zeta(\gamma)\ \gamma^{-k} \qquad \textit{where}\ \ \zeta(\gamma) \triangleq \left(\max_{|\mu|=\gamma}\ |f(\mu)|\right). \tag{III.47}$$

where γ is a positive constant such that $1 < \gamma < 1/\vartheta$. The function $f(\mu)$ as defined in (III.61) *is used to obtain a bound for $\zeta(\gamma)$. The function $\zeta(\gamma)$ is bounded as*

$$\zeta(\gamma) \leq B(\gamma) = exp\left\{\frac{\gamma}{8\,\varrho}\ \sum_{j=1}^{\nu}\frac{b_j^2}{\lambda_j}(1+\gamma-\frac{\lambda_j\gamma}{\varrho})^{-1}\right\}\prod_{j=1}^{\nu}\left(1-\gamma+\frac{\lambda_j\gamma}{\varrho}\right)^{-1/2} \tag{III.48}$$

Proof. The proof is given in Appendix III.H.6. ∎

Note that $B(\gamma)$ is finite and independent of k. The bound of the form (III.47) can also be obtained when $0 < \gamma \leq 1$, however, this is not useful for our case. Using (III.47), the series in (III.46) can now be bounded as

$$\begin{aligned} I_2 &\leq \frac{B(\gamma)}{\pi} \sum_{k=0}^{\infty} \gamma^{-k} \left(\varrho/\sin^2\theta\right)^{-\nu/2} \\ &\qquad \times \left(\frac{(\varrho/\sin^2\theta)}{1+(\varrho/\sin^2\theta)}\right)^{k+\nu/2} \\ &< \frac{B(\gamma)}{\pi} \left(\sum_{k=0}^{\infty} \gamma^{-k}\right) \left(\varrho^{-\nu/2}(\sin^2\theta)^{\nu/2}\right) \qquad \text{(III.49)} \\ &< \frac{B(\gamma)}{\pi} \left(\frac{\gamma}{(\gamma-1)}\right) \left(\varrho^{-\nu/2}\right) < \infty. \qquad \text{(III.50)} \end{aligned}$$

The inequality in (III.49) follows by noting that $\left(\frac{(\varrho/\sin^2\theta)}{1+(\varrho/\sin^2\theta)}\right) < 1$. The inequality in (III.50) is obtained first by noting that $(\sin^2\theta) \leq 1$, and second by simplifying the sum in (III.49) as $\left(\sum_{k=0}^{\infty} \gamma^{-k}\right) = (1-(1/\gamma))^{-1}$. Note that this series converges, since γ is greater than unity. The upper bound on I_2 in (III.50) is finite and is independent of $\theta \in [0, \frac{\pi}{2}]$, thus, the interchange of summation and integration order in the second last equality in (III.44) is justified.

We next show that η is zero. By defining $\mathbf{y} \triangleq \mathbf{U}^T C_{\mathbf{aa}}^{-1/2} \boldsymbol{\mu}_{\mathbf{a}}$, $\mathbf{b}_1 \triangleq \mathbf{U}^T C_{\mathbf{aa}}^{1/2} \mathbf{g}$, and $\mathbf{f} \triangleq \sqrt{2}\, C^{-1/2}\mathbf{h}$, the vector $\mathbf{b}$ (see after (III.27)) can be written as $\mathbf{b} = (\mathbf{b}_1 + 2\boldsymbol{\Lambda}\mathbf{y})$, and η can be rewritten as

$$\begin{aligned} \eta &= \left[\boldsymbol{\mu}_{\mathbf{a}}^T \mathbf{A} \boldsymbol{\mu}_{\mathbf{a}} + \boldsymbol{\mu}_{\mathbf{a}}^T \mathbf{g} + e - \frac{1}{4}\mathbf{b}^T \boldsymbol{\Lambda}^{-1} \mathbf{b}\right] \\ &= \Bigg[\mathbf{y}^T \boldsymbol{\Lambda} \mathbf{y} + \mathbf{y}^T \mathbf{b}_1 + \mathbf{f}^H \mathbf{f} \\ &\qquad - \frac{1}{4}(\mathbf{b}_1 + 2\boldsymbol{\Lambda}\mathbf{y})^T \boldsymbol{\Lambda}^{-1} (\mathbf{b}_1 + 2\boldsymbol{\Lambda}\mathbf{y})\Bigg] \\ &= \left[\mathbf{f}^H \mathbf{f} - \frac{1}{4}\mathbf{b}_1^T \boldsymbol{\Lambda}^{-1} \mathbf{b}_1\right]. \qquad \text{(III.51)} \end{aligned}$$

The third equality in (III.51) follows due to cancellation of terms. Using the property $\mathbf{F}(\mathbf{XY}) = \mathbf{F}(\mathbf{X})\mathbf{F}(\mathbf{Y})$ and $\mathbf{F}(\mathbf{X}^H) = \mathbf{F}(\mathbf{X})^T$ (see Appendix III.H.7), we have

$$\begin{aligned} \mathbf{A} &= 2\,\mathbf{F}(\mathbf{Y}^H C^{-1/2} C^{-1/2}\mathbf{Y}) \\ &= 2\,\mathbf{F}(\mathbf{Y}^H C^{-1/2})\mathbf{F}(C^{-1/2}\mathbf{Y}) \\ &= (\hat{\mathbf{A}})^T\hat{\mathbf{A}}, \end{aligned} \tag{III.52}$$

where $\hat{\mathbf{A}} \triangleq \sqrt{2}\,\mathbf{F}(C^{-1/2}\mathbf{Y})$. Using this, and substituting for $\mathbf{g}$, $\mathbf{b}_1$ can be written as

$$\begin{aligned} \mathbf{b}_1 &= 2\sqrt{2}\,\mathbf{U}^T C_{\mathbf{aa}}^{1/2} \\ &\qquad \times \left[\mathrm{Re}(\mathbf{Y}^H C^{-1/2}\,\mathbf{f})^T,\ \mathrm{Img}(\mathbf{Y}^H C^{-1/2}\,\mathbf{f})^T\right]^T \\ &= 2\,\mathbf{U}^T C_{\mathbf{aa}}^{1/2}\left(\sqrt{2}\,\mathbf{F}(\mathbf{Y}^H C^{-1/2})\right) \\ &\qquad\qquad \times \left[\mathrm{Re}(\mathbf{f})^T,\ \mathrm{Img}(\mathbf{f})^T\right]^T \\ &= 2\,\mathbf{U}^T C_{\mathbf{aa}}^{1/2}(\hat{\mathbf{A}})^T\ \left[\mathrm{Re}(\mathbf{f})^T,\ \mathrm{Img}(\mathbf{f})^T\right]^T \\ &= 2\,\mathbf{\Lambda}^{1/2}\,\mathbf{V}^T\ \left[\mathrm{Re}(\mathbf{f})^T,\ \mathrm{Img}(\mathbf{f})^T\right]^T \end{aligned} \tag{III.53}$$

where $\mathbf{U}\mathbf{\Lambda}^{1/2}\,\mathbf{V}^T$ denotes the singular value decomposition of $C_{\mathbf{aa}}^{1/2}(\hat{\mathbf{A}})^T$. The second equality in (III.53) is obtained by using the property $\left(\mathbf{F}(\mathbf{X})\left[\mathrm{Re}(\mathbf{f})^T,\ \mathrm{Img}(\mathbf{f})^T\right]^T = \left[\mathrm{Re}(\mathbf{Xf})^T,\ \mathrm{Img}(\mathbf{Xf})^T\right]^T\right)$. Upon substituting (III.53) in the last equality of (III.51), we get η equal to zero.

Note that ν is the rank of the matrix $C_{\mathbf{aa}}^{1/2}\mathbf{A}C_{\mathbf{aa}}^{1/2}$. Since $C_{\mathbf{aa}} = \frac{1}{2}\,\mathbf{F}(C_{\widehat{\boldsymbol{\beta}_1}^{(1)}\widehat{\boldsymbol{\beta}_1}^{(1)}})$, we can write $C_{\mathbf{aa}}^{1/2}\mathbf{A}C_{\mathbf{aa}}^{1/2} = \mathbf{F}(\mathbf{S})$, where

$$\mathbf{S} \triangleq \left(\left(C_{\widehat{\boldsymbol{\beta}_1}^{(1)}\widehat{\boldsymbol{\beta}_1}^{(1)}}\right)^{1/2}\left(\mathbf{Y}^H C^{-1}\mathbf{Y}\right)\left(C_{\widehat{\boldsymbol{\beta}_1}^{(1)}\widehat{\boldsymbol{\beta}_1}^{(1)}}\right)^{1/2}\right).$$

Denoting $\widetilde{\mathbf{U}}\widetilde{\mathbf{\Lambda}}\widetilde{\mathbf{U}}^H$ as the evd of $\mathbf{S}$, we obtain $C_{\mathbf{aa}}^{1/2}\mathbf{A}C_{\mathbf{aa}}^{1/2} = \mathbf{F}(\widetilde{\mathbf{U}}\widetilde{\mathbf{\Lambda}}\widetilde{\mathbf{U}}^H) = \mathbf{F}(\widetilde{\mathbf{U}})\ \mathbf{F}(\widetilde{\mathbf{\Lambda}})\ \mathbf{F}(\widetilde{\mathbf{U}})^T$, which represents the evd of $C_{\mathbf{aa}}^{1/2}\mathbf{A}C_{\mathbf{aa}}^{1/2}$. Thus, the rank of $C_{\mathbf{aa}}^{1/2}\mathbf{A}C_{\mathbf{aa}}^{1/2}$ is twice the rank of $\mathbf{S}$, which implies that ν is an even number. Since $\mathbf{S}$ is full rank in our case (equal to PM), we have $\nu = 2PM$.

III.H.5 Proof of Lemma 7

The series result in (III.40) is obtained following the approach in [73]. For completeness, the detailed steps of the derivation are presented. The series in (III.40) is obtained with the help of the following two expansions [74]

$$ln(1+x) = \sum_{k=1}^{\infty}(-1)^{k+1}\frac{x^k}{k} \quad \text{and} \quad (1-x)^{-1} = \sum_{k=0}^{\infty} x^k \qquad \text{for } |x|<1. \text{ (III.54)}$$

We start out with the mgf in (III.38) and bring it into a form such that the series expansion in (III.54) can be applied and the condition for their applicability is satisfied, i.e., $|x|<1$ must be satisfied to use the expansion for $ln(1+x)$ and $(1-x)^{-1}$. From the mgf expression in (III.38), the mgf of the random variable $Q(\widetilde{\mathbf{a}})-\eta$ can be written as

$$\varphi_{Q(\widetilde{\mathbf{a}})-\eta}(s) \;=\; exp\Big\{\frac{s}{4}\,\mathbf{b}^T\mathbf{\Lambda}^{-1}\mathbf{b}+\frac{s^2}{2}\sum_{j=1}^{\nu}b_j^2(1-2s\lambda_j)^{-1}\Big\}\prod_{j=1}^{\nu}(1-2s\lambda_j)^{-1/2}. \tag{III.55}$$

The above expression can be further simplified as

$$\begin{aligned}\varphi_{Q(\widetilde{\mathbf{a}})-\eta}(s) &= exp\Big\{\frac{s}{4}\sum_{j=1}^{\nu}\frac{b_j^2}{\lambda_j}-\frac{s}{4}\sum_{j=1}^{\nu}\frac{b_j^2}{\lambda_j}(-2s\lambda_j)(1-2s\lambda_j)^{-1}\Big\}\\ &\qquad\qquad\times\prod_{j=1}^{\nu}(1-2s\lambda_j)^{-1/2}\\ &= exp\Big\{\frac{s}{4}\sum_{j=1}^{\nu}\frac{b_j^2}{\lambda_j}(1-2s\lambda_j)^{-1}\Big\}\prod_{j=1}^{\nu}(1-2s\lambda_j)^{-1/2}\\ &= exp\Big\{\frac{-1}{8}\sum_{j=1}^{\nu}\frac{b_j^2}{\lambda_j^2}+\frac{1}{8}\sum_{j=1}^{\nu}\frac{b_j^2}{\lambda_j^2}(1-2s\lambda_j)^{-1}\Big\}\prod_{j=1}^{\nu}(1-2s\lambda_j)^{-1/2}. \end{aligned}\tag{III.56}$$

To further simplify (III.56), we carry out the following separate simplification for the terms that appear in (III.56)

$$\begin{aligned}\prod_{j=1}^{\nu}(1-2s\lambda_j)^{-1/2} &= \prod_{j=1}^{\nu}(1-2s\lambda_j+2s\varrho-2s\varrho)^{-1/2}\\ &= (1-2s\varrho)^{-\nu/2}\prod_{j=1}^{\nu}\left(1+\frac{-2s\varrho}{1-2s\varrho}\frac{\lambda_j}{\varrho}-\frac{-2s\varrho}{1-2s\varrho}\right)^{-1/2}\end{aligned}\tag{III.57}$$

Note that in the set of equations in (III.57), an arbitrary positive constant ϱ is introduced. Let us now define the following quantities

$$\mu(s) = \frac{-2s\varrho}{1-2s\varrho}, \quad \text{and} \quad \alpha_j = 1 - \frac{\lambda_j}{\varrho}. \tag{III.58}$$

Equation (III.57) can now be expressed in terms of $\mu(s)$ and α_j as follows

$$\prod_{j=1}^{\nu}(1-2s\lambda_j)^{-1/2} \;=\; (1-2s\varrho)^{-\nu/2}\prod_{j=1}^{\nu}(1-\mu(s)\alpha_j)^{-1/2}. \tag{III.59}$$

Similarly, we have the following

$$\begin{aligned}(1-2s\lambda_j)^{-1} &= (1-2s\lambda_j+2s\varrho-2s\varrho)^{-1} \\ &= (1-2s\varrho)^{-1}\left(1-\frac{2s\lambda_j}{1-2s\varrho}+\frac{2s\varrho}{1-2s\varrho}\right)^{-1} \\ &= (1-\mu(s))\,(1-\mu(s)\alpha_j)^{-1}. \end{aligned} \tag{III.60}$$

Substituting (III.59) and (III.60) back in (III.56), the mgf can be written as

$$\begin{aligned}\varphi_{Q(\tilde{\mathbf{a}})-\eta}(s) &= exp\left\{\frac{-1}{8}\sum_{j=1}^{\nu}\frac{b_j^2}{\lambda_j^2}+\frac{1}{8}\sum_{j=1}^{\nu}\frac{b_j^2}{\lambda_j^2}(1-\mu(s))\,(1-\mu(s)\alpha_j)^{-1}\right\} \\ &\qquad \times\; (1-2s\varrho)^{-\nu/2}\prod_{j=1}^{\nu}(1-\mu(s)\alpha_j)^{-1/2} \\ &= (1-2s\varrho)^{-\nu/2}\;\; f(\mu(s)) \end{aligned}$$

$$\begin{aligned}\text{where} \qquad f(\mu(s)) &= exp\left\{\frac{-1}{8}\sum_{j=1}^{\nu}\frac{b_j^2}{\lambda_j^2}+\frac{1}{8}\sum_{j=1}^{\nu}\frac{b_j^2}{\lambda_j^2}(1-\mu(s))\,(1-\mu(s)\alpha_j)^{-1}\right\} \\ &\qquad \times\prod_{j=1}^{\nu}(1-\mu(s)\alpha_j)^{-1/2}. \end{aligned} \tag{III.61}$$

We now wish to obtain a series expansion for (III.61) and also identify the region of convergence (ROC) of this series. To apply the two series expansion stated earlier in (III.54), we take the natural logarithm of $f(\mu(s))$ in (III.61) and obtain

$$\begin{aligned}ln(f(\mu(s))) &= \frac{-1}{8}\sum_{j=1}^{\nu}\frac{b_j^2}{\lambda_j^2}+\frac{1}{8}\sum_{j=1}^{\nu}\frac{b_j^2}{\lambda_j^2}(1-\mu(s))\,(1-\mu(s)\alpha_j)^{-1} \\ &\qquad -\frac{1}{2}\sum_{j=1}^{\nu}ln(1-\mu(s)\alpha_j). \end{aligned} \tag{III.62}$$

Under the condition

$$|\alpha_j\ \mu(s)| < 1, \qquad \text{for all } j, \tag{III.63}$$

we are now ready to apply the series expansion in (III.54) to (III.62) where $x = \alpha_j\ \mu(s)$. Thus, applying these expansions in (III.62), we have

$$\begin{aligned} ln(f(\mu(s))) &= \frac{-1}{8}\sum_{j=1}^{\nu}\frac{b_j^2}{\lambda_j^2} + \frac{1}{8}\sum_{j=1}^{\nu}\frac{b_j^2}{\lambda_j^2}(1-\mu(s))\left(\sum_{k=0}^{\infty}(\mu(s)\alpha_j)^k\right) \\ &\quad -\frac{1}{2}\sum_{j=1}^{\nu}\left(-\sum_{k=1}^{\infty}(\mu(s)\alpha_j)^k\frac{1}{k}\right). \end{aligned} \tag{III.64}$$

Further simplifying (III.64), we have

$$\begin{aligned} &ln(f(\mu(s))) \\ &= \frac{-1}{8}\sum_{j=1}^{\nu}\frac{b_j^2}{\lambda_j^2} + \frac{1}{8}\sum_{j=1}^{\nu}\frac{b_j^2}{\lambda_j^2}\left(\sum_{k=0}^{\infty}(\mu(s)\alpha_j)^k - \sum_{k=0}^{\infty}\mu(s)^{k+1}\alpha_j^k\right) \\ &\qquad +\frac{1}{2}\sum_{j=1}^{\nu}\left(\sum_{k=1}^{\infty}(\mu(s)\alpha_j)^k\frac{1}{k}\right) \\ &= \frac{-1}{8}\sum_{j=1}^{\nu}\frac{b_j^2}{\lambda_j^2} + \frac{1}{8}\sum_{j=1}^{\nu}\frac{b_j^2}{\lambda_j^2}\left(1+\sum_{k=1}^{\infty}(\mu(s)^k\alpha_j^k) - \sum_{k=1}^{\infty}\mu(s)^k\alpha_j^{k-1}\right) \\ &\qquad +\frac{1}{2}\sum_{j=1}^{\nu}\left(\sum_{k=1}^{\infty}(\mu(s)\alpha_j)^k\frac{1}{k}\right) \\ &= \frac{1}{8}\sum_{j=1}^{\nu}\frac{b_j^2}{\lambda_j^2}\left(\sum_{k=1}^{\infty}\mu(s)^k(\alpha_j^k - \alpha_j^{k-1})\right) + \frac{1}{2}\sum_{j=1}^{\nu}\left(\sum_{k=1}^{\infty}(\mu(s)\alpha_j)^k\frac{1}{k}\right) \\ &= \sum_{k=1}^{\infty}\frac{\mu(s)^k}{k}\left(\frac{k}{8}\sum_{j=1}^{\nu}\frac{b_j^2}{\lambda_j^2}\left(\alpha_j^k - \alpha_j^{k-1}\right) + \frac{1}{2}\sum_{j=1}^{\nu}\alpha_j^k\right) \\ &= \sum_{k=1}^{\infty}\frac{\mu(s)^k}{k}\chi_k \qquad \text{where } \chi_k \triangleq \left(\frac{k}{8}\sum_{j=1}^{\nu}\frac{b_j^2}{\lambda_j^2}\left(\alpha_j^k - \alpha_j^{k-1}\right) + \frac{1}{2}\sum_{j=1}^{\nu}\alpha_j^k\right). \end{aligned} \tag{III.65}$$

Using the definition of α_j in (III.58), we simplify the coefficients χ_k as

$$\begin{aligned}\chi_k &= \left(\frac{k}{8}\sum_{j=1}^{\nu}\frac{b_j^2}{\lambda_j^2}(1-\frac{\lambda_j}{\varrho})^{k-1}\left(1-\frac{\lambda_j}{\varrho}-1\right)+\frac{1}{2}\sum_{j=1}^{\nu}(1-\frac{\lambda_j}{\varrho})^k\right)\\ &= \left(-\frac{k}{8\,\varrho}\sum_{j=1}^{\nu}\frac{b_j^2}{\lambda_j}(1-\frac{\lambda_j}{\varrho})^{k-1}+\frac{1}{2}\sum_{j=1}^{\nu}(1-\frac{\lambda_j}{\varrho})^k\right) \qquad \text{for } k\geq 1. \end{aligned} \tag{III.66}$$

Thus, we have found the series expansion for the ln of $f(\mu(s))$ as given in (III.65) and the coefficients for the expansion in (III.66). Note that the ROC of the series for the natural logarithm (ln) of $f(\mu(s))$ is

$$\begin{aligned} & |\alpha_j\ \mu| < 1 \quad \text{for all } j \\ \text{or equivalently} \quad & |\alpha_j||\mu| < 1 \quad \text{for all } j \\ \text{or equivalently} \quad & |\mu| \ < \ 1\ /\left(\max_j\left(|\alpha_j|\right)\right) \end{aligned} \tag{III.67}$$

and the radius of convergence is $\left[1/\left(\max_j\left(|\alpha_j|\right)\right)\right]$. Note that in (III.67) the dependence of $\mu(s)$ on s is suppressed. So far ϱ is any positive constant, now let ϱ also satisfy $\varrho > \lambda_{\max} > 0$. With this we can see that $\vartheta \stackrel{\Delta}{=} \max_j\left(|\alpha_j|\right) = \max_j\left(\left|1-\frac{\lambda_j}{\varrho}\right|\right)$ is always less than unity, i.e. $\vartheta < 1$. In fact $0 < \vartheta < 1$.

We next wish to obtain a power series for $f(\mu(s))$. The following approach is used to obtain a series for $f(\mu(s))$ [73]. Consider the following power series

$$ln(g(\omega)) = \chi_0 + \sum_{k=1}^{\infty}\frac{\omega^k}{k}\chi_k \tag{III.68}$$

that is uniformly convergent in some region of ω, then we can write

$$g(\omega) = exp\,\{\,\chi_0 + \sum_{k=1}^{\infty}\frac{\omega^k}{k}\chi_k\,\}. \tag{III.69}$$

Since $exp(x) = \sum_{n=0}^{\infty}\frac{x^n}{n!}$, we can simplify the right hand side of (III.69) and write it as a power series in ω. The region of convergence of this power series would be the same as that for $ln(g(\omega))$ in (III.68). Let us express the right hand side of (III.69) as the following power series (say)

$$g(\omega) = \sum_{k=0}^{\infty}\xi_k\ \omega^k. \tag{III.70}$$

We next need to specify the coefficients ξ_k in (III.70) and also relate it to the coefficients χ_k in (III.69).

For this we carry out term by term differentiation of the power series in (III.68). This is allowed from Theorem 7 stated in Appendix III.H.8. Thus, differentiating both sides of (III.68) we have

$$\frac{d\, ln(g(\omega))}{d\,\omega} = \sum_{k=1}^{\infty} \omega^{k-1} \chi_k$$

$$\Rightarrow \qquad \frac{d\, g(\omega)}{d\,\omega} = g(\omega) \sum_{k=1}^{\infty} \omega^{k-1} \chi_k$$

$$\Rightarrow \qquad \sum_{k=1}^{\infty} \xi_k \; k \; \omega^{k-1} = \left(\sum_{k=0}^{\infty} \xi_k \omega^k \right) \left(\sum_{k=1}^{\infty} \omega^{k-1} \chi_k \right) \tag{III.71}$$

$$\Rightarrow \qquad \sum_{k=1}^{\infty} \xi_k \; k \; \omega^{k-1} = \left(\sum_{l=1}^{\infty} \sum_{m=0}^{\infty} \xi_m \; \chi_l \; \omega^{l+m-1} \right). \tag{III.72}$$

The equation (III.71) is obtained by substituting for $g(\omega)$ from (III.70) and performing differentiation (allowed from Theorem 7 in Appendix III.H.8). By comparing the coefficients of ω^{k-1} on both sides of (III.72), the coefficients ξ_k and χ_k can be related. Consider the r.h.s of the equality in (III.72) and let $l + m = k$. Since $l \in [1, \infty)$ and $m \in [0, \infty)$, we have $k \in [1, \infty)$. Also, note that $0 \leq m \leq k-1$, thus we can write

$$\Rightarrow \qquad \sum_{k=1}^{\infty} \xi_k \; k \; \omega^{k-1} = \left(\sum_{k=1}^{\infty} \sum_{m=0}^{k-1} \xi_m \; \chi_{k-m} \; \omega^{k-1} \right). \tag{III.73}$$

The coefficients ξ_k and χ_k can be related as

$$\xi_k \;\; = \;\; \frac{1}{k} \sum_{r=0}^{k-1} \chi_{k-r} \;\; \xi_r \qquad , k \geq 1. \tag{III.74}$$

Also, $g(0) = \xi_0$ and $ln(g(0)) = \chi_0$. This implies that $\xi_0 = e^{\chi_0}$.

Using the above approach, we can now write a power series for $f(\mu(s))$. We have already shown that the natural logarithm (ln) of $f(\mu(s))$ can be written as a power

series in (III.65). By identifying function $f(.)$ as $g(.)$, and $\mu(s)$ as ω in (III.68) and (III.70), we can write a power series for $f(\mu(s))$ as

$$f(\mu(s)) = \sum_{k=0}^{\infty} \xi_k \mu(s)^k \tag{III.75}$$

where coefficient ξ_k and χ_k are given in (III.74) and (III.66), respectively. Since $\chi_0 = 0$ in our case (see (III.65)), we have $\xi_0 = e^{\chi_0} = 1$. Thus, the coefficients ξ_k and χ_k are completely specified.

The ROC of this series is the same as in (III.67). Note that in (III.75), s, and therefore $\mu(s)$, take on real values. Since $\mu(s)$ is function of s, we wish to identify the interval of s for which the condition in (III.67) is satisfied. This interval, denoted as S_1, can be shown to be

$$S_1 = \left(-\infty, \frac{1}{2\varrho(1+\vartheta)}\right) \cup \left(\frac{1}{2\varrho(1-\vartheta)}, \infty\right). \tag{III.76}$$

The derivation of (III.76) is presented at the end of this section.

With (III.61) and (III.75), the mgf can be written as

$$\varphi_{Q(\widetilde{\mathbf{a}})-\eta}(s) \;\; = \;\; (1-2s\varrho)^{-\nu/2} \left(\sum_{k=0}^{\infty} \xi_k \mu(s)^k\right). \tag{III.77}$$

Note that the expression in (III.77) has to be real valued. Since the series inside the parenthesis in (III.77) is real for real s, the term $(1-2s\varrho)^{-\nu/2}$ must be real for odd ν. This implies that $(1-2s\varrho)$ must be positive, which then puts an additional constraint on s (in addition to s being in S_1). According to this constraint, s must belong to the interval S_2 defined as $S_2 = \left(-\infty, \frac{1}{2\varrho}\right)$. Note that with this constraint the condition in (III.38) is satisfied as well. Thus, the expression in (III.40) is valid for all s that lies in the intersection of S_1 and S_2. This intersection interval is $\left(-\infty, \frac{1}{2\varrho\,(1+\vartheta)}\right)$, implying that (III.40) is valid for any real s such that $s < \frac{1}{2\varrho\,(1+\vartheta)}$.

Substituting for $\mu(s)$ in (III.77) results in the desired expression in (III.40).

1) Derivation of (III.76) - *ROC:*

The region of convergence in the 's' plane of the series in (III.75) is the set of all s for which condition in (III.67), reproduced below, is satisfied.

$$|\mu(s)| \quad < \quad 1/\vartheta \qquad \text{where } \vartheta = \max_j \left(|\alpha_j|\right) \tag{III.78}$$

Recall that $\vartheta < 1$ as shown earlier. Since we are interested in real s, we wish to identify the interval in s such that (III.78) is satisfied.

Substituting $\mu(s) = (-2s\varrho)/(1-2s\varrho)$ in (III.78), we have

$$\left|\frac{-2s\varrho}{1-2s\varrho}\right| \quad < \quad 1/\vartheta. \tag{III.79}$$

Simplifying (III.79), we get the following two inequalities

$$\frac{1}{2\varrho\ (1-\vartheta)} < s \quad \text{and} \quad s < \frac{1}{2\varrho\ (1+\vartheta)}. \tag{III.80}$$

Thus, (III.80) are the intervals in s for which (III.78) is satisfied. We denote this interval as $S_1 = \left(-\infty, \frac{1}{2\varrho(1+\vartheta)}\right) \cup \left(\frac{1}{2\varrho(1-\vartheta)}, \infty\right)$.

III.H.6 Proof of Lemma 8

For simplicity, we divide the proof of lemma 8 in two parts. The first part deals with establishing (III.47) and the second part deals with proving (III.48).

1) Proof of (III.47) - *Bounding the coefficients:*

Consider a power series expansion

$$g(\omega) = \sum_{k=0}^{\infty} \beta_k \, \omega^k \tag{III.81}$$

that uniformly converges for all complex ω that satisfies

$$|\omega| < R, \tag{III.82}$$

where R is the radius of convergence (see also Theorem 8 in Appendix III.H.8). The authors in [76] bound the coefficients, β_k, using the Cauchy's integral formula. It is worthwhile to point out that [76] states only the final result, i.e., the bound of the form (III.47), and omits the proof of it in [76]. For clarity, the complete derivation is presented below.

Note that the coefficients β_k in (III.81) can be written as

$$\beta_k \; = \left(\frac{g^{(k)}(\omega)}{k!} \right) \Bigg|_{\omega=0}. \tag{III.83}$$

The region of convergence $|\omega| < R$ symbolizes a circular disk with center at the origin and radius R as shown in Figure III.7. Now consider a circular path $C : |\omega| = \alpha$, where $0 < \alpha < R$. This circle is shown as dashed circle in Figure III.7. Now consider the following formula

$$h^{(k)}(z_o) = \frac{k!}{2\pi i} \oint_{\widetilde{C}} \frac{h(z)}{(z - z_o)^{k+1}} dz \tag{III.84}$$

where $h^{(k)}(z_o)$ denotes the kth derivative of the function $h(z)$ evaluated at z_o. Equation (III.84) is from Theorem 9 in Appendix III.H.8. Note that function $g(\omega)$ is analytic from Theorem 6 in Appendix III.H.8, therefore, Theorem 9 of Appendix III.H.8 can be applied to further simplify (III.83). Thus, applying (III.84) to the r.h.s of (III.83) over the circular path C and noting that ω_o corresponding to z_o in (III.84) is zero, we have

$$\beta_k = \frac{1}{2\pi i} \oint_{C:|\omega|=\alpha} \frac{g(\omega)}{\omega^{k+1}} d\omega. \tag{III.85}$$

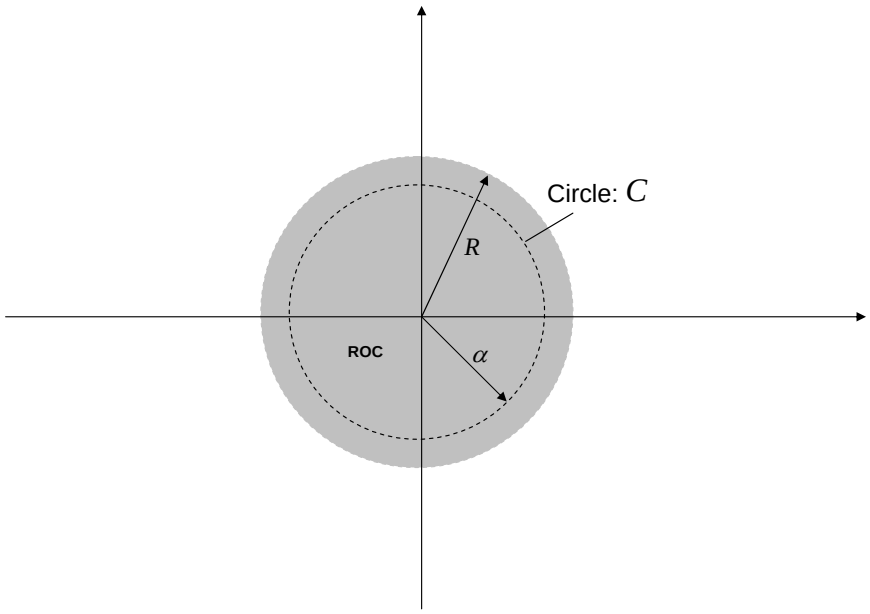

Figure III.7: ROC in the complex plane.

The absolute value of the coefficients β_k can be written as

$$|\beta_k| = \left| \frac{1}{2\pi i} \oint_{C:|\omega|=\alpha} \frac{g(\omega)}{\omega^{k+1}} d\omega \right|. \tag{III.86}$$

Now consider the following change of variable in (III.86), $\omega = \alpha e^{i\phi}$ where $\phi \in [0, 2\pi]$, we have

$$\begin{aligned} |\beta_k| &= \left| \frac{1}{2\pi i} \int_0^{2\pi} \frac{g(\alpha e^{i\phi})}{\alpha^{k+1}\ e^{i\phi(k+1)}}\ \alpha i e^{i\phi}\ d\phi \right|. \\ &= \frac{1}{2\pi\alpha^k}\ \left| \int_0^{2\pi} \frac{g(\alpha e^{i\phi})}{e^{i\phi k}}\ d\phi \right|. \end{aligned} \tag{III.87}$$

We now bound the coefficients β_k as follows:

$$\begin{aligned} |\beta_k| &\leq \frac{1}{2\pi\alpha^k}\ \int_0^{2\pi} \left| \frac{g(\alpha e^{i\phi})}{e^{i\phi k}} \right| d\phi \\ &\leq \frac{1}{2\pi\alpha^k}\ \int_0^{2\pi} \max_{\phi\in[0,2\pi]} \left| \frac{g(\alpha e^{i\phi})}{e^{i\phi k}} \right| d\phi \\ &= \frac{1}{\alpha^k}\ \max_{\phi\in[0,2\pi]} \left| \frac{g(\alpha e^{i\phi})}{e^{i\phi k}} \right| \\ &= \frac{1}{\alpha^k}\ \max_{|\omega|=\alpha} |g(\omega)|. \end{aligned} \tag{III.88}$$

(III.88) is the desired bound for the coefficients β_k.

Now note that the power series in (III.75) is of the same form as (III.81) with the radius of convergence being $1/\vartheta$, where $\vartheta = \left(\max_j (|\alpha_j|)\right)$. For simplicity of notation and without causing any confusion, we drop the dependence of $\mu(s)$ on s in (III.75) for the time being, and rewrite (III.75) as

$$f(\mu) = \sum_{k=0}^{\infty} \xi_k \, \mu^k \qquad \text{where } |\mu| < 1/\vartheta. \tag{III.89}$$

Consider a constant γ such that $1 < \gamma < 1/\vartheta$, and a circular path $C_1 : |\mu| = \gamma$. Recall that $\vartheta < 1$. We can bound the coefficients ξ_k in the same manner as (III.88) by applying the Cauchy's integral formula over the circle C_1. The region $|\mu| < 1/\vartheta$ and the circle C_1 is similar to the grey disk region and circle C in Figure III.7 by simply replacing R by $1/\vartheta$ and α by γ.

Thus, we can bound the coefficients ξ_k of the power series in (III.89) as

$$|\xi_k| \leq \frac{1}{\gamma^k} \left(\max_{|\mu|=\gamma} |f(\mu)|\right). \tag{III.90}$$

Note that the bound of the form (III.90) is also valid for $0 < \gamma \leq 1$, however, this is not useful for our case.

2) Proof of inequality (III.48) - *Bounding* $\zeta(\gamma)$*:*

For simplicity of notation and without causing any confusion, we drop the dependence of $\mu(s)$ on s in the expression for the function $f(\mu(s))$ in (III.61) for the time being, and rewrite it as

$$\begin{aligned} f(\mu) &= exp\left\{ \frac{-1}{8} \sum_{j=1}^{\nu} \frac{b_j^2}{\lambda_j^2} + \frac{1}{8} \sum_{j=1}^{\nu} \frac{b_j^2}{\lambda_j^2} (1-\mu) \left(1-\mu\alpha_j\right)^{-1} \right\} \prod_{j=1}^{\nu} \left(1-\mu\alpha_j\right)^{-1/2} \\ &= exp\left\{ \frac{-1}{8} \sum_{j=1}^{\nu} \frac{b_j^2}{\lambda_j^2} (1-\alpha_j) \left(\mu(1-\mu\alpha_j)^{-1}\right) \right\} \prod_{j=1}^{\nu} \left(1-\mu\alpha_j\right)^{-1/2}. \end{aligned} \tag{III.91}$$

We wish to find a bound on $\zeta(\gamma) = \left(\max_{|\mu|=\gamma} |f(\mu)|\right)$. To carry out this maximization, we first note that $\alpha_j < 1$ (see the definition of α_j in (III.58) and then refer to the discussion

after equation (III.67)), thus, we have

$$\zeta(\gamma) = \left(\max_{|\mu|=\gamma} |f(\mu)|\right) = \left(\max_{|\mu|=\gamma} \left| exp\left\{ \frac{-1}{8} \sum_{j=1}^{\nu} \frac{b_j^2}{\lambda_j^2} (1-\alpha_j) \left(\mu(1-\mu\alpha_j)^{-1}\right) \right\} \times \prod_{j=1}^{\nu} (1-\mu\alpha_j)^{-1/2} \right| \right). \quad \text{(III.92)}$$

We can now bound $\zeta(\gamma)$ as follows,

$$\begin{aligned}
\zeta(\gamma) &\leq \left(\max_{|\mu|=\gamma} \left| exp\left\{ \frac{-1}{8} \sum_{j=1}^{\nu} \frac{b_j^2}{\lambda_j^2} (1-\alpha_j) \left(\mu(1-\mu\alpha_j)^{-1}\right) \right\} \right| \right) \\
&\quad \times \left(\max_{|\mu|=\gamma} \prod_{j=1}^{\nu} \left| (1-\mu\alpha_j)^{-1/2} \right| \right) \\
&= \left(\max_{|\mu|=\gamma} \prod_{j=1}^{\nu} \left| exp\left\{ \frac{-b_j^2}{8\,\lambda_j^2} (1-\alpha_j) \left(\mu(1-\mu\alpha_j)^{-1}\right) \right\} \right| \right) \\
&\quad \times \left(\max_{|\mu|=\gamma} \prod_{j=1}^{\nu} \left| (1-\mu\alpha_j)^{-1/2} \right| \right) \\
&\leq \left(\prod_{j=1}^{\nu} \max_{|\mu|=\gamma} \left| exp\left\{ \frac{b_j^2}{8\,\lambda_j^2} (1-\alpha_j) \left(-\mu(1-\mu\alpha_j)^{-1}\right) \right\} \right| \right) \\
&\quad \times \left(\prod_{j=1}^{\nu} \max_{|\mu|=\gamma} \left| (1-\mu\alpha_j)^{-1/2} \right| \right). \quad \text{(III.93)}
\end{aligned}$$

The above maximization is equivalent to the following

$$\begin{aligned}
&= \left(\prod_{j=1}^{\nu} exp\left\{ \frac{b_j^2}{8\,\lambda_j^2} (1-\alpha_j) \left[\max_{|\mu|=\gamma} Re\left(-\mu(1-\mu\alpha_j)^{-1}\right) \right] \right\} \right) \\
&\quad \times \left(\prod_{j=1}^{\nu} \max_{|\mu|=\gamma} \left| (1-\mu\alpha_j)^{-1/2} \right| \right). \quad \text{(III.94)}
\end{aligned}$$

In Equation (III.94) the following identity is used: $|e^z| = |e^{(x+iy)}| = |e^x e^{iy}| = |e^x||e^{iy}| = e^x = e^{Re(z)}$ (where $z = x + iy$ is a complex number). The Equation (III.94) can be further simplified as

$$\begin{aligned}
&= \left(\prod_{j=1}^{\nu} exp\left\{ \frac{b_j^2}{8\,\lambda_j^2} (1-\alpha_j) \left[\max_{|\mu|=\gamma} Re\left(-\mu(1-\mu\alpha_j)^{-1}\right) \right] \right\} \right) \\
&\quad \times \left(\prod_{j=1}^{\nu} \left(\max_{|\mu|=\gamma} \left| (1-\mu\alpha_j)^{-1} \right| \right)^{1/2} \right). \quad \text{(III.95)}
\end{aligned}$$

Next, we consider

$$\begin{aligned}
\left(\max_{|\mu|=\gamma} Re\ \left(-\mu(1-\mu\alpha_j)^{-1}\right)\right) &= \left(\max_{|\mu|=\gamma} Re\ \left(1-(1-\mu\alpha_j)^{-1}\right)\frac{1}{\alpha_j}\right) \\
&= \frac{1}{\alpha_j}\left(1-\min_{|\mu|=\gamma} Re\left[(1-\mu\alpha_j)^{-1}\right]\right) \\
&= \frac{1}{\alpha_j}\left(1-\min_{\phi\in[0,2\pi]} Re\left[\left(1-\gamma e^{i\phi}\alpha_j\right)^{-1}\right]\right) \\
&= \frac{1}{\alpha_j}\left(1-(1+\gamma\alpha_j)^{-1}\right) \\
&= \left(\gamma(1+\gamma\alpha_j)^{-1}\right), \qquad \text{(III.96)}
\end{aligned}$$

and similarly,

$$\left(\max_{|\mu|=\gamma} \left|(1-\mu\alpha_j)^{-1}\right|\right) = \left(\max_{\phi\in[0,2\pi]} \left|(1-\gamma e^{i\phi}\alpha_j)^{-1}\right|\right) = (1-\gamma\alpha_j)^{-1}. \quad \text{(III.97)}$$

Using (III.96) and (III.97) in (III.92), we have

$$\zeta(\gamma) \leq \ exp\left\{\frac{1}{8}\sum_{j=1}^{\nu}\frac{b_j^2}{\lambda_j^2}(1-\alpha_j)\left(\gamma(1+\gamma\alpha_j)^{-1}\right)\right\}\prod_{j=1}^{\nu}(1-\gamma\alpha_j)^{-1/2}. \qquad \text{(III.98)}$$

Substituting for α_j from (III.58) in (III.98), the final bound for $\zeta(\gamma)$ can be written as

$$\zeta(\gamma) \leq exp\left\{\frac{\gamma}{8\varrho}\sum_{j=1}^{\nu}\frac{b_j^2}{\lambda_j}\left(\left(1+\gamma-\frac{\gamma\lambda_j}{\varrho}\right)^{-1}\right)\right\}\prod_{j=1}^{\nu}\left(1-\gamma+\frac{\gamma\lambda_j}{\varrho}\right)^{-1/2}. \quad \text{(III.99)}$$

This is the desired expression in (III.48).

III.H.7 Properties of the function: $\mathbf{F}(\mathbf{X})$

Property 1) $\mathbf{F}(\mathbf{XY}) = \mathbf{F}(\mathbf{X})\mathbf{F}(\mathbf{Y})$:

Let $\mathbf{X} = \mathbf{X}_1 + i\mathbf{X}_2$ and $\mathbf{Y} = \mathbf{Y}_1 + i\mathbf{Y}_2$, then $\mathbf{XY} = \mathbf{X}_1\mathbf{Y}_1 - \mathbf{X}_2\mathbf{Y}_2 + i\mathbf{X}_1\mathbf{Y}_2 + i\mathbf{X}_2\mathbf{Y}_1$.

Now consider,

$$\mathbf{F}(\mathbf{X})\mathbf{F}(\mathbf{Y}) = \begin{bmatrix} \mathbf{X}_1 & -\mathbf{X}_2 \\ \mathbf{X}_2 & \mathbf{X}_1 \end{bmatrix} \begin{bmatrix} \mathbf{Y}_1 & -\mathbf{Y}_2 \\ \mathbf{Y}_2 & \mathbf{Y}_1 \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1\mathbf{Y}_1 - \mathbf{X}_2\mathbf{Y}_2 & -\mathbf{X}_1\mathbf{Y}_2 - \mathbf{X}_2\mathbf{Y}_1 \\ \mathbf{X}_1\mathbf{Y}_2 + \mathbf{X}_2\mathbf{Y}_1 & \mathbf{X}_1\mathbf{Y}_1 - \mathbf{X}_2\mathbf{Y}_2 \end{bmatrix}. \tag{III.100}$$

Note that the r.h.s of (III.100) is $\mathbf{F}(\mathbf{XY})$.

Property 2) $\mathbf{F}(\mathbf{X}^H) = \mathbf{F}(\mathbf{X})^T$:

We have

$$\mathbf{F}(\mathbf{X}^H) = \mathbf{F}(\mathbf{X}_1^T - i\mathbf{X}_2^T) = \begin{bmatrix} \mathbf{X}_1^T & \mathbf{X}_2^T \\ -\mathbf{X}_2^T & \mathbf{X}_1^T \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1 & -\mathbf{X}_2 \\ \mathbf{X}_2 & \mathbf{X}_1 \end{bmatrix}^T = \mathbf{F}(\mathbf{X})^T, \tag{III.101}$$

establishing the result.

Property 3) $\mathbf{F}(\mathbf{X}) \left[\mathrm{Re}(\mathbf{f})^T \;\; \mathrm{Img}(\mathbf{f})^T\right]^T = \left[\mathrm{Re}(\mathbf{Xf})^T \;\; \mathrm{Img}(\mathbf{Xf})^T\right]^T$:

We have

$$\mathbf{F}(\mathbf{X}) \left[\mathrm{Re}(\mathbf{f})^T \;\; \mathrm{Img}(\mathbf{f})^T\right]^T = \begin{bmatrix} \mathbf{X}_1 \,\mathrm{Re}(\mathbf{f}) - \mathbf{X}_2 \,\mathrm{Img}(\mathbf{f}) \\ \mathbf{X}_2 \,\mathrm{Re}(\mathbf{f}) + \mathbf{X}_1 \,\mathrm{Img}(\mathbf{f}) \end{bmatrix}. \tag{III.102}$$

Using the following

$$\mathbf{Xf} \;\; = \;\; \mathbf{X}_1\mathrm{Re}(\mathbf{f}) - \mathbf{X}_2\mathrm{Img}(\mathbf{f}) + i\left[\mathbf{X}_1\mathrm{Img}(\mathbf{f}) + \mathbf{X}_2\mathrm{Re}(\mathbf{f})\right], \tag{III.103}$$

the vector $\left[\mathrm{Re}(\mathbf{Xf})^T \;\; \mathrm{Img}(\mathbf{Xf})^T\right]^T$ can be simplified to the r.h.s of (III.102).

III.H.8 Some useful results

A power series in powers of $z - z_o$ is a series of the form

$$\sum_{n=0}^{\infty} a_n (z - z_o)^n = a_o + a_1(z - z_o) + a_2(z - z_o)^2 + \cdots \tag{III.104}$$

where z is a complex variable, $a_o, a_1, \cdots$ are complex (or real) constants, called the coefficients of the series, and z_o is a complex (or real) constant, called the center of the series.

Theorem 6 *(Analytic functions. Their derivatives)*
A power series with a nonzero radius of convergence R represents an analytic function at every point interior to its circle of convergence. The derivatives of this function are obtained by differentiating the original series term by term. All the series thus obtained have the same radius of convergence as the original series. Hence, by the first statement, each of them represents an analytic function.

Theorem 7 *(Termwise differentiation of a power series)*
The derived series of a power series has the same radius of convergence as the original series.

The derived series in the above means the power series obtained by termwise differentiation of (III.104).

Theorem 8 *(Uniform convergence of power series)*
A power series of the form (III.104) *with a nonzero radius of convergence R is uniformly convergent in every circular disk $|z - z_o| \leq r$ of radius $r < R$.*

Theorem 9 *(Derivatives of an analytic function)*
If $h(z)$ is analytic in a domain D, then it has derivatives of all orders in D, which are then also analytic functions in D. The values of these derivatives at point z_o in D are given by the formula

$$h^{(k)}(z_o) = \frac{k!}{2\pi i} \oint_{\tilde{C}} \frac{h(z)}{(z - z_o)^{k+1}} dz \tag{III.105}$$

where $h^{(k)}(z_o)$ denotes the kth derivative of the function $h(z)$ evaluated at z_o, and $\widetilde{C}$ is any simple closed path in D that encloses z_o and whose full interior belongs to D; and we integrate counterclockwise around $\widetilde{C}$.

IV

Performance Analysis of QAM Multiple Antenna Systems at 60 GHz

IV.A Introduction

The FCC has allocated a large bandwidth, spanning several gigahertz, for unlicensed use around 60 GHz [5]. This spectral band has opened up new possibilities for realizing short range, high rate wireless communications [79]. There is an abundance of applications requiring access over short range together with data rates in excess of hundreds of megabit-per-second, such as a high speed home link between a DVD player and an HDTV, transmission of pictures between a digital camera and PC server, video-conferencing and wireless connections between home appliances, laptops, and printers [80].

The communications band at 60GHz is particularly attractive primarily because of the availability of 7GHz of contiguous spectrum (57-64 GHz frequency band), and the attenuation characteristics of the millimeter waves. At 60GHz, there is a large free space attenuation of 68 dB at a distance of 1m [6], [7]. Considering even larger distances (several km), there is an additional attenuation due to atmospheric oxygen of 15.1 dB/km [8], thus, making the band ideally suited for short range communications. When considering indoor applications, the millimeter wave (mmW) transmitter would

typically be communicating to a receiver located within the same home (in a residential or office environment). It is therefore desirable to characterize the propagation loss incurred at mmW frequencies.

Measurement results have revealed that the mmW signals are considerably attenuated while passing through a wall. For example, transmission loss through a typical wall with a 4 inch gap between wall boards of 5/8 inch thickness and 2.5% moisture content can be as high as 35 dB [81] at 60 GHz. This is advantageous, since the multipath components would be mostly confined to a single room, and multipaths that exhibit large delays would be significantly attenuated. Compared to the signal strength received over the direct path (or the line-of-sight path) between the transmitter and receiver, the multipath arriving after being reflected from an adjoining room would be reduced by about 38dB for frequencies above 8-12 GHz [81]. This also implies that the interference level experienced from a similar mmW device placed in adjacent rooms would be quite low. The multipath delay spread would be significantly limited to several nanoseconds [7] due to the severe attenuation and path-loss.

The range of a mmW device can be improved by the use of directional antennas [7]. However, antenna obstruction (e.g., by either a moving object or a person) can lead to significant loss in received signal power due to poor diffraction of the millimeter waves. In order to address the design issues related to a communications system operating in the 60GHz band, it is important to first characterize the channel behavior and identify a channel model that is both realistic and yet simple enough to be usable. The channel modeling subgroup of the IEEE 802.15.3c Task Group (TG) has undertaken this task and has recently submitted its recommendations for a mmW channel model based on measurement results [7]. The proposed model accounts for the clustering phenomenon [7], [82]. The cluster model used is the extension of the one proposed by Saleh-Valenzuela (S-V model) [82], [83]. The channel model has a line-of-sight (LOS) component and a non-line-of-sight (NLOS) component. The multipath components arrive in clusters. There are several clusters, the number of which is dictated by the room superstructure (like walls, furniture, doors), and each ray (or the multipath component)

within a cluster is associated with an amplitude gain, delay and angle of arrival (AoA). The channel model has been simplified to ignore the correlation between the time of arrival and AoA statistics, and both an unobstructed LOS component and a static channel are assumed.

As a result of these propagation and bandwidth considerations, the 60 GHz band is an attractive solution for short range, high rate data communications. The mmW system has good coexistence properties because of the large attenuation and path-loss, and offers no interference to systems in the microwave band due to large frequency difference. Higher transmission speeds, greater than 1gigabit-per-second, can be supported by a mmW system that has to conform to a simple spectrum mask [84]. In addition, due to its shorter wavelength, the mmW band offers the possibility of smaller components and antennas, thus reducing the overall size of these devices.

In order to satisfy the high data rate requirements (in excess of gigabit-per-second) of the supported applications, the 60GHz band must be effectively utilized. One way to increase the spectral efficiency would be to employ a higher order modulation such as M-ary quadrature amplitude modulation (M-QAM). In this work, we consider a M-QAM based communication system operating at 60GHz with a single antenna at the transmitter and an antenna array at the receiver side [85]. Due to the attenuation characteristics of the millimeter waves and large bandwidth, the received signal is comprised of large number of low-energy multipath components. In addition, there is co-channel interference from a similar mmW devices placed in the same room. In order to capture the multipath energy, and to suppress both the intersymbol interference and multi-user interference, we consider the use of a spatio-temporal receiver. Each antenna element at the receiver array utilizes an equalizer. The equalizer weights are determined jointly across all the receive antennas. We investigate the performance of this system through the analytically derived expression for the bit error rate (BER). The channel model used in our work is the one proposed by the IEEE TG.

The rest of the chapter is organized as follows. In Sections IV.B and IV.C, the system model is presented, where the transmitted signal, the mmW IEEE channel

model and the receiver structure are described. In Section IV.D, we obtain an exact BER expression for QAM signalling in the presence of ISI and MAI. Numerical results and discussion are presented in Section IV.E. Finally, we state the conclusions in Section IV.F.

IV.B System Model

IV.B.1 Transmitted Signal and Channel Model

We consider a single-carrier system operating at 60 GHz where a single antenna with directional radiation pattern is employed at the transmitter and an antenna array with Q omni elements is used at the receiver. A higher order modulation scheme such as M-QAM is used and a total of K users is assumed. A block diagram for the system is shown in Figure IV.1. In the Figure, $g(t)$, represents the impulse response of a wave shaping filter and satisfies the Nyquist criterion. The $Q \times P$ weight matrix, $\mathbf{W}$, represents the spatial and temporal processing at the receiver side. A linear equalizer with P taps is employed at each of the receive antennas.

We assume a frequency selective fading slowly fading channel. The channel model used is the one proposed by the IEEE 802.15.3c Task Group [7]. The proposed model accounts for clustering phenomenon [7], and has a line-of-sight component and a non-line-of-sight component. Using this model, the complex low-pass equivalent impulse response of the kth user's channel as seen at the qth receive antenna can be written as

$$h_{k,q}(t) \;=\; \widetilde{\beta}_{k,q}\,\delta(t) + \sum_{l=0}^{L-1}\sum_{i=0}^{K_l-1} \widetilde{a}_{k,q,l,i}\,\delta(t - T_{k,l} - \widetilde{\tau}_{k,l,i}). \tag{IV.1}$$

The parameters appearing in (IV.1) are explained below.

- L is the total number of clusters.
- K_l is the total number of rays in the lth cluster.
- $T_{k,l}$ is the lth cluster delay.

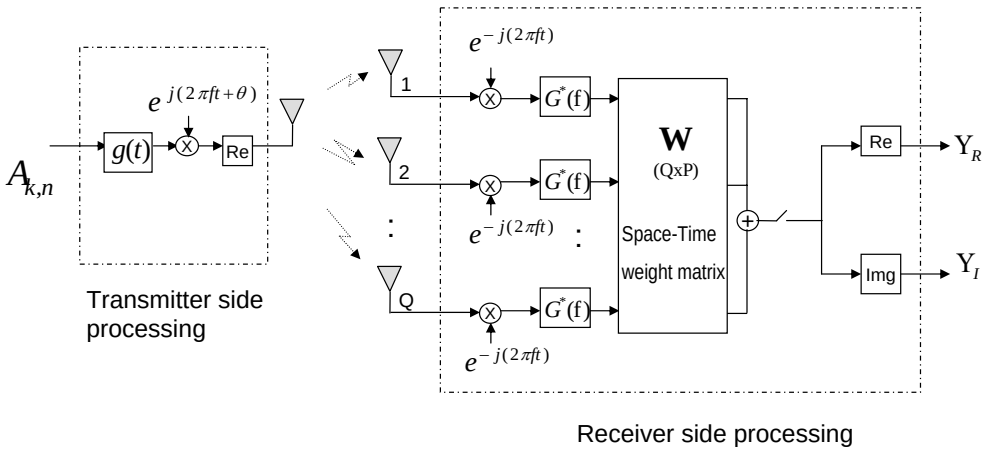

Figure IV.1: System block diagram

- $\widetilde{a}_{k,q,l,i}$, $\widetilde{\tau}_{k,l,i}$ are the complex gain and delay, respectively, of the ith ray of the lth cluster. Delay $\widetilde{\tau}_{k,l,i}$ is measured relative to the cluster arrival time $T_{k,l}$.
- $\widetilde{\beta}_{k,q}$ is the multipath gain of the LOS component.

The amplitude of the multipath gain $\widetilde{a}_{k,q,l,i}$ is modeled by the log normal distribution. We assume that the temporal correlation between various multipaths is negligible. As a result, $\widetilde{a}_{k,q,l,i}$ for all k, l and i are assumed independent random variables. However, for different values of q, the $\widetilde{a}_{k,q,l,i}$ are spatially correlated. The spatial correlation model is discussed later in Section IV.E. The power delay profile of the multipaths is given by

$$E[|\widetilde{a}_{k,q,l,i}|^2] = e^{-T_{k,l}/\Gamma}\ e^{-\widetilde{\tau}_{k,l,i}/\gamma} e^{-\widetilde{k}[1-\bar{\delta}(i)]} \tag{IV.2}$$

where

- Γ is the cluster decay factor.
- γ is the ray decay factor.
- $\widetilde{k}$ is the Rician effect in each cluster [7] and $\bar{\delta}$ is the Kronecker delta.

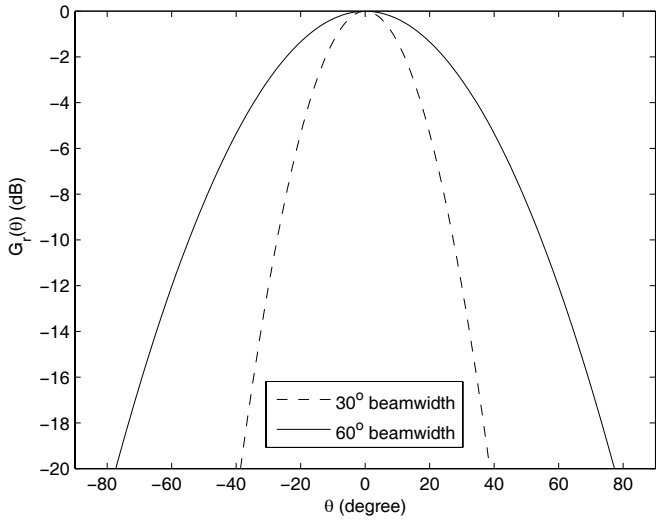

Figure IV.2: Directional antenna pattern

Note that the channel model in (IV.1) implicitly accounts for the transmitter directional radiation pattern. The effect of the directional radiation pattern is captured in the parameter values for the channel model [7]. The radiation pattern considered in [7] is given by

$$G_r(\theta) = G_o exp(-\kappa\theta^2) \tag{IV.3}$$

where G_o is the maximum gain, and κ is determined by the half-power beamwidth θ_{-3dB} as follows [86]

$$\kappa = \frac{4log(2)}{\theta^2_{-3dB}}. \tag{IV.4}$$

An example of the radiation pattern for -3dB bandwidth of 30^o and 60^o is shown in Figure IV.2.

We assume that the multipath arrival times $T_{k,l} + \widetilde{\tau}_{k,l,i}$ are spaced at least a symbol interval apart and can take on arbitrary values (not necessarily integer multiples of symbol interval). The transmitted signal of the kth user is given by

$$\widetilde{y}_k(t) = \mathrm{Re}(s_k(t)\ e^{j2\pi ft}) \tag{IV.5}$$

$$\text{where} \qquad s_k(t) = \sqrt{P_k} \sum_{n=-\infty}^{\infty} A_{k,n} g(t - nT), \tag{IV.6}$$

$A_{k,n}$ is the transmitted complex data sequence (M-QAM symbol), T is the symbol duration, and f is the carrier frequency. After passing through the channel the low-pass equivalent of the received signal at the qth antenna can be written as

$$\begin{aligned} r_q(t) &= \left(\sum_{k=1}^{K} \widetilde{\beta}_{k,q} s_k(t - \hat{\tau}_k) \right. \\ &\quad \left. + \sum_{k=1}^{K} \sum_{l=0}^{L-1} \sum_{i=0}^{K_l - 1} \widetilde{a}_{k,q,l,i} s_k \left(t - T_{k,l} - \widetilde{\tau}_{k,l,i} - \hat{\tau}_k\right) + n_q(t) \right) \end{aligned} \tag{IV.7}$$

where $\hat{\tau}_k$ is the delay of the kth user. The delay of the desired user ($k = 1$) is assumed tracked, i.e., $\hat{\tau}_1 = 0$ but for the interfering users, $k = 2, \cdots, K$, $\hat{\tau}_k$ is a random variable uniformly distributed over $[0, T]$. The thermal noise component, $n_q(t)$, is modeled as zero mean complex Gaussian process with two sided spectral density of N_o, and is assumed spatially uncorrelated. We assume that the signal is narrowband with respect to the antenna due to the fact that the fractional bandwidth of the signal is much less than unity.

We next assume that a single resolvable ray of the lth cluster arrives at the time delay $(T_{k,l} + \widetilde{\tau}_{k,l,i})$, then (IV.7) can be written as

$$\begin{aligned} r_q(t) &= \sum_{k=1}^{K} \alpha_{k,0,q} s_k(t - \hat{\tau}_k) + \sum_{k=1}^{K} \sum_{j=1}^{J-1} \alpha_{k,j,q} s_k \left(t - \tau_{k,j} - \hat{\tau}_k\right) + n_q(t) \\ &= \sum_{k=1}^{K} \sum_{j=0}^{J-1} \alpha_{k,j,q} s_k \left(t - \tau_{k,j} - \hat{\tau}_k\right) + n_q(t) \end{aligned} \tag{IV.8}$$

where $\tau_{k,j} \triangleq T_{k,l} + \widetilde{\tau}_{k,l,i}$ is the delay of the jth ray which is the ith ray in the lth cluster, and $J - 1 = (K_0 + K_1 + \cdots + K_{L-1})$ is the total number of rays in all the clusters. Also, $\alpha_{k,0,q} \triangleq \widetilde{\beta}_{k,q}$, $\alpha_{k,j,q} \triangleq \widetilde{a}_{k,q,j}$, and $\widetilde{a}_{k,q,j}$ is renaming of the multipath gain $\widetilde{a}_{k,q,l,i}$ where j corresponds to the ith ray of the lth cluster. The delay of the first multipath arrival in (IV.8) is set to zero.

After matched filtering with $g(-t)$, and substituting for $s_k(t)$ from (IV.6) in (IV.8), we have

$$\begin{aligned}\widetilde{r}_q(t) &= \sum_{k=1}^{K}\sqrt{P_k}\sum_{n=-\infty}^{\infty}A_{k,n}\sum_{j=0}^{J-1}\alpha_{k,j,q}x(t-nT-\tau_{k,j}-\hat{\tau}_k) \\ &\quad +b_q(t) \end{aligned} \tag{IV.9}$$

where $b_q(t) \stackrel{\Delta}{=} n_q(t) \odot g(-t)$ and $x(t)$ is the raised cosine pulse with a roll off factor ϱ and is given by

$$\begin{aligned} x(t) &= g(t) \odot g(-t) \\ &= \left[\frac{sin(\pi t/T)}{\pi t/T}\right]\left[\frac{cos(\pi \varrho t/T)}{1-4\varrho^2t^2/T^2}\right]. \end{aligned} \tag{IV.10}$$

Let us consider the following definition

$$\begin{aligned} y_{k,q}(t) &\stackrel{\Delta}{=} \sum_{j=0}^{J-1}\alpha_{k,j,q}x(t-\tau_{k,j}) \qquad \text{for } k=1,\cdots,K \\ &= \mathbf{x}_k(t)^T\boldsymbol{\alpha}_{k,q} \end{aligned} \tag{IV.11}$$

where $\mathbf{x}_k(t) \stackrel{\Delta}{=} [x(t-\tau_{k,0}), x(t-\tau_{k,1}), \cdots, x(t-\tau_{k,J-1})]$, and $\boldsymbol{\alpha}_{k,q} \stackrel{\Delta}{=} [\alpha_{k,0,q}, \alpha_{k,1,q}, \cdots, \alpha_{k,J-1,q}]$. Using the above definition, (IV.9) can be written as

$$\begin{aligned}\widetilde{r}_q(t) &= \sqrt{P_1}\sum_{n=-\infty}^{\infty}A_{1,n}y_{1,q}(t-nT) \\ &\quad +\sum_{k=2}^{K}\sqrt{P_k}\sum_{n=-\infty}^{\infty}A_{k,n}y_{k,q}(t-nT-\hat{\tau}_k)+b_q(t). \end{aligned} \tag{IV.12}$$

We consider a T-spaced linear equalizer at each of the receive antennas with a total number of taps $P = P_f + P_b$, where P_f are the total number of taps after the center tap and $P_b - 1$ are the total number of taps before the center tap. Sampling the matched filtered signal at $t = lT$ and stacking up in a vector form the signal at P taps at the qth antenna, we have

$$\bar{\mathbf{r}}_q(lT) \stackrel{\Delta}{=} [\widetilde{r}_q((l-(P_b-1))T), \cdots, \widetilde{r}_q(lT)\cdots, \widetilde{r}_q((l+P_f)T)]^T. \tag{IV.13}$$

Thus, the jth (for $j = 1, \cdots, P$) element of the vector $\bar{\mathbf{r}}_q(lT)$ can be written as

$$\begin{aligned}(\bar{\mathbf{r}}_q(lT))_j &= \widetilde{r}_q((l+j-P_b)T) \\ &= \sqrt{P_1}A_{1,l}\; y_{1,q}((j-P_b)T) + \sqrt{P_1}\sum_{\substack{m=-\infty \\ m\neq 0}}^{\infty} A_{1,l-m}\; y_{1,q}((m+j-P_b)T) \\ &\quad + \sum_{k=2}^{K}\sqrt{P_k}\sum_{m=-\infty}^{\infty} A_{k,l-m}\; y_{k,q}((m+j-P_b)T - \hat{\tau}_k) \\ &\quad + b_q((l+j-P_b)T). \end{aligned} \tag{IV.14}$$

The second, third and the fourth terms in (IV.14) are the contributions due to the ISI, MAI and the thermal noise, respectively, at the jth (for $j = 1, \cdots, P$) equalizer tap of the qth antenna. Combining the stacked up signal vector $\bar{\mathbf{r}}_q(lT)$ with the tap weights $\mathbf{w}_q$ we have

$$\begin{aligned} r_{q,l} &= \mathbf{w}_q^H \bar{\mathbf{r}}_q(lT) \\ &= \sqrt{P_1}A_{1,l}\; \mathbf{w}_q^H \mathbf{z}_{1,q} + \sqrt{P_1}\sum_{\substack{m=-\infty \\ m\neq 0}}^{\infty} A_{1,l-m}\; \mathbf{w}_q^H \widetilde{\mathbf{z}}_{1,q,m} \\ &\quad + \sum_{k=2}^{K}\sqrt{P_k}\sum_{m=-\infty}^{\infty} A_{k,l-m}\; \mathbf{w}_q^H \mathbf{z}_{k,q,m} + \mathbf{w}_q^H \mathbf{b}_{q,l} \end{aligned} \tag{IV.15}$$

where the vectors $\mathbf{z}_{1,q}$, $\widetilde{\mathbf{z}}_{1,q,m}$ and $\mathbf{z}_{k,q,m}$ are given below in terms of their jth component

$$\begin{aligned} (\mathbf{z}_{1,q})_j &\triangleq y_{1,q}((j-P_b)T) &&= \mathbf{x}_k((j-P_b)T)^T \boldsymbol{\alpha}_{1,q} \\ (\widetilde{\mathbf{z}}_{1,q,m})_j &\triangleq y_{1,q}((m+j-P_b)T) &&= \mathbf{x}_k((m+j-P_b)T)^T \boldsymbol{\alpha}_{1,q} \\ (\mathbf{z}_{k,q,m})_j &\triangleq y_{k,q}((m+j-P_b)T - \hat{\tau}_k) &&= \mathbf{x}_k((m+j-P_b)T - \hat{\tau}_k)^T \boldsymbol{\alpha}_{k,q} \\ (\mathbf{b}_{q,l})_j &\triangleq b_q((l+j-P_b)T). \end{aligned}$$

Finally, combining the signal from all the receive antennas we have

$$\begin{aligned} \widetilde{r}_l &\triangleq \sum_{q=1}^{Q} r_{q,l} = \mathbf{w}^H \mathbf{z}_l \\ \text{where } \mathbf{z}_l &\triangleq \sqrt{P_1}A_{1,l}\; \mathbf{z}_1 + \sqrt{P_1}\sum_{\substack{m=-\infty \\ m\neq 0}}^{\infty} A_{1,l-m}\; \widetilde{\mathbf{z}}_{1,m} \\ &\quad + \sum_{k=2}^{K}\sqrt{P_k}\sum_{m=-\infty}^{\infty} A_{k,l-m}\; \mathbf{z}_{k,m} + \mathbf{b}_l \end{aligned} \tag{IV.16}$$

where the following vector definitions are used $\mathbf{w} \triangleq [\mathbf{w}_1^T, \mathbf{w}_2^T, \cdots, \mathbf{w}_Q^T]^T$, $\mathbf{z}_1 \triangleq [\mathbf{z}_{1,1}^T, \mathbf{z}_{1,2}^T, \cdots, \mathbf{z}_{1,Q}^T]^T$, $\widetilde{\mathbf{z}}_{1,m} \triangleq [\widetilde{\mathbf{z}}_{1,1,m}^T, \widetilde{\mathbf{z}}_{1,2,m}^T, \cdots, \widetilde{\mathbf{z}}_{1,Q,m}^T]^T$, $\mathbf{z}_{k,m} \triangleq [\mathbf{z}_{k,1,m}^T, \mathbf{z}_{k,2,m}^T, \cdots, \mathbf{z}_{k,Q,m}^T]^T$, and $\mathbf{b}_l \triangleq [\mathbf{b}_{1,l}^T, \mathbf{b}_{2,l}^T, \cdots, \mathbf{b}_{Q,l}^T]^T$. Note that $\widetilde{r}_l$ in (IV.16) is complex valued output test statistics. Therefore, $\mathrm{Re}(\widetilde{r}_l)$ is the inphase channel output and $\mathrm{Img}(\widetilde{r}_l)$ is the quadrature channel output. Due to symmetry, we only consider the inphase component for the probability of error analysis. The inphase component can be written as

$$\begin{aligned}\mathrm{Re}(\widetilde{r}_l) &= \sqrt{P_1}\,\mathrm{Re}(A_{1,l}\mathbf{w}^H\mathbf{z}_1) + \sqrt{P_1}\sum_{\substack{m=-\infty\\ m\neq 0}}^{\infty}\mathrm{Re}(A_{1,l-m}\mathbf{w}^H\widetilde{\mathbf{z}}_{1,m}) \\ &\quad + \sum_{k=2}^{K}\sqrt{P_k}\sum_{m=-\infty}^{\infty}\mathrm{Re}(A_{k,l-m}\mathbf{w}^H\mathbf{z}_{k,m}) + \mathrm{Re}(\mathbf{w}^H\mathbf{b}_l). \qquad \text{(IV.17)}\end{aligned}$$

Consider the following identity:

$$\mathrm{Re}(\mathbf{a}^H\mathbf{b}\;c) = [\mathrm{Re}(\mathbf{a})^T, \mathrm{Img}(\mathbf{a})^T]\begin{bmatrix}\mathrm{Re}(\mathbf{b}), & -\mathrm{Img}(\mathbf{b})\\ \mathrm{Img}(\mathbf{b}), & \mathrm{Re}(\mathbf{b})\end{bmatrix}\begin{bmatrix}\mathrm{Re}(c)\\ \mathrm{Img}(c)\end{bmatrix} \qquad \text{(IV.18)}$$

where $\mathbf{a}$ and $\mathbf{b}$ are complex vectors, and c is a complex scalar. Using (IV.18) in (IV.17), we have

$$\begin{aligned}\mathrm{Re}(\widetilde{r}_l) &= \sqrt{P_1}\,y_1A_{1,l}^{(r)} + \sqrt{P_1}\,y_2A_{1,l}^{(i)} + \sqrt{P_1}\sum_{\substack{m=-\infty\\ m\neq 0}}^{\infty}(\widetilde{y}_{1,m}A_{1,l-m}^{(r)} + \widetilde{y}_{2,m}A_{1,l-m}^{(i)}) \\ &\quad + \sum_{k=2}^{K}\sqrt{P_k}\sum_{m=-\infty}^{\infty}(y_{1,k,m}A_{k,l-m}^{(r)} + y_{2,k,m}A_{k,l-m}^{(i)}) + \mathrm{Re}(\mathbf{w}^H\mathbf{b}_l) \qquad \text{(IV.19)}\end{aligned}$$

where

$$[y_1, y_2] \triangleq [\mathrm{Re}(\mathbf{w})^T, \mathrm{Img}(\mathbf{w})^T]\begin{bmatrix}\mathrm{Re}(\mathbf{z}_1), & -\mathrm{Img}(\mathbf{z}_1)\\ \mathrm{Img}(\mathbf{z}_1), & \mathrm{Re}(\mathbf{z}_1)\end{bmatrix}$$

$$[\widetilde{y}_{1,m}, \widetilde{y}_{2,m}] \triangleq [\mathrm{Re}(\mathbf{w})^T, \mathrm{Img}(\mathbf{w})^T]\begin{bmatrix}\mathrm{Re}(\widetilde{\mathbf{z}}_{1,m}), & -\mathrm{Img}(\widetilde{\mathbf{z}}_{1,m})\\ \mathrm{Img}(\widetilde{\mathbf{z}}_{1,m}), & \mathrm{Re}(\widetilde{\mathbf{z}}_{1,m})\end{bmatrix}$$

$$[y_{1,k,m}, y_{2,k,m}] \triangleq [\mathrm{Re}(\mathbf{w})^T, \mathrm{Img}(\mathbf{w})^T]\begin{bmatrix}\mathrm{Re}(\mathbf{z}_{k,m}), & -\mathrm{Img}(\mathbf{z}_{k,m})\\ \mathrm{Img}(\mathbf{z}_{k,m}), & \mathrm{Re}(\mathbf{z}_{k,m})\end{bmatrix}$$

and $A_{1,l}^{(r)}$ and $A_{1,l}^{(i)}$ are the real and imaginary parts, respectively, of the complex data symbol $A_{1,l}$. Similarly, $A_{k,l-m}^{(r)}$, and $A_{k,l-m}^{(i)}$ denotes the real and imaginary parts of the complex symbol $A_{k,l-m}$ for $k = 2, \cdots, K$. To simplify notation we further reduce (IV.19) to

$$\mathrm{Re}(\widetilde{r}_l) \;=\; \sqrt{P_1}\; y_1 A_{1,l}^{(r)} + \sqrt{P_1}\; y_2 A_{1,l}^{(i)} \; + x_1 + x_2 + x_3 \tag{IV.20}$$

where $x_1 \stackrel{\Delta}{=} \sqrt{P_1} \sum_{\substack{m=-\infty \\ m\neq 0}}^{\infty} (\widetilde{y}_{1,m} A_{1,l-m}^{(r)} + \widetilde{y}_{2,m} A_{1,l-m}^{(i)})$,
$x_2 \stackrel{\Delta}{=} \sum_{k=2}^{K} \sqrt{P_k} \sum_{m=-\infty}^{\infty} (y_{1,k,m} A_{k,l-m}^{(r)} + y_{2,k,m} A_{k,l-m}^{(i)})$, and $x_3 \stackrel{\Delta}{=} \mathrm{Re}(\mathbf{w}^H \mathbf{b}_l)$.

IV.C Optimum Weights

We use the MMSE criterion to determine the weights $\mathbf{w}$, where we condition on the channel of all the users and delays of the interfering users, i.e., we condition on the set $B = \{\mathbf{z}_1, \widetilde{\mathbf{z}}_{1,m}, \mathbf{z}_{k,m}\}$. Since $\widetilde{r}_l$ in (IV.16) is the estimate of the complex symbol $A_{1,l}$, using the orthogonality principle we have

$$E[(\widetilde{r}_l - A_{1,l})\mathbf{z}_l^H | B] = 0 \tag{IV.21}$$

where $\mathbf{z}_l$ is given in (IV.16). Simplifying above we have

$$\mathbf{w} \;=\; \sqrt{P_1} c_1 \mathbf{\Psi}^{-1} \mathbf{z}_1 \tag{IV.22}$$

where the correlation matrix $\mathbf{\Psi} \stackrel{\Delta}{=} E[\mathbf{z}_l \mathbf{z}_l^H | B]$ and the average symbol power $c_1 \stackrel{\Delta}{=} E[|A_{1,l}|^2]$ in (IV.22) are given below

$$\mathbf{\Psi} \;=\; c_1 \left(P_1\; \mathbf{z}_1 \mathbf{z}_1^H \; + P_1 \sum_{\substack{m=-\infty \\ m\neq 0}}^{\infty} \widetilde{\mathbf{z}}_{1,m}\, \widetilde{\mathbf{z}}_{1,m}^H + \sum_{k=2}^{K} P_k \sum_{m=-\infty}^{\infty} \mathbf{z}_{k,m} \mathbf{z}_{k,m}^H \right) + \mathbf{R}. \tag{IV.23}$$

In the above, $\mathbf{R} \stackrel{\Delta}{=} E[\mathbf{b}_l \mathbf{b}_l^H]$. The cross terms in (IV.23) are zero because of the independence of the data of the desired and the interfering users.

We evaluate the average symbol power c_1 for the case of a rectangular M-QAM where $M = M_H M_V$, and M_H is the number of symbols for the horizontal PAM and M_V is the number of symbols for the vertical PAM. Then it can be shown that

$$c_1 = E[|A_{1,l}|^2] = E[(A_{1,l}^{(r)})^2 + (A_{1,l}^{(i)})^2] = \left[\frac{1}{3}(M_H^2 + M_V^2 - 2)\right]. \tag{IV.24}$$

Finally, the covariance matrix $\mathbf{R}$ is computed as shown below

$$\mathbf{R} \triangleq E[\mathbf{b}_l \mathbf{b}_l^H] = \mathbf{I}_Q \otimes \mathbf{R}_1 \tag{IV.25}$$

where $\mathbf{I}_Q$ is $Q \times Q$ identity matrix and $\otimes$ denotes the Kronecker product. The correlation matrix $\mathbf{R}_1$ in (IV.25) is defined as $\mathbf{R}_1 \triangleq E[\mathbf{b}_{q,l}\mathbf{b}_{q,l}^H]$ any $q = 1, \cdots, Q$. The ijth $(i, j = 1, \cdots, P)$ component of the matrix $\mathbf{R}_1$ is given by

$$\begin{aligned}
(\mathbf{R}_1)_{i,j} &= E[(\mathbf{b}_{q,l})_i (\mathbf{b}_{q,l}^*)_j] \\
&= E\left[\left(\int_{-\infty}^{\infty} n_q((l+i-P_b)T - \varepsilon) g(-\varepsilon) d\varepsilon\right)\right. \\
&\quad \left. \times \left(\int_{-\infty}^{\infty} n_q((l+j-P_b)T - \zeta)^* g(-\zeta) d\zeta\right)\right] \\
&= x((i-j)T) N_o.
\end{aligned} \tag{IV.26}$$

IV.D Probability of Error Analysis

Consider the real part of the output test statistics in (IV.20). Note that the third, fourth and the fifth terms in (IV.20) are due to ISI, MAI, and the thermal noise, respectively. In this work, we are assume that the total number of interfering users is small, therefore, the Gaussian approximation for MAI, valid for large number of users, does not hold in our work. Also, the ISI term is Non-Gaussian.

Next, note that $\Upsilon_R \triangleq \mathrm{Re}(\widetilde{r}_l)$ conditioned on the terms $\{y_1, A_{1,l}^{(r)}, y_2, A_{1,l}^{(i)}, x_1, x_2\}$ is Gaussian. Its conditional distribution can, thus, be written as $\sim N(\mu_R, \sigma_R^2)$ where the mean and variance are

$$\begin{aligned}
\mu_R &\triangleq \sqrt{P_1}\, y_1 A_{1,l}^{(r)} + \sqrt{P_1}\, y_2 A_{1,l}^{(i)} + x_1 + x_2 \\
\sigma_R^2 &\triangleq E[\mathrm{Re}(\mathbf{w}^H \mathbf{b}_l)^2] = \frac{1}{2}\mathbf{w}^H E[\mathbf{b}_l \mathbf{b}_l^H]\mathbf{w} = \frac{1}{2}\mathbf{w}^H \mathbf{R} \mathbf{w}.
\end{aligned} \tag{IV.27}$$

IV.D.1 Derivation for the conditional BER for M-QAM

In this section we derive the conditional BER, where the conditioning is on the channel of all the users, delays of the interfering users, ISI and the MAI, i.e., we condition on the terms $\{y_1, y_2, x_1, x_2\}$. Also, note that the derivation in this section, as well as those in Sections IV.D.2 and IV.D.3, hold for any arbitrary weight vector $\mathbf{w}$. Consider a rectangular M-QAM configuration. Let the horizontal PAM (H-PAM) be of dimension M_H and vertical PAM (V-PAM) be of dimension M_V, then $M = M_H M_V$. The total number of bits transmitted are $log_2(M) = log_2(M_H) + log_2(M_V)$. Let us denote $G \triangleq log_2(M)$, $G_H \triangleq log_2(M_H)$, and $G_V \triangleq log_2(M_H)$. Figure IV.7 shows the square 16-QAM constellation. The complex symbols (shown by the black dots in the figure) and the corresponding gray encoded bit assignment are shown in the figure. For the bit assignment shown in the Figure IV.7 for 16-QAM, the first and the third bits are transmitted through the inphase bit stream and the second and the fourth bits are transmitted through the quadrature bit stream.

Let us denote the inphase symbols by $\{s_{m_r}^{(r)}\}_{m_r=1}^{M_H}$ and the corresponding *Gray encoded* G_H bits by $(\widetilde{g}_{m_r,G_H}, \widetilde{g}_{m_r,G_H-1}, \cdots, \widetilde{g}_{m_r,1})$, where $\widetilde{g}_{m_r,G_H}$ is the most significant bit (MSB) and $\widetilde{g}_{m_r,1}$ is the least significant bit (LSB). Also, $s_{m_r}^{(r)}$ takes on values $\{\pm d, \pm 3d, \cdots, \pm(M_H - 1)d\}$. Similarly, let $s_{m_i}^{(i)}$ denote the quadrature symbols that takes values $\{\pm d, \pm 3d, \cdots, \pm(M_V - 1)d\}$ and has *Gray encoded* G_V bits given by $(\hat{g}_{m_i,G_V}, \hat{g}_{m_i,G_V-1}, \cdots, \hat{g}_{m_i,1})$. For the QAM constellation, the complex symbols are of the form

$$s_m = s_{m_r}^{(r)} + i s_{m_i}^{(i)}, \qquad \text{for } m = 1, \cdots, M. \tag{IV.28}$$

The Gray code for s_m is obtained from the Gray code for the horizontal and vertical PAM.

In Figure IV.7 for 16-QAM, in terms of the notation defined above, $(\widetilde{g}_{m_r,2}, \widetilde{g}_{m_r,1})$ are the inphase bits and $(\hat{g}_{m_i,2}, \hat{g}_{m_i,1})$ are the quadrature bits for the mth symbol point in the constellation. Note that the inphase and the quadrature components of the output test statistics are demodulated independently. Therefore, the bit error prob-

ability of these two components can be computed separately, infact, for a square QAM constellation the BER for the two components would be the same. Thus, in the following we will concentrate on the BER for the horizontal PAM. The BER for the vertical PAM for a different dimension can be computed similarly.

Consider the horizontal PAM (for example, see Figure IV.7 for 4-PAM). The total number of inphase bits are G_H. We will compute the BER for each of these bits. Note that our derivation for BER is more general compared to past works and can be reduced to the special cases considered before [89], [90]. Unlike the work in [89], [90], we derive closed-form expression for QAM BER that is valid for any QAM size and any distribution for the output test statistics. For clarity, detailed steps for derivation (starting from the basic principles [87]) are presented.

Next, we define the complex signal sets

$$\begin{aligned} S_{1,j} &\triangleq \{\text{all } s_m \text{ with bit } \widetilde{g}_{m_r,j} = 1\} \text{ and} \\ S_{0,j} &\triangleq \{\text{all } s_m \text{ with bit } \widetilde{g}_{m_r,j} = 0\} \qquad \text{for } j = 1, \cdots, G_H. \end{aligned}$$

Note that the cardinality of these sets are $|S_{1,j}| = |S_{0,j}| = M/2$. Using these definitions, the probability of error for the jth (for $j = 1, \cdots, G_H$) bit can be written as

$$\begin{aligned} & P_{e,c}^{H}(j) \\ &= \frac{1}{2|S_{1,j}|} \sum_{\text{all } s_{m'} \in S_{1,j}} Prob\Big\{ \Upsilon_R \in I_{H,m_r} \text{ , for all } s_{m_r}^{(r)} \text{ such that } s_m \in S_{0,j} \\ &\qquad\qquad \Big| s_{m'} \text{ was transmitted} \Big\} \\ &+ \frac{1}{2|S_{0,j}|} \sum_{\text{all } s_{m'} \in S_{0,j}} Prob\Big\{ \Upsilon_R \in \ I_{H,m_r}, \text{ for all } s_{m_r}^{(r)} \text{ such that } s_m \in S_{1,j} \\ &\qquad\qquad \Big| \ s_{m'} \text{ was transmitted} \Big\}. \end{aligned} \tag{IV.29}$$

In (IV.29), I_{H,m_r} is the interval associated with $s_{m_r}^{(r)}$ on the I-axis (or H-PAM) given by $I_{H,m_r} = [s_{m_r}^{(r)} - d, s_{m_r}^{(r)} + d]$ except for the boundary points for which the intervals as $(-\infty, -(M_H - 2)d]$ and $[(M_H - 2)d, \infty)$.

Let $f_{\Upsilon_R|\widetilde{S}}(x|s_{m'})$ denote the conditional pdf of Υ_R conditioned on a set $\widetilde{S}$, then

(IV.29) can be written as

$$P_{e,c}^{H}(j) = \frac{1}{M}\left[\sum_{s_{m'}\in S_{1,j}}\sum_{s_m\in S_{0,j}}\int_{I_{H,m_r}} f_{\Upsilon_R|\widetilde{S}}(x|s_{m'})\,dx + \sum_{s_{m'}\in S_{0,j}}\sum_{s_m\in S_{1,j}}\int_{I_{H,m_r}} f_{\Upsilon_R|\widetilde{S}}(x|s_{m'})dx\right]. \quad \text{(IV.30)}$$

For the bit assignment shown in Figure IV.7 for 16 QAM and similar bit assignments for higher QAM and from the properties of Gray code, it can be shown that the number of disjoint intervals over which integration is performed in the first term in (IV.30) is $M_H/(2^{j+1})$, and for the second term it is $(M_H/(2^{j+1}))+1$ for $j = 1, \cdots, G_H - 1$. Note that the intervals for the integral in the second term include the two boundary intervals, i.e., $(-\infty, (-M_H + 2^j)d]$ and $[(M_H - 2^j)d, \infty)$. Thus, we can write (IV.30) as

$$\begin{aligned} P_{e,c}^{H}(j) = & \\ & \frac{1}{M}\sum_{s_{m_i}^{(i)}\in\bar{S}}\left[\sum_{s_{m_r}^{(r)}\in\hat{S}_{1,j}}\sum_{l=1}^{M_H/(2^{j+1})}\int_{L_l}^{U_l} f_{\Upsilon_R|\widetilde{S}}(x|s_{m_r}^{(r)}, s_{m_i}^{(i)})\,dx\right. \\ & + \sum_{s_{m_r}^{(r)}\in\hat{S}_{0,j}}\left(\left\{\sum_{l=1}^{(M_H/(2^{j+1}))-1}\int_{U_l}^{L_{l+1}} f_{\Upsilon_R|\widetilde{S}}(x|s_{m_r}^{(r)}, s_{m_i}^{(i)})dx\right\}\right. \\ & \left.\left. + \int_{-\infty}^{(-M_H+2^j)d} f_{\Upsilon_R|\widetilde{S}}(x|s_{m_r}^{(r)}, s_{m_i}^{(i)})dx + \int_{(M_H-2^j)d}^{\infty} f_{\Upsilon_R|\widetilde{S}}(x|s_{m_r}^{(r)}, s_{m_i}^{(i)})dx\right)\right] \end{aligned} \quad \text{(IV.31)}$$

for $j = 1, \cdots, G_H - 1$. In (IV.31), we have replaced the first summation over the complex symbols in the first and the second terms in (IV.30) with two separate summations over the inphase and the quadrature symbols. Also, $L_l = (-M_H + 2^j(4l-3))d$ and $U_l = (-M_H + 2^j(4l-1))d$. The sets $\bar{S}$, $\hat{S}_{1,j}$ and $\hat{S}_{0,j}$ are given by $\bar{S} = \{(-M_V + 2j - 1)d\}_{j=1}^{M_V}$, $\hat{S}_{1,j} \triangleq \{\text{all } s_{m_r}^{(r)} \text{ with bit } \widetilde{g}_{m_r,j} = 1\}$ and $\hat{S}_{0,j} \triangleq \{\text{all } s_{m_r}^{(r)} \text{ with bit } \widetilde{g}_{m_r,j} = 0\} =$

$\hat{S} - \hat{S}_{1,j}$, where $\hat{S} = \{(-M_H + 2j - 1)d\}_{j=1}^{M_H}$. For $j = G_H$, (IV.30) can be simplified to

$$P_{e,c}^{H}(G_H) = \frac{1}{M} \sum_{s_{m_i}^{(i)} \in \bar{S}} \left[\sum_{s_{m_r}^{(r)} \in \hat{S}_{1,G_H}} \left(\int_0^{\infty} f_{\Upsilon_R|\widetilde{S}}(x|s_{m_r}^{(r)}, s_{m_i}^{(i)})\, dx + \int_{-\infty}^{0} f_{\Upsilon_R|\widetilde{S}}(x| - s_{m_r}^{(r)}, -s_{m_i}^{(i)})\, dx \right) \right]. \quad \text{(IV.32)}$$

In our case, Υ_R conditioned on the terms $\widetilde{S} \triangleq \{y_1, A_{1,l}^{(r)}, y_2, A_{1,l}^{(i)}, x_1, x_2\}$ is Gaussian $\sim N(\mu_R, \sigma_R^2)$ (see before (IV.27)). Thus, its conditional density, condition on the set $\widetilde{S}$, can be written as

$$f_{\Upsilon_R|\widetilde{S}}(x) = \frac{1}{\sqrt{2\pi}\sigma_R} e^{\frac{-(x-\mu_R)^2}{2\sigma_R^2}}. \quad \text{(IV.33)}$$

Thus, we have

$$\int_{L_l}^{U_l} f_{\Upsilon_R|\widetilde{S}}(x|s_{m_r}^{(r)}, s_{m_i}^{(i)})\, dx = Q\left(\frac{L_l - \mu_R}{\sigma_R}\right) - Q\left(\frac{U_l - \mu_R}{\sigma_R}\right) \quad \text{(IV.34)}$$

where $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{\frac{-y^2}{2}} dy$. Substituting (IV.34) in (IV.31), we have

$$\begin{aligned} P_{e,c}^{H}(j) = & \\ & \frac{1}{M} \sum_{s_{m_i}^{(i)} \in \bar{S}} \left[\sum_{s_{m_r}^{(r)} \in \hat{S}_{1,j}} \sum_{l=1}^{M_H/(2^{j+1})} \left[Q\left(\frac{L_l - \mu_R(s_{m_r}^{(r)}, s_{m_i}^{(i)})}{\sigma_R}\right) - Q\left(\frac{U_l - \mu_R(s_{m_r}^{(r)}, s_{m_i}^{(i)})}{\sigma_R}\right) \right] \right. \\ & + \sum_{s_{m_r}^{(r)} \in \hat{S}_{0,j}} \left(\left\{ \sum_{l=1}^{(M_H/(2^{j+1}))-1} \left[Q\left(\frac{U_l - \mu_R(s_{m_r}^{(r)}, s_{m_i}^{(i)})}{\sigma_R}\right) - Q\left(\frac{L_{l+1} - \mu_R(s_{m_r}^{(r)}, s_{m_i}^{(i)})}{\sigma_R}\right) \right] \right\} \right. \\ & \left. \left. + Q\left(\frac{(M_H - 2^j)d + \mu_R(s_{m_r}^{(r)}, s_{m_i}^{(i)})}{\sigma_R}\right) + Q\left(\frac{(M_H - 2^j)d - \mu_R(s_{m_r}^{(r)}, s_{m_i}^{(i)})}{\sigma_R}\right) \right) \right]. \end{aligned} \quad \text{(IV.35)}$$

Note that the conditional mean μ_R depends on the transmitted inphase and quadrature symbols $(s_{m_r}^{(r)}, s_{m_i}^{(i)})$, this is made explicit in (IV.35) by the functional form $\mu_R(s_{m_r}^{(r)}, s_{m_i}^{(i)})$. The conditional mean μ_R (conditioned on the set $\widetilde{S}$ (see before (IV.33))) is given in (IV.27). Using the notation developed in this section, we rewrite μ_R as a function of the transmitted signal $(s_{m_r}^{(r)}, s_{m_i}^{(i)})$ as

$$\mu_R(s_{m_r}^{(r)}, s_{m_i}^{(i)}) = y_1 s_{m_r}^{(r)} + y_2 s_{m_i}^{(i)} + x_1 + x_2 \quad \text{(IV.36)}$$

where $d = \sqrt{P_1}$. Let $x_{1,2} \triangleq x_1 + x_2$, then using this definition and substituting (IV.36) in (IV.35) we have

$$P_{e,c}^{H}(j) =$$

$$\frac{1}{M}\sum_{s_{m_i}^{(i)}\in\bar{S}}\left[\sum_{s_{mr}^{(r)}\in\hat{S}_{1,j}}\sum_{l=1}^{M_H/(2^{j+1})}\left[Q\left(\frac{L_l - y_1 s_{mr}^{(r)} - y_2 s_{m_i}^{(i)} - x_{1,2}}{\sigma_R}\right) - Q\left(\frac{U_l - y_1 s_{mr}^{(r)} - y_2 s_{m_i}^{(i)} - x_{1,2}}{\sigma_R}\right)\right]\right.$$

$$+\sum_{s_{mr}^{(r)}\in\hat{S}_{0,j}}\left[\left\{\sum_{l=1}^{(M_H/(2^{j+1}))-1}\left[Q\left(\frac{U_l - y_1 s_{mr}^{(r)} - y_2 s_{m_i}^{(i)} - x_{1,2}}{\sigma_R}\right) - Q\left(\frac{L_{l+1} - y_1 s_{mr}^{(r)} - y_2 s_{m_i}^{(i)} - x_{1,2}}{\sigma_R}\right)\right]\right\}\right.$$

$$\left.\left.+Q\left(\frac{(M_H - 2^j)d + y_1 s_{mr}^{(r)} + y_2 s_{m_i}^{(i)} + x_{1,2}}{\sigma_R}\right) + Q\left(\frac{(M_H - 2^j)d - y_1 s_{mr}^{(r)} - y_2 s_{m_i}^{(i)} - x_{1,2}}{\sigma_R}\right)\right]\right]. \tag{IV.37}$$

From (IV.32), (IV.34) and (IV.36), we have

$$P_{e,c}^{H}(G_H) =$$

$$\frac{1}{M}\sum_{s_{m_i}^{(i)}\in\bar{S}}\left[\sum_{s_{mr}^{(r)}\in\hat{S}_{1,G_H}}\left[Q\left(\frac{-y_1 s_{mr}^{(r)} - y_2 s_{m_i}^{(i)} - x_{1,2}}{\sigma_R}\right) + Q\left(\frac{-y_1 s_{mr}^{(r)} - y_2 s_{m_i}^{(i)} + x_{1,2}}{\sigma_R}\right)\right]\right]. \tag{IV.38}$$

The BER for the M-QAM can be written as

$$P_{e,c} = \frac{1}{G}\left(\sum_{j=1}^{G_H} P_{e,c}^{H}(j) + \sum_{j=1}^{G_V} P_{e,c}^{V}(j)\right). \tag{IV.39}$$

IV.D.2 The conditional BER for 16-QAM

In this section, we will specialize the results derived in the previous section for the case of 16 QAM. Note that the BER for 16 QAM derived in this section is conditioned on the variables $\{y_1, y_2, x_1, x_2\}$. Based on the BER expression in (IV.39), the BER for square 16-QAM, shown in Figure IV.7, can be written. For 16-QAM we have: $M = 16$, $G = 4$, $G_H = G_V = 2$, and $P_{e,c}^{H} = P_{e,c}^{V}$. Substituting these in (IV.39), the BER is given by

$$P_{e,c} = \frac{1}{2}\left(\sum_{j=1}^{2} P_{e,c}^{H}(j)\right). \tag{IV.40}$$

From Figure IV.7, we identify the sets:

$\hat{S}_{1,2} = \{\text{all } s_{m_r}^{(r)} \text{ with bit } \widetilde{g}_{m_r,2} = 1\} = \{-3d, -d\}$

$\hat{S}_{0,2} = \{\text{all } s_{m_r}^{(r)} \text{ with bit } \widetilde{g}_{m_r,2} = 0\} = \{3d, d\}$

$\hat{S}_{1,1} = \{\text{all } s_{m_r}^{(r)} \text{ with bit } \widetilde{g}_{m_r,1} = 1\} = \{3d, -3d\}$

$\hat{S}_{0,1} = \{\text{all } s_{m_r}^{(r)} \text{ with bit } \widetilde{g}_{m_r,1} = 0\} = \{d, -d\}$ and $\bar{S} = \{(-M_V + 2j - 1)d\}_{j=1}^{M_V} = \{-3d, -d, d, 3d\}$. $P_{e,c}^{H}(2)$ is obtained by evaluating (IV.38) over $\bar{S}$ and $\hat{S}_{1,2}$ given above and setting $M = 16$. $P_{e,c}^{H}(1)$ is obtained by simplifying (IV.37), and is given below

$$P_{e,c}^{H}(1) = \frac{1}{16} \sum_{s_{m_i}^{(i)} \in \bar{S}} \left[\sum_{s_{m_r}^{(r)} \in \hat{S}_{1,1}} \left[Q\left(\frac{-2d - y_1 s_{m_r}^{(r)} - y_2 s_{m_i}^{(i)} - x_{1,2}}{\sigma_R} \right) - Q\left(\frac{2d - y_1 s_{m_r}^{(r)} - y_2 s_{m_i}^{(i)} - x_{1,2}}{\sigma_R} \right) \right] + \sum_{s_{m_r}^{(r)} \in \hat{S}_{0,1}} \left(Q\left(\frac{2d + y_1 s_{m_r}^{(r)} + y_2 s_{m_i}^{(i)} + x_{1,2}}{\sigma_R} \right) + Q\left(\frac{2d - y_1 s_{m_r}^{(r)} - y_2 s_{m_i}^{(i)} - x_{1,2}}{\sigma_R} \right) \right) \right]. \quad \text{(IV.41)}$$

Further simplifications result by noting that $y_2 = 0$ for the MMSE receiver.

Note that the BER expression in general can be written in terms of finite sum over the $Q(.)$ function. The above BER is conditioned on the channel of all the users and delays of the interfering users, the ISI and the MAI (i.e., terms y_1, y_2, x_1, x_2). Thus, the argument of the $Q(.)$ in the BER expression depends on $x_{1,2}$ (see after (IV.36)). Next, we uncondition the BER with respect to the ISI and MAI.

IV.D.3 BER unconditioned over ISI and MAI

In order to uncondition the BER derived in section IV.D.1 with respect to the ISI and the MAI (note that the conditioning on the channel of all the users and delays of the interfering users remain), we adopt an approximate Fourier series expression for the $Q(.)$ function given in [88], obtained using the Chernoff bound and the Fourier series representation. According to [88], the $Q(.)$ function can be approximated as

$$Q(z) = \frac{1}{2} - \frac{2}{\pi} \sum_{\substack{j=-\infty \\ j \text{ odd}}}^{\infty} \frac{exp(-j^2\omega^2/2)\, exp(ij\omega z)}{2ij} + \xi(z) \quad \text{(IV.42)}$$

where $i = \sqrt{-1}$ and $\xi(z)$ is the approximation error and is bounded by [88]

$$|\xi(z)| \leq Q\left(\frac{T_e}{2} - |z|\right) \tag{IV.43}$$

and $\omega = 2\pi/T_e$ is the Fourier series frequency and T_e is the Fourier series period [88]. Note that the bound in (IV.43) is a decreasing function of T_e. Thus, the approximation error in (IV.42) (i.e., the second term in (IV.42)) has a tight upper bound when the following are satisfied: $|z| << T_e/2$ and T_e is large, for this case, the error in approximating the $Q(.)$ function with the series in the first term in (IV.42) is small.

Recall that the conditional BER expression in general can be written in terms of a finite sum over the $Q(.)$ function, where each $Q(.)$ has the form $Q(\widetilde{a}\widetilde{x} - \widetilde{a}x_1 - \widetilde{a}x_2)$, and where $\widetilde{a} \triangleq 1/\sigma_R$ and $\widetilde{x} \triangleq \varsigma_m - y_1 s_{m_r}^{(r)} - \ y_2 s_{m_i}^{(i)}$ (see, for example, (IV.37)). Thus, we will average this general form of the $Q(.)$ over x_1 and x_2. Using (IV.42), we have

$$\begin{aligned} Q(\widetilde{a}\widetilde{x} - \widetilde{a}x_1 - \widetilde{a}x_2) &= \frac{1}{2} - \frac{2}{\pi} \sum_{\substack{j=-\infty \\ j \text{ odd}}}^{\infty} \frac{exp(-j^2\omega^2/2)}{2ij}\, exp(ij\omega\widetilde{a}(\widetilde{x} - x_1 - x_2)) \\ &\quad + \xi(\widetilde{a}\widetilde{x} - \widetilde{a}x_1 - \widetilde{a}x_2). \end{aligned} \tag{IV.44}$$

Next we average (IV.44) w.r.t to x_1, yielding

$$\begin{aligned} Q_1(\widetilde{x}, x_2) &\triangleq E_{x_1}[Q(\widetilde{a}\widetilde{x} - \widetilde{a}x_1 - \widetilde{a}x_2)] \\ &= \frac{1}{2} - \frac{2}{\pi} \sum_{\substack{j=-\infty \\ j \text{ odd}}}^{\infty} \frac{exp(-j^2\omega^2/2)}{2ij} exp(ij\omega\widetilde{a}(\widetilde{x} - x_2)) E_{x_1}[\, exp(-x_1 ij\omega\widetilde{a})] \\ &\quad + E_{x_1}[\xi(\widetilde{a}\widetilde{x} - \widetilde{a}x_1 - \widetilde{a}x_2)]. \end{aligned} \tag{IV.45}$$

Note that the order of expectation operator and the infinite summation in (IV.45) commutes. This can be proved as follows: In order to show that the expectation and the infinite summation order can be swapped, we need to prove the following:

$$\widetilde{L} \triangleq \sum_{\substack{j=-\infty \\ j \text{ odd}}}^{\infty} E_{x_1}\left[\left|\frac{exp(-j^2\omega^2/2)}{2ij} exp(ij\omega\widetilde{a}(\widetilde{x} - x_2))\, exp(-x_1 ij\omega\widetilde{a})\right|\right] < \infty. \tag{IV.46}$$

Simplifying (IV.46), we get

$$\begin{aligned}\widetilde{L} &< \sum_{\substack{j=-\infty \\ j \text{ odd}}}^{\infty} \frac{exp(-j^2\omega^2/2)}{2|j|} = 2\sum_{\substack{j=1 \\ j \text{ odd}}}^{\infty} \frac{exp(-j^2\omega^2/2)}{2j} \\ &< \sum_{\substack{j=1 \\ j \text{ odd}}}^{\infty} exp(-j^2\omega^2/2) < \infty. \end{aligned} \tag{IV.47}$$

The series in (IV.47) converges. Thus, the order of expectation and the infinite summation in (IV.45) can be swapped.

Let us define $f(j) \triangleq \frac{exp(-j^2\omega^2/2)}{2ij}$, $\rho(\widetilde{x}, x_2) \triangleq E_{x_1}[\xi(\widetilde{a}\widetilde{x} - \widetilde{a}x_1 - \widetilde{a}x_2)]$, $c_1 \triangleq j\omega\widetilde{a}$, and $\phi_{x_1}(c_1) \triangleq E_{x_1}[\ exp(-x_1 ij\omega\widetilde{a})] = E_{x_1}[\ exp(-x_1 ic_1)]$. Then we can write (IV.45) as

$$Q_1(\widetilde{x}, x_2) = \frac{1}{2} - \frac{2}{\pi}\sum_{\substack{j=-\infty \\ j \text{ odd}}}^{\infty} f(j)e^{(ic_1(\widetilde{x}-x_2))}\ \phi_{x_1}(c_1) + \rho(\widetilde{x}, x_2). \tag{IV.48}$$

Note that x_1, defined after (IV.20), is given in term of the random data symbols $D_1 \triangleq \{A^{(r)}_{1,l-m}, A^{(i)}_{1,l-m}\}^{\infty}_{\substack{m=-\infty \\ m\neq 0}}$. Using this definition, we obtain $\phi_{x_1}(c_1)$ as shown below:

$$\begin{aligned}&\phi_{x_1}(c_1) \\ &= E_{\left\{A^{(r)}_{1,l-m}, A^{(i)}_{1,l-m}\right\}^{\infty}_{\substack{m=-\infty \\ m\neq 0}}}\left[exp\left(-ic_1\left[\sqrt{P_1}\sum_{\substack{m=-\infty \\ m\neq 0}}^{\infty}(\widetilde{y}_{1,m}A^{(r)}_{1,l-m} + \widetilde{y}_{2,m}A^{(i)}_{1,l-m})\right]\right)\right].\end{aligned} \tag{IV.49}$$

Further simplifying, we have

$$\begin{aligned}\phi_{x_1}(c_1) &= E_{\left\{A^{(r)}_{1,l-m}\right\}^{\infty}_{\substack{m=-\infty, \\ m\neq 0}}}\left[exp\left(-ic_1\sqrt{P_1}\left[\sum_{\substack{m=-\infty \\ m\neq 0}}^{\infty}\left(\widetilde{y}_{1,m}A^{(r)}_{1,l-m}\right)\right]\right)\right] \\ &\quad \times E_{\left\{A^{(i)}_{1,l-m}\right\}^{\infty}_{\substack{m=-\infty, \\ m\neq 0}}}\left[exp\left(-ic_1\sqrt{P_1}\left[\sum_{\substack{m=-\infty \\ m\neq 0}}^{\infty}(\widetilde{y}_{2,m}A^{(i)}_{1,l-m})\right]\right)\right]\end{aligned}$$

$$
\begin{aligned}
&= \left[\prod_{\substack{m=-\infty \\ m\neq 0}}^{\infty} E_{A_{1,l-m}^{(r)}}\left[exp\Big(-ic_1\sqrt{P_1}\widetilde{y}_{1,m}A_{1,l-m}^{(r)}\Big)\right]\right] \\
&\quad \times \left[\prod_{\substack{m=-\infty \\ m\neq 0}}^{\infty} E_{A_{1,l-m}^{(i)}}\left[exp\Big(-ic_1\sqrt{P_1}\widetilde{y}_{2,m}A_{1,l-m}^{(i)}\Big)\right]\right] \\
&= \left(\prod_{\substack{m=-\infty \\ m\neq 0}}^{\infty} \phi_{x_1}^{(1)}(c_1,m)\right)\left(\prod_{\substack{m=-\infty \\ m\neq 0}}^{\infty} \phi_{x_1}^{(2)}(c_1,m)\right) \qquad \text{(IV.50)}
\end{aligned}
$$

where $\phi_{x_1}^{(1)}(c_1,m) \triangleq E_{A_{1,l-m}^{(r)}}\left[\, exp\Big(-ic_1\sqrt{P_1}\,\widetilde{y}_{1,m}A_{1,l-m}^{(r)}\Big)\right]$ and $\phi_{x_1}^{(2)}(c_1,m) \triangleq E_{A_{1,l-m}^{(i)}}\left[\, exp\Big(-ic_1\sqrt{P_1}\,\widetilde{y}_{2,m}A_{1,l-m}^{(i)}\Big)\right]$. Letting $\gamma_{1,m} \triangleq c_1\sqrt{P_1}\,\widetilde{y}_{1,m}$, we have

$$
\begin{aligned}
\phi_{x_1}^{(1)}(c_1,m) &= E_{A_{1,l-m}^{(r)}}\left[\, exp\Big(-i\gamma_{1,m}A_{1,l-m}^{(r)}\Big)\right] \\
&= \frac{1}{M_H}\,[\, exp(i\gamma_{1,m}) + exp(-i\gamma_{1,m}) + exp(3i\gamma_{1,m}) + exp(-3i\gamma_{1,m}) + \\
&\qquad \cdots + exp(M_H-1)i\gamma_{1,m}) + exp(-(M_H-1)i\gamma_{1,m})] \\
&= \frac{1}{M_H}\left[\sum_{\widetilde{m}=1}^{M_H} e^{(M_H+1-2\widetilde{m})i\gamma_{1,m}}\right] = \frac{1}{M_H}\left[\frac{sin(M_H\gamma_{1,m})}{sin(\gamma_{1,m})}\right]. \qquad \text{(IV.51)}
\end{aligned}
$$

Similarly, we have

$$
\phi_{x_1}^{(2)}(c_1,m) = \frac{1}{M_V}\left[\frac{sin(M_V\gamma_{2,m})}{sin(\gamma_{2,m})}\right] \qquad \text{(IV.52)}
$$

where $\gamma_{2,m} \triangleq c_1\sqrt{P_1}\,\widetilde{y}_{2,m}$. Substituting (IV.51) and (IV.52) in (IV.50), we have

$$
\begin{aligned}
\phi_{x_1}(c_1) &= \left(\prod_{\substack{m=-\infty \\ m\neq 0}}^{\infty} \frac{1}{M_H}\left[\frac{sin(M_H\gamma_{1,m})}{sin(\gamma_{1,m})}\right]\right) \\
&\quad \times \left(\prod_{\substack{m=-\infty \\ m\neq 0}}^{\infty} \frac{1}{M_V}\left[\frac{sin(M_V\gamma_{2,m})}{sin(\gamma_{2,m})}\right]\right). \qquad \text{(IV.53)}
\end{aligned}
$$

Next we average Q_1 in (IV.48) w.r.t to x_2 (note that x_1 and x_2 are independent):

$$\begin{aligned} Q_2(\widetilde{x}) &\triangleq E_{x_2}[Q_1(\widetilde{x}, x_2)] \\ &= \frac{1}{2} - \frac{2}{\pi} \sum_{\substack{j=-\infty \\ j \text{ odd}}}^{\infty} f(j) e^{(ic_1\widetilde{x})}\ \phi_{x_1}(c_1) E_{x_2}[e^{(-ic_1x_2)}] + E_{x_2}[\rho(\widetilde{x}, x_2)] \\ &= \frac{1}{2} - \frac{2}{\pi} \sum_{\substack{j=-\infty \\ j \text{ odd}}}^{\infty} f(j) e^{(ic_1\widetilde{x})}\ \phi_{x_1}(c_1)\phi_{x_2}(c_1) + E_{x_2}[\rho(\widetilde{x}, x_2)] \end{aligned} \tag{IV.54}$$

where $\phi_{x_2}(c_1) \triangleq E_{x_2}[\ e^{(-x_2ic_1)}]$. Let $\rho(\widetilde{x}) \triangleq E_{x_2}[\rho(\widetilde{x}, x_2)]$. Using the similar approach as above, we can show that

$$\begin{aligned} \phi_{x_2}(c_1) &= \prod_{k=2}^{K} \left\{ \left(\prod_{\substack{m=-\infty \\ m\neq 0}}^{\infty} \frac{1}{M_H} \left[\frac{sin(M_H\gamma_{1,k,m})}{sin(\gamma_{1,k,m})} \right] \right) \right. \\ &\quad \left. \times \left(\prod_{\substack{m=-\infty \\ m\neq 0}}^{\infty} \frac{1}{M_V} \left[\frac{sin(M_V\gamma_{2,k,m})}{sin(\gamma_{2,k,m})} \right] \right) \right\} \end{aligned} \tag{IV.55}$$

where $\gamma_{1,k,m} \triangleq c_1\sqrt{P_k}y_{1,k,m}$ and $\gamma_{2,k,m} \triangleq c_1\sqrt{P_k}y_{2,k,m}$. Thus, (IV.54) is the average BER (averaged over x_1, x_2). In the final form, (IV.54) can be written as

$$Q_2(\widetilde{x}) = \frac{1}{2} - \frac{2}{\pi} \sum_{\substack{j=1 \\ j \text{ odd}}}^{\infty} \frac{exp(-j^2\omega^2/2)}{j}\ sin(j\omega\widetilde{x}/\sigma_R)\ \phi_{x_1}(j\omega/\sigma_R)\phi_{x_2}(j\omega/\sigma_R) + \rho(\widetilde{x}) \tag{IV.56}$$

where we have substituted for $f(j) = \frac{exp(-j^2\omega^2/2)}{2ij}$, $c_1 = j\omega\widetilde{a}$, $\widetilde{a} = 1/\sigma_R$, and used the relation $sin(x) = (e^{ix} - e^{-ix})/2i$.

Next we give a bound on the approximation error $\rho(\widetilde{x})$ in (IV.56). Note that $\rho(\widetilde{x})$ is defined after (IV.54), and is repeated below:

$$\rho(\widetilde{x}) = E_{x_2}[\rho(\widetilde{x}, x_2)], \qquad \text{where } \rho(\widetilde{x}, x_2) = E_{x_1}[\xi(\widetilde{a}\widetilde{x} - \widetilde{a}x_1 - \widetilde{a}x_2)]. \tag{IV.57}$$

Bounding $|\rho(\widetilde{x})|$, we have

$$
\begin{aligned}
|\rho(\widetilde{x})| &= |E_{x_2}[E_{x_1}[\xi(\widetilde{a}\widetilde{x} - \widetilde{a}x_1 - \widetilde{a}x_2)]]| \\
&= \left| E_{\left\{ \substack{A_{k,l-m}^{(r)}, \\ A_{k,l-m}^{(i)}} \right\}_{m=-\infty}^{\infty}} \left[E_{\left\{ \substack{A_{1,l-m}^{(r)}, \\ A_{1,l-m}^{(i)}} \right\}_{\substack{m=-\infty \\ m\neq 0}}^{\infty}} \left[\xi\left(\widetilde{a}\left(\widetilde{x} \right.\right.\right.\right. \right. \\
&\qquad \left. - \left[\sqrt{P_1} \sum_{\substack{m=-\infty \\ m\neq 0}}^{\infty} (\widetilde{y}_{1,m} A_{1,l-m}^{(r)} + \widetilde{y}_{2,m} A_{1,l-m}^{(i)}) \right] \right) \\
&\qquad \left. \left. \left. - \left[\sum_{k=2}^{K} \sqrt{P_k} \sum_{m=-\infty}^{\infty} (y_{1,k,m} A_{k,l-m}^{(r)} + y_{2,k,m} A_{k,l-m}^{(i)}) \right] \right) \right) \right] \right] \Bigg| \qquad \text{(IV.58)}
\end{aligned}
$$

$$
\begin{aligned}
&\leq \max_{\substack{\left\{ X_{1,m}^{(r)}, X_{1,m}^{(i)} \right\}_{\substack{m=-\infty, \\ m\neq 0}}^{\infty}, \\ \left\{ X_{k,m}^{(r)}, X_{k,m}^{(i)} \right\}_{m=-\infty}^{\infty}}} \left[\xi\left(\widetilde{a}\left\{ \widetilde{x} - \left[\sqrt{P_1} \sum_{\substack{m=-\infty \\ m\neq 0}}^{\infty} (\widetilde{y}_{1,m} X_{1,m}^{(r)} + \widetilde{y}_{2,m} X_{1,m}^{(i)}) \right] \right.\right. \right. \\
&\qquad \left. \left. \left. - \left[\sum_{k=2}^{K} \sqrt{P_k} \sum_{m=-\infty}^{\infty} (y_{1,k,m} X_{k,m}^{(r)} + y_{2,k,m} X_{k,m}^{(i)}) \right] \right\} \right) \right] \qquad \text{(IV.59)}
\end{aligned}
$$

$$
\begin{aligned}
&\leq \max_{\substack{\left\{ X_{1,m}^{(r)}, X_{1,m}^{(i)} \right\}_{\substack{m=-\infty, \\ m\neq 0}}^{\infty} \\ \left\{ X_{k,m}^{(r)}, X_{k,m}^{(i)} \right\}_{m=-\infty}^{\infty}}} \left[Q\left(\frac{T_e}{2} - \widetilde{a}\left\{ \left| \widetilde{x} - \left[\sqrt{P_1} \sum_{\substack{m=-\infty \\ m\neq 0}}^{\infty} (\widetilde{y}_{1,m} X_{1,m}^{(r)} + \widetilde{y}_{2,m} X_{1,m}^{(i)}) \right] \right.\right. \right. \\
&\qquad \left. \left. \left. - \left[\sum_{k=2}^{K} \sqrt{P_k} \sum_{m=-\infty}^{\infty} (y_{1,k,m} X_{k,m}^{(r)} + y_{2,k,m} X_{k,m}^{(i)}) \right] \right| \right\} \right) \right] \qquad \text{(IV.60)}
\end{aligned}
$$

$$
\begin{aligned}
&\leq \max_{\substack{\left\{ X_{1,m}^{(r)}, X_{1,m}^{(i)} \right\}_{\substack{m=-\infty, \\ m\neq 0}}^{\infty} \\ \left\{ X_{k,m}^{(r)}, X_{k,m}^{(i)} \right\}_{m=-\infty}^{\infty}}} Q\left(\frac{T_e}{2} - \widetilde{a}\left\{ |\widetilde{x}| + \sqrt{P_1} \sum_{\substack{m=-\infty \\ m\neq 0}}^{\infty} (|\widetilde{y}_{1,m}||X_{1,m}^{(r)}| + |\widetilde{y}_{2,m}||X_{1,m}^{(i)}|) \right.\right. \\
&\qquad \left. \left. + \sum_{k=2}^{K} \sqrt{P_k} \sum_{m=-\infty}^{\infty} (|y_{1,k,m}||X_{k,m}^{(r)}| + |y_{2,k,m}||X_{k,m}^{(i)}|) \right\} \right) \qquad \text{(IV.61)}
\end{aligned}
$$

$$= \quad Q\left(\frac{T_e}{2} - \mu\right)$$

$$\text{where} \;\; \mu \stackrel{\Delta}{=} \widetilde{a}\Bigg\{|\widetilde{x}| + \sqrt{P_1} \sum_{\substack{m=-\infty \\ m\neq 0}}^{\infty} \left(|\widetilde{y}_{1,m}|(M_H - 1) + |\widetilde{y}_{2,m}|(M_V - 1)\right)$$

$$+ \sum_{k=2}^{K} \sqrt{P_k} \sum_{m=-\infty}^{\infty} \left(|y_{1,k,m}|(M_H - 1) + |y_{2,k,m}|(M_V - 1)\right)\Bigg\} \quad \text{(IV.62)}$$

where (IV.58) is obtained by substituting for x_1 and x_2, and noting that the r.v.s x_1 and x_2 are defined in term of the random data symbols $D_1 = \{A^{(r)}_{1,l-m}, A^{(i)}_{1,l-m}\}^{\infty}_{\substack{m=-\infty \\ m\neq 0}}$, and $D_2 \stackrel{\Delta}{=} \{A^{(r)}_{k,l-m}, A^{(i)}_{k,l-m}\}^{\infty}_{m=-\infty}$, respectively.

The maximization in (IV.59) is over integers $\{X^{(r)}_{1,m}, X^{(i)}_{1,m}\}^{\infty}_{\substack{m=-\infty \\ m\neq 0}}$, $\{X^{(r)}_{k,m}, X^{(i)}_{k,m}\}^{\infty}_{m=-\infty}$, where $X^{(r)}_{1,m}$ and $X^{(r)}_{k,m}$ take values from the set $\{-M_H + 2j - 1\}^{M_H}_{j=1}$, and $X^{(i)}_{1,m}$, $X^{(i)}_{k,m}$ take values from the set $\{-M_V + 2j - 1\}^{M_V}_{j=1}$.

The inequality in (IV.60) follows from (IV.43) (with T_e chosen such that following are satisfied: $\mu << T_e/2$, and T_e is large). Equation (IV.62) is obtained by simplifying the maximization operation in the previous step (i.e., in (IV.61)). For small approximation error, the value of T_e is chosen such that following are satisfied: $\mu << T_e/2$, and T_e is large. Note that μ is finite. This can be seen by first noting that the quantities $\widetilde{y}_{1,m}, \widetilde{y}_{2,m}, y_{1,k,m}, y_{2,k,m}$ in μ are defined in terms of the pulse $x(t)$, where the subscript m in these terms appears inside the pulse, in the form $x(mT)$. This means that the infinite sum (with summation index m) in μ is finite (see (IV.10)).

In order to use the series in (IV.42) for obtaining the BER results numerically, the infinite series must be truncated. The truncated series can be written as

$$Q(x) \quad = \quad \frac{1}{2} - \frac{2}{\pi} \sum_{\substack{j=1 \\ j \text{ odd}}}^{\widetilde{T}} \frac{exp(-j^2\omega^2/2)\; sin(j\omega x)}{j} + \xi(x) + \eta_{\widetilde{T}}(x) \qquad \text{(IV.63)}$$

where $\eta_{\widetilde{T}}(x)$ is the error in truncating the series and is bounded by [88]

$$|\eta_{\widetilde{T}}(x)| \leq \frac{\sqrt{2\pi}T_e}{\pi^2\widetilde{T}}\; Q\left(\frac{2\pi\widetilde{T}}{T_e}\right). \qquad \text{(IV.64)}$$

For the choice of $\widetilde{T}$, such that $\widetilde{T} \geq T_e$, the truncation error is very small. For example, for $\widetilde{T} = T_e$, the right side of (IV.64) takes the value $\frac{\sqrt{2\pi}}{\pi^2}\; Q\,(2\pi) = 4.21 \times 10^{-11}$.

Thus, truncating the series in the averaged $Q(.)$ function in (IV.56), we have

$$Q_2(\widetilde{x}) = \frac{1}{2} - \frac{2}{\pi} \sum_{\substack{j=1 \\ j \text{ odd}}}^{\widetilde{T}} \frac{exp(-j^2\omega^2/2)}{j} sin(j\omega\widetilde{x}/\sigma_R)\phi_{x_1}(j\omega/\sigma_R)\phi_{x_2}(j\omega/\sigma_R) + \rho(\widetilde{x}) + \eta_{\widetilde{T}}(\widetilde{x}). \quad \text{(IV.65)}$$

Thus, averaging the $Q(.)$ function, with argument $(\widetilde{a}\widetilde{x} - \widetilde{a}x_1 - \widetilde{a}x_2)$, over x_1 and x_2, results in the expression given in (IV.65). Using this expression in the BER results in (IV.37), gives the final expression for the BER for the system. Note that this final expression is conditioned on the channel of all the users and the delays of the interfering users.

IV.E Numerical Results

Numerical results are obtained by averaging the conditional BER (conditioned on the channel of all the users and the delays of the interfering users) obtained in Section IV.D over different channel realizations and the delay of the interfering users. We consider an indoor NLOS operating environment, where the LOS is obstructed or there is no direct path between the mmW transmitter and the receiver. An example of such an operating environment is an office setting where furnishing typically includes multiple desks, chairs, computers, bookshelves etc. The IEEE 802.15.3c TG has carried out measurements for several indoor environments and has proposed channel models that characterize these underlying environments [7].

In general, the parameter values for the channel models proposed in [7] depend on the beamwidth of the transmit antenna and the direction in which the transmit power is launched. For channel model CM4, which is for the NLOS office environment, three separate sets of measurement data are available, based on three different beamwidths of the transmitter. We denote the three CM4 models as CM4.1, CM4.2 and CM4.3. The parameter values for these models are given in Table IV.1 [7], where Γ and γ are the cluster and ray decay factors, respectively, $\widetilde{k}$ is the Rician effect in each cluster

Table IV.1: Parameter values for the NLOS CM4 models

	CM4.1 Tx-360°	CM4.2 Tx-30°	CM4.3 Tx-omni
$\Gamma[ns]$	109.2	134	19.44
$\gamma[ns]$	67.9	59	0.42
$\sigma_c[dB]$	3.24	4.37	1.82
$\sigma_r[dB]$	5.54	6.66	1.88
$\sigma_\theta[degree]$	60.2	22.2	9.1
$\widetilde{k}[dB]$	19	19.2	10

[7] defined in Section IV.B.1, σ_c and σ_r are the cluster and ray lognormal standard deviations, respectively, and σ_θ is the angle spread.

For obtaining the numerical results, we consider the following system parameter values: the data rate is 1 Gbps, the number of users, K, is 2, and the roll off factor, ϱ, is 0.2. At the receiver, we consider the same antenna array model as described in the previous chapters, i.e., we consider a fixed length array with fixed effective aperture. In our results, the total length of the antenna array is constrained to 2λ, where λ is the carrier wavelength.

We consider a single antenna with directional radiation pattern at the transmitter. The transmitter beamwidth, for the channel models CM4.1, CM4.2 and CM4.3, is assumed to be the same as given in Table IV.1, above. The spatial correlation between the antennas at the receive array depends on the mean AoA of the cluster, antenna spacing, angle spread, and the power azimuth profile (density function of the AoA). Based on the recommendations in [7], the Laplacian density for the AoA is adopted. For this case, the correlation between the antenna elements is given in [91], which is employed in our results. Furthermore, we assume that all channel models have the same number of clusters and the same number of rays within the cluster, and normalize the power in the

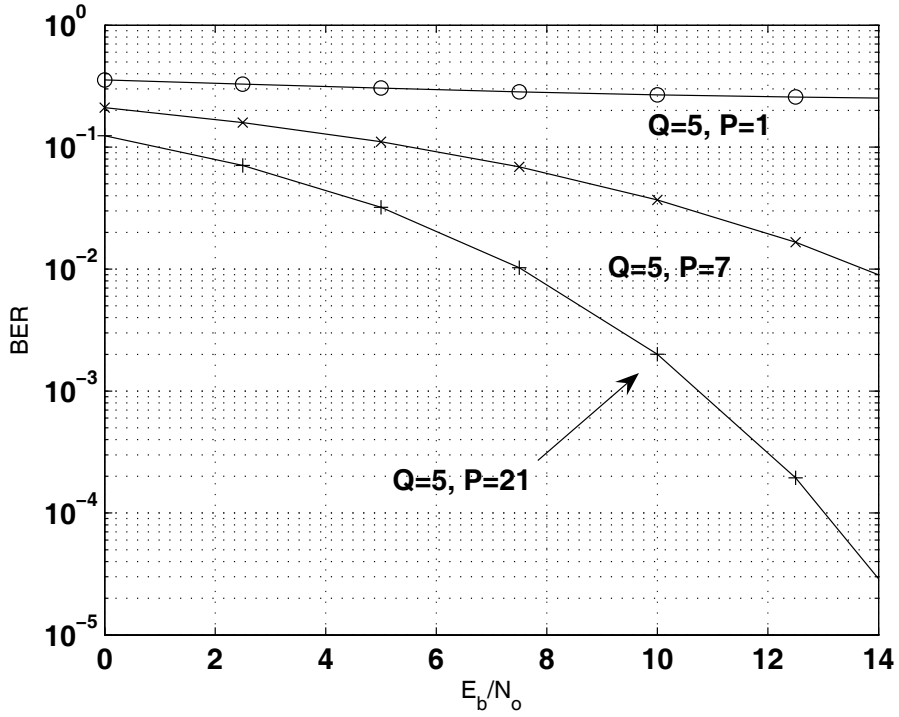

Figure IV.3: BER plot for 4-QAM when $Q = 5$ and the number of temporal taps P takes on values $\{1, 7, 21\}$. The channel model is CM4.1.

multipath components. The total number of clusters considered is 2, and the number of rays within a cluster is 4, unless specified otherwise. The mean AoA of the two clusters of the desired user is assumed at $[0^\circ, 20^\circ]$ and that of the interfering user at $[10^\circ, 30^\circ]$.

In Figure IV.3, we plot the system performance for 4-QAM for a varying number of temporal taps P and fixed number of antennas $Q = 5$. The channel model used is CM4.1. The angle spread (AS) for the two users is fixed to 60.2°, same as the value specified in Table IV.1. Since the interfering and the desired users coexist in the same environment (for example, the same room) and experience similar type of the scattering, it is reasonable to assume that the AS for the two users is the same. Note that the performance gets better as P is increased, where P being unity corresponds to the case when no equalization is performed. Clearly, the performance is limited by the ISI, in addition to the MAI, and the system performance can be improved significantly by the use of an equalizer. Similar behavior is observed when the channel model CM4.2 is used. Figure

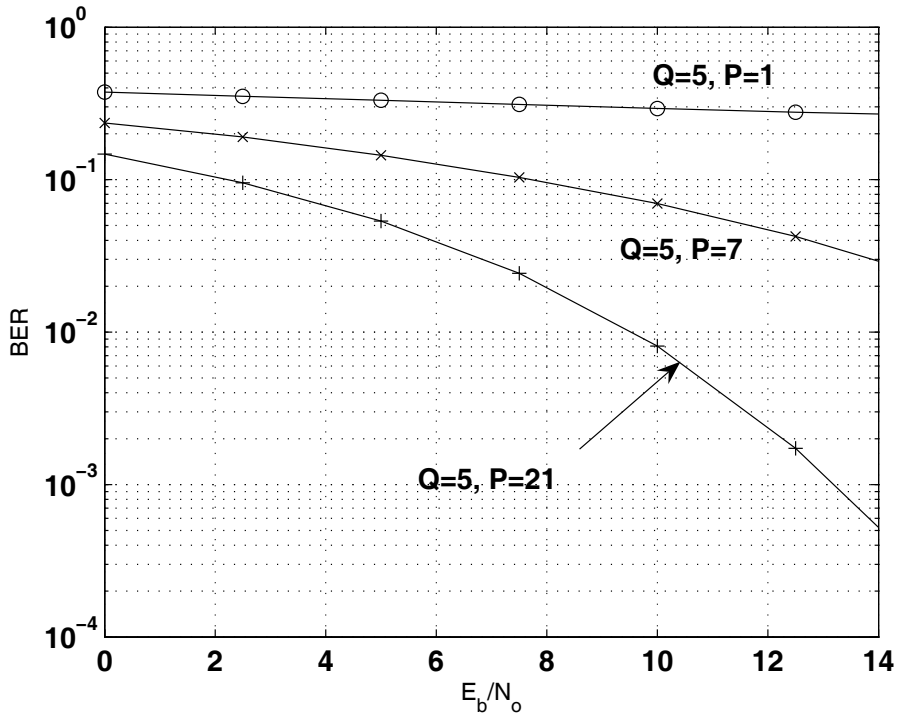

Figure IV.4: BER plot for 4-QAM when $Q = 5$ and the number of temporal taps P takes on values $\{1, 7, 21\}$. The channel model is CM4.2.

IV.4, shows the performance for CM4.2, with the same set of parameters as above.

Figure IV.5 shows the performance for CM4.1 when the number of antennas Q is 5, P is 21, and the angular spread is varied. The AS takes on values $\{5°, 10°, 15°, 20°\}$ in Figure IV.5. Note that the spatial correlation depends on the AS, as mentioned before. A larger AS corresponds to the case of a richer scattering environment, therefore, a lower correlation between the antennas. Thus, spatial correlation decreases as AS is increased due to reduced correlation between the signals received at different antennas. This has the effect of providing better spatial diversity gains. On the other hand, as the AS of the interferer increases, the ability to suppress interference through beamforming is reduced. Thus, for a given Q and P values and interferer location, the performance is determined by the net effect of these two factors. Note that in Figure IV.5, the performance improves as AS is increased mainly due to improved diversity.

In Figure IV.6, the performance of channel models CM4.2 and CM4.3 for an

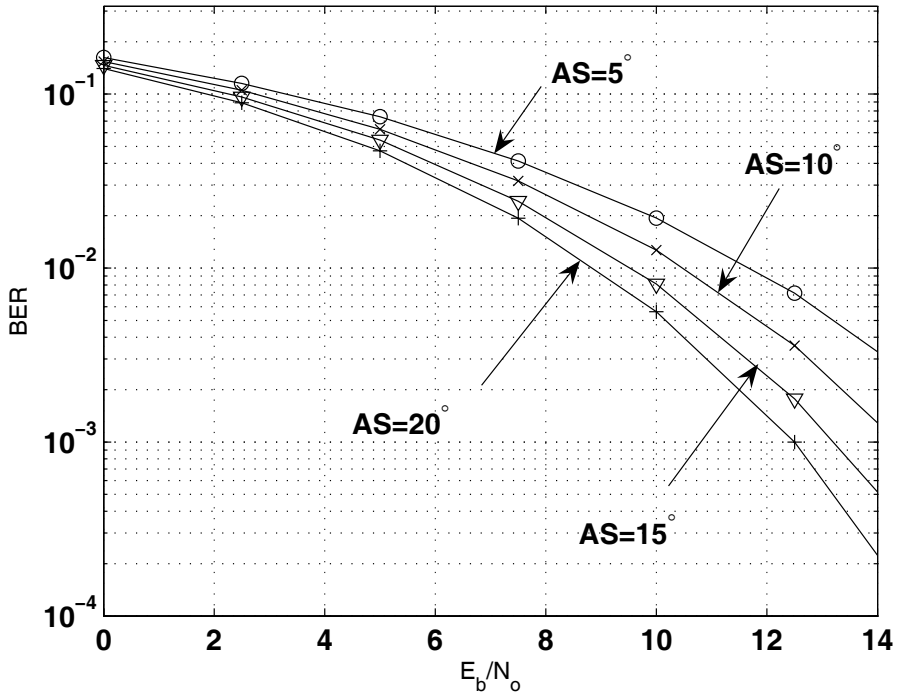

Figure IV.5: BER plot for 4-QAM when $Q = 5$, $P = 21$ and AS is varied $\{5°, 10°, 15°, 20°\}$ for channel model CM4.1

AS value of $15°$ is shown. The number of antennas Q is 5 and the number of temporal taps P is 21. Note that CM4.3 performs better than CM4.2. For the given Q and P, the relative performance of the channel models is determined by a combination of the three factors; residual ISI, MAI and the total available diversity. Note that CM4.2 has a much slower decaying multipath intensity profile when compared to the CM4.3. This can be seen from the corresponding Γ and γ values in Table IV.1. This, thus, provides larger multipath diversity gain for CM4.2, as the multipath coefficients with larger delays are associated with higher SNR, when compared to CM4.3. Also, CM4.2 would experience larger ISI levels. Figure IV.6, shows the net effect of these factors for the given Q and P values. CM4.3 outperforms primarily due to lower interference level (ISI and MAI) when compared to CM4.2.

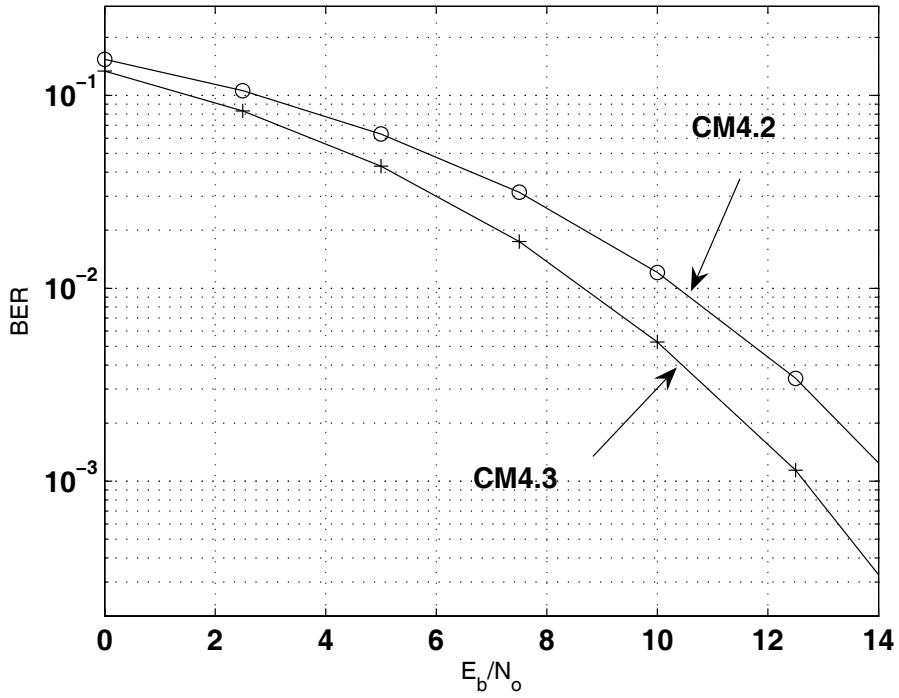

Figure IV.6: BER plot for channel models CM4.2 and CM4.3 for 4-QAM, $Q = 5$, $P = 21$ and AS 15°.

IV.F Conclusion

In this work, the performance of a 60 GHz communication system employing multiple antenna equalization scheme to suppress both the intersymbol interference and multiple-access interference was analyzed. In order to increase spectral efficiency, a higher order modulation format based upon M-QAM was used. An analytical expression for the bit error rate was derived for the case of non-Gaussian intersymbol interference and multiple access interference. A realistic IEEE channel model was used for this analysis. Joint spatial and temporal processing was shown to improve the system performance significantly. The combined effects of interference suppression and spatial correlation on the system performance was then studied.

IV.G Acknowledgement

This chapter is, in part, a reprint of the material as it appears in:

1. P. Nagvanshi, E. Masry and L.B. Milstein, "Performance Analysis of QAM Multiple Antenna Systems at 60GHz ," *IEEE Transactions on Communications,* to be submitted.

The dissertation author was the primary researcher and author, and the co-authors listed in this publication directed and supervised the research which forms the basis for this chapter.

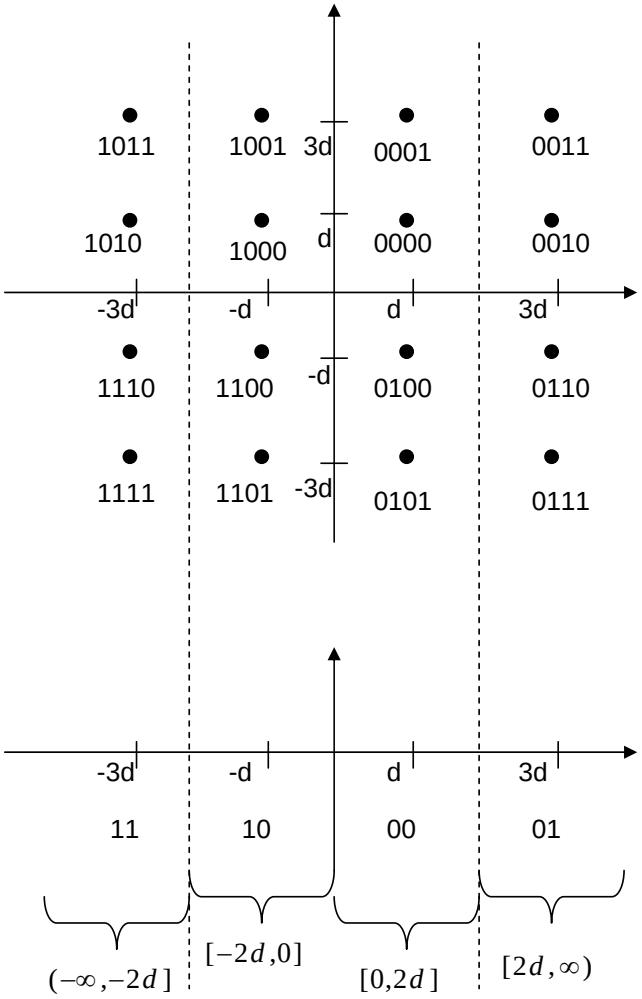

Figure IV.7: 16QAM and 4PAM constellations

V

Conclusion

Ultra wideband and millimeter wave are emerging technologies that hold great promise to satisfy the need for short range, high speed wireless communication systems. Such systems are envisioned to pervade the home and office environments, and facilitate interconnection of various consumer electronic devices, which require data rates in excess of tens of megabits per second and often rely on wireline communication. Effective utilization of the spatial domain through the use of MIMO techniques will further enhance the viability of UWB and mmW for such wireless consumer electronic devices.

The significantly higher occupied bandwidth of UWB and mmW systems provides immense advantages in terms of higher data rates, while at the same time presenting new challenges such as susceptibility to interferers and possibly complex radio design. The success of these new technologies will depend, to a significant extent, on the success of scientists and engineers in overcoming these new challenges, while taking advantage of the higher data rates offered by these systems. These advantages of UWB and mmW should come at little or no extra cost to the end user for these technologies to be successful and pervasive.

This research has focused on furthering the goal of realizing robust and reliable UWB and mmW communication systems which provide unprecedented data rates. The effectiveness of multiple antenna techniques in improving reliability and overall performance was investigated. The performance of these systems was studied in a real

world environment.

We first studied the performance of a MIMO beamforming system. In the presence of both imperfect channel estimation and antenna correlation, an optimal MIMO beamforming scheme was proposed in Chapter II. The performance of this scheme was analyzed through a closed-form expression for the probability of error, and the combined effects of channel estimation errors and diversity on the system performance were studied. It was assumed that spatially-correlated fading exists either at the transmitter or at the receiver. We showed that, with imperfect channel estimation, the performance is identical if there is independent fading at, say L, transmitters and correlated fading at, say M, receivers, compared to the case when there is correlated fading at M transmitters and independent fading at L receivers. We then focused on the case of independent fading at the transmitter and correlated fading at the receiver. Due to spatial constraints, we assumed that a fixed length linear array is deployed at the receiver, where the effective aperture of the array is fixed. For such an array, it was shown that an optimal number of receive antenna elements exists, and that the optimum value depends on the available transmit and receive diversity, the quality of the estimates and the array length.

It was also shown that for a given number of transmit and receive antennas, as the correlation between the channel coefficients increases due to poor scattering, the system performance degrades. This indicated that the impact of reduced diversity order on the performance was more pronounced than the benefits of improved channel estimation accuracy as the fade correlation increases due to poor scattering.

In Chapter III, the performance of a DS-CDMA UWB system with multiple antennas at the receiver was evaluated. Due to operation in a dense wireless environment, we assumed the presence of multiple interfering UWB devices, as well as interference from other narrowband systems. A spatio-temporal receiver was designed that first forms an estimate of the channel in the presence of NBI and MAI, and then uses it to optimally combine the multipath components such that the output signal-to-interference-plus-noise-ratio is maximized. An exact closed-form expression for the probability of error was derived for the case when correlation between the NBI during

the channel estimation phase and the data detection phase is negligible. The effects of channel estimation accuracy, interference suppression, and spatial correlation on the performance were studied. At the receiver, we assumed the same antenna array model as in Chapter II, i.e., we assumed a fixed length array with fixed effective array aperture. For this antenna model, it was shown that, for fixed maximum diversity, there is an optimal combination of the number of antennas and the number of temporal taps that achieves the best performance. This optimal combination depends on the signal-to-interference ratio and the quality of the estimates.

Chapter IV focused on the performance analysis of a 60 GHz communication system. A multiple antenna equalization scheme was employed to suppress both the intersymbol interference and multiple access interference. In order to increase spectral efficiency, a higher order modulation format based upon M-QAM was used. The performance of this system through an analytically derived expression for the bit error rate was investigated. A realistic IEEE channel model [7] was used for this analysis. Joint spatial and temporal processing was shown to improve the system performance significantly. The combined effects of interference suppression and spatial correlation on the system performance was then studied.

Bibliography

[1] P. McDermott-Wells, "What is Bluetooth," *IEEE Potentials*, vol. 23, no. 5, pp. 33-35, Dec 2004.

[2] D. Porcino and W. Hirt, "Ultra-wideband radio technology: potential and challenges ahead," *IEEE Commun. Mag.*, vol. 41, no. 7, pp. 66-74, July 2003.

[3] FCC report, "Revision of part 15 of the commission's rules regarding the ultra-wideband transmission," *ET Docket 98-153*, April 2002.

[4] L. Yang and G. B. Giannakis, "Ultra-wideband communications: an idea whose time has come," *IEEE Signal Processing Magazine*, vol. 21, no. 6, pp. 26-54, Nov. 2004.

[5] FCC report, "Amendment of parts 2, 15 and 97 of the commission's rules to permit use of radio frequencies above 40 GHz for new radio applications", ET Docket 94-124, December 1995.

[6] "Revision of TSV model for LOS desktop channel environments", IEEE 802.15-06/0297r3, September 2006.

[7] "TG3c channel modeling sub-committee final report", IEEE 802.15-06/0195r5, November 2006.

[8] "Generalization and parameterization of the mmwave channel models ", IEEE 802.15-05/0261r1, May 2005.

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

[10] G. J. Foschini and M. J. Gans,"On Limits of Wireless Communications in a Fading Environment when Using Multiple Antennas,"*Wireless Personal Communications*, vol. 6, pp. 311–335, March 1998.

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

[12] P. Amihood, E. Masry, L. B. Milstein and J. G. Proakis, "Performance Analysis of a Pre-BLAST-DFE Technique for MISO Channels With Decentralized Receivers," *IEEE Transactions on Communications*, vol. 55, no. 7, pp. 1385-1396, July 2007.

[13] P. Amihood, E. Masry, L. B. Milstein and J. G. Proakis, "Analysis of a MISO Pre-BLAST-DFE Technique for Decentralized Receivers," in the *Asilomar Conference on Signals, Systems and Computers*, Nov. 2006.

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

[15] P.A. Dighe, R.K. Mallik and S.S. Jamuar, "Analysis of transmit-receive diversity in Rayleigh fading," *IEEE Transactions on Communications*, vol. 51, pp. 694-703, April 2003.

[16] L. Tong, B.M. Sadler and M. Dong, "Pilot-assisted wireless transmissions: general model, design criteria, and signal processing," *IEEE Signal Processing magazine*, vol. 21, pp. 12-25, Nov. 2004.

[17] M. Kang and M.S. Alouini, "Impact of correlation on the capacity of MIMO channels," *IEEE International Conference on Communications* , vol. 4, pp. 2623-2627, May 2003.

[18] S.A. Jafar and A. Goldsmith, "Multiple-antenna capacity in correlated Rayleigh fading with channel covariance information," *IEEE Trans. on Wireless Communications*, vol. 4, pp. 990-997, May 2005.

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

[20] S.A. Jafar and A. Goldsmith, "Transmitter optimization and optimality of beamforming for multiple antenna systems," *IEEE Trans. on Wireless Communications*, vol. 3, pp. 1165-1175, july 2004.

[21] B. Hassibi and B.M. Hochwald, "How much training is needed in multiple-antenna wireless links?," *IEEE Transactions on Information Theory*, vol. 49, pp. 951-963, April 2003.

[22] J.H. Kotecha and A.M. Sayeed, "Transmit signal design for optimal estimation of correlated MIMO channels," *IEEE Trans. on Signal Processing*, vol. 52, pp. 546-557, Feb. 2004.

[23] S. Serbetli, S. Bethanabhotla and A. Yener, "The effect of channel estimation on transceiver design for MIMO systems with QoS constraints," *Proc. 38th Ann. Conf. on Information Sciences and Sys.*, March 2004.

[24] S. Serbetli and A. Yener, "Transmission strategies for correlated MIMO links with imperfect channel estimates," *IEEE International Conf. on Communications*, vol. 2, pp. 767-771, May 2005.

[25] S. Serbetli and A. Yener, "MMSE transmitter design for correlated MIMO systems with imperfect channel estimates: power allocation trade-offs," *IEEE Trans. on Wireless Communications*, vol. 5, pp. 2295-2304, August 2006.

[26] P. Nagvanshi, E. Masry and L. B. Milstein, "The effect of channel estimation errors on transceiver optimization for correlated MIMO Rayleigh channels," *Proc. 39th Annual Conference on Information Sciences and Systems*, March 2005.

[27] P. Nagvanshi, E. Masry and L.B. Milstein, "Optimum transmit-receive beamforming with noisy channel estimates for correlated MIMO Rayleigh channels," accepted for publication in the *IEEE Transactions on Communications.*

[28] A. Abdi and M. Kaveh, "A space-time correlation model for multielement antenna systems in mobile fading channels," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 3, pp. 550-560, April 2002.

[29] K. Yu and B. Ottersten, "Models for MIMO propagation channels, A review," *Wiley Journal on Wireless Comm. and Mobile Computing*, 2(7):653-666, Nov. 2002. Special Issue on Adaptive Antennas and MIMO systems.

[30] S.M. Kay, *Fundamentals of statistical signal processing: estimation theory*, Prentice Hall PTR, 1993.

[31] H. Lütkepohl, *Handbook of matrices*, John Wiley & Sons, West Sussex, England, 1996.

[32] R.A. Horn and C.R. Johnson, *Matrix Analysis*, Cambridge University Press, 1999.

[33] M. Abramowitz and I.A. Stegun, *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, ninth printing, 1970.

[34] M. Kang and M.S. Alouini, "Largest eigenvalue of complex wishart matrices and performance analysis of MIMO MRC systems," *IEEE Journal Sel. Areas in Comm.*, vol. 21, pp. 418-426, April. 2003.

[35] M. Kang and M.S. Alouini, "Quadratic forms in complex gaussian matrices and performance analysis of MIMO systems with cochannel interference," *IEEE Trans. Wireless Comm.*, vol. 3, pp. 418-431, March 2004.

[36] A.T. James, "Distributions of matrix variates and latent roots derived from normal samples," *The Annals of Mathematical Statistics*, vol. 35, no. 2, pp. 475-501, June 1964.

[37] C.G. Khatri, "On the moments of traces of two matrices in three situations for complex multivariate normal populations," *Sankhya: The Indian Journal of Statistics: Series A*, vol. 32, pp. 65-80, 1970.

[38] M.A. Golberg, "The derivative of a determinant," *The American Math. Monthly* , vol. 79, pp. 1124-1126, 1972.

[39] H. Gao and P.J. Smith, "A determinant representation for the distribution of quadratic forms in complex normal vectors," *Journal of Multivariate Analysis*, vol. 73, pp. 155-165, 2000.

[40] A.S.Y. Poon, R.W. Brodersen and D.N.C. Tse, "Degrees of freedom in multiple-antenna channels: a signal space approach," *IEEE Trans. on Information Theory*, vol. 51, pp. 523-536, Feb. 2005.

[41] J.D. Kraus, *Antennas*, 2nd edition, New York: McGraw-Hill, 1988.

[42] K. Fujimoto and J.R. James, *Mobile antenna systems handbook*, 2nd edition, Artech House, 2001.

[43] M.S. Alouini and M.K. Simon, "An MGF-based performance analysis of generalized selection combining over Rayleigh fading channels," *IEEE Trans. on Communications*, vol. 48, pp. 401-415, March 2000.

[44] Gupta and Nagar, *Matrix variate distributions*, John Wiley & Sons, West Sussex, England, 1996.

[45] C.G. Khatri, "On certain distribution problems on positive definite quadratic functions in normal vectors," *The Annals of Mathematical Statistics*, vol. 37, pp. 468-479, April 1966.

[46] C.G. Khatri, "Non-central distributions of ith largest characteristic roots of three matrices concerning complex multivariate normal populations," *Annals of the Institute of Statistical Mathematics* , vol. 21, pp. 23-32, 1969.

[47] Q. Li and L. A. Rusch, "Multiuser detection for DS-CDMA UWB in the home environment," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 9, pp. 1701-1711, Dec. 2002.

[48] N. Boubaker and K. B. Letaief, "MMSE multipath diversity combining for multi-access TH-UWB in the presence of NBI," *IEEE Trans. Wireless Commun.*, vol. 5, no. 4, pp. 712-719, April 2006.

[49] C. R. C. M da Silva and L. B. Milstein, "The effects of narrowband interference on UWB communication systems with imperfect channel estimation," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 4, pp. 717-723, April 2006.

[50] J. Wang and L. B. Milstein, "Multicarrier CDMA overlay for ultra-wideband communications," *IEEE Transactions on Communications*, vol. 52, no. 10, pp. 1664-1669, Oct. 2004.

[51] P. Nagvanshi, E. Masry and L.B. Milstein, "Performance of Multiple Antenna DS-CDMA UWB System With Noisy Channel Estimates and Narrow-Band Interference ," *IEEE Journal of Selected Topics in Signal Processing,* vol. 1, no. 3, pp. 470 - 482, Oct. 2007.

[52] P. Nagvanshi, E. Masry and L.B. Milstein, "Performance Analysis of Multiple Antenna DS-CDMA UWB System with Noisy Channel Estimates and Narrow-Band Interference ," *IEEE Military Communications Conference (MILCOM)*, pp. 1-7, Oct. 2006.

[53] "Merger#2 proposal DS-CDMA", *IEEE 802.15-03/334r3*, Sept. 2003.

[54] M. Z. Win and R. A. Scholtz, "Impulse radio: how it works," *IEEE Commun. Lett.*, vol. 2, no. 2, pp. 36-38, Feb. 1998.

[55] S. Stroh, "Wideband: multimedia unplugged," *IEEE Spectr.*, vol. 40, issue 9, pp. 23-27, Sept. 2003.

[56] S. Roy, J. R. Foerster, V. S. Somayazulu and D. G. Leeper, "Ultrawideband radio design: the promise of high-speed, short-range wireless connectivity," *Proc. IEEE*, vol. 92, no. 2, pp. 295-311, Feb. 2004.

[57] A. Batra, J. Balakrishnan, G. R. Aiello, J. R. Foerster and A. Dabak, "Design of a Multiband OFDM system for realistic UWB channel environments," *IEEE Trans. Microw. Theory Tech.*, vol. 52, no. 9, pp. 2123-2138, Sept. 2004.

[58] P. Amihood, E. Masry, L. B. Milstein and J. G. Proakis, "Performance Analysis of High Data Rate MIMO Systems in Frequency Selective Fading Channels," accepted in *IEEE Transactions on Information Theory*, Jan 2007.

[59] P. Amihood, E. Masry, L. B. Milstein and J. G. Proakis, "Asymptotic Performance of Multicode MIMO Systems in Frequency Selective Fading Channels," *IEEE MILCOM*, Vol. 5, pp. 3036 - 3041, Oct. 2005.

[60] J. S. Thompson, P. M. Grant and B. Mulgrew, "Smart antenna arrays for CDMA systems," *IEEE Pers. Commun.*, vol. 3, pp. 16-25, Oct. 1996.

[61] J. H. Winters, "Smart antennas for wireless systems," *IEEE Pers. Commun.*, vol. 5, pp. 23-27, Feb. 1998.

[62] M. Biagi and E. Baccarelli, "A simple multiple-antenna ultra wide band transceiver scheme for 4th generation WLAN," *in Proc. IEEE Vehicular Tech. Conf.*, vol. 3, no. 6, pp. 1903-1907, Oct. 2003.

[63] H. Liu, R. C. Qiu and Z. Tian, "Error performance of pulse-based ultra-wideband MIMO systems over indoor wireless channels," *IEEE Trans. Wireless Commun.*, vol. 4, no.6, pp. 2939-2944, Nov. 2005.

[64] M. Sabattini, E. Masry and L. B. Milstein, "Beamforming for Interference Mitigation in TH Impulse Radio UWB Systems," *in Proc. IEEE Personal, Indoor and Mobile Radio Comm. Conf. (PIMRC)*, vol. 2, pp. 921-925, Sept. 2005.

[65] L. Yang and G. B. Giannakis, "Analog space-time coding for multiantenna ultra-wideband transmissions," *IEEE Trans. Commun.*, vol. 52, no.3, pp. 507-517, March 2004.

[66] W. P. Siriwongpairat, W. Su, M. Olfat and K. J. R. Liu, "Multiband-OFDM MIMO coding framework for uwb communication systems," *IEEE Trans. Signal Process.*, vol. 54, no.1, pp. 214-224, Jan. 2006.

[67] V. P. Tran and A. Sibille, "Spatial multiplexing in UWB MIMO communications," *in IEE Electronics Letters*, vol. 42, no. 16, pp. 931-932, Aug. 2006.

[68] J. Keignart, C. A. Rjeily, C. Delaveaud and N. Daniele, "UWB SIMO channel measurements and simulations," *IEEE Trans. Microw. Theory Tech.*, vol. 54, no. 4, pp. 1812-1819, April 2006.

[69] D. Cassioli, M. Z. Win and A. F. Molisch, "The ultrawide bandwidth indoor channel: from statistical model to simulations," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 6, pp. 1247-1257, Aug. 2002.

[70] M. Nakagami, "The m-distribution-A general formula of intensity distribution of rapid fading," *Statistical Methods in Radio Wave Propagation*, U.K: Pergamon, 1960.

[71] D. Cassioli, M. Z. Win, F. Vatalaro and A. F. Molisch, "Performance of low-complexity rake reception in a realistic UWB channel," *in Proc. IEEE ICC*, pp. 763-767, 2002.

[72] C. Prettie, D. Cheung, L. Rusch and M. Ho, "Spatial correlation of UWB signals in a home environment," *IEEE Conf. on Ultra Wideband Systems and Technologies*, pp. 65-69, May 2002.

[73] A. M. Mathai and S. B. Provost, *Quadratic Forms in Random Variables*, Marcel Dekker Inc., 1992.

[74] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products*, Academic Press, 2000.

[75] J. D. Kraus, *Antennas*, 2nd edition, New York: McGraw-Hill, 1988.

[76] S. Kotz, N. L. Johnson and D. W. Boyd, "Series representations of distributions of quadratic forms in normal variables I. Central case," *The Annals of Mathematical Statistics*, vol. 38, pp. 823-837, June 1967.

[77] H. L. Van Trees, *Optimum Array Processing: Part IV of Detection, Estimation and Modulation Theory*, John Wiley and Sons, Inc., New York, 2002.

[78] A. S. Y. Poon, R. W. Brodersen and D. N. C. Tse, "Degrees of freedom in multiple-antenna channels: a signal space approach," *IEEE Trans. Inf. Theory*, vol. 51, pp. 523-536, Feb. 2005.

[79] P. Smulders, "Exploiting the 60 GHz band for local wireless multimedia access: prospects and future directions," *IEEE Communications Magazine*, vol. 40, pp. 140-147, Jan. 2002.

[80] "Application of mmW-based PHYs", IEEE 802.15-04/118r0, March 2004.

[81] "Through wall propagation upto 60 GHz band ", IEEE 802.15-06/0371r0, August 2006.

[82] A. Saleh and R. Valenzuela, "A Statistical Model for Indoor Multipath Propagation," IEEE J. Selected Areas Communications, Vol. SAC-5, No. 2, pp. 128-137, Feb. 1987.

[83] Q. Spencer, B. D. Jeffs, M. A. Jensen and A. L. Swindlehurst, "Modeling the Statistical Time and Angle of Arrival Characteristics," IEEE J. Sel. Areas Commun., Vol., No. 3, pp. 347-360, Mar. 2000.

[84] " TG3c system requirements", IEEE 802.15-05/0353r8, November 2006.

[85] P. Nagvanshi, E. Masry and L.B. Milstein, "Performance Analysis of QAM Multiple Antenna Systems at 60GHz," *IEEE Trans. on Communications,* to be submitted.

[86] " Reference antenna model with side lode for TG3c evaluation", IEEE 802.15-06/474r0, November 2006.

[87] J.G. Proakis, *Digital Communications*, McGraw-Hill, third edition, 1995.

[88] N.C. Beaulieu, "The evaluation of error probabilities for intersymbol and cochannel interference," *IEEE Trans. Commun.*, vol. 39, no.12, pp. 1740-1749, Dec. 1991.

[89] Y. Ma, R. Schober and D. Zhang, "Exact BER for M-QAM with MRC and Imperfect Channel Estimation in Rician Fading Channels," *IEEE Trans. on Commun.*, vol. 6, pp. 926-936, March 2007.

[90] M.S. Raju, R. Annavajjala and A. Chockalingam, "BER analysis of QAM on fading channels with transmit diversity," *IEEE Trans. on Commun.*, vol. 5, pp. 481-486, March 2006.

[91] L. Schumacher and B. Raghothaman, "Closed-form expressions for the correlation coefficient of directive antennas impinged by a multimodal truncated Laplacian PAS," *IEEE Transactions on Wireless Communications*, vol. 4, pp. 1351-1359, July 2005.