Title
Performance analysis of interference suppression techniques for multiple antenna systems

Permalink
https://escholarship.org/uc/item/1g5015h3

Author
Amihood, Patrick

Publication Date
2007

Peer reviewed|Thesis/dissertation
UNIVERSITY OF CALIFORNIA, SAN DIEGO

Performance Analysis of Interference Suppression Techniques for Multiple Antenna Systems

A dissertation submitted in partial satisfaction of the requirements for the degree
Doctor of Philosophy

in

Electrical Engineering (Communication Theory and Systems)

by

Patrick Amihood

Committee in charge:

Professor Laurence B. Milstein, Chair
Professor Elias Masry, Co-Chair
Professor John G. Proakis, Co-Chair
Professor William S. Hodgkiss
Professor Alon Orlitsky

2007
The dissertation of Patrick Amihood is approved, and it is acceptable in quality and form for publication on microfilm:

Co-Chair

Co-Chair

Co-Chair

Chair

University of California, San Diego

2007
# TABLE OF CONTENTS

Signature Page ................................................................. iii

Table of Contents ............................................................. iv

List of Abbreviations ............................................................ vii

List of Figures ................................................................. viii

Acknowledgements .............................................................. x

Vita and Publications ........................................................... xiii

Abstract of the Dissertation .................................................. xv

1 Introduction ................................................................. 1

2 Performance Analysis of High Data Rate MIMO Systems in Frequency Selective Fading Channels ................................................................. 7
   2.1 Introduction ............................................................ 7
   2.2 System Model and Problem Statement ................................ 9
   2.3 Asymptotic Analysis and Error Probability .......................... 15
   2.4 Randomly spaced multipath arrivals .................................. 18
   2.5 Performance Analysis ................................................ 19
   2.6 Conclusion ............................................................ 26
   2.7 Acknowledgements .................................................... 26

3 Performance Analysis of a Pre-BLAST-DFE Technique for MISO Channels with Decentralized Receivers ................................................................. 28
   3.1 Introduction ............................................................ 28
   3.2 Tomlinson-Harashima Precoding operating over SISO Channels with ISI .................................................. 30
      3.2.1 Modulo Operation ............................................... 30
      3.2.2 Channel Model .................................................. 31
      3.2.3 Signal Model .................................................... 31
      3.2.4 The Effect of Tomlinson-Harashima Precoding on the Probability of Error .................................................. 38
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AWGN</td>
<td>Additive White Gaussian Noise</td>
</tr>
<tr>
<td>BER</td>
<td>Bit Error Rate</td>
</tr>
<tr>
<td>BLAST</td>
<td>Bell Labs Layered Space-Time Architecture</td>
</tr>
<tr>
<td>BPSK</td>
<td>Binary Phase-Shift Keying</td>
</tr>
<tr>
<td>cdf</td>
<td>cumulative distribution function</td>
</tr>
<tr>
<td>CDMA</td>
<td>Code Division Multiple Access</td>
</tr>
<tr>
<td>CSI</td>
<td>Channel State Information</td>
</tr>
<tr>
<td>dB</td>
<td>decibels, $10 \log_{10}(\cdot)$</td>
</tr>
<tr>
<td>DFE</td>
<td>Decision Feedback Equalizer</td>
</tr>
<tr>
<td>HS-DSCH</td>
<td>High Speed-Downlink Shared Channel</td>
</tr>
<tr>
<td>ISI</td>
<td>Inter-Symbol Interference</td>
</tr>
<tr>
<td>MF</td>
<td>Matched-Filter</td>
</tr>
<tr>
<td>MIMO</td>
<td>Multiple-Input Multiple-Output</td>
</tr>
<tr>
<td>MISO</td>
<td>Multiple-Input Single-Output</td>
</tr>
<tr>
<td>MMSE</td>
<td>Minimum Mean-Squared Error</td>
</tr>
<tr>
<td>OVSF</td>
<td>Orthogonal Variable Spreading Factor</td>
</tr>
<tr>
<td>pdf</td>
<td>probability density function</td>
</tr>
<tr>
<td>PSK</td>
<td>Phase-Shift Keying</td>
</tr>
<tr>
<td>QPSK</td>
<td>Quadrature Phase-Shift Keying</td>
</tr>
<tr>
<td>SIR</td>
<td>Signal-to-Interference Ratio</td>
</tr>
<tr>
<td>SINR</td>
<td>Signal-to-Interference-and-Noise Ratio</td>
</tr>
<tr>
<td>SNR</td>
<td>Signal-to-Noise Ratio</td>
</tr>
<tr>
<td>SISO</td>
<td>Single-Input Single-Output</td>
</tr>
<tr>
<td>ZF</td>
<td>Zero-Forcing</td>
</tr>
<tr>
<td>$(\cdot)^H$</td>
<td>matrix/vector Hermitian</td>
</tr>
<tr>
<td>$(\cdot)^T$</td>
<td>matrix/vector Transpose</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure 1.1: MIMO channels ......................................................... 2

Figure 2.1: System model for multicode MIMO system. ..................... 10
Figure 2.2: RAKE receiver for the $i^{th}$ transmit antenna at the $j^{th}$ receive antenna. ................................................................. 10
Figure 2.3: Despreader for the $i^{th}$ transmit antenna. ...................... 10
Figure 2.4: Numerical results for $N_{t} = 2, 4, N_{c} = 1, 2, 4, N_{r} = 2$, and $L = 32$. .................................................. 19
Figure 2.5: Numerical results for $N_{t} \times N_{c} = 6, N_{r} = 1, L = 32, K = 2$. ................................................................. 20
Figure 2.6: Numerical results for $N_{r} \times K = 6, N_{t} = 2, N_{c} = 1, L = 32$. ................................................................. 21
Figure 2.7: Numerical results of a system with equally and randomly spaced multipath arrivals for $N_{t} = 2, N_{r} = 1, L = 32, K = 2$. ................................................................. 22
Figure 2.8: Numerical results for $N_{t} = 2, N_{r} = 1, L = 32$, and $N_{c} = 1, 2, 4$, with and without self-interference. ................................................................. 23

Figure 3.1: System model for Tomlinson-Harashima precoding operating over a SISO ISI channel. .................................................. 31
Figure 3.2: Histograms and kernel densities of $\text{Re}\{x(n)\}$ with $N = 1000, M = 4, 16$, and $L = 2, 4$. ............................................. 34
Figure 3.3: Histograms and kernel densities of $\text{Re}\{x(n)\}$ with $N = 10000, M = 4, 16$, and $L = 2, 4$. ............................................. 35
Figure 3.4: Normalized estimation error of the second moment of $x(n)$ with $N = 1000, M = 4, 16$, and $L = 2, 4$. ............................................. 36
Figure 3.5: Normalized estimation error of the second moment of $x(n)$ with $N = 10000, M = 4, 16$, and $L = 2, 4$. ............................................. 37
Figure 3.6: Graphical representation of Tomlinson-Harashima precoding operating over a SISO channel with ISI. .................................................. 39
Figure 3.7: System model for a pre-BLAST-DFE technique operating over a MISO frequency selective fading channel with decentralized receivers. ....... 44
Figure 3.8: Normalized estimation error of the second moment of $x_k(n)$ with $L = 1, M = 4, N_T = N_U = 4$, and $N = 1000$. ............................................. 47
Figure 3.9: The probability of symbol error as a function of the SNR-per-bit for Tomlinson-Harashima precoding operating over a SISO channel with ISI, and with $M = 4$ and $L = 2$. ............................................. 52
Figure 3.10: The probability of symbol error as a function of the SNR-per-bit with $N_T = N_U = 2, M = 4$, and $L = 1$ and 2. ............................................. 53
Figure 3.11: The probability of symbol error as a function of the SNR-per-bit with $N_T = N_U = 4$, $M = 4$, and $L = 1$ and 2. .......................... 53

Figure 3.12: The probability of symbol error as a function of the SNR-per-bit with $N_U = 2$, $M = 4$, $L = 1$, and $N_T = 2, 3, 4$. .......................... 54

Figure 3.13: The probability of symbol error as a function of the SNR-per-bit for $N_T = N_U = 2$, $M = 4$, and $L = 1$. .......................... 54

Figure 3.14: The probability of symbol error as a function of the SNR-per-bit for $N_T = N_U = 2$, $N_T = N_U = 4$, $N_T = N_U = 6$, $M = 4$, and $L = 1$. .......................... 55

Figure 3.15: Performance of ZF linear precoding with $N_T = N_U = 4, 6, 10$. .......................... 59

Figure 3.16: Performance of MMSE linear precoding with $N_T = N_U = 4, 6, 10$. .......................... 59

Figure 3.17: Comparison of the QR decomposition and the linear precoders with $N_T = N_U = 4$. .......................... 60

Figure 3.18: Comparison of the QR decomposition and the linear precoders with $N_T = N_U = 6$. .......................... 60

Figure 4.1: System model of Tomlinson-Harashima precoding and the QR decomposition technique operating over a MISO frequency-flat fading channel with decentralized receivers using the imperfect channel estimation matrix at the transmitter. .......................... 69

Figure 4.2: The probability of symbol error as a function of the SNR-per-bit. Here, $N_T = N_U = 2$ and the variance of the estimation error takes on the values 0, 0.001, 0.01, and 0.1. .......................... 79

Figure E.1: Regions of integration for matched filtering in the case of randomly spaced multipath arrivals. .......................... 102
I feel an enormous sense of gratitude towards my advisors, Professors Elias Masry, Laurence Milstein and John Proakis. In the dark forest that is research, they have handed me the compass, given me sustenance, and, in more than one occasion, lifted me up to my feet. I would like to thank Professor Masry for not accepting anything but the highest standards of rigor and thought. I am honored by the time he has spent sharing his clarity of mind. I would like to thank Professor Milstein for his constant and unwavering support throughout so many years, without which I could not have completed this dissertation. I have greatly benefited from his keen insight and his technical knowledge. His professionalism, efficiency, and kindness are virtues that I hope to learn and carry with me beyond my time at UCSD. I would like to thank Professor Proakis for guiding me throughout my research. I am indebted to his continued support. He has taught me to keep my eye on the big picture.

I would also like to thank Professor James Zeidler for his support and guidance towards the end of my dissertation. I would like to thank Professors William Hodgkiss and Alon Orlitsky for serving on my committee. I would like to thank Professor Paul Siegel for serving on the committee of my preliminary exam. I would like to thank the faculty at UCSD, especially Professors Pamela Cosman, Alon Orlitsky, Bhaskar Rao, Jack Wolf, and Ken Zeger, for teaching and inspiring communication theory.

I would like to thank my fellow colleagues at UCSD for their friendship and support, in particular Adam Anderson, Ramesh Annavajjala, Arun Batra, Zhongren Cao, Michael Cheng, Lucy Chong, Ashay Dhamdere, Anand Dhulipala, Amir Djahanshadi, Jenny Fang, Ben Farber, Mishal Al-Gharabally, Kirti Gupta, Junsheng Han, Laddan Hashemian, Greg Horwitz, Jilei Hou, Aditya Jagannatham, Avinash Jain, Mohammed Jalloh, Brian Jessup, Nikola Jevtic, Jittra Jootar, Sandeep Kanumuri, Solmaz Alipour Kivi, Gert Lanckriet, Andrew Ling, Mingkui Liu, Steve Meyer, Preeti Nagvanshi, Yoav Nebat, Maziar Nezhad, Pavan Nuggehalli, Jason Palmer, Farzad Parvaresh, Henry Pfister, Kevin Quirk, Sriram Ramabhadran, Naomi Ann Suba Ramos, Rey Ramirez, Nate Ricklin, Matteo Sabattini, Sajama, Narayana Santhanam, Mahasweta Sarkar, Jim Sifferlen, Claudio Da Silva, Durgaprasad Shamain, Joseph Soriaga, Yiannis Spyropoulos, Vikram Srinivasan, Kostas Stamatiou, Anand Subramaniam, Haichang Sui, Mohammad Hossein Taghavi, Zeinab Taghavi, Kai Tang, Steve Thompson, Hugo Tullberg, Douglas
Turnbull, Krishnamurthy Viswanathan, Brandon Zeidler, Yan Zhang, Wenhua Zhao, and Li Zhu.

I would like to thank Kameswari Chebrolu and Bhaskaran Raman for their generosity and kindness.

I would like to thank Vijay Chellappa, who has been with me since the beginning.

I would like to thank David Wipf, Luke Barrington, Chandra Murthy, David Kauchak, Chris Ellison, Kevin Tetz, Antoni Chan, and that stingray north of Scripps Pier, for explorations into our mortality that succeeded in putting even this dissertation into perspective. I would like to thank David Wipf especially, who has a knack of helping me put things into perspective, and who started riding the waves with me many summers ago.

I would like to thank Qinghua Zhao, who was there when I needed her most. Her friendship has been a constant source of support without which I could not have finished this dissertation.

Most of all, I would like to thank my mother, who provided all the love that allowed me to be who I am today.

Chapter 2 is, in part, a reprint of material published as “Asymptotic Performance of Multicode MIMO Systems in Frequency Selective Fading Channels,” P. Amihood, E. Masry, L. B. Milstein, and J. G. Proakis, in Proceedings of the IEEE Military Communications Conference (MILCOM), Atlantic City, New Jersey, USA, vol. 5, pp. 3036 - 3041, October 17 - 20, 2005, and a reprint of material that has been accepted for publication as “Performance Analysis of High Data Rate MIMO Systems in Frequency Selective Fading Channels,” P. Amihood, E. Masry, L. B. Milstein, and J. G. Proakis, IEEE Transactions on Information Theory, August 2007. The dissertation author was the primary researcher and author of these publications.

Chapter 3 is, in part, a reprint of material published as “Analysis of a MISO Pre-BLAST-DFE Technique for Decentralized Receivers,” P. Amihood, E. Masry, L. B. Milstein, and J. G. Proakis, in Proceedings of the 40th Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA, October 29 - November 1, 2006, and a reprint of material published as “Performance Analysis of a Pre-BLAST-DFE Technique for MISO Channels with Decentralized Receivers,” P. Amihood, E. Masry,
L. B. Milstein, and J. G. Proakis, *IEEE Transactions on Communications*, vol. 55, pp. 1385-1396, July 2007. The dissertation author was the primary researcher and author of these publications.

Chapter 4 is, in part, a reprint of material that is in preparation for submission under the title “The Effects of Channel Estimation Errors on a Nonlinear Precoder for MISO Channels with Decentralized Receivers,” P. Amihood, E. Masry, L. B. Milstein, and J. G. Proakis. The dissertation author was the primary researcher and author of this publication.

Appendices A, B, C, D, and E, are, in part, a reprint of material that has been accepted for publication as “Performance Analysis of High Data Rate MIMO Systems in Frequency Selective Fading Channels,” P. Amihood, E. Masry, L. B. Milstein, and J. G. Proakis, *IEEE Transactions on Information Theory*, August 2007. The dissertation author was the primary researcher and author of this publication.

Appendix F is, in part, a reprint of material published as “Performance Analysis of a Pre-BLAST-DFE Technique for MISO Channels with Decentralized Receivers,” P. Amihood, E. Masry, L. B. Milstein, and J. G. Proakis, *IEEE Transactions on Communications*, vol. 55, pp. 1385-1396, July 2007. The dissertation author was the primary researcher and author of this publication.

This work was supported in part by the Office of Naval Research under grant N-00014-02-1-0001, the Multidisciplinary University Research Initiative (MURI) under the Office of Naval Research grant N-00014-00-1-0564, and the US Army Research Office under the MURI grant number W911NF-04-1-0224.
VITA

1977  Born, São Paulo, Brazil

2000  B.A., Physics, magna cum laude
       Cornell University
       Ithaca, NY, USA

2003  M.S., Electrical Engineering
       (Communication Theory and Systems)
       University of California, San Diego
       La Jolla, CA, USA

2000-2007  Graduate Student Research Assistant
            Department of Electrical and Computer Engineering
            University of California, San Diego
            La Jolla, CA, USA

2007  Ph.D., Electrical Engineering
       (Communication Theory and Systems)
       University of California, San Diego
       La Jolla, CA, USA

PUBLICATIONS


ABSTRACT OF THE DISSERTATION

Performance Analysis of Interference Suppression Techniques for Multiple Antenna Systems

by

Patrick Amihood

Doctor of Philosophy in Electrical Engineering (Communication Theory and Systems)

University of California San Diego, 2007

Professor Laurence B. Milstein, Chair

Professor Elias Masry, Co-Chair

Professor John G. Proakis, Co-Chair

In this dissertation, we consider the performance analysis of interference suppression techniques for multiple antenna systems. In the first part of the dissertation, the performance of multicode direct sequence spread spectrum Multiple-Input Multiple-Output (MIMO) systems in the presence of frequency selective fading is evaluated. We derive the asymptotic distribution of the multi-antenna interference when the processing gain is sufficiently large. The probability of error is derived for the conventional RAKE receiver, and its performance is compared for various system configurations. We consider system tradeoffs for both fixed rate and fixed diversity. For a fixed total data rate, we demonstrate the advantage of decreasing the number of transmit antennas while increasing the number of codes, and for a fixed total diversity order, we demonstrate the advantage of decreasing the number of RAKE taps while increasing the number of receive antennas.

In the second part of this dissertation, interference suppression is achieved by precoding at the transmitter. The performance of a pre-BLAST-DFE technique with...
Tomlinson-Harashima precoding and decentralized receivers, operating over Multiple-Input Single-Output (MISO) frequency selective fading channels, is evaluated. First, we derive the probability of symbol error for Tomlinson-Harashima precoding operating over Single-Input Single-Output (SISO) channels with Intersymbol Interference (ISI). We then generalize this result to MISO frequency selective fading channels with decentralized receivers by using the QR decomposition technique. The effect of optimal ordering of the decentralized receivers, which minimizes the total transmit energy, is investigated. We obtain a closed-form expression for the probability density function of the squared diagonal elements of the upper-right triangular matrix belonging to the optimal QR decomposition when two transmit antennas and two receivers are used. We also provide simulations to corroborate the analytical results.

Finally, we consider the effects of channel estimation errors on the performance of a system employing Tomlinson-Harashima precoding and the QR decomposition, operating over MISO frequency-flat fading channels with decentralized receivers. The QR decomposition of the channel matrix is employed to arrive at an equivalent channel where successive interference cancellation at the transmitter can be used to remove the effect of the multiuser interference.
Introduction

Motivated by the need for high data rate communications, MIMO systems exploit the spatial dimension to achieve high spectral efficiency and significant diversity advantage [1–4]. Much focus has been given on the design of both transmitter and receiver algorithms to reduce the interference caused by the superposition of the signals transmitted from the multiple antennas. In the presence of frequency selective fading in MIMO channels, it is also necessary to perform temporal equalization to remove intersymbol interference.

In general, there are three different channels considered in the literature: the point-to-point MIMO channel (also known as the single-user channel (MIMO-SU)), the multipoint-to-point MIMO channel (also known as the multiaccess or uplink channel (MIMO-MA)), and the point-to-multipoint MIMO channel (also known as the broadcast or downlink channel (MIMO-BC)). Figure 1.1 presents the three different MIMO channels. In the MIMO-SU channel, joint processing is possible at both the transmitter and receiver. In the MIMO-MA channel, joint processing at the receiver is optimal. In the MIMO-BC channel, joint processing at the transmitter is optimal, whereas joint processing at the receiver is not only suboptimal, but also too complex and, hence, very costly.

In this dissertation, we consider the performance analysis of interference suppression techniques for MIMO-SU and MIMO-BC channels. In MIMO-MA channels, the receiver has co-located antennas and can be designed to suppress the spatial and temporal interference [5]. In MIMO-BC channels, the receivers are decentralized and
Figure 1.1: MIMO channels
the transmitter can be designed to aid in suppressing interference using techniques that are similar to those used in MIMO-MA channels. Recent work has considered the duality between MIMO-MA and MIMO-BC channels [6, 7].

The first part of this dissertation considers the MIMO-SU channel. When the channel is known at the transmitter, capacity results motivate signalling along the eigen-modes of the channel [1]. When the channel is unknown at the transmitter, different approaches have been developed. Either the goal is to maximize throughput, in which case we transmit at the largest rate possible and rely on the receiver to decorrelate the signals [2,3,8,9], or the goal is to minimize the probability of error, in which case we sacrifice the transmission rate in order to further exploit the available spatial diversity [10–13]. It is possible to design transmit-receive algorithms that strike a balance between these two approaches. Optimal diversity-rate tradeoffs have been shown to exist [14,15] and currently much research attention is devoted to finding practical algorithms that achieve this optimal tradeoff [16–18].

Chapter 2 considers the performance of a multicode direct sequence spread spectrum system with a conventional RAKE receiver operating over frequency selective fading MIMO channels. Walsh-Hadamard codes are used to transmit multiple streams of data in parallel at each antenna. The composite signal at each antenna is then scrambled by a pseudorandom sequence. While at a given transmit antenna the multiple streams are orthogonal to one another, the signals at the outputs of different antennas interfere, since they are not perfectly orthogonal. The multipath channel introduces further inter-antenna interference, as well as inter-chip interference. The difficulty of decorrelating the superimposed signals is alleviated by using a different spreading sequence on each transmit antenna. To simplify receiver design, no joint processing between the different receive antennas is considered. However, since the receive antennas are co-located, performance benefits from receive diversity. The system model described in Chapter 2 corresponds to the basic physical layer structure of the High Speed Downlink Shared Channel (HS-DSCH) for MIMO systems as specified in the 3GPP [19,20].

In deriving the probability of error, it is convenient to make the assumption that the interference due to the other transmitting antennas is Gaussian when the number of transmitting antennas is large. In [21], it is shown that for a large number of antennas, the multi-antenna interference is asymptotically Gaussian. However, in practice, the
number of antennas may be small while the processing gain can be relatively large. Therefore, we establish the asymptotic distribution of the multi-antenna interference when the processing gain is sufficiently large.

Complexity and power consumption considerations at the receiver may limit the possibility for spatial processing and temporal equalization. For example, this may be the case in MIMO-BC channels where receivers are simple mobile units or in the nodes of wireless sensor networks. Therefore, there is need to unburden the receiver and perform these operations at the transmitter. The second part of this dissertation considers the MIMO-BC channel.

In MIMO-BC channels, the transmitter, which has co-located antennas, performs the spatial encoding by a technique called precoding [22–24]. Precoding is a form of pre-equalization at the transmitter. When accurate channel state information (CSI) is available at the transmitter, precoding can successfully eliminate the interference due to the interfering transmit antennas as well as intersymbol interference due to the presence of frequency selective fading.

An introduction to the MIMO-BC channel can be found in [25]. In general, there are two classes of pre-processing available to the transmitter: linear and nonlinear processing. Linear processing techniques include the duals to the ZF and MMSE receivers in MIMO-SU channels. Nonlinear processing techniques include the dual to the DFE as well as vector precoding [26,27], which are inspired by the application of dirty paper coding [28]. The ‘writing on dirty paper’ result, due to Costa [28], shows that if the transmitter knows the additive interference that is present at the receiver, the transmitter can use optimal coding together with interference pre-subtraction to transmit at the same rate as if there was no interference. The application of Costa precoding to the MIMO-BC channel is established in [29]. Other techniques [30–34] include lattice-reduction-aided precoders which avoid the complexity of sphere encoding required by vector precoders, yet achieve similar performance.

In Chapter 3, we analyze the performance of a pre-BLAST-DFE technique for MISO channels with decentralized receivers operating over frequency selective fading channels. The MISO channel with decentralized receivers is the same as the MIMO-BC channel except that each user has only one receive antenna. We use the QR decomposition technique to enable successive interference cancellation. Consequently, neither
multiuser interference nor ISI is present at the receivers. However, performance is degraded by the increase in power inherent in the pre-BLAST-DFE technique.

The pre-BLAST-DFE technique is presented in [35], which provides simulation results and employs the Gaussian approximation, ignoring the effect of Tomlinson-Harashima precoding [36, 37]. In this dissertation, we provide analytical results which include the effect of Tomlinson-Harashima precoding as well as the statistics of the QR decomposition. Furthermore, we provide analytical results when the optimal ordering of the decentralized receivers is employed to minimize the total transmit energy. We obtain a closed-form expression for the probability density function of the squared diagonal elements of the upper-right triangular matrix belonging to the optimal QR decomposition when two transmit antennas and two receivers are used. The ordered QR decomposition is compared to the unordered QR decomposition, as well as the ZF and MMSE linear precoders.

The performance of communication systems operating over MIMO-BC channels is dependent on the accuracy and availability of CSI at the transmitter. Imperfect CSI may be caused, for example, by noisy channel estimates, feedback delay, and finite rate quantization [38, 39]. Depending on the channel coherence time, the CSI may consist, for example, of imperfect estimates of the channel coefficients or long-term estimates of the channel correlation matrix [40].

In Chapter 4, we analyze the effects of channel estimation errors on the performance of a system employing Tomlinson-Harashima precoding and the QR decomposition for MISO channels with decentralized receivers operating over frequency-flat fading channels. The system considered in Chapter 4 is the same as that presented in Chapter 3, except for the fact that we do not consider ordering of the decentralized receivers. Due to imperfect channel estimation, which causes a mismatch between the precoder and the channel, multiuser interference is present at the receivers and must be accounted for in the derivation of the probability of symbol error. The channel is estimated in the uplink with training sequences transmitted by each user. The channel coherence time is assumed to be large enough so that these estimates may be used in the downlink.

Previous work on precoding with imperfect channel estimation includes [41], which attempts to derive the probability of symbol error for a system similar to that considered in Chapter 4. In [41], matrix differentials are applied to derive the precoding
matrices by assuming the channel estimation error is small. Indeed, we show that the variance of the entries of the channel estimation error matrix is inversely proportional to the length of the training sequences. However, there is no guarantee that actual realizations of the entries of the channel estimation error are small and, therefore, it is incorrect to use matrix differentials and their identities. The work in [43] and [44] considers receiver equalization with the QR decomposition approach, but both make various approximations related to the dependency of the channel estimation matrix and its QR decomposition.

We arrive at our results by different techniques, effectively avoiding having to establish the dependencies of various random variables, by initially deriving the conditional probability of symbol error, conditioned on the elements of the actual channel matrix as well as the elements of the channel estimation error matrix. The conditional probability of symbol error is then averaged by numerical integration over the densities of these random variables, leading to results that coincide with simulations.

Finally, a summary and possible extensions are presented in Chapter 5.
Performance Analysis of High Data Rate MIMO Systems in Frequency Selective Fading Channels

2.1 Introduction

Multiple antennas can improve the performance of wireless communication systems by increasing the data rate and achieving spatial diversity to combat fading [3]. In addition, the use of multicode direct sequence spread spectrum techniques [45], motivated by application to future cellular data communications, allows even higher data rates. While multiple transmit antennas provide multiple spatial channels, the transceiver must be designed to reduce the interchannel interference among the spatial channels. It is therefore important to understand the statistics of the multi-antenna interference.

In this chapter, we derive the performance of a multicode direct sequence spread spectrum system with a conventional RAKE receiver operating over frequency selective fading MIMO channels. While we use the same system model as in [21], we generalize the channel model from two multipath components to an arbitrary number of multipath components. Walsh-Hadamard codes are used to transmit multiple streams of data in
parallel at each antenna. The composite signal at each antenna is then scrambled by a pseudorandom sequence. While at a given transmit antenna the multiple streams are orthogonal to one another, the signals at the outputs of different antennas interfere, since they are not perfectly orthogonal. The multipath channel introduces further inter-antenna interference, as well as inter-chip interference.

In deriving the probability of error, it is convenient to make the assumption that the interference due to the other transmitting antennas is Gaussian when the number of transmitting antennas is large. In [21], the Lindeberg-Lévy version of the central-limit theorem [46] is applied to show that, for a large number of antennas, the multi-antenna interference is asymptotically Gaussian. However, in practice, the number of antennas may be small. For example, in the 3rd Generation Partnership Project (3GPP), no more than four antennas are supported [19]. In this case, the Gaussian assumption may not hold. In contrast, the processing gain is relatively large, up to 512 as specified in the 3GPP [20]. Here, we establish the asymptotic distribution of the multi-antenna interference when the processing gain is sufficiently large.

We remark that the underlying process is not stationary. We obtain an expression for the conditional asymptotic variance, derive the conditional characteristic function, and establish the conditional asymptotic normality of the multi-antenna interference when the processing gain is large. The derivations are quite involved technically, as will be seen from the proof of Theorems 2.1 and 2.2.

We show in particular that the conditional variance converges for large processing gain to a constant that is independent of the terms we condition over. First, this specifies the rate of convergence of the central limit theorem. Second, the fact that the conditional asymptotic variance is independent of the desired user’s sequence and the interfering data means that we do not need to uncondition over a large number of terms that would have made the computation of the probability of error intractable.

There are various papers in the literature that have considered the asymptotic analysis of MIMO systems. The work in [47] considers the asymptotic statistics of the eigenvalues of the covariance of a channel matrix that is designed to include the effect of scattering objects. We consider the asymptotic distribution of the multi-antenna interference in a spread system and derive our results from first principles. The work in [48] considers the convergence of the SIR of the MMSE receiver as the processing gain
gets arbitrarily large. However, this result does not prove the asymptotic normality of our test statistic. Reference [49] presents a valuable technique for using spreading with MIMO systems. However, it does not characterize the multi-antenna interference in the model that we are considering. To the best of our knowledge, our work is the first to establish rigorously the conditional asymptotic normality of the test statistic.

The system model and problem statement are described in Section 2.2. The asymptotic normality of the multi-antenna interference is established in Section 2.3, where we also derive the probability of error. The effect of randomly spaced multipath arrivals is considered in Section 2.4, the performance analysis is given in Section 2.5 and conclusions are made in Section 2.6.

### 2.2 System Model and Problem Statement

The system model, shown in Figures 2.1, 2.2, and 2.3, consists of $N_t$ transmit antennas and $N_r$ receive antennas. At each transmit antenna, $N_c$ Walsh-Hadamard codes of length $L$ are used to transmit multiple streams of data. There is a different symbol on each branch of each transmit antenna. We assume each antenna transmits the same number of data streams. Since the transmit antennas are co-located, they are assumed to be perfectly synchronized in order to simplify the analysis. The chip duration is $T_{ch}$ and the symbol duration is given by the chip duration times the processing gain, $T_s = T_{ch} \times L$. The composite signal at the $i^{th}$ antenna is then spread by a real sequence of length $N \gg L$. Each spreading sequence is periodic with period much longer than the processing gain. This enables us to model the sequences as random, where each chip in a sequence is independently determined and the sequences are mutually independent. The modulation considered is BPSK.

The low-pass-equivalent signal transmitted on the $i^{th}$ antenna is

$$x_i(t) = p_i(t) \sum_{l=1}^{N_c} s_{i,l}(t)d_{i,l}(t).$$

The spreading sequence is given by $p_i(t) = \sum_{n=-\infty}^{\infty} p_i(n)P_{T_{ch}}(t-nT_{ch})$, where $P_{T_{ch}}(t) = \frac{1}{\sqrt{T_{ch}}}$ for $0 \leq t \leq T_{ch}$ and equals zero otherwise, and the $p_i(n) \in \{-1, 1\}$ are elements of a pseudonoise sequence of length $N$. The different data streams on a given antenna are orthogonalized by $s_{i,l}(t) = \sum_{n=-\infty}^{\infty} s_{i,l}(n)P_{T_{ch}}(t-nT_{ch})$, where the $s_{i,l}(n) \in \{-1, 1\}$ are
Figure 2.1: System model for multicode MIMO system.

Figure 2.2: RAKE receiver for the $i^{th}$ transmit antenna at the $j^{th}$ receive antenna.

Figure 2.3: Despreader for the $i^{th}$ transmit antenna.
real-valued nonrandom elements of a Walsh-Hadamard code of length $L$.

The data is given by

$$d_{i,l}(t) = \sum_{n=-\infty}^{\infty} d_{i,l}\left(\left\lfloor \frac{n}{L} \right\rfloor \right) P_{T_{ch}}(t - nT_{ch}),$$

where the $d_{i,l}(\left\lfloor \frac{n}{L} \right\rfloor) \in \{-1,1\}$ are i.i.d. in $l$ for fixed $n$. For fixed $l$, and different $n$, they are dependent.

The channel between the $i^{th}$ transmit antenna and the $j^{th}$ receive antenna is described by

$$h_{i,j}(t) = \sum_{k=0}^{K-1} g_{i,j,k} \delta(t - kT_{ch}),$$

where $K$ is the number of resolvable paths and it is assumed that $(K - 1)T_{ch}$ is much smaller than the symbol duration $T_s$.

The gain parameter between the $i^{th}$ transmit antenna and the $j^{th}$ receive antenna for the $k^{th}$ path is defined as $g_{i,j,k}$. The gain parameters are complex-valued Gaussian random variables, each with zero mean and variance $E[|g_{i,j,k}|^2] = \frac{1}{KN_r}$, for all $i$, $j$, and $k$. The real and imaginary parts of the gain parameters are independent with equal variance.

The gain parameters are normalized by the number of paths and the number of receive antennas. By normalizing to the number of receive antennas, it is assumed that the average receive power at each antenna is the same. When the length of the antenna array is fixed, increasing the number of elements decreases the effective aperture of each element. This holds, for example, with a collinear array of dipoles or an array of microstrip antennas [50].

The complex valued low-pass-equivalent AWGN is zero-mean with two-sided spectral density $N_0$.

We assume that the channels between each transmit and receive antenna pair have identical delays so we have synchronous reception. The assumption of synchronous reception, in conjunction with the use of spreading sequences that are much longer than the processing gain, adheres to the technical specifications of the 3GPP [19]. The system model described in Section 2.2 corresponds to the basic physical layer structure of the High Speed Downlink Shared Channel (HS-DSCH) for MIMO systems, which proposes
the use of common OVSF codes and different long spreading sequences for the different transmit antennas.

In addition to the assumption that synchronization has been achieved following an acquisition procedure, synchronous reception is also valid when the differential time delay between each transmit and receive antenna pair is negligible compared to a chip interval. For example, consider a system with $N_t = 2$, $N_r = 1$, an antenna spacing of $\lambda/2$ m, and a carrier frequency of $f_c = 1900$ MHz. Let the distance between one transmit and receive antenna pair be 10 m. Then the differential delay is approximately $\tau \approx 1.67 \times 10^{-12}$ s. The chip interval used in the 3GPP is $T_{ch} = \frac{1}{3.86 \times 10^9}$ s $= 2.6 \times 10^{-7}$ s, which is 5 orders of magnitude larger than the differential delay.

If one were not to adhere to the standard and assume randomly spaced multipath arrivals, we provide an outline of the changes in the analysis and results in Section 2.4.

The received signal at each antenna is matched filtered and processed by $N_t$ RAKE receivers, each with perfect knowledge of the channel between the receive antenna and the corresponding transmit antenna. The RAKE receiver is generalized and optimized in [51–53]. The complex low-pass-equivalent received signal at the $j^{th}$ antenna is

$$y_j(t) = \sum_{i=1}^{N_t} x_i(t) * h_{i,j}(t) + w_j(t),$$

where $w_j(t)$ is the AWGN process. Chip matched filtering gives

$$y_j(n) = \sum_{i=1}^{N_t} \sum_{k=0}^{K-1} g_{i,j,k} p_i(n-k) \sum_{l=1}^{N_r} s_{i,l}(n-k) d_{i,l} \left( \left\lfloor \frac{n-k}{L} \right\rfloor \right) + w_j(n),$$

where the noise samples, $\{w_j(n)\} = \left\{ \frac{1}{\sqrt{T_{ch}}} \int_{nT_{ch}}^{(n+1)T_{ch}} w_j(t) \, dt \right\}$, are zero-mean uncorrelated complex Gaussian random variables with variance $E[|w_j(n)|^2] = N_0 \triangleq \sigma_w^2$.

Without loss of generality, let the desired bit be $d_{1,1}(0)$. The interference due to the $i^{th}$ transmit antenna at the output of the matched filter at the $j^{th}$ antenna is

$$I_{A}^{(i,j)}(n) = \sum_{k=0}^{K-1} g_{i,j,k} p_i(n-k) \sum_{l=1}^{N_r} s_{i,l}(n-k) d_{i,l} \left( \left\lfloor \frac{n-k}{L} \right\rfloor \right),$$

and the interference at the $j^{th}$ antenna due to all interfering antennas can be written as

$$I_{A}^{(j)}(n) = \sum_{i=2}^{N_t} I_{A}^{(i,j)}(n).$$
Then the received signal at the output of the matched filter at the \( j \)th antenna at time \( n \) is

\[
y_j(n) = \sum_{k=0}^{K-1} g_{1,j,k} p_1(n-k) \sum_{l=1}^{N_r} s_{1,l}(n-k) d_{1,l} \left( \left\lfloor \frac{n-k}{L} \right\rfloor \right) + I_A^{(j)}(n) + w_j(n). \tag{2.1}
\]

The output of the RAKE and despreader is

\[
Z_{1,1}(L) = \frac{1}{L} \sum_{n=K-1}^{L+(K-1)-1} s_{1,1}(n-(K-1)) \sum_{j=1}^{N_r} p_1(n-(K-1)) \sum_{k=0}^{K-1} g_{1,j,k}^* y_j(n-(K-1-k)) \\
- \frac{1}{L} \sum_{n=0}^{L-1} s_{1,1}(n) \sum_{j=1}^{N_r} p_1(n) \sum_{k=0}^{K-1} g_{1,j,k}^* y_j(n+k). \tag{2.2}
\]

Substituting \( y_j(n) \) of (2.1) in (2.2), we have

\[
Z_{1,1}(L) = \frac{1}{L} \sum_{n=0}^{L-1} s_{1,1}(n) \sum_{j=1}^{N_r} p_1(n) \sum_{k_1=0}^{K-1} \sum_{k_2=0}^{K-1} g_{1,j,k_1}^* g_{1,j,k_2} p_1(n+k_1-k_2) \\
\cdot \sum_{l=1}^{N_r} s_{1,l}(n+k_1-k_2) d_{1,l} \left( \left\lfloor \frac{n+k_1-k_2}{L} \right\rfloor \right) + Z_{A,L} + Z_{W,L}, \tag{2.3}
\]

where

\[
Z_{A,L} = \frac{1}{L} \sum_{n=0}^{L-1} s_{1,1}(n) \sum_{j=1}^{N_r} p_1(n) \sum_{k=0}^{K-1} g_{1,j,k}^* f_A^{(j)}(n+k)
\]

and

\[
Z_{W,L} = \frac{1}{L} \sum_{n=0}^{L-1} s_{1,1}(n) \sum_{j=1}^{N_r} p_1(n) \sum_{k=0}^{K-1} g_{1,j,k}^* w_j(n+k)
\]

are the contributions due to the multi-antenna interference and AWGN, respectively.

It is possible to further simplify (2.3) by exploiting the orthogonality of the Walsh-Hadamard sequences. The sequences align when \( k_1 = k_2 \). We then have

\[
Z_{1,1}(L) = \frac{1}{L} \sum_{j=1}^{N_r} \sum_{k_1=0}^{K-1} \sum_{k=0}^{K-1} \sum_{l=1}^{N_r} |g_{1,j,k_1}|^2 d_{1,l}(0) \sum_{n=0}^{L-1} s_{1,1}(n) s_{1,l}(n) + Z_{S,L} + Z_{A,L} + Z_{W,L} \\
= \sum_{j=1}^{N_r} \sum_{k=0}^{K-1} |g_{1,j,k}|^2 d_{1,1}(0) + Z_{S,L} + Z_{A,L} + Z_{W,L}, \tag{2.4}
\]
where we have used the orthogonality of the Walsh-Hadamard codes to eliminate the \( N_c - 1 \) interfering signals on the same antenna as the desired signal, and where

\[
Z_{S,L} \triangleq \frac{1}{L} \sum_{n=0}^{L-1} s_{1,1}(n) \sum_{j=1}^{N_r} p_1(n) \sum_{k_1=0}^{K-1} g_{1,j,k_1}^* \sum_{k_2=0}^{K-1} g_{1,j,k_2} \cdot p_1(n + k_1 - k_2) \sum_{l=1}^{N_c} s_{i,l}(n + k_1 - k_2) d_{i,l} \left( \left\lfloor \frac{n + k_1 - k_2}{L} \right\rfloor \right).
\]

Note that \( Z_{S,L} \) contains the self-interference of the desired signal as well as the other signals on the same transmit antenna.

Finally, \( T_{1,1}(L) = \Re\{Z_{1,1}(L)\} \) is compared to zero.

To find the distribution of the decision variable \( T_{1,1}(L) \), we must first find the distribution of the multi-antenna interference.

Expanding \( Z_{A,L} \) gives

\[
Z_{A,L} = \frac{1}{L} \sum_{n=0}^{L-1} s_{1,1}(n) \sum_{j=1}^{N_r} p_1(n) \sum_{k_1=0}^{K-1} g_{1,j,k_1}^* \sum_{i=2}^{N_t} \sum_{k_2=0}^{K-1} g_{i,j,k_2} \sum_{l=1}^{N_c} s_{i,l}(n + k_1 - k_2) d_{i,l} \left( \left\lfloor \frac{n + k_1 - k_2}{L} \right\rfloor \right).
\]

Setting \( b(n) \triangleq s_{1,1}(n)p_1(n) \) and \( q_i(n) \triangleq \sum_{l=1}^{N_c} s_{i,l}(n) d_{i,l} \left( \left\lfloor \frac{n}{L} \right\rfloor \right) \), and rearranging terms gives

\[
Z_{A,L} = \sum_{j=1}^{N_r} \sum_{i=2}^{N_t} \left[ K-1 \sum_{k_1=0}^{K-1} g_{1,j,k_1}^* g_{i,j,k_1} \sum_{n=0}^{L-1} b(n)p_1(n)q_i(n) \right] \triangleq Z_L^{(i)}(0)
\]

\[
+ \sum_{k_1=0}^{K-1} \sum_{k_2=0}^{K-1} g_{i,j,k_1}^* g_{i,j,k_2} \sum_{n=0}^{L-1} b(n)p_1(n + k_1 - k_2)q_i(n + k_1 - k_2) \right] \triangleq Z_L^{(i)}(k_1 - k_2)
\]

Let \( m = k_1 - k_2 \). Since the sequences \( \{p_i(n)\}, \{d_{i,l} \left( \left\lfloor \frac{n}{L} \right\rfloor \right) \}, \{g_{i,j,k}\} \), are each independent in \( i \), and the \( \{Z_L^{(i)}(m)\} \) are not a function of \( j \), it suffices to establish the joint asymptotic normality of \( \{Z_L^{(i)}(m)\} \), for fixed \( i \).
2.3 Asymptotic Analysis and Error Probability

We first consider the asymptotic normality of the multi-antenna interference. To establish the joint asymptotic normality of \( \{ Z_L^{(i)}(m) \} \), for fixed \( i \), we use the Cramér device [54]. Let \( \{ \alpha_m; m = -(K - 1), \ldots, K - 1 \} \) be arbitrary real numbers and set

\[
Y_{A,L}^{(i)} = \sum_{m=-(K-1)}^{K-1} \alpha_m Z_L^{(i)}(m). \tag{2.5}
\]

We next establish the statistics of \( \{ Y_{A,L}^{(i)} \} \) when \( L \) is large.

**Theorem 2.1**

The \( \{ \sqrt{L} Y_{A,L}^{(i)} \} \) are independent in \( i \) for each fixed \( L \). Conditioned on the desired user’s spreading sequence \( \{ p_1(n) \} \), and the interfering data \( \{ d_{i,l}(0) \} \), \( \sqrt{L} Y_{A,L}^{(i)} \) is conditionally asymptotically normal with zero mean and variance \( N_c \sum_{m=-(K-1)}^{K-1} \alpha_m^2 \), as \( L \to \infty \).

**Proof.** See Appendix A. □

**Corollary 1**

The \( \{ Z_L^{(i)}(m) \} \) are independent in \( i \) for each fixed \( L \). The random variables \( \{ \sqrt{L} Z_L^{(i)}(m); m = -(K - 1), \ldots, K - 1 \} \), conditioned on \( \{ p_1(n) \} \) and \( \{ d_{i,l}(0) \} \), are conditionally asymptotically independent normal with zero means and identical variance \( N_c \), as \( L \to \infty \).

We apply a similar analysis to the component of the test statistic due to AWGN. Write

\[
Z_{W,L} = \sum_{j=1}^{N_w} \sum_{k=0}^{K-1} g_{1,j,k}^* \frac{1}{L} \sum_{n=0}^{L-1} s_{1,1}(n) p_1(n) w_j(n + k). \]

Since the \( \{ w_j(n) \} \) are independent in \( j \), it suffices to consider the joint asymptotic normality of the \( \{ Z_{W,L}^{(j)}(k) \} \) for fixed \( j \). We again use the Cramér device. Let \( \{ \alpha_k; k = 0, \ldots, K - 1 \} \) be arbitrary real numbers, and set

\[
Y_{W,L}^{(j)} = \sum_{k=0}^{K-1} \alpha_k Z_{W,L}^{(j)}(k). \]

**Theorem 2.2**

The \( \{ \sqrt{L} Y_{W,L}^{(j)} \} \) are independent in \( j \) for each fixed \( L \). Conditioned on the desired user’s
spreading sequence \( \{ p_1(n) \} \), \( \sqrt{L}Y_{W,L}^{(j)} \) is conditionally asymptotically normal with zero mean and variance \( \sigma_w^2 \sum_{k=0}^{K-1} \alpha_k^2 \), as \( L \to \infty \).

**Proof.** See Appendix C. ■

**Corollary 2**

The \( \{ Z_{W,L}^{(j)}(k) \} \) are independent in \( k \) and \( j \) for each fixed \( L \). Also, \( Z_{W,L} \), conditioned on \( \{ p_1(n) \} \), is conditionally independent of \( Z_{A,L} \). Each of the \( \sqrt{L}Z_{W,L}^{(j)}(k), k = 0, \ldots, K - 1 \), conditioned on \( \{ p_1(n) \} \), is conditionally asymptotically normal with zero mean and variance \( \sigma_w^2 \), as \( L \to \infty \).

The self-interference due to multipath, \( Z_{S,L} \), is assumed to be negligible to simplify the analysis. For \( N_t \gg 1 \), the interference due to the other transmit antennas will dominate the self-interference. Even for small \( N_t \), it is possible to show that the contribution of the self-interference to the probability of error is negligible compared to the multi-antenna interference.

Therefore, \( Z_{1,1}(L) \) of (2.4) can be written as

\[
Z_{1,1}(L) \approx \sum_{j=1}^{N_r} \sum_{k=0}^{K-1} |g_{1,j,k}|^2 d_{1,1}(0) + \sum_{j=1}^{N_r} N_t \sum_{i=2}^{K-1} \sum_{k_1=0}^{K-1} \sum_{k_2=0}^{K-1} g_{1,j,k_1}^* g_{i,j,k_2}^* Z_L^{(i)}(k_1 - k_2)
\]

\[
+ \sum_{j=1}^{N_r} \sum_{k=0}^{K-1} g_{1,j,k}^* Z_{W,L}^{(j)}(k),
\]

where the approximation is due to the assumption that the self-interference is negligible.

We now state our principal result.

**Theorem 2.3**

The decision variable \( T_{1,1}(L) = \Re\{Z_{1,1}(L)\} \), conditioned on \( \{ g_{i,j,k} \} \), \( \{ p_1(n) \} \), and \( \{ d_{i,l}(0) \} \), is a conditionally asymptotically normal random variable with conditional mean

\[
m_T = \sum_{j=1}^{N_r} \sum_{k=0}^{K-1} |g_{1,j,k}|^2 d_{1,1}(0),
\]

and conditional variance \( \sigma_T^2(L) \), where

\[
L \sigma_T^2(L) \triangleq L \Var\left[ T_{1,1}(L) \bigg| \{ g_{i,j,k} \}, \{ p_1(n) \}, \{ d_{i,l}(0) \} \right]
\]

\[
\to \frac{1}{2} \left( N_c \sum_{L} g_{1,j_1,k_1}^* g_{1,j_1,k_1} g_{1,j_2,k_2} g_{1,j_2,k_2}^* + \sigma_w^2 \sum_{j=1}^{N_r} \sum_{k=0}^{K-1} |g_{1,j,k}|^2 \right) \triangleq \tilde{\sigma}_T^2, \quad (2.6)
\]
with probability one as \( L \to \infty \), and where

\[
\sum_{\nu} = \sum_{i=2}^{N_t} \sum_{j_1=1}^{N_r} \sum_{j_2=1}^{N_r} \sum_{k_1=0}^{K-1} \sum_{k_2=0}^{K-1} \sum_{k'_1=0}^{K-1} \sum_{k'_2=0}^{K-1} \cdot
\]

Then the conditional variance of \( T_{1,1}(L) \) can be approximated as

\[
\sigma^2_T(L) \approx \frac{1}{L} \tilde{\sigma}^2_T.
\]

Note that the asymptotic conditional variance \( \tilde{\sigma}^2_T \) does not depend either on \( \{p_1(n)\} \) or \( \{d_{i,l}(0)\} \).

Assume \( d_{1,1}(0) = 1 \). The conditional probability of bit error can be approximated by Theorem 2.3 for large \( L \):

\[
\Pr \left( T_{1,1}(L) < 0 \bigg| \{g_{i,j,k}\}, \{p_1(n)\}, \{d_{i,l}(0)\} \right) \approx Q \left( \frac{m_T \sqrt{L}}{\tilde{\sigma}_T} \right),
\]

where \( Q(\cdot) \) is the Gaussian Q-function [55]. The unconditional probability of bit error, \( P_e \), can then be approximated, for large \( L \), by \( \hat{P}_e \), so that \( P_e \approx \hat{P}_e \) and we write

\[
\hat{P}_e = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} \ldots
\]

\[
\sqrt{2L} \sum_{j=1}^{N_r} \sum_{k=0}^{K-1} |x_{1,j,k}|^2
\]

\[
\cdot Q \left( \left( \sum_{\nu} x_{1,j_1,k_1}^* x_{1,j_1,k_1'} + \sigma^2_w \sum_{j=1}^{N_r} \sum_{k=0}^{K-1} |x_{1,j,k}|^2 \right)^{1/2} \right)
\]

\[
\cdot p_{g_{1,1,0}}(x_{1,1,0}) \ldots p_{g_{N_t,N_r,K-1}}(x_{N_t,N_r,K-1}) \ dx_{1,1,0} \ldots \ dx_{N_t,N_r,K-1}.
\]

Evaluating multi-dimensional integrals is computationally intensive, particularly over infinite regions. Here, the number of function evaluations increases as the \((N_tN_rK)^{th}\) power of the number needed to evaluate a one-dimensional integral. For example, selecting typical parameters such as \( N_t = 4 \), \( N_r = 2 \), and \( K = 2 \) results in a 16-dimensional integral.

The multidimensional integral in (2.7) is evaluated by importance sampling [56]. Since the region of integration is large, and the integrand is small over most of the multidimensional space, then approximating the distribution and sampling accordingly gives more accurate results for a given number of samples than does blind sampling.
2.4 Randomly spaced multipath arrivals

All analyses and results above assume equally spaced multipath arrivals, but the same techniques can be applied when the multipath arrivals are random. The channel is now described by

\[ h_{i,j}(t) = \sum_{k=0}^{K-1} g_{i,j,k} \delta(t - kT_{ch} - \tau_{i,j}), \]

where \( \tau_{i,j} \) is the delay between transmit antenna \( i \) and receive antenna \( j \). The \( \{ \tau_{i,j} \} \) are i.i.d. in \( i \) and \( j \), and are each uniformly distributed random variables in \( [0, T_{ch}] \).

In designing the chip matched filter, we need to take into account the differential delay from the desired transmit antenna \( (i = 1) \) to receive antenna \( j \). In Appendix E, we derive the asymptotic statistics of the multi-antenna interference and condition also on the \( \{ \tau_{i,j} \} \). The end result is an additional term, \( K_i(j_1, j_2) \), multiplying the channel fades in the expression for the asymptotic conditional variance of the multi-antenna interference (2.6). This term is given by

\[
K_i(j_1, j_2) \triangleq \begin{cases} 
\frac{\tau_{i,j_1} T_{ch}}{T_{ch}} + \left( 1 - \frac{\tau_{i,j_1}}{T_{ch}} \right) \left( 1 - \frac{\tau_{i,j_2}}{T_{ch}} \right) & \tau_{1,j_1} > \tau_{i,j_1} \text{ and } \tau_{1,j_2} > \tau_{i,j_2} \text{ or } \\
\left( 1 - \frac{\tau_{i,j_1}}{T_{ch}} \right) \left( 1 - \frac{\tau_{i,j_2}}{T_{ch}} \right) & \tau_{1,j_1} < \tau_{i,j_1} \text{ and } \tau_{1,j_2} < \tau_{i,j_2} \text{ otherwise } 
\end{cases}
\]

Therefore, the probability of error in the case of randomly spaced multipath arrivals is

\[
\tilde{P}_e = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \int_0^{T_{ch}} \cdots \int_0^{T_{ch}} \left( \frac{\sqrt{2L} \sum_{j=1}^{N_r} \sum_{k=0}^{K-1} |x_{1,j,k}|^2}{N_c \sum_{\nu} x_{1,j_1,k_1,x_{1,j_2,k_2}^* K_i(j_1, j_2) + \sigma_w^2 \sum_{j=1}^{N_r} \sum_{k=0}^{K-1} |x_{1,j,k}|^2} \right)^{1/2} \cdot Q \cdot p_{\tau_{1,0}}(\tau_{1,0}) \cdots p_{\tau_{N_r,N_r,K-1}}(\tau_{N_r,N_r,K-1}) \cdot p_{\tau_{N_r,N_r,K}}(\tau_{N_r,N_r,K}) \cdot dx_{1,0} \cdots dx_{N_r,N_r,K-1} \cdot d\tau_{1,1} \cdots d\tau_{N_r,N_r}. \tag{2.8}
\]
2.5 Performance Analysis

In this section, we consider performance as a function of the system parameters. We also consider system tradeoffs for both fixed diversity and fixed data rate. We first present numerical results. These are obtained by numerically integrating expressions (2.7) and (2.8). Then, based on an approximation to the probability of error given in (2.7), we derive analytical insights which allow us to predict the system behavior shown in the numerical results. Note that we are using multiple transmit antennas and multiple codes to increase the data rate, while using multiple receive antennas and the RAKE receiver to achieve diversity.

Figure 2.4 shows results for $N_t = 2$ and $N_r = 2$. As expected, the performance degrades when $N_t = 4$ compared to $N_t = 2$, since there are two additional interfering transmit antennas. As we increase the number of signals on each transmit antenna from $N_c = 1$ to $N_c = 4$, we also observe a degradation in performance. The curves are compared to 4th order diversity, since the diversity order is, in general, equal to $N_r \times K$.

In Figure 2.5, we keep the product $N_t \times N_c$ equal to six, and we see that of the various combinations, $N_t = 1$ and $N_c = 6$ yields the best results. As we decrease the number of transmit antennas, the multi-antenna interference decreases, and this results in an improvement in the probability of error because of the orthogonality of the Walsh-
Figure 2.5: Numerical results for $N_t \times N_c = 6$, $N_r = 1$, $L = 32$, $K = 2$.

Figure 2.6 considers the tradeoff between the number of RAKE taps with the number of receive antennas for a fixed order of diversity. In this case, $N_t = 2$, $N_r \times K = 6$, and the best performance is achieved when $N_r = 6$ and $K = 1$. This is due to the fact that the multipath interference increases as the number of paths, $K$, increases. For example, when $N_r = 6$ and $K = 1$, each receive antenna detects the desired signal together with one interfering path from the interfering transmit antenna. There is no multipath interference due to either the desired antenna or the interfering antenna. Alternatively, when $N_r = 1$ and $K = 6$, each tap on the RAKE receiver sees not only the multipath interference from the desired antenna, but also all six interfering paths from the interfering antenna. In general, the number of interfering terms grows as $(K - 1) + (N_t - 1) \times N_r \times K^2$. This is because, on every receive antenna, each of the $K$ paths of the desired signal sees $K$ interfering paths from each transmit antenna.

The system with randomly spaced multipath arrivals performs better than the system with equally spaced multipath arrivals. When the multipath arrivals are randomly spaced, the different delays randomize the multi-antenna interference. This can
also be seen by noting that
\[ 0 \leq K_i(j_1, j_2) < 1, \forall i, j_1, j_2. \]  
(2.9)

This term multiplies the scaled conditional variance of the multiple-antenna interference, which is a positive quantity, and appears in the denominator of the conditional SNR in (2.8). Setting the delays to be identically zero gives \( K_i(j_1, j_2) = 1 \), \( \forall i, j_1, j_2 \), and the result reduces to the case of equally spaced multipath arrivals. Numerical results in Figure 2.7 indicate marginal improvement in performance for a system with randomly spaced multipath arrivals compared to a system with equally spaced multipath arrivals for \( N_t = 2, N_c = 1, N_r = 1, L = 32 \), and \( K = 2 \).

The numerical results above ignore the self-interference, \( Z_{S,L} \). We show that even for small \( N_t = 2 \), the contribution of the self-interference term to the probability of error is negligible. If we assume \( Z_{S,L} \), conditioned on \( \{g_{i,j,k}\} \) and \( \{p_1(n)\} \), is a conditionally asymptotically normal random variable, as \( L \to \infty \), and, furthermore, we assume \( N_t = 2, N_r = 1, K = 2 \), and the \( \{g_{i,j,k}\} \) are real, then we can write a simplified expression for
Figure 2.7: Numerical results of a system with equally and randomly spaced multipath arrivals for $N_t = 2$, $N_r = 1$, $L = 32$, $K = 2$.

the unconditional probability of error, given as

$$
\tilde{P}_e = \frac{8}{\pi^2} \int_0^\infty \int_0^\infty \int_0^\infty \int_0^\infty e^{-(x^2+y^2+w^2+z^2)} \cdot dx dy dw dz.
$$

In this case, the conditional variance of the self-interference term is $N_c (2x^2y^2)$, appearing in the denominator of the Q-function. Figure 2.8 plots $\tilde{P}_e$, with and without the conditional variance of the self-interference. It can be seen that the self-interference is negligible, even for $N_t = 2$. When the desired signal is in a deep fade, the self-interference is also in a deep fade. In this case, the multi-antenna interference is a more dominant contribution to the probability of error than is the self-interference. Alternatively, when the self-interference is large relative to the multi-antenna interference, the desired signal is also large. In this case, the contribution to the probability of error is small.

We now show analytically how performance varies with the system parameters:
Since the Q-function is monotonic, we consider system behavior as a function of its argument, the expected value of the conditional SNR. Since the denominator of the conditional SNR is both real and positive, then, using a first-order Taylor series approximation, as in [58, Theorem 4, p. 181], it is possible to make an approximate performance analysis based on the ratio of the expected value of the numerator to the expected value of the denominator. For convenience, we consider the square of the conditional SNR, which gives

$$\tilde{P}_e \geq Q \left( E \left[ \frac{\sqrt{2L} \sum_{j=1}^{N_r} \sum_{k=0}^{K-1} |g_{1,j,k}|^2}{N_c \sum_\nu g_{1,j_1,k_1} g_{i,j_1,k_1'} g_{1,j_2,k_2} g_{i,j_2,k_2'} + \sigma_w^2 \sum_{j=1}^{N_r} \sum_{k=0}^{K-1} |g_{1,j,k}|^2} \right]^{1/2} \right).$$

(2.10)

Since the $Q(\cdot)$ function is monotonic, we consider system behavior as a function of its argument, the expected value of the conditional SNR. Since the denominator of the conditional SNR is both real and positive, then, using a first-order Taylor series approximation, as in [58, Theorem 4, p. 181], it is possible to make an approximate performance analysis based on the ratio of the expected value of the numerator to the expected value of the denominator.

For convenience, we consider the square of the conditional SNR, which gives

$$E \left[ \left( \frac{\sqrt{2L} \sum_{j=1}^{N_r} \sum_{k=0}^{K-1} |g_{1,j,k}|^2}{N_c \sum_\nu g_{1,j_1,k_1} g_{i,j_1,k_1'} g_{1,j_2,k_2} g_{i,j_2,k_2'} + \sigma_w^2 \sum_{j=1}^{N_r} \sum_{k=0}^{K-1} |g_{1,j,k}|^2} \right)^2 \right].$$

$$\approx E \left[ \left( \frac{\sqrt{2L} \sum_{j=1}^{N_r} \sum_{k=0}^{K-1} |g_{1,j,k}|^2}{N_c \sum_\nu g_{1,j_1,k_1} g_{i,j_1,k_1'} g_{1,j_2,k_2} g_{i,j_2,k_2'} + \sigma_w^2 \sum_{j=1}^{N_r} \sum_{k=0}^{K-1} |g_{1,j,k}|^2} \right)^2 \right].$$

(2.11)
In the following, recall that the gain parameters are complex-valued Gaussian random variables, each with zero mean and variance \( E[|g_{i,j,k}|^2] = \frac{1}{KN_r} \), for all \( i, j, \) and \( k \). The real and imaginary parts of the gain parameters are independent with equal variance \( \sigma^2 = \frac{1}{2KN_r} \), and we denote \( g_{i,j,k}^{(R)} \triangleq \Re\{g_{i,j,k}\} \) and \( g_{i,j,k}^{(I)} \triangleq \Im\{g_{i,j,k}\} \).

First, we evaluate the numerator in (2.11), giving

\[
E \left[ \left( \sqrt{2L} \sum_{j=1}^{N_r} \sum_{k=0}^{K-1} |g_{1,j,k}|^2 \right)^2 \right] = 2L \sum_{j_1=1}^{N_r} \sum_{k_1=0}^{K-1} \sum_{j_2=1}^{N_r} \sum_{k_2=0}^{K-1} E \left[ |g_{1,j_1,k_1}|^2 |g_{1,j_2,k_2}|^2 \right] \tag{2.12}
\]

If \( j_1 = j_2 \) and \( k_1 = k_2 \) then \( E \left[ |g_{1,j_1,k_1}|^2 |g_{1,j_2,k_2}|^2 \right] = 8\sigma^4 \). Otherwise, if \( j_1 \neq j_2 \) or \( k_1 \neq k_2 \), then \( E \left[ |g_{1,j_1,k_1}|^2 |g_{1,j_2,k_2}|^2 \right] = 4\sigma^4 \). Substituting in (2.12), we have

\[
E \left[ \left( \sqrt{2L} \sum_{j=1}^{N_r} \sum_{k=0}^{K-1} |g_{1,j,k}|^2 \right)^2 \right] = 2L \left( N_rK \cdot 8\sigma^4 + (N_r^2K^2 - N_rK) \cdot 4\sigma^4 \right).
\]

Since \( \sigma^4 = \frac{1}{4N_r^2K^2} \), we have

\[
E \left[ \left( \sqrt{2L} \sum_{j=1}^{N_r} \sum_{k=0}^{K-1} |g_{1,j,k}|^2 \right)^2 \right] = 2L \left( 1/(N_rK) + 1 \right).
\]

Now, we evaluate the denominator in (2.11). Note that the second term is given by

\[
E \left[ \sigma_w^2 \sum_{j=1}^{N_r} \sum_{k=0}^{K-1} |g_{1,j,k}|^2 \right] = \sigma_w^2 \sum_{j=1}^{N_r} \sum_{k=0}^{K-1} E \left[ |g_{1,j,k}|^2 \right] = \sigma_w^2 \sum_{j=1}^{N_r} \sum_{k=0}^{K-1} 2\sigma^2
\]

\[
= \sigma_w^2 \sum_{j=1}^{N_r} \sum_{k=0}^{K-1} 2/(2N_rK) = \sigma_w^2.
\]

The first term in the denominator of (2.11) gives

\[
E \left[ N_c \sum_{\nu} g_{1,j_1,k_1}^* g_{i,j_1,k_1}' g_{1,j_2,k_2} g_{i,j_2,k_2}' \right] = N_c \sum_{\nu} E \left[ g_{1,j_1,k_1}^* g_{i,j_1,k_1}' g_{1,j_2,k_2} g_{i,j_2,k_2}' \right]
\]

\[
= N_c \sum_{i=2}^{N_r} \sum_{j_1=1}^{N_r} \sum_{j_2=1}^{N_r} \sum_{k_1=0}^{K-1} \sum_{k_2=0}^{K-1} \sum_{k_1'=0}^{K-1} \sum_{k_2'=0}^{K-1} \sum_{k_1-k_1'=k_2-k_2'} E \left[ g_{1,j_1,k_1}^* g_{i,j_1,k_1}' g_{1,j_2,k_2} g_{i,j_2,k_2}' \right].
\]

Using the facts that

\[
E \left[ g_{1,j_1,k_1}^* g_{i,j_1,k_1}' g_{1,j_2,k_2} g_{i,j_2,k_2}' \right] = E \left[ g_{1,j_1,k_1}^* g_{i,j_1,k_1}' \right] E \left[ g_{1,j_2,k_2} g_{i,j_2,k_2}' \right] + E \left[ g_{1,j_1,k_1}^* g_{i,j_1,k_1}' \right] E \left[ g_{1,j_2,k_2} g_{i,j_2,k_2}' \right] + E \left[ g_{1,j_1,k_1}^* g_{i,j_1,k_1}' \right] E \left[ g_{1,j_2,k_2} g_{i,j_2,k_2}' \right],
\]
that the \( \{g_{1,j_1,k_1}\} \) are independent of the \( \{g_{i,j_2,k_2}\} \), and that \( E[g_{i,j,k}] = 0, \forall i, j, k \), we have

\[
E \left[ N_c \sum_{\nu} g_{1,j_1,k_1}^* g_{i,j_1,k_1'} g_{1,j_2,k_2} g_{i,j_2,k_2'} \right] \\
= N_c \sum_{i=2}^{N_t} \sum_{j_1=1}^{N_r} \sum_{k_1=0}^{K-1} \sum_{k_1'=0}^{K-1} \sum_{k_2=0}^{K-1} \sum_{k_2'=0}^{K-1} \sum_{k_1'-k_2'=0} E \left[ g_{1,j_1,k_1}^* g_{i,j_1,k_1'} g_{1,j_2,k_2} g_{i,j_2,k_2'} \right].
\]

The only non-zero contributions occur when \( j_1 = j_2, k_1 = k_2, \) and \( k_1' = k_2' \). Therefore,

\[
E \left[ N_c \sum_{\nu} g_{1,j_1,k_1}^* g_{i,j_1,k_1'} g_{1,j_2,k_2} g_{i,j_2,k_2'} \right] = N_c \sum_{i=2}^{N_t} \sum_{j_1=1}^{N_r} \sum_{k_1=0}^{K-1} \sum_{k_1'=0}^{K-1} E \left[ |g_{1,j_1,k_1}|^2 \right] E \left[ |g_{i,j_1,k_1'}|^2 \right]
= N_c \sum_{i=2}^{N_t} \sum_{j_1=1}^{N_r} \sum_{k_1=0}^{K-1} \sum_{k_1'=0}^{K-1} (2\sigma^2)(2\sigma^2)
= N_c \sum_{i=2}^{N_t} \sum_{j_1=1}^{N_r} \sum_{k_1=0}^{K-1} \sum_{k_1'=0}^{K-1} 4(1/(4N_c^2K^2))
= \frac{N_c(N_t-1)}{N_r}.
\]

Substituting into (2.11), we have

\[
E \left[ \left( \sqrt{2L} \sum_{j=1}^{N_r} K=0 \sum_{k=0}^{K-1} |g_{1,j,k}|^2 \right)^2 \right] \\
E \left[ N_c \sum_{\nu} g_{1,j_1,k_1}^* g_{i,j_1,k_1'} g_{1,j_2,k_2} g_{i,j_2,k_2'} + \sigma_w^2 \sum_{j=1}^{N_r} K=0 \sum_{k=0}^{K-1} |g_{1,j,k}|^2 \right] = \frac{2L(\frac{1}{N_cK} + 1)}{N_c(N_t-1) + \sigma_w^2}
\triangleq f(L, N_c, N_t, N_r, K).
\]

(2.13)

Now we can see how the performance varies with the system parameters. For increasing \( L, f(L, N_c, N_t, N_r, K) \) increases, and performance improves. For increasing \( N_c \) or \( N_t, f(L, N_c, N_t, N_r, K) \) decreases, and performance degrades. This behavior is exhibited in Figure 2.4.

Equation (2.13) allows us to predict system performance in terms of the various tradeoffs discussed earlier. It is easy to see that for a fixed order of diversity \( N_r \times K \), we achieve overall better performance by decreasing the number of RAKE taps \( K \) while increasing the number of receive antennas \( N_r \).
For a fixed total number of streams $N_c \times N_t$, the asymmetry predicted by the term $N_c(N_t - 1)$ in the denominator of $f(L, N_c, N_t, N_r, K)$ shows mathematically why it is better decreasing the number of transmit antennas while increasing the number of codes. This corresponds to the numerical results, where we showed improved performance by taking advantage of the orthogonality of the Walsh-Hadamard sequences. This result applies to the system and channel model that we have presented. If channel state information is available at the transmitter and we apply transmit precoding [22, 57], it may be more advantageous to exploit the spatial dimension rather than code orthogonality.

2.6 Conclusion

We have analysed the performance of a multicode direct sequence spread spectrum system operating over a frequency selective fading MIMO channel. We obtained an expression for the asymptotic variance and established the asymptotic normality of the multi-antenna interference for sufficiently large processing gain. This was used to derive an expression for the probability of bit error which we evaluated numerically.

We showed that the system with randomly spaced multipath arrivals yields better performance compared to the system with equally spaced multipath arrivals. Also, for a fixed total data rate, we demonstrated the advantage of decreasing the number of transmit antennas while increasing the number of codes, and for a fixed total diversity order, we demonstrated the advantage of decreasing the number of RAKE taps while increasing the number of receive antennas.

2.7 Acknowledgements

This chapter is, in part, a reprint of material published as “Asymptotic Performance of Multicode MIMO Systems in Frequency Selective Fading Channels,” P. Amihood, E. Masry, L. B. Milstein, and J. G. Proakis, in Proceedings of the IEEE Military Communications Conference (MILCOM), Atlantic City, New Jersey, USA, vol. 5, pp. 3036 - 3041, October 17 - 20, 2005, and a reprint of material that has been accepted for publication as “Performance Analysis of High Data Rate MIMO Systems in Frequency Selective Fading Channels,” P. Amihood, E. Masry, L. B. Milstein, and J. G. Proakis,
IEEE Transactions on Information Theory, August 2007. The dissertation author was the primary researcher and author of these publications.
Performance Analysis of a Pre-BLAST-DFE Technique for MISO Channels with Decentralized Receivers

3.1 Introduction

In a multiuser MIMO system [25], the design of receivers with low complexity and reduced power consumption is complicated by the fact that the receivers may not jointly process the received signals and may operate over frequency selective fading channels. A given receiver must consequently detect the desired signal among all other interfering signals. It must also perform temporal equalization. Precoding [22] at the transmitter can unburden the receivers by performing joint spatial and temporal equalization. Several precoding algorithms have been proposed for the multiuser MIMO downlink [25]. For example, those that involve linear processing include the transmit duals to the zero-forcing and MMSE receivers. Those that involve nonlinear processing include the successive interference cancellation approach [35], and vector precoding [26, 27].

In this chapter, we analyze the performance of a pre-BLAST-DFE technique for MISO channels with decentralized receivers operating over frequency selective fad-
We use the QR decomposition technique to enable successive interference cancellation. Consequently, neither multiuser interference nor ISI is present at the receivers. However, performance is degraded by the increase in power inherent in the pre-BLAST-DFE technique.

The pre-BLAST-DFE technique is presented in [35], which provides simulation results and employs the Gaussian approximation, ignoring the effect of Tomlinson-Harashima precoding. In this chapter, we provide analytical results which include the effect of Tomlinson-Harashima precoding as well as the statistics of the QR decomposition. Furthermore, we provide analytical results when the optimal ordering of the decentralized receivers is employed to minimize the total transmit energy. This is compared to the performance of the unordered QR decomposition.

For simplicity, we employ one receive antenna at each user terminal. Therefore, each link between the transmitter and a given receiver is a MISO channel. The MISO system with decentralized receivers employed in this chapter is equivalent to the MIMO system in [35]. This is due to the fact that, in [35], no joint processing is performed at the receiver. Therefore, the results in this chapter also apply to a MIMO point-to-point system, where the receiver does not have the power necessary to equalize the channel. This may be the case in MIMO ad-hoc networks [59, 60], where two communicating nodes can come to an agreement on whether the transmitter or the receiver will perform equalization, based on the power availability of each node. Whether it is advantageous for the transmitter to perform equalization as opposed to the receiver is a question that can be posed by considering the costs associated with the receiver feeding back an estimate of the channel to the transmitter. On the other hand, the benefits of equalizing at the receiver can be considered in light of the costs associated with using imperfectly detected symbols in order to perform successive interference cancellation.

In Section 3.2, we consider SISO channels with ISI employing Tomlinson-Harashima precoding at the transmitter. We find a closed-form expression for the probability of symbol error. Then, in Section 3.3, we consider MISO frequency selective fading channels with decentralized receivers. We generalize the result of Section 3.2 in Section 3.3, by using the QR decomposition technique. By obtaining the statistics of the QR decomposition, we find an expression for the unconditional probability of symbol error, which includes the effect of Tomlinson-Harashima precoding, and we compare this to the
Gaussian approximation. We also present analytical results when the optimal ordering of the decentralized receivers is employed. In Section 3.5, we compare the ordered QR decomposition to various linear precoders. In Section 3.4, we present simulation and numerical results. A summary is given in Section 3.6.

3.2 Tomlinson-Harashima Precoding operating over SISO Channels with ISI

3.2.1 Modulo Operation

The modulo operation is defined as
\[ m_y(x) \triangleq x - y \left\lfloor \frac{x + y/2}{y} \right\rfloor = x - yz, \]
where \( y > 0 \), and \( z = \left\lfloor \frac{x + y/2}{y} \right\rfloor \) is the unique integer such that \( m_y(x) \in [-y/2, y/2) \). If the difference of two numbers \( x_1 \) and \( x_2 \) is integrally divisible by \( y \), then \( x_1 \) and \( x_2 \) are congruent modulo \( y \).

**Proposition 3.1**
Whenever \( i \) is an integer, \( m_y(x - yi) = m_y(x) \), where \( y > 0 \).

**Proof.**
\[
m_y(x - yi) = x - yi - y \left\lfloor \frac{x - yi + 1}{y} \right\rfloor \\
= x - yi - y \left\lfloor \frac{x}{y} + \frac{1}{2} \right\rfloor + yi \\
= m_y(x).
\]

If the input to the modulo operator is complex, then the operation is performed separately on the real and imaginary parts, so that
\[
m_y(x) = \{\text{Re}\{x\} - y \left\lfloor (\text{Re}\{x\} + y/2)/y \right\rfloor \} \\
+ j \{\text{Im}\{x\} - y \left\lfloor (\text{Im}\{x\} + y/2)/y \right\rfloor \} = x - yz,
\]
where \( \text{Re}\{x\} \) and \( \text{Im}\{x\} \) denote the real and imaginary parts of \( x \), respectively, \( z = z^{(R)} + j z^{(I)} \),
\[
z^{(R)} \triangleq \left\lfloor (\text{Re}\{x\} + y/2)/y \right\rfloor ,
\]
and
\[
z^{(I)} \triangleq \left\lfloor (\text{Im}\{x\} + y/2)/y \right\rfloor .
\]
Figure 3.1: System model for Tomlinson-Harashima precoding operating over a SISO ISI channel.

3.2.2 Channel Model

The channel is given by \( h(t) = \sum_{l=0}^{L-1} h_l \delta(t - lT) \), where \( L \) is the number of paths and \( T \) is the symbol duration. We normalize \( h_0 = 1 \), and assume the random variables \( \{h_l\} \) are complex and known at the transmitter. A quasi-static fading model is assumed, where the \( \{h_l\} \) are fixed and do not change with time in a given frame. A discrete-time representation of the channel is used by the feedback filter at the transmitter, where \( h(n) \triangleq h(t)|_{nT} = \sum_{l=0}^{L-1} h_l \delta(n - l) \).

3.2.3 Signal Model

The system model is shown in Figure 3.1. We precode the \( M \)-QAM symbol \( d(n) = d_c(n) + j d_s(n) \), where \( d_c(n) \) and \( d_s(n) \) are the in-phase and quadrature data signals, respectively, each having values in \( A = \{\pm 1, \pm 3, \ldots, \pm (\sqrt{M} - 1)\} \). We assume square constellations, where \( M = 2^b \) and \( b \) is even, so that the signal constellation is equivalent to two \( \sqrt{M} \)-PAM signals on quadrature carriers.

Knowing the channel at the transmitter, as well as the previously transmitted symbols, we can pre-subtract the ISI. When the channel adds interference to the transmitted signal, the net interference at the receiver is zero. However, since the \( \{h_l\} \) are arbitrary, Tomlinson-Harashima precoding is employed to constrain the transmitted power. This is done by using a modulo operator whose output is guaranteed to fall within the boundaries of the constellation region.

Applying the modulo operation to the data signal with the pre-subtracted ISI
where

\begin{align}
    x(n) & \triangleq m_2 \sqrt{M} \left( d(n) - \sum_{l=1}^{L-1} h_l x(n - l) \right) \\
    &= d(n) - \sum_{l=1}^{L-1} h_l x(n - l) - 2 \sqrt{M} z_x(n), 
\end{align}

(3.1)

and

\begin{align}
    z_x(n) & \triangleq z_{xR}(n) + j z_{xI}(n), \\
    z_{xR}(n) & \triangleq \left[ \text{Re}\{d(n) - \sum_{l=1}^{L-1} h_l x(n - l)\} + \sqrt{M}\right]/(2\sqrt{M}), \\
    z_{xI}(n) & \triangleq \left[ \text{Im}\{d(n) - \sum_{l=1}^{L-1} h_l x(n - l)\} + \sqrt{M}\right]/(2\sqrt{M}).
\end{align}

After power scaling to achieve the required SNR, the transmitted signal is given by

\[ s(n) = px(n). \]

The continuous-time transmitted signal is given by

\[ s(t) = \text{Re}\{\tilde{s}(t)e^{j \omega_c t}\}, \]

where \( \omega_c \) is the carrier frequency, \( \tilde{s}(t) = \sum_{m=-\infty}^{\infty} s(m)g(t - mT) \), and

\[ g(t) = \begin{cases} 
1/\sqrt{T}, & 0 \leq t < T \\
0, & \text{otherwise}
\end{cases}. \]

(3.2)

The received signal is \( r(t) = \text{Re}\{\tilde{r}(t)e^{j \omega_c t}\} \). After frequency down-conversion, the output of the matched filter, sampled at time \( (n+1)T \), is given by

\[ \hat{r}(n) = \frac{1}{\sqrt{T}} \int_{(n+1)T}^{nT} \tilde{r}(t) \, dt \]

\[ = \frac{1}{\sqrt{T}} \int_{nT}^{(n+1)T} h(t) \ast \tilde{s}(t) \, dt + \frac{1}{\sqrt{T}} \int_{nT}^{(n+1)T} n_w(t) \, dt, \]

\[ \triangleq S_T(n) \triangleq N_T(n), \]

(3.3)

where \( \tilde{r}(t) = h(t) \ast \tilde{s}(t) + n_w(t) \), and \( n_w(t) \) is a zero-mean, complex white Gaussian noise process, with two-sided spectral density \( N_0 \). The signal component in (3.3) can be expressed as

\[ S_T(n) = \frac{1}{\sqrt{T}} \sum_{l=0}^{L-1} h_l \sum_{m=-\infty}^{\infty} \int_{(n+1)T-mT-lT}^{nT-mT-lT} s(m)g(t) \, dt \]

\[ = \sum_{l=0}^{L-1} h_l s(n - l), \]

(3.4)
since the integral is non-zero only when \( m = n - l \). The contribution due to AWGN, \( \{N_T(n)\} \), is a sequence of i.i.d. zero-mean circularly symmetric complex Gaussian random variables with covariance given by \( \frac{1}{2}E[N_T(n_1)N_T^*(n_2)] = N_0 \delta_{n_1,n_2} \). Substituting (3.4) in (3.3), we have

\[
\hat{r}(n) = s(n) + \sum_{l=1}^{L-1} h_l s(n - l) + N_T(n)
= p(d(n) - 2\sqrt{M}z_x(n)) + N_T(n),
\]

where the second equality is obtained by substituting for \( s(n) \) and, in turn, for \( x(n) \), given by (3.1), and cancelling the ISI terms. After power scaling and modulo filtering, we have

\[
y(n) = m_2\sqrt{M} \left( \frac{\hat{r}(n)}{p} \right) = m_2\sqrt{M} \left( d(n) + \hat{N}_T(n) \right)
= d(n) + \hat{N}_T(n) - 2\sqrt{M}z_y(n), \tag{3.5}
\]

where the second equality is due to Proposition 3.1, the contribution due to AWGN is given by \( \hat{N}_T(n) \triangleq N_T(n)/p, z_y(n) \triangleq z^{(R)}_y(n) + jz^{(I)}_y(n), \)

\[
z^{(R)}_y(n) \triangleq \left[ (\text{Re}\{d(n) + \hat{N}_T(n)\} + \sqrt{M})/(2\sqrt{M}) \right],
\]

and

\[
z^{(I)}_y(n) \triangleq \left[ (\text{Im}\{d(n) + \hat{N}_T(n)\} + \sqrt{M})/(2\sqrt{M}) \right].
\]

Now, computing the average energy-per-symbol of the transmitted signal, we have

\[
E_{av}^{(s)} = \frac{1}{2} E \left[ \int_0^T \left| \sum_{m=-\infty}^{\infty} s(m)g(t - mT) \right|^2 dt \right]
= \frac{1}{2} E \left[ \int_0^T |s(0)|^2 g^2(t) dt \right] = \frac{p^2}{2} E \left[ |x(0)|^2 \right],
\]

where we have used the fact that \( g(t) \) is non-zero only in \([0,T)\), and consequently the integral is non-zero only when \( m = 0 \). Note that the real and imaginary parts of \( x(n) \) are independent. Motivated by the analysis in [22], they are each expected to converge to a uniform distribution in \([-\sqrt{M}, \sqrt{M})\) for large \( L \). This conjecture is confirmed by simulation results in Figures 3.2 - 3.5. We consider the empirical distribution of \( x(n) \) for various parameters as well as the normalized estimation error defined by \( f(n) \triangleq \)
Figure 3.2: Histograms and kernel densities of Re\{x(n)\} with N = 1000, M = 4, 16, and L = 2, 4.

\[ \left( \frac{1}{n} \sum_{j=1}^{n} |x(n)|^2 - \frac{2M}{3} \right) \left( \frac{2M}{3} \right)^{-1}, \quad \text{for } n = 1, \ldots, N, \] where N is the maximum sample size. In each simulation, we used a single realization of the channel.

Figures 3.2 and 3.3 show that the distribution of Re\{x(n)\} becomes more uniform for increasing M, L, and especially N. This also applies to Im\{x(n)\} although we have not included the figures for space consideration. Figures 3.4 and 3.5 illustrate f(n) as a function of n, and the figures show that, while the corresponding distributions of the previous figures may not look uniform for small values of N, M, and L, the empirical second moment still approaches the correct value even for the small values of N = 1000, L = 2, and M = 2.

Therefore, we approximate the second moment of x(n) by \( E \left[ |x(n)|^2 \right] \approx 2M/3. \) Substituting in the expression for the average energy-per-symbol, we have

\[ E_{av}^{(s)} \approx p^2 M/3. \] (3.6)

Figure 3.6 is a graphical representation of this scheme with L = 2. In order to
Figure 3.3: Histograms and kernel densities of $\text{Re}\{x(n)\}$ with $N = 10000$, $M = 4$, 16, and $L = 2, 4$. 
Figure 3.4: Normalized estimation error of the second moment of $x(n)$ with $N = 1000$, $M = 4, 16$, and $L = 2, 4$. 
Figure 3.5: Normalized estimation error of the second moment of $x(n)$ with $N = 10000$, $M = 4, 16$, and $L = 2, 4$. 
simplify the diagrammatical representation of this scheme, we assume, only in Figure 3.6, that \( h_1 \) is real, and \( d(n) \) is also real, belonging to a 2-PAM constellation.

### 3.2.4 The Effect of Tomlinson-Harashima Precoding on the Probability of Error

The test statistic (3.5) for symbol \( n \) can be written as

\[
y(n) = y_R(n) + jy_I(n),
\]

where

\[
y_R(n) = \text{Re} \{ y(n) \} = m_{2\sqrt{M}} \{ \text{Re} \{ d(n) + \hat{N}_T(n) \} \}
\]

\[
y_I(n) = \text{Im} \{ y(n) \} = m_{2\sqrt{M}} \{ \text{Im} \{ d(n) + \hat{N}_T(n) \} \}
\]

and we have used the fact that the modulo operator operates separately on the real and imaginary parts of its input. The contributions due to AWGN, for symbol \( n \), are written as

\[
\hat{N}_T(R)(n) = \text{Re} \{ \hat{N}_T(n) \} \quad \text{and} \quad \hat{N}_T(I)(n) = \text{Im} \{ \hat{N}_T(n) \}.
\]

Since \{\hat{N}_T(n)\} is a sequence of i.i.d. zero-mean circularly symmetric complex Gaussian random variables, each with variance \( N_0/p^2 \), then the \{\hat{N}_T(R)(n)\} and \{\hat{N}_T(I)(n)\} are both sequences of i.i.d. zero-mean real Gaussian random variables with variance \( N_0/p^2 \). This is due to the fact that the powers of the in-phase and quadrature components are equal to the power of the original bandpass process. Then

\[
\text{Var} \left[ \hat{N}_T(R)(n) \right] = N_0/p^2 \triangleq \sigma^2,
\]

and

\[
\text{Var} \left[ \hat{N}_T(I)(n) \right] = \sigma^2.
\]

Also, the in-phase and quadrature noise processes are independent.

Since the signals in the phase-quadrature components given by (3.7) and (3.8) are perfectly separated by coherent detection, then the probability that the \( M \)-QAM symbol is correct is equal to the product of the probabilities that the corresponding \( \sqrt{M} \)-PAM symbols are correct. Since the test statistics for the phase-quadrature components are identically distributed, the following derivation for the in-phase component is also valid for the quadrature component.

Let \( \hat{d}_c(n) \) be an estimate of \( d_c(n) \), which is given by

\[
\hat{d}_c(n) = q \quad \text{if} \quad q - 1 \leq y^{(R)}(n) < q + 1,
\]
The 2-PAM constellation and its boundary region is shown below.

The constellation is replicated throughout the entire real line.

We select \( d(n) = 1 \) (represented by the square) and add the feedback term \(-h_1d(n-1)\).

We then apply the modulo operation to restrict the transmitted signal to be inside the boundary region of the original constellation.

Ignoring power scaling and AWGN, the output of the channel is

This is the congruent point to the desired point. Applying the modulo operation gives a perfect estimate of the transmitted data (represented by the triangle).

Figure 3.6: Graphical representation of Tomlinson-Harashima precoding operating over a SISO channel with ISI.
for $q = \pm 1, \ldots, \pm(\sqrt{M} - 1)$. The probability of a correct decision when $d_c(n) = q$, assuming equally likely transmitted symbols, is

$$P_c = \frac{1}{\sqrt{M}} \sum_{q \in A} Pr\{\hat{d}_c(n) = q | d_c(n) = q\}. \quad (3.11)$$

When $\hat{d}_c(n) = q$ then, by (3.10), $y^{(R)}(n) \in [q - 1, q + 1)$, which gives

$$Pr\{\hat{d}_c(n) = q | d_c(n) = q\} = Pr\{q - 1 \leq y^{(R)}(n) < q + 1 | d_c(n) = q\} = Pr\{q - 1 \leq m_2\sqrt{M}(q + \hat{N}_T^{(R)}(n)) < q + 1 | d_c(n) = q\},$$

where we have substituted for $y^{(R)}(n)$ using (3.7). Equation (3.12) specifies the conditional probability of the received signal falling in the decision region of all points congruent to $q$. Recall that a point $y$ in a replicated constellation is congruent to another point $x$ in the original constellation if $m_2\sqrt{M}(y) = x$. This implies that, for a correct decision when $d_c(n) = q$, the AWGN must be contained within the periodic interval $[2\sqrt{Mi} - q - 1, 2\sqrt{Mi} - q + 1)$, where $i \in \mathbb{Z}$. This argument is described in [61]. We provide explicit derivations as follows: From Equation (3.12), we expand $m_2\sqrt{M}(q + \hat{N}_T^{(R)}(n))$ by writing

$$m_2\sqrt{M}(q + \hat{N}_T^{(R)}(n)) = q + \hat{N}_T^{(R)}(n) - 2\sqrt{Mi} \quad (3.13)$$

if $2\sqrt{Mi} - \sqrt{M} \leq q + \hat{N}_T^{(R)}(n) < 2\sqrt{Mi} + \sqrt{M}$, for $i \in \mathbb{Z}$. We note the domains are disjoint. Using (3.13), we have

$$\{q - 1 \leq m_2\sqrt{M}(q + \hat{N}_T^{(R)}(n)) < q + 1\} = \bigcup_{i=-\infty}^{\infty} \{S_1^{(i)} \cap S_2^{(i)}\},$$

where $S_1^{(i)} \equiv \{2\sqrt{Mi} - 1 \leq \hat{N}_T^{(R)}(n) < 2\sqrt{Mi} + 1\}$ and

$$S_2^{(i)} \equiv \{2\sqrt{Mi} - \sqrt{M} - q \leq \hat{N}_T^{(R)}(n) < 2\sqrt{Mi} + \sqrt{M} - q\}.$$ 

Since for every $q = \pm 1, \ldots, \pm(\sqrt{M} - 1)$, the set $S_1^{(i)}$ is contained in the set $S_2^{(i)}$, then

$$\{q - 1 \leq m_2\sqrt{M}(q + \hat{N}_T^{(R)}(n)) < q + 1\} = \bigcup_{i=-\infty}^{\infty} \{2\sqrt{Mi} - 1 \leq \hat{N}_T^{(R)}(n) < 2\sqrt{Mi} + 1\}.$$
Then the probability of the union of disjoint events is given by

$$Pr\{d_c(n) = q \mid d_c(n) = q\} = \sum_{i=-\infty}^{\infty} Pr\left\{2\sqrt{M}i - 1 \leq N_T^{(R)}(n) < 2\sqrt{M}i + 1\right\} ,$$

(3.14)

which is independent of $d_c(n) = q$. Substituting (3.14) into (3.11) gives

$$P_c = \sum_{i=-\infty}^{\infty} Pr\left\{2\sqrt{M}i - 1 < N_T^{(R)}(n) < 2\sqrt{M}i + 1\right\}$$

$$= \sum_{i=-\infty}^{\infty} \left[2\Phi\left((2\sqrt{M}i + 1)/\sigma\right) - 1\right] ,$$

where $\Phi(x) \triangleq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt$, and we have used the fact that $\Phi(-x) = 1 - \Phi(x)$.

Finally, since the probability of a correct decision for the quadrature component is equal to the probability of a correct decision for the in-phase component, the probability of symbol error for $M$-QAM, including the effect of Tomlinson-Harashima precoding, is given by

$$P_s^{(TH)}(E) = 1 - \left\{\sum_{i=-\infty}^{\infty} \left[2\Phi\left((2\sqrt{M}i + 1)/\sqrt{p^2/N_0}\right) - 1\right]\right\}^2 ,$$

(3.15)

where we have substituted for $\sigma$, which is defined in (3.9). Let $\gamma_b \triangleq E_{av}^{(b)}/N_0$, where $E_{av}^{(b)}$ is the average transmitted energy-per-bit. Then, using (3.6) and the fact that $E_{av}^{(s)} = (\log_2 M)E_{av}^{(b)}$, the probability of symbol error can be expressed as

$$P_s^{(TH)}(E) = 1 - \left\{\sum_{i=-\infty}^{\infty} \left[2\Phi\left((2\sqrt{M}i + 1)/\sqrt{3(\log_2 M)/M\gamma_b}\right) - 1\right]\right\}^2 .$$

(3.16)

### 3.2.5 Gaussian Approximation

Ignoring the effect of the modulo operation at the receiver, (3.7) and (3.8) are equal to $y^{(R)}(n) = d_c(n) + N_T^{(R)}(n)$ and $y^{(I)}(n) = d_s(n) + N_T^{(I)}(n)$. This simplifies the computation of the probability of error, since now $y(n)$ is a conditional Gaussian random variable. This Gaussian approximation is valid when $\gamma_b$ is large and, consequently, when the thermal noise component of the test statistic is less likely to effect the modulo operation.

Since the phase-quadrature components are perfectly separated by coherent detection, the probability of error for $M$-QAM is determined from the probability of error
of $\sqrt{M}$-PAM [55]. Since the distance between adjacent signal points is 2, and an error occurs in only one direction when either one of the two outside points is transmitted, then the probability of symbol error for the one-dimensional $\sqrt{M}$-PAM constellation is given by

$$P_{\sqrt{M}} = \frac{\sqrt{M} - 1}{\sqrt{M}} \Pr \left\{ |N_T(n)| > 1 \right\}$$

$$= 2 \frac{\sqrt{M} - 1}{\sqrt{M}} \Phi \left( -\frac{1}{\sigma} \right) = 2 \frac{\sqrt{M} - 1}{\sqrt{M}} \Phi \left( -\sqrt{\frac{P^2}{N_0}} \right).$$

The probability of a correct decision for $M$-QAM is given by $P_c = (1 - P_{\sqrt{M}})^2$, and the probability of symbol error for $M$-QAM, expressed in terms of $\gamma_b$, is given by

$$P_s^{(GA)}(E) = 1 - P_c = 1 - \left( 1 - 2 \frac{\sqrt{M} - 1}{\sqrt{M}} \Phi \left( -\sqrt{\frac{3(\log_2 M)}{M \gamma_b}} \right) \right)^2. \quad (3.17)$$

3.3 Tomlinson-Harashima Precoding Operating over MISO Frequency Selective Fading Channels with Decentralized Receivers

Let the number of transmit antennas at the transmitter be equal to $N_T$, the number of users be equal to $N_U$, and the number of receive antennas at each user be equal to one. In addition to ISI, the signals intended for different users interfere with each other. This is the motivation behind using the QR decomposition of the channel matrix to arrive at an equivalent channel where successive interference cancellation can be used to remove the effect of the multiuser interference in addition to removing the ISI.

3.3.1 Channel Model

The channel between the $i^{th}$ transmit antenna and the receive antenna at user $j$ is given by $h_{j,i}(t) = \sum_{l=0}^{L-1} h_{j,i}^{(l)} \delta(t - lT)$, where $L$ is the number of paths between a given transmit and a given receive antenna, $T$ is the symbol duration, and $h_{j,i}^{(l)}$ is the channel coefficient of the $l^{th}$ path. The coefficients $\{h_{j,i}^{(l)}\}$ are known at the transmitter and are realizations of i.i.d. zero-mean circularly symmetric complex Gaussian random variables.
with variance

\[ E[|h_{j,i}^{(l)}|^2] = \frac{1}{L} \triangleq \sigma_{h_{j,i}}^2, \quad \forall \, i, j, \text{ and } l. \quad (3.18) \]

The coefficients of the \( l^{th} \) path between the \( N_T \) transmitting antennas and the \( N_U \) receiving users can be represented by an \( N_U \times N_T \) matrix \( H^{(l)} \), where \( [H^{(l)}]_{j,i} = h_{j,i}^{(l)} \), \( i = 1, \ldots, N_T, j = 1, \ldots, N_U \).

### 3.3.2 Signal Model

The QR decomposition is given by \( (H^{(0)})^H = QR \), where \((\cdot)^H\) is the Hermitian operator, \( Q \) is an \( N_T \times N_U \) matrix such that \( QQ^H = I \), and \( R \) is an \( N_U \times N_U \) upper-right triangular matrix. The entries of \( H^{(0)} \) are random variables. Consequently, \( Q \) and \( R \) are also random matrices with statistics determined in \([62], [63, Lemma 2.1]\). The channel matrix \( H^{(0)} \) is assumed known at the transmitter, as are the matrices \( Q \) and \( R \), since the QR decomposition is performed at the transmitter. The signal to be transmitted is precoded with the matrix \( W = QA \), where \( A \) is a diagonal matrix with entries \( [A]_{i,i} = 1/r_{i,i}, \quad i = 1, \ldots, N_U \). The \( \{r_{i,i}\} \) are the diagonal entries of the matrix \( R \), and can be shown to be real and positive \([62], [63, Lemma 2.1]\). Therefore, the effective channel is given by \( H^{(0)}W = (R^HQ^H)(QA) = R^HA \), which is a lower triangular matrix with unit diagonal. As a result, user \( j \) sees interference due only to users \( 1, \ldots, j-1 \). In order for the effective channel matrix, \( H^{(0)}W \), to be of full rank \( N_U \), it is necessary that \( N_T \geq N_U \).

Successive interference cancellation is implemented with a feedback filter given by

\[ B = [I - H^{(0)}W, -H^{(1)}W, -H^{(2)}W, \ldots, -H^{(L-1)}W], \quad (3.19) \]

where the matrix \( I - H^{(0)}W \) is used to cancel the interference due to the other users in the current symbol interval, and \( -H^{(1)}W, \ldots, -H^{(L-1)}W \) are used to cancel the remaining interference.

The output of the modulo operators for the \( n^{th} \) symbol vector is given by

\[ x(n) = m_{2\sqrt{M}} (d(n) + B\hat{x}(n)) = d(n) + B\hat{x}(n) - 2\sqrt{M}z_x(n), \quad (3.20) \]
where the modulo operator operates on each element of the vector in its argument, \( \mathbf{x}(n) \) is the \( N_U \times 1 \) vector at the output of the modulo operator, 

\[
\mathbf{\hat{x}}(n) = \left[ \mathbf{x}(n)^T, \mathbf{x}(n-1)^T, \ldots, \mathbf{x}(n-(L-1))^T \right]^T,
\]

\((\cdot)^T\) is the transpose operator, \( \mathbf{d}(n) \) is the \( N_U \times 1 \) data vector,

\[
\mathbf{z}_x(n) = \left[ z_{x,1}(n), \ldots, z_{x,N_U}(n) \right]^T,
\]

where 

\[
z_{x,j}(n) \equiv z_{x,j}^{(R)}(n) + \mathrm{j} z_{x,j}^{(I)}(n),
\]

\[
z_{x,j}^{(R)}(n) \equiv \left[ (\text{Re}\left\{ [\mathbf{d}(n) + \mathbf{B}\hat{x}(n)]_j \right\} + \sqrt{M})/(2\sqrt{M}) \right],
\]

\[
z_{x,j}^{(I)}(n) \equiv \left[ (\text{Im}\left\{ [\mathbf{d}(n) + \mathbf{B}\hat{x}(n)]_j \right\} + \sqrt{M})/(2\sqrt{M}) \right],
\]

and \( [\mathbf{v}]_j \) is the \( j^{th} \) element of \( \mathbf{v} \).

The \( j^{th} \) element of the data vector is given by \( d_j(n) = d_{c,j}(n) + j d_{s,j}(n) \), where \( d_{c,j}(n) \) and \( d_{s,j}(n) \) are the in-phase and quadrature data signals, respectively, with the same properties as defined in Section 3.2.3.

The system diagram is shown in Figure 3.7. Before multiplying by \( \mathbf{W} \), the signals are multiplied by a power scaling matrix \( \mathbf{P} \), used to satisfy the SNR requirement. If all users require the same output SNR from their respective de-modulator, then \( \mathbf{P} = p \mathbf{I} \), where \( p \) is a constant. Then, if \( \mathbf{s}(n) \) is the vector of transmitted signals, and \( \mathbf{x}(n) \) is the vector of precoded data symbols, \( \mathbf{s}(n) = \mathbf{W}\mathbf{P}\mathbf{x}(n) \).
The low-pass equivalent transmitted signal vector is given by
\[
\tilde{s}(t) = \sum_{m=-\infty}^{\infty} s(m)g(t-mT),
\]
where \(g(t)\) is defined in (3.2). The low-pass equivalent received signal vector, with the \(j^{th}\) element corresponding to the received signal at the \(j^{th}\) user, is given by
\[
\tilde{r}(t) = \sum_{l=0}^{L-1} H^{(l)}\tilde{s}(t-lT) + n_w(t),
\]
where \(n_w(t)\) is a \(N_U \times 1\) noise vector with independent elements. Each element is a zero-mean complex white Gaussian noise process with two-sided spectral density \(N_0\).

The output of the matched filters, sampled at time \((n+1)T\), is given by
\[
\hat{r}(n) = \frac{1}{\sqrt{T}} \int_{nT}^{(n+1)T} \tilde{r}(t) \, dt = \sum_{l=0}^{L-1} H^{(l)}s(n-l) + N_T(n),
\]
where \(N_T(n) \triangleq \frac{1}{\sqrt{T}} \int_{nT}^{(n+1)T} n_w(t) \, dt\). Multiplying (3.21) by \(P^{-1} = (1/p)I\) gives
\[
P^{-1}\hat{r}(n) = H^{(0)}Wx(n) + \sum_{l=1}^{L-1} H^{(l)}Wx(n-l) + \hat{N}_T(n)
= x(n) + (H^{(0)}W-I)x(n) + \sum_{l=1}^{L-1} H^{(l)}Wx(n-l) + \hat{N}_T(n),
\]
where we used the fact that \(s(n) = WPx(n)\), we added and subtracted the term \(x(n)\), and we defined \(\hat{N}_T(n) \triangleq P^{-1}N_T(n)\). Substituting for \(x(n)\) and \(B\), defined in (3.20) and (3.19), respectively, it is possible to show that
\[
P^{-1}\hat{r}(n) = d(n) + \hat{N}_T(n) - 2\sqrt{M}z_{y,j}(n).
\]
From (3.22), it can be seen that both the interference due to the other users and the ISI are cancelled perfectly. The vector of test statistics for the \(n^{th}\) symbol vector is given by
\[
y(n) = m_{2\sqrt{M}}(P^{-1}\hat{r}(n)).
\]
The test statistic for user \(j\) and symbol \(n\) is given by
\[
y_j(n) = m_{2\sqrt{M}}(\frac{\hat{r}_j(n)}{p}) = m_{2\sqrt{M}}(d_j(n) + \hat{N}_{T,j}(n))
= d_j(n) + \hat{N}_{T,j}(n) - 2\sqrt{M}z_{y,j}(n),
\]
where the second equality is due to Proposition 3.1, \(z_{y,j}(n) = z_{y,j}^{(R)}(n) + j z_{y,j}^{(I)}(n)\),
\[
z_{y,j}^{(R)}(n) \triangleq \left[\text{Re}\{d_j(n) + \hat{N}_{T,j}(n)\} + \sqrt{M}\right]/(2\sqrt{M})
\]
\[
z_{y,j}^{(I)}(n) \triangleq \left[\text{Im}\{d_j(n) + \hat{N}_{T,j}(n)\} + \sqrt{M}\right]/(2\sqrt{M})
\]
\[ z_{y,j}^{(I)}(n) \triangleq \left[ \operatorname{Im}(d_j(n) + \hat{N}_{T,j}(n)) + \sqrt{M}\right]/(2\sqrt{M}) \]

and \( \{\hat{N}_{T,j}(n)\} \triangleq \{N_{T,j}(n)/p\} \) is a sequence of i.i.d. zero-mean circularly symmetric complex Gaussian random variables with covariance \( \frac{1}{p}E[\hat{N}_{T,j}(n_1)\hat{N}_{T,j}(n_2)] = \frac{N_0}{p^2}\delta_{n_1,n_2} \).

Note that the \( \{y_j(n)\} \) are i.i.d. for different \( n \). Since the \( \{d_j(n)\} \) and the \( \{\hat{N}_{T,j}(n)\} \) are both i.i.d. for different \( j \) and fixed \( n \), by (3.23), the \( \{y_j(n)\} \) are also i.i.d. for different \( j \) and fixed \( n \).

The total average transmit energy is given by

\[
E_T = \frac{1}{2}E \left[ \operatorname{tr}(s(n)s^H(n)) \right] = \frac{p^2}{2}E \left[ \operatorname{tr}(x(n)x^H(n)W^H W) \right] = \frac{p^2}{2}E \left[ \operatorname{tr}(x(n)x^H(n)A^H A) \right],
\]

where we have used the fact that \( s(n) = WPx(n) \), \( P = pI \), \( W = QA \), the trace of a product of two square matrices is independent of the order of the multiplication, and \( Q \) is a unitary matrix. By matrix multiplication, it is possible to show that 
\[
\operatorname{tr}(x(n)x^H(n)A^H A) = \sum_{j=1}^{N_U} |x_j(n)|^2/r_{j,j}^2,
\]
so that the average transmitted energy-per-symbol is given by

\[
E_{av}^{(s)} = E_T/N_U = \left( \frac{p^2}{2N_U} \right) \sum_{j=1}^{N_U} E[|x_j(n)|^2]/r_{j,j}^2.
\]

(3.24)

### 3.3.3 The Effect of Tomlinson-Harashima Precoding on the Probability of Error

The probability of error is computed from the test statistic given in (3.23), which, for fixed \( j \) and \( p \), has the same distribution as the test statistic given in (3.5). Therefore, the probability of symbol error, for fixed \( j \) and \( p \), is given in (3.15), and can be written in terms of the average transmitted energy-per-symbol, given in (3.24). For \( j = 1 \) and \( L = 1 \), \( x_1(n) = d_1(n) \), so that \( E[|x_1(n)|^2] = E[|d_1(n)|^2] = 2(M - 1)/3 \), since the data symbols are assumed to be equally likely and are derived from a discrete distribution.

For \( j = 1 \) and \( L > 1 \), and for \( j = 2, \ldots, N_U \), and any \( L \), we approximate the second moment of \( x_j(n) \) by \( E[|x_j(n)|^2] \approx 2M/3 \). This is motivated by the fact that the real and imaginary parts of \( x_j(n) \) are independent and, by the analysis in [22], are
Figure 3.8: Normalized estimation error of the second moment of $x_k(n)$ with $L = 1$, $M = 4$, $N_T = N_U = 4$, and $N = 1000$.

Each expected to converge to a uniform distribution in $[-\sqrt{M}, \sqrt{M})$ for large $j$ or $L$. This approximation of the second moment is also valid for small $j$ or $L$, as can be seen by simulation results of the normalized estimation error defined by $f_j(n) \triangleq \left( \frac{1}{n} \sum_{j=1}^{n} |x_j(n)|^2 - \frac{2M}{3} \right)^{-1}$, for $n = 1, \ldots, N$, where $N$ is the maximum sample size. Figure 3.8 illustrates $f_j(n)$ as a function of $n$, for $j = 2$ and $N_U$, with $L = 1$, $M = 4$, $N_T = N_U = 4$, and $N = 1000$. Even for $j = 2$, the normalized estimation error approaches zero for a relatively small sample size of $n = 100$.

In summary, we approximate the second moment of $x_j(n)$, when $j = 1$, by

$$E[|x_1(n)|^2] \approx 2(M - \delta_{L,1})/3,$$

where the approximation is exact when $L = 1$, and we approximate the second moment of $x_j(n)$, when $j = 2, \ldots, N_U$, by

$$E[|x_j(n)|^2] \approx 2M/3.$$

Substituting the approximation for the second moment of $x_j(n)$ in (3.24), we have

$$E_{av}^{(s)} = p^2 f_{M,L,N_U}(\{r_{j,j}^2\}), \quad (3.25)$$

where, for convenience, we define

$$f_{M,L,N_U}(\{r_{j,j}^2\}) \triangleq \frac{1}{3N_U} \left[ \frac{(M - \delta_{L,1})}{r_{1,1}^2} + \sum_{j=2}^{N_U} \frac{M}{r_{j,j}^2} \right].$$
Solving for \( p^2 \), substituting into the probability of symbol error in (3.15), and using the fact that \( E_{av}^{(a)} = (\log_2 M) E_{av}^{(b)} \) and \( \gamma_b \equiv E_{av}^{(b)} / N_0 \), the conditional probability of symbol error for each user, conditioned on the \( \{r_{j,j}\} \), is given by

\[
P_s^{(TH)}(E|\{r_{j,j}\}) = 1 - \left\{ \sum_{i=-\infty}^{\infty} \left[ 2\Phi\left( (2\sqrt{Mi} + 1) \sqrt{\frac{\log_2 M \gamma_b}{f_{M,L,N_U}(\{r_{j,j}\})}} - 1 \right) \right] \right\}^2.
\]

### 3.3.4 Gaussian Approximation

The Gaussian approximation, which we used to derive the probability of symbol error for a SISO channel in Section 3.2.5, and which is employed in [35], can be used to approximate the test statistic in (3.23) as a conditional Gaussian random variable. Then, from the average transmitted energy-per-symbol in (3.24), the conditional probability of error for each user, conditioned on the \( \{r_{j,j}\} \), is given by

\[
P_s^{(GA)}(E|\{r_{j,j}\}) = 1 - \left( 1 - 2\sqrt{M - 1} \frac{1}{\sqrt{M}} \Phi\left( -\sqrt{\frac{(\log_2 M \gamma_b)}{f_{M,L,N_U}(\{r_{j,j}\})}} \right) \right)^2. \tag{3.26}
\]

### 3.3.5 Unconditional Probability of Error

From [62], [63, Lemma 2.1], it is possible to show that the diagonal elements of the matrix \( \mathbf{R} \) are independent and the pdf of \( v_j \equiv r_{j,j}^2 \), \( j = 1, \ldots, N_U \), is given by

\[
f_{v_j}(t_j) = \begin{cases} 
\frac{t_j^{N_T-j} \exp(-t_j/\sigma_h^2)}{(\sigma_h^2)^{N_T-j+1} \Gamma(N_T-j+1)} & t_j \geq 0, \\
0 & t_j < 0,
\end{cases}
\]

for \( j \in 1, \ldots, N_U \).

The conditional probability of symbol error, given by (3.26), is averaged over the statistics of the \( \{r_{j,j}\} \). This gives the unconditional probability of symbol error, which includes the effect of the modulo operations, and is given by

\[
P_s^{(TH)}(E) = \int_{0}^{\infty} \cdots \int_{0}^{\infty} \prod_{j_1=1}^{N_U} \frac{N_{T-j_1+1} \Gamma(N_{T-j_1})}{u_{j_1}} e^{-Lu_{j_1}} d u_{j_1} \\
\cdot \left( 1 - \left\{ \sum_{i=-\infty}^{\infty} \left[ 2\Phi\left( (2\sqrt{Mi} + 1) \sqrt{\frac{\log_2 M \gamma_b}{f_{M,L,N_U}(\{u_j\})}} - 1 \right) \right] \right\}^2 \right). \tag{3.28}
\]
3.3.6 Optimal Ordering of Decentralized Receivers

As indicated in [35], the ordering of the decentralized receivers affects the construction of the channel matrix $H^{(0)}$. Since there are $N_U$ receivers, there exist $N_U!$ possible column permutations of the matrix $(H^{(0)})^H$, and one QR decomposition associated with each permutation. This affects the construction of the matrix $A$, which in turn affects the total transmit energy. In order to minimize the total transmit energy, it is necessary to search over all possible column permutations of $(H^{(0)})^H$. Note that this search can be simplified by the methods described in [64] and the references therein. Here, we derive the optimal statistics associated with the column permutation of $(H^{(0)})^H$ that minimizes the total transmit energy.

Let $M \triangleq (H^{(0)})^H$. We denote each column permutation of $M$ by the index $i = 1, \ldots, N_U!$, and the matrix corresponding to the $i^{th}$ permutation by $M^{(i)} = Q^{(i)}R^{(i)}$. We write the diagonal elements of $R^{(i)}$ as $r^{(i)}_{1,1}, \ldots, r^{(i)}_{N_U,N_U}$, and rewrite (3.25) as a function of $i$, giving

$$E^{(i)} \triangleq \frac{p^2}{3N_U} \left[ (M - \delta_{L,1}) \left( r^{(i)}_{1,1} \right)^{-2} + \sum_{j=2}^{N_U} M \left( r^{(i)}_{j,j} \right)^{-2} \right]. \quad (3.29)$$

We select the index $i = i^*$ that minimizes the simpler expression given by $c^{(i)} \triangleq \sum_{j=1}^{N_U} \left( r^{(i)}_{j,j} \right)^{-2}$. That is, we desire to find the joint density of the squared diagonal elements of $R^{(i^*)}$, namely $\left( r^{(i^*)}_{1,1} \right)^2, \ldots, \left( r^{(i^*)}_{N_U,N_U} \right)^2$, where $i^* = \arg\min_{i=1,\ldots,N_U!} c^{(i)}$, and $R^{(i^*)}$ is given by the QR decomposition of $M^{(i^*)} = Q^{(i^*)}R^{(i^*)}$.

The random variables $\left\{ \left( r^{(i^*)}_{j,j} \right)^2 \right\}_{j=1}^{N_U}$ cannot be explicitly written as a function of the elements of the matrix $M$. The optimal permutation index, $i^*$, is, in fact, itself a random variable. We proceed by conditioning on $i^*$. Note $i^*$ can be written as $i^* = i \iff \bigcap_{j=1}^{N_U!} \{ c^{(i)} < c^{(j)} \} \triangleq S_i$, where $i \in \{1, \ldots, N_U!\}$. Then the joint distribution of the
\[
\left\{ \left( r_{j,j}^{(i^*)} \right)^2 \right\}_{j=1}^{NU} \text{ is given by}
\]
\[
Pr \left\{ \left\{ \left( r_{j,j}^{(i^*)} \right)^2 \leq x_j \right\}_{j=1}^{NU} \right\}_{i=1}^{NU!} = \bigcup_{i=1}^{NU!} Pr \left\{ \left\{ \left( r_{j,j}^{(i^*)} \right)^2 \leq x_j \right\}_{j=1}^{NU} \mid i^* = i \right\} Pr \{ i^* = i \}
\]
\[
= \bigcup_{i=1}^{NU!} Pr \left\{ \left\{ \left( r_{j,j}^{(i)} \right)^2 \leq x_j \right\}_{j=1}^{NU} \mid S_i \right\} Pr \{ S_i \}
\]
\[
= \bigcup_{i=1}^{NU!} Pr \left\{ \left\{ \left( r_{j,j}^{(i)} \right)^2 \leq x_j \right\}_{j=1}^{NU} \cap S_i \right\}.
\]

Since the \( \left\{ r_{j,j}^{(i)} \right\}_{j=1}^{NU} \) are identically distributed for fixed \( j \) and different \( i \), then
\[
Pr \left\{ \left\{ \left( r_{j,j}^{(i)} \right)^2 \leq x_j \right\}_{j=1}^{NU} \cap S_i \right\} = Pr \left\{ \left\{ \left( r_{j,j}^{(1)} \right)^2 \leq x_j \right\}_{j=1}^{NU} \cap S_1 \right\}, \forall i = 1, \ldots, NU!.
\]
(3.30)

Substituting (3.30) in (3.30), we have
\[
Pr \left\{ \left\{ \left( r_{j,j}^{(i^*)} \right)^2 \leq x_j \right\}_{j=1}^{NU} \right\} = Pr \left\{ \left\{ \left( r_{j,j}^{(1)} \right)^2 \leq x_j \right\}_{j=1}^{NU} \cap S_1 \right\} .
\]
(3.31)

Although the \( \left\{ r_{j,j}^{(1)} \right\}_{j=1}^{NU} \) are a function of the \( NT \times NU \) elements of the matrix \( M^{(1)} \) corresponding to permutation \( i = 1 \), we proceed by more simply expressing the \( \left\{ r_{j,j}^{(1)} \right\}_{j=1}^{NU} \) as a function of the \( NU(NU+1)/2 \) inner-products of the column vectors of \( M^{(1)} \), which are given by \( m_j = [m_{1,j}, \ldots, m_{NT,j}]^T, j = 1, \ldots, NU \). This simplification allows the computation of (3.31) by integrating over the joint density of the inner-products.

The QR decomposition \( M^{(1)} = Q^{(1)}R^{(1)} \) can be computed by the Gram-Schmidt orthogonalization procedure. Let \( u_1 = m_1, u_j = m_j - \sum_{k=1}^{j-1} (e_k^H m_j) e_k, j = 2, \ldots, NU \), where \( e_j = u_j/|u_j|, j = 1, \ldots, NU \), \( e_k^H m_j \) is the projection of \( m_j \) onto \( e_k \), and \( |u_j| = \sqrt{\sum_{i=1}^{NT} |u_{i,j}|^2} \) is the \( l^2 \)-norm of vector \( u_j \). Alternatively, we can write \( m_1 = u_1, m_j = u_j + \sum_{k=1}^{j-1} (e_k^H m_j) e_k, j = 2, \ldots, NU \). The \( \{m_j\} \) can be written in matrix
form as

\[
M^{(1)} = [m_1, \ldots, m_{NU}]
\]

\[
= [e_1, \ldots, e_{NU}] \Delta Q^{(1)} = [u_1 \mid (e_1^H m_2) \ (e_1^H m_3) \ \cdots \mid 0 \ | u_2 \mid (e_2^H m_3) \ \cdots \mid 0 \ 0 \ | u_3 \mid \ \cdots \mid \cdots \mid \cdots ] \Delta R^{(1)}.
\]

Therefore, the diagonal elements of the upper-right triangular matrix, \( R^{(1)} \), are given by

\[
 r^{(1)}_{j,j} = |u_j| = \left[ |m_j|^2 - \sum_{k=1}^{j-1} |e_k^H m_j|^2 \right]^{1/2},
\]

for \( j = 1, \ldots, NU \), and where we have used the fact that the \( \{e_k\} \) are orthonormal. Note that the \( \{r^{(1)}_{j,j}\}_{j=1}^{NU} \) are a function only of the inner products of the column vectors of \( M^{(1)} \). When the entries of \( M^{(1)} \) are Gaussian distributed, the joint distribution of the inner products of the column vectors of \( M^{(1)} \) is given by the Wishart distribution.

Define \( W \triangleq (M^{(1)})^H M^{(1)} \), where the entries of \( W \) are the inner products of the column vectors of \( M^{(1)} \). Note that \( w_{i,j} = w^{(R)}_{i,j} \) for \( i < j \), \( i = 1, \ldots, NU \), \( j = 1, \ldots, NU \). The off-diagonal elements can be written as \( w^{(R)}_{i,j} = w^{(I)}_{i,j} \) for \( i \neq j \), \( i = 1, \ldots, NU \), \( j = 1, \ldots, NU \), where \( w^{(R)}_{i,j} \triangleq \text{Re}\{w_{i,j}\} \) and \( w^{(I)}_{i,j} \triangleq \text{Im}\{w_{i,j}\} \). Since the entries of \( M^{(1)} \) are complex circularly symmetric i.i.d. zero-mean Gaussian random variables each with variance \( \sigma_h^2 \), then the \( \{w_{i,i} \cup \{w^{(R)}_{i,j} , w^{(I)}_{i,j} , i < j \} \} \) are Wishart distributed with joint density given by [65]

\[
 f_{w_{1,1}, w^{(R)}_{1,i}, w^{(I)}_{1,i}, w_{NU,NU}} (a_{1,1}, a^{(R)}_{1,2}, a^{(I)}_{1,2}, \ldots, a_{NU,NU}) = \frac{|A|^{NT-NU} \exp \left( -\text{tr} \left( \Sigma^{-1} A \right) \right)}{\pi^{NU(NU-1)/2} |\Sigma|^{NT} \prod_{i=1}^{NU} \Gamma(N_T + 1 - i)}, \quad A \succeq 0,
\]

where \( \Sigma = \sigma_h^2 I \), \( |A| = \det(A) \), and \( A \succeq 0 \) denotes \( A \) is Hermitian positive semi-definite.

Subsequently, (3.31) can be obtained by integrating the Wishart density, given by (3.33), over the region described by \( \left\{ (r^{(1)}_{j,j})^2 \leq x_j \right\}_{j=1}^{NU} \cap S_1 \). In order to compute
Figure 3.9: The probability of symbol error as a function of the SNR-per-bit for Tomlinson-Harashima precoding operating over a SISO channel with ISI, and with \( M = 4 \) and \( L = 2 \).

The joint density of the \( \left\{ \left( r_{j,j}^{(i)} \right)^2 \right\}_{j=1}^{N_U} \), we must take the partial derivative of the joint distribution with respect to the \( \{ x_j \} \). This is complicated by the fact that the region of integration is a function of the \( \{ x_j \} \), and that the limits of integration may not easily be expressed as an explicit function of these variables. Therefore, in Appendix F, we solve for the joint density of the \( \left\{ \left( r_{j,j}^{(i)} \right)^2 \right\}_{j=1}^{N_U} \) when \( N_T = N_U = 2 \), giving

\[
\frac{\partial}{\partial x_2 \partial x_1} P\left\{ \left( r_{1,1}^{(i)} \right)^2 \leq x_1, \left( r_{1,2}^{(i)} \right)^2 \leq x_2 \right\} = \frac{2x_1}{\sigma_h^4} e^{-\frac{1}{\sigma_h^2} (x_1 + \max(x_1, x_2))}, \tag{3.34}
\]

for any \( x_1 > 0 \) and \( x_2 > 0 \). To obtain the unconditional probability of symbol error when the optimal ordering of the decentralized receivers is employed with \( N_T = N_U = 2 \), the conditional probability of symbol error, given by (3.26), is averaged over (3.34), giving

\[
P_s^{(TH, \text{opt})}(E) = \int_0^\infty \int_0^\infty 2L^2 u_1 e^{-L(u_1 + \max(u_1, u_2))} P_s^{(TH)}(E|u_1, u_2) \, du_1 \, du_2. \tag{3.35}
\]

### 3.4 Numerical and Simulation Results

Figure 3.9 illustrates the probability of symbol error as a function of the SNR-per-bit for the system in Section 3.2, where we consider a SISO channel with \( L \) paths. The simulation results coincide with the analytical results, which are derived from the
Figure 3.10: The probability of symbol error as a function of the SNR-per-bit with $N_T = N_U = 2$, $M = 4$, and $L = 1$ and 2.

Figure 3.11: The probability of symbol error as a function of the SNR-per-bit with $N_T = N_U = 4$, $M = 4$, and $L = 1$ and 2.
Figure 3.12: The probability of symbol error as a function of the SNR-per-bit with $N_U = 2$, $M = 4$, $L = 1$, and $N_T = 2, 3, \text{and } 4$.

Figure 3.13: The probability of symbol error as a function of the SNR-per-bit for $N_T = N_U = 2$, $M = 4$, and $L = 1$. 
expression for $P_{s}^{(TH)}(E)$, given by (3.16). The constellation size is equal to $M = 4$. There are $L = 2$ paths, where the first path is non-random and its gain is normalized to unity, and the second path is a realization of a complex circularly symmetric Gaussian random variable with zero mean and unit variance. The channel realization is fixed for 1,000 symbols, and the probability of symbol error is averaged over 10,000 channels. The figure includes analytical results for $P_{s}^{(GA)}(E)$, given by (3.17), which improves as $\gamma_b$ increases, since the affect of the noise on the modulo operation becomes negligible.

Figure 3.10 illustrates the probability of symbol error as a function of the SNR-per-bit for the system described in Section 3.3. The QR decomposition is employed, but it is not optimized to minimize the total transmit energy, as described in Section 3.3.6. Here, the QR decomposition is performed on each realization of the Hermitian of the channel matrix, without permuting its columns. Results for the optimized QR decomposition are considered in Figures 3.13 and 3.14.

In Figure 3.10, $N_T = N_U = 2$, $M = 4$, and $L = 1$ and $L = 2$. For each $L$, results are obtained by simulation, by averaging the probability of symbol error over different channel realizations, and also by numerical integration. The simulation results are obtained by transmitting 1,000 data symbols over each of 10,000 channel realizations. Averaging the probability of symbol error over different channel realizations is accomplished by computing the $\{r_{jj}\}$ for each channel realization, substituting them
into $P_s^{(TH)}(E\{r_{jj}\})$ and $P_s^{(GA)}(E\{r_{jj}\})$, given by (3.26) and (3.26), respectively, and averaging over all channel realizations. The analytical results are evaluated by numerically integrating each of the expressions for the conditional probability of symbol error, $P_s^{(TH)}(E\{r_{jj}\})$ and $P_s^{(GA)}(E\{r_{jj}\})$, over the probability density function of the conditioning parameters. Unconditioning $P_s^{(TH)}(E\{r_{jj}\})$ results in $P_s^{(TH)}(E)$, given by (3.28). The results from averaging $P_s^{(TH)}(E\{r_{jj}\})$ over all channel realizations, and the results from numerically integrating $P_s^{(TH)}(E\{r_{jj}\})$ to obtain $P_s^{(TH)}(E)$, correspond closely to the simulation results. As in the SISO case, the performance degradation compared to the Gaussian approximation is the penalty associated with Tomlinson-Harashima precoding. Also, the performance difference between the curves for $L = 1$ and $L = 2$ is the price for not making constructive use of the ISI.

Figure 3.11 shows a similar set of curves as in Figure 3.10 but for $N_T = N_U = 4$. The performance is marginally better than when $N_T = N_U = 2$. This improvement in performance is due to the increase in spatial diversity when $N_T = N_U = 4$, as opposed to $N_T = N_U = 2$. The benefit of spatial diversity comes at the expense of making and feeding back estimates of the channel to the transmitter, as well as the expense of the increase in complexity when using a larger number of transmit antennas.

Figure 3.12 shows a similar set of curves as in Figure 3.10 but for $N_U = 2$, $L = 1$, and $N_T = 2, 3, \text{and} 4$. By fixing $N_U = 2$, the number of users or, equivalently, the number of data streams at the transmitter, is kept constant. As expected, increasing the number of transmit antennas, $N_T$, increases the received power and the spatial diversity. This can be seen in the expression for the probability of error, $P_s^{(TH)}(E\{r_{jj}\})$, given by (3.26), which is a function of the $\{r_{jj}\}$. The degrees of freedom of each of the $r_{jj}$, $j = 1, \ldots, N_U$, increases with $N_T$.

Figure 3.13 illustrates the probability of symbol error as a function of the SNR-per-bit when both an unordered QR decomposition and the optimal ordering are employed with $N_T = N_U = 2$, $M = 4$, and $L = 1$. The analytical result, with optimal ordering, given by (3.35), is evaluated by numerical integration. The analytical results are in agreement with the simulation results, which are obtained by transmitting over 100,000 different channel realizations. As expected, minimizing the total transmit energy by selecting the optimal order of the decentralized receivers results in improved performance compared to an unordered QR decomposition.
Figure 3.14 illustrates simulation results of the probability of symbol error as a function of the SNR-per-bit with \( N_T = N_U = 2 \), \( N_T = N_U = 4 \), and \( N_T = N_U = 6 \). As in Figure 3.13, \( M = 4 \), and \( L = 1 \). As \( N_T \) and \( N_U \) increase, the difference in performance between using the optimal ordering and using the unordered case also increases.

### 3.5 Comparison with Linear Precoders

In this section, we compare the ordered QR decomposition with linear precoders for a flat fading channel. Therefore, the output of the matched filters, sampled at time \((n + 1)T\), is given by (3.21) and, with \( L = 1 \), reduces to

\[
\hat{r}(n) = H(0)s(n) + N_T(n).
\]

To simplify notation, we drop the index \( n \) and the superscript on the channel matrix, giving

\[
\hat{r} = Hs + N_T.
\]

The transmitted signal is given by \( s = Wd \), where \( W \) is the precoding matrix and \( d \) is the same data vector as defined earlier, except we removed the time index.

The matrix that eliminates the MAI at each receiver is generally given by the Moore-Penrose pseudo-inverse

\[
H^+ = H^H( HH^H)^{-1}.
\]

Hence, the precoding matrix is

\[
W = \alpha H^+,
\]

where \( \alpha \) is a scale factor that is selected to satisfy the total transmitted power allocation, i.e., \( \| Wd \|^2 = P \). Thus, the precoding matrix in (3.36) allows the individual users to recover their desired symbols without any interference from the signals transmitted to the other users. We also observe that in the special case where \( N_U = N_T \), \( W = \alpha H^{-1} \), so that the precoding matrix is proportional to the inverse channel matrix. This constitutes a zero-forcing equalizer implemented at the transmitter.

In contrast to the previous figures, the remaining figures in this section show the symbol error probability plotted as a function of the total transmitted signal power over
all antennas divided by $N_0$. Figure 3.15 illustrates the error rate performance of the zero-forcing precoder obtained via Monte Carlo simulation for $N_U = N_T = 4, 6$ and $10$ and QPSK modulation. We observe that the error rate increases with an increase in the number of users.

The performance of a zero-forcing precoder generally suffers from ill-conditioning of the channel matrix $H$. This is the major drawback with the zero-forcing precoder.

If we relax the condition that the interference be zero at all the receivers, the performance degradation can be reduced. This can be accomplished by using the linear MSE criterion in the design of the precoding matrix $W$. Thus, we select $W$ to minimize the cost function

$$J(W, \alpha) = \arg \min_{\alpha, W} E \left\| \frac{1}{\alpha} (HWd + N_T) - d \right\|^2$$

subject to the transmitted power allocation $\|Wd\|^2 = P$, where the expectation in (3.37) is taken over the noise and signal statistics. The solution to the MSE criterion is the precoding matrix

$$W = \alpha H^H (HH^H + \beta I)^{-1},$$

where $\alpha$ is the scale factor that is selected to satisfy the power allocation and $\beta$ is defined as a loading factor, which when selected as $\beta = N_U / P$ maximizes the signal-to-interference-plus-noise ratio (SINR) at the receiver [26].

The error rate performance of the MMSE linear precoder obtained by Monte Carlo simulation in a frequency nonselective Rayleigh fading channel is illustrated in Figure 3.16 for $N_U = N_T = 4, 6$ and $10$. We observe that the error rate performance improves slightly as the number of users $N_U$ increases, and that it exceeds the performance of the zero-forcing precoder.

Figure 3.17 shows a comparison of the error rate performance of the linear $ZF$ and MMSE precoding methods with the ordered $QR$ decomposition for QPSK modulation with $L = 1$, and $N_U = N_T = 4$. Figure 3.18 shows a similar comparison for $N_U = N_T = 6$. We observe that the performance of the $QR$ decomposition method is better than that of the linear precoders at high SNRs, but poorer at low SNRs. However, the improvement in performance of the $QR$ decomposition method at high SNRs should be weighted against the significantly higher computational complexity compared with the linear MMSE precoder.
Figure 3.15: Performance of ZF linear precoding with $N_T = N_U = 4, 6, 10$.

Figure 3.16: Performance of MMSE linear precoding with $N_T = N_U = 4, 6, 10$. 
Figure 3.17: Comparison of the QR decomposition and the linear precoders with $N_T = N_U = 4$.

Figure 3.18: Comparison of the QR decomposition and the linear precoders with $N_T = N_U = 6$. 
3.6 Summary

We have derived the probability of symbol error for a system operating over SISO channels with ISI and employing Tomlinson-Harashima precoding. Using the QR decomposition, we extended this result to MISO frequency selective fading channels with decentralized receivers. We obtained the statistics of the QR decomposition, which enabled us to find an expression for the unconditional probability of symbol error. Furthermore, we investigated the effect of the optimal ordering of the decentralized receivers, obtaining a closed-form expression for the desired probability density function when two transmit antennas and two receivers are employed. Simulation and numerical results demonstrated the accuracy of the analysis, which quantified the penalty associated with Tomlinson-Harashima precoding.

3.7 Acknowledgements

The Effects of Channel Estimation Errors on a Nonlinear Precoder for MISO Channels with Decentralized Receivers

4.1 Introduction

In this chapter, we analyze the effects of channel estimation errors on the performance of a system employing Tomlinson-Harashima precoding and the QR decomposition for MISO channels with decentralized receivers operating over frequency-flat fading channels. We use the QR decomposition technique to enable successive interference cancellation. Due to imperfect channel estimation which causes a mismatch between the precoder and the channel, multiuser interference is present at the receivers and must be accounted for in the derivation of the probability of symbol error. In addition, as in the case when the channel is known, performance is degraded by the increase in power inherent in the precoding technique. The system considered in this chapter is the same as the pre-BLAST-DFE technique that is presented in Chapter 3 and [35], except for the fact that we do not consider ordering of the decentralized receivers.

The model for channel estimation used in this chapter is similar to the model
used in [41]. In the uplink transmission, geographically separated users transmit training sequences which are used by the receiver with co-located antennas to estimate the channel. Assuming time division duplex (TDD) and a large enough channel coherence time such that the channel does not change significantly between uplink and downlink modes, the channel estimates formed in the uplink transmission can be employed in the downlink.

The work in [41] attempts to derive the probability of symbol error for the same system considered in this paper, except that, in [41], the number of transmit antennas is equal to the number of receive antennas whereas, in this paper, we also consider the case when the number of transmit antennas is greater than the number of receive antennas. In [41], matrix differentials are applied to derive the precoding matrices, which are formed from realizations of the channel by assuming the channel estimation error, $\Delta H$, is small. Indeed, the entries of $\Delta H$ are normally distributed and, in this paper, we show that the entries are i.i.d. with variance proportional to $1/N_{Tr}$, where $N_{Tr}$ is the length of the training vectors used for channel estimation. For large enough $N_{Tr}$, which is a design parameter, the variance of the entries of $\Delta H$ can be made small. However, there is no guarantee that actual realizations of the entries of $\Delta H$ are small, and therefore, it is unclear whether the analysis in [41] can be rigorously justified.

Furthermore, there are several errors in the analysis in [41] associated with Lemma 1. This Lemma is used in [41] to show that the contribution of the channel estimation error $\Delta H$ to the test statistic is asymptotically Gaussian. Unfortunately, Lemma 1 is a misquotation of Theorem 3.3A in [42]: This latter theorem, which deals with nonlinear transformations of asymptotically Gaussian random vectors, requires that 1) the differential of the transformation is nonzero, 2) the value of this differential must be computed explicitly in order to obtain the asymptotic variance/covariance matrix of the transformed variables. There is no evidence that these two steps have been carried out in [41].

We arrive at our results by different techniques, effectively avoiding having to establish the dependencies of various random variables by initially deriving the conditional probability of symbol error, conditioned on the elements of the actual channel matrix as well as the elements of the channel estimation error matrix. The conditional probability of symbol error is then averaged by integration over the densities of these
random variables, leading to results that coincide with simulations.

Both [43] and [44] consider receiver equalization with the QR decomposition approach, and both use the same channel estimation model employed in this chapter. The outage probability due to channel estimation is computed in [43] by approximating one of the matrices derived from the QR decomposition of the imperfect channel estimation matrix by one of the matrices derived from the QR decomposition of the actual channel matrix. This is convenient, since the actual channel matrix is independent of the error estimation matrix. The same approximation is made in [44]. Furthermore, in deriving the probability of symbol error, the dependency between the channel estimation matrix and its QR decomposition is implicitly ignored in [44]. Our approach does not make any of these approximations.

Other works consider different types of CSI at the transmitter. For example, [39] employs a low-rate channel to feed back partial CSI to the transmitter, which uses the information to update an average of the channel covariance matrix. Residual linear equalization is performed at the receiver to compensate for the mismatch between the precoder and the actual channel. The work in [40] considers only long-term CSI, in the form of correlation matrices of the fading and additive noise, is available at the transmitter, and proceeds to optimize linear and nonlinear precoders based on the MMSE criterion.

In Section 4.2, we describe the system, including the model for channel estimation in the uplink transmission and the signal model in the downlink transmission. In Section 4.3, we derive the probability of symbol error, including the effect of channel estimation. In Section 4.4, we present simulation and numerical results. A summary is given in Section 4.5.

4.2 System Model

Let the number of transmit antennas at the transmitter be equal to \( N_T \), the number of users be equal to \( N_U \), and the number of receive antennas at each user be equal to unity. The signals intended for different users interfere with each other. This is the motivation behind using the QR decomposition of the channel matrix to arrive at an equivalent channel where successive interference cancellation can be used to remove
the effect of the multiuser interference.

First we consider channel estimation which is performed in the uplink transmission. Then we derive the signal model in the downlink transmission.

### 4.2.1 Channel Estimation

The channel coefficient between the \(i\)th transmit antenna and the receive antenna of user \(j\) is given by \(h_{j,i}\). The coefficients \(\{h_{j,i}\}\) are realizations of i.i.d. zero-mean circularly symmetric complex Gaussian random variables with variance

\[
\sigma_h^2 \triangleq E[|h_{j,i}|^2], \quad \forall i \text{ and } j,
\]  

where we normalize \(\sigma_h^2 = 1\). The coefficients between the \(N_T\) transmitting antennas and the \(N_U\) receiving users can be represented by an \(N_U \times N_T\) matrix \(H\), where \([H]_{j,i} = h_{j,i}, i = 1, \ldots, N_T, j = 1, \ldots, N_U\).

To form the channel estimates, we extend the derivation in [41] to show that, by properly constructing the training vectors used for channel estimation as in [43] and [44], the channel estimation error matrix can be modelled as a matrix of i.i.d. complex circularly symmetric Gaussian random variables.

In the uplink transmission, \(N_U\) users transmit to the \(N_T\) receiving antennas. The \(N_U\) training vectors are of length \(N_{Tr}\) and are used to form the maximum-likelihood estimate of the channel matrix. These vectors are each of dimension \(1 \times N_{Tr}\) so that the training vector transmitted from the \(i\)th user is given by \(v_i = [v_{i,1}, \ldots, v_{i,N_{Tr}}]\), \(i = 1, \ldots, N_U\). We form the \(N_U \times N_{Tr}\) training matrix \(V = [v_1^T, \ldots, v_{N_U}^T]^T\) and normalize \(\text{tr}(VV^H) = N_U N_{Tr}\). The received vector and noise vector corresponding to the \(N_{Tr}\) transmitted vectors of dimension \(N_U \times 1\) are grouped into matrices of dimension \(N_T \times N_{Tr}\) given by \(Y = [y_1, \ldots, y_{N_{Tr}}]\) and \(N = [n_1, \ldots, n_{N_{Tr}}]\), respectively. Then we can write

\[
Y = HV + N.
\]

The entries of the matrix \(N\) are i.i.d. complex circularly symmetric Gaussian random variables with zero-mean and variance \(\sigma_n^2\).

Theorem 7.5 in [66] derives the Maximum-Likelihood Estimate (MLE) for the
linear model in 4.2, which is given by

$$\hat{H} = YV^H(VV^H)^{-1}$$

$$= HVV^H(VV^H)^{-1} + NV^H(VV^H)^{-1}$$

$$= H + \Delta H,$$

where $\Delta H \triangleq NV^H(VV^H)^{-1}$ is the channel estimation error matrix.

Note that to obtain the MLE of $H$, we multiply $Y$ by the left pseudoinverse of $V$.

To invert $VV^H$, the rows of $V$ must be linearly independent. Consequently, we require $N_{Tr} \geq N_U$.

Set $M = [m_1, \ldots, m_{NU}] \triangleq V^H(VV^H)^{-1}$, and denote the $i^{th}$ row of $N$ and $\Delta H$ by $n_i^{(r)}$ and $\Delta h_i^{(r)}$, respectively, where the notation $(\cdot)^{(r)}$ is used to denote row vector. Then we can explicitly write $\Delta H$ in terms of these vectors as

$$\Delta H = NM = \begin{bmatrix} n_1^{(r)} \\ \vdots \\ n_{NU}^{(r)} \end{bmatrix} = \begin{bmatrix} n_1^{(r)} & \cdots & n_{NU}^{(r)} \end{bmatrix} \begin{bmatrix} \Delta h_1^{(r)} \\ \vdots \\ \Delta h_{NU}^{(r)} \end{bmatrix}. $$

From the above, it is easy to see that since the rows of $N$ are uncorrelated then the rows of $\Delta H$ are uncorrelated, and that

$$\Delta h_i^{(r)} = n_i^{(r)}M, \quad i = 1, \ldots, N_{Tr}. $$

To construct the $N_U \times N_{Tr}$ signal matrix $V$ that results in $\Delta H$ having i.i.d. entries, consider minimizing the trace of the covariance of the rows of $\Delta H$ subject to the energy constraint $\text{tr}(VV^H) = N_U N_{Tr}$. The covariance of each row of $\Delta H$ is given by

$$E[(\Delta h_i^{(r)})^H \Delta h_j^{(r)}] = E[M^H(n_i^{(r)})^H n_j^{(r)} M]$$

$$= \sigma_n^2 M^H M \delta_{i,j} $$

$$= \sigma_n^2 (V^H(VV^H)^{-1} V^H(VV^H)^{-1}) \delta_{i,j} $$

$$= \sigma_n^2 (VV^H)^{-1} VV^H(VV^H)^{-1} \delta_{i,j} $$

$$= \sigma_n^2 (VV^H)^{-1} \delta_{i,j}. $$

Using the SVD, we can write $V^H = \Phi \Sigma \Psi^H$, where $\Phi$ is $N_{Tr} \times N_U$ with orthonormal columns, $\Psi$ is $N_U \times N_U$ unitary, and $\Sigma$ is $N_U \times N_U$ diagonal with real nonnegative
\(i^{th}\) diagonal entry \(\sigma_i\) (not to be confused with noise variance \(\sigma_n^2\)). Then the covariance is given by
\[
E[(\Delta h_i^{(r)})^H \Delta h_j^{(r)}] = \sigma_n^2 (VV^H)^{-1} \delta_{i,j} = \sigma_n^2 \Psi \Sigma^{-2} \Psi^H \delta_{i,j}.
\]
Computing the trace of the covariance for \(i = j\), we have
\[
\text{tr}\{E[(\Delta h_i^{(r)})^H \Delta h_i^{(r)}]\} = \sigma_n^2 \text{tr}\{\Psi \Sigma^{-2} \Psi^H\} = \sigma_n^2 \text{tr}\{\Sigma^{-2}\} = \sigma_n^2 \sum_{k=1}^{N_U} \sigma_k^{-2},
\]
where we have used the cyclic property of the trace operation [67] and the fact that \(\Psi\) is unitary.

From this, we can see that the trace is proportional to \(\sum_{k=1}^{N_U} \sigma_k^{-2}\). Subject to the energy constraint \(\text{tr}(VV^H) = \sum_{k=1}^{N_U} \sigma_k^2 = N_U N_{Tr}\), the trace is minimized when the \(N_U\) singular values are equal [43], that is \(\sigma_k = \sqrt{N_{Tr}}, \forall k\). Therefore, we select a training matrix \(V\) such that its SVD is given by \(V^H = \sqrt{N_{Tr}} \Phi \Psi^H\), where \(\Phi\) is \(N_{Tr} \times N_U\) with orthonormal columns and \(\Psi\) is \(N_U \times N_U\) unitary. This results in covariance given by
\[
E[(\Delta h_i^{(r)})^H \Delta h_j^{(r)}] = \sigma_n^2 (N_{Tr} \Phi \Phi^H \Phi \Psi^H)^{-1} \delta_{i,j} = \frac{\sigma_n^2}{N_{Tr}} I \delta_{i,j}.
\]
Consequently, the channel estimation error matrix \(\Delta H\) can be modeled as a matrix of i.i.d. complex circularly symmetric Gaussian random variables with zero-mean and variance \(\sigma_e^2\), which becomes smaller as the training length increases.

### 4.2.2 Signal Model

Now, having obtained the channel estimates, we consider the downlink transmission. The QR decomposition of the channel matrix is given by \(H^H = QR\). However, at the transmitter, we compute the QR decomposition of the channel estimation matrix giving \((\hat{H})^H = \hat{Q} \hat{R}\), where \(\hat{Q}\) is an \(N_T \times N_U\) matrix such that \(\hat{Q} \hat{Q}^H = I\), and \(\hat{R}\) is an \(N_U \times N_U\) upper-right triangular matrix. The entries of \(\hat{H}\) are random variables. Consequently, \(\hat{Q}\) and \(\hat{R}\) are also random matrices with statistics determined in [62], [63, Lemma 2.1]. The signal to be transmitted is precoded with the matrix \(\hat{W} = \hat{Q} \hat{A}\), where \(\hat{A}\) is a
diagonal matrix with entries \([\hat{A}_{i,i} = 1/\hat{r}_{i,i}, i = 1,\ldots, N_U]\). The \(\{\hat{r}_{i,i}\}\) are the diagonal entries of the matrix \(\hat{R}\), and can be shown to be real and positive [62], [63, Lemma 2.1].

Precoding with the matrix \(\hat{W}\) results in an effective channel given by \(H\hat{W} = (R^HQ^T)(\hat{Q}\hat{A})\), which is not necessarily a lower triangular matrix. In the case when the channel is known, an effective channel that is represented by a lower triangular matrix implies that user \(j\) sees interference due only to users \(1,\ldots, j-1\), which is removed by successive interference cancellation at the transmitter. In the case of imperfect channel estimation, successive interference cancellation for user \(j\) does not perfectly remove the interference due to users \(1,\ldots, j-1\). In addition, since the effective channel is not lower triangular, interference is also present due to users \(j+1,\ldots, N_U\).

Recall that when the channel is known, the feedforward precoding \(N_T \times N_U\) matrix \(W\) is designed so that the \(N_U \times N_U\) effective channel matrix \(HW\) is lower triangular with unit diagonal, as shown in Chapter 3. Therefore, \(HW\) is of full rank \(N_U\). This is only possible if \(N_T \geq N_U\).

Successive interference cancellation is implemented with a feedback filter given by

\[
\tilde{B} = I - \hat{H}\hat{W}.
\]

The output of the modulo operators for the \(n^{th}\) symbol vector is given by

\[
x(n) = m_{2\sqrt{M}} \left( d(n) + \hat{B}x(n) \right) = d(n) + \hat{B}x(n) - 2\sqrt{M}z_x(n),
\]

where \(x(n)\) is the \(N_U \times 1\) vector at the output of the modulo operators, \(d(n)\) is the \(N_U \times 1\) data vector, \(z_x(n) = [z_{x,1}(n), \ldots, z_{x,N_U}(n)]^T\), \(z_{x,j}(n) = z_{x,j}^{(R)}(n) + j z_{x,j}^{(I)}(n)\),

\[
z_{x,j}^{(R)}(n) \triangleq \left\lfloor \frac{\text{Re}\left\{ [d(n) + \hat{B}x(n)]_j \right\} + \sqrt{M}}{2\sqrt{M}} \right\rfloor,
\]

\[
z_{x,j}^{(I)}(n) \triangleq \left\lfloor \frac{\text{Im}\left\{ [d(n) + \hat{B}x(n)]_j \right\} + \sqrt{M}}{2\sqrt{M}} \right\rfloor,
\]

and \([v]_j\) is the \(j^{th}\) element of \(v\). The modulo operation is defined as \(m_y(x) \triangleq x - y \left\lfloor \frac{x+y/2}{y} \right\rfloor\). If the input to the modulo operator is complex, then the operation is performed separately on the real and imaginary parts. If the input to the modulo operator is a vector, then the modulo operator operates on each element of the vector in its argument.
Figure 4.1: System model of Tomlinson-Harashima precoding and the QR decomposition technique operating over a MISO frequency-flat fading channel with decentralized receivers using the imperfect channel estimation matrix at the transmitter.

The $j^{th}$ element of the data vector is given by $d_j(n) = d_{c,j}(n) + j d_{s,j}(n)$, where $d_{c,j}(n)$ and $d_{s,j}(n)$ are the in-phase and quadrature data signals, respectively, each having values in

$$A = \{ \pm 1, \pm 3, \ldots, \pm (\sqrt{M} - 1) \}. \quad (4.4)$$

We assume square $M$-QAM constellations where $M = 2^b$ and $b$ is even, so that the signal constellation is equivalent to two $\sqrt{M}$-PAM signals on quadrature carriers.

The system diagram is shown in Figure 4.1. Before multiplying by $\hat{W}$, the signals are multiplied by a power scaling matrix $P = pI$ used to satisfy the SNR requirement, where $p$ is a constant. Then, if $s(n)$ is the vector of transmitted signals, and $x(n)$ is the vector of precoded data symbols, $s(n) = \hat{W}P x(n)$.

The low-pass equivalent transmitted signal vector is given by

$$\tilde{s}(t) = \sum_{m=-\infty}^{\infty} s(m) g(t - mT),$$

where $T$ is the symbol duration and

$$g(t) = \begin{cases} 1/\sqrt{T}, & 0 \leq t < T \\ 0, & \text{otherwise} \end{cases}.$$ 

The low-pass equivalent received signal vector, with the $j^{th}$ element corresponding to the received signal at the $j^{th}$ user, is given by $\tilde{r}(t) = \mathbf{H}\tilde{s}(t) + \mathbf{n}_w(t)$, where $\mathbf{n}_w(t)$ is a $N_U \times 1$ noise vector with independent elements. Each element is a zero-mean complex white Gaussian noise process with two-sided spectral density $N_0$. 
The output of the matched filters, sampled at time \((n + 1)T\), is given by

\[
\hat{r}(n) = \frac{1}{\sqrt{T}} \int_{nT}^{(n+1)T} \tilde{r}(t) \, dt = \mathbf{Hs}(n) + \mathbf{N}_T(n),
\]

where \(\mathbf{N}_T(n) \triangleq \frac{1}{\sqrt{T}} \int_{nT}^{(n+1)T} \mathbf{n}_w(t) \, dt\). Multiplying (4.5) by \(\mathbf{P}^{-1} = \frac{1}{p} \mathbf{I}\) gives

\[
\mathbf{P}^{-1}\hat{r}(n) = \mathbf{H}\hat{\mathbf{W}}x(n) + \mathbf{N}_T(n)
\]

\[
= x(n) + (\mathbf{H}\hat{\mathbf{W}} - \mathbf{I})x(n) + \mathbf{N}_T(n),
\]

where we added and subtracted the term \(x(n)\), and \(\{\mathbf{N}_{nT,j}(n)\}\) is a sequence of i.i.d. zero-mean circularly symmetric complex Gaussian random variables with covariance

\[
\frac{1}{2}E[\mathbf{N}_{nT,j}(n_1)\mathbf{N}_{nT,j}(n_2)] = \frac{N_0}{p^2} \delta_{n_1,n_2}.
\]

Substituting for \(x(n)\), we have

\[
\mathbf{P}^{-1}\hat{r}(n) = \mathbf{d}(n) + \hat{\mathbf{B}}x(n) - 2\sqrt{M}\mathbf{x}(n) + (\mathbf{H}\hat{\mathbf{W}} - \mathbf{I})x(n) + \mathbf{N}_T(n).
\]

Substituting for \(\hat{\mathbf{B}}\), gives

\[
\mathbf{P}^{-1}\hat{r}(n) = \mathbf{d}(n) + (\mathbf{I} - \hat{\mathbf{H}}\hat{\mathbf{W}})x(n) - 2\sqrt{M}\mathbf{x}(n) + (\mathbf{H}\hat{\mathbf{W}} - \mathbf{I})x(n) + \mathbf{N}_T(n)
\]

\[
= \mathbf{d}(n) + (\mathbf{H} - \hat{\mathbf{H}})\mathbf{W}x(n) + \mathbf{N}_T(n) - 2\sqrt{M}\mathbf{x}(n)
\]

\[
= \mathbf{d}(n) + \Delta\mathbf{H}\mathbf{W}x(n) + \mathbf{N}_T(n) - 2\sqrt{M}\mathbf{x}(n)
\]

\[
= \mathbf{d}(n) + \mathbf{e}(n) + \mathbf{N}_T(n) - 2\sqrt{M}\mathbf{x}(n),
\]

where the contribution due to the estimation error is given by the vector

\[
\mathbf{e}(n) \triangleq \Delta\mathbf{H}\mathbf{W}x(n). \tag{4.6}
\]

The vector of test statistics for the \(n^{th}\) symbol vector is given by

\[
\mathbf{y}(n) = m_{2\sqrt{M}}(\mathbf{P}^{-1}\hat{r}(n)).
\]

The test statistic for user \(j\) and symbol \(n\) is given by

\[
y_{j}(n) = m_{2\sqrt{M}}\left(\frac{\hat{r}_{j}(n)}{p}\right)
\]

\[
= m_{2\sqrt{M}}\left(d_{j}(n) + \mathbf{e}_{j}(n) + \mathbf{N}_{T,j}(n) - 2\sqrt{M}\mathbf{z}_{x,j}(n)\right)
\]

\[
= m_{2\sqrt{M}}(d_{j}(n) + \mathbf{e}_{j}(n) + \mathbf{N}_{T,j}(n))
\]

\[
= d_{j}(n) + \mathbf{e}_{j}(n) + \mathbf{N}_{T,j}(n) - 2\sqrt{M}\mathbf{z}_{y,j}(n), \tag{4.7}
\]
where the third equality is due to Proposition 3.1 in Chapter 3, \( z_{y,j}(n) = z_{y,j}^{(R)}(n) + j z_{y,j}^{(I)}(n) \),

\[
z_{y,j}^{(R)}(n) \triangleq \left[ \text{Re}\{d_j(n) + \hat{N}_{T,j}(n)\} + \sqrt{M}/(2\sqrt{M}) \right],
\]

and

\[
z_{y,j}^{(I)}(n) \triangleq \left[ \text{Im}\{d_j(n) + \hat{N}_{T,j}(n)\} + \sqrt{M}/(2\sqrt{M}) \right].
\]

Note that the \( \{y_j(n)\} \) are i.i.d. for different \( n \). Since the \( \{d_j(n)\} \) and the \( \{\hat{N}_{T,j}(n)\} \) are both i.i.d. for different \( j \) and fixed \( n \), by (4.7), the \( \{y_j(n)\} \) are also i.i.d. for different \( j \) and fixed \( n \).

It is possible to show, by a similar derivation as in Chapter 3, that the average transmitted energy-per-symbol is given by

\[
E_{av}^{(s)} = p^2 f_{M,N_U}(\{\tilde{r}_{j,j}^2\}),
\]

where, for convenience, we define

\[
f_{M,N_U}(\{\tilde{r}_{j,j}^2\}) \triangleq \frac{M}{3N_U} \sum_{j=1}^{N_U} \frac{1}{\tilde{r}_{j,j}^2}.
\]

### 4.3 Probability of Error

The test statistic for user \( j \) is given by \( y_j(n) \) in Equation (4.7) and can be written as

\[
y_j(n) \triangleq y_j^{(R)}(n) + j y_j^{(I)}(n),
\]

where

\[
y_j^{(R)}(n) \triangleq \text{Re}\{y_j(n)\}
\]

\[
= m_{2\sqrt{M}}(\text{Re}\{d_j(n) + e_j(n) + \hat{N}_{T,j}(n)\})
\]

\[
= d_{e,j}(n) + e_j^{(R)}(n) + \hat{N}_{T,j}^{(R)}(n) - 2\sqrt{M} z_{y,j}^{(R)}(n),
\]

\[
y_j^{(I)}(n) \triangleq \text{Im}\{y_j(n)\}
\]

\[
= m_{2\sqrt{M}}(\text{Im}\{d_j(n) + e_j(n) + \hat{N}_{T,j}(n)\})
\]

\[
= d_{s,j}(n) + e_j^{(I)}(n) + \hat{N}_{T,j}^{(I)}(n) - 2\sqrt{M} z_{y,j}^{(I)}(n),
\]

and we have used the fact that the modulo operator operates separately on the real and imaginary parts of its input. We have also defined \( e_j^{(R)}(n) \triangleq \text{Re}(e_j(n)) \) and \( e_j^{(I)}(n) \triangleq \text{Im}(e_j(n)) \). The contributions due to AWGN, for symbol \( n \), are written as \( \hat{N}_{T,j}^{(R)}(n) \triangleq \)
Re\{\tilde{N}_{T,j}(n)\} and \tilde{N}^{(I)}_{T,j}(n) \triangleq \text{Im}\{\tilde{N}_{T,j}(n)\}. Since \{\tilde{N}_{T,j}(n)\} is a sequence of i.i.d. zero-mean circularly symmetric complex Gaussian random variables, each with variance \(N_0/p^2\), then the \{\tilde{N}^{(R)}_{T,j}(n)\} and \{\tilde{N}^{(I)}_{T,j}(n)\} are both independent sequences of i.i.d. zero-mean real Gaussian random variables with variance
\[
\sigma^2 \triangleq N_0/p^2. \tag{4.11}
\]

The vector \(e(n)\) is independent of the noise \(\tilde{N}_T(n)\) but not necessarily independent of the data \(d(n)\), since \(x(n)\) is a function of \(d(n)\). We rewrite \(x(n)\) in a form that shows its dependency on \(\hat{H}\) explicitly. From Equation (4.3), we can write
\[
x(n) = m_{2\sqrt{M}} \left( d(n) + \hat{B}x(n) \right) = m_{2\sqrt{M}} \left( d(n) + (I - \hat{H}\hat{W})x(n) \right),
\]
where
\[
I - \hat{H}\hat{W} = \begin{bmatrix}
0 & 0 & \cdots & 0 \\
\hat{r}^*_{1,2}/\hat{r}_{1,1} & 0 & \cdots & \\
\vdots & \ddots & \ddots & 0 \\
\hat{r}^*_{1,NU}/\hat{r}_{1,1} & \cdots & \hat{r}^*_{NU-1,NU}/\hat{r}_{NU-1,NU-1}
\end{bmatrix}.
\]
Substituting, we have
\[
x_j(n) = m_{2\sqrt{M}} \left( d_j(n) - \sum_{i=1}^{i-1} \frac{\hat{r}^*_{i,j}}{\hat{r}_{i,i}} x_i(n) \right) .
\]
Therefore, \(x(n)\) depends on \(d(n)\) and \(\hat{H}\).

Also, \(\hat{W}\) is dependent on \(\hat{H}\), which is a function of \(\Delta H\). In order to avoid establishing the dependencies between \(d(n)\), \(x(n)\), \(\hat{W}\), and \(\hat{H}\), we can condition on \(H\), \(\Delta H\), and \(d(n)\).

Let \(\hat{d}_{c,j}(n)\) be an estimate of \(d_{c,j}(n)\), which is given by
\[
\hat{d}_{c,j}(n) = \text{Quant} \left( y_j^{(R)}(n) \right), \tag{4.12}
\]
where \(y_j^{(R)}(n)\) is given in (4.9). The function \(\text{Quant}(\cdot)\) is an \(\sqrt{M}\)-level uniform quantizer given by
\[
\text{Quant}(x) = \begin{cases} 
-(\sqrt{M} - 1) & \text{if } x < -\sqrt{M} \\
q & \text{if } q - 1 \leq x < q + 1, \\
\sqrt{M} - 1 & \text{if } x \geq \sqrt{M}
\end{cases}
\]
where \( q \) is an integer from the set \( A \) defined in (4.4).

Similarly, let \( \hat{d}_{s,j}(n) \) be an estimate of \( d_{s,j}(n) \), which is given by

\[
\hat{d}_{s,j}(n) = \text{Quant} \left( y_j^{(I)}(n) \right),
\]

(4.13)

where \( y_j^{(I)}(n) \) is given in (4.10).

For convenience, we define the conditioning set of random variables as

\[
S \triangleq \{ \{ h_{i_1,i_2}^{(0)} \}_{i_1=1,...,N_U, i_2=1,...,N_T}, \{ \Delta h_{i_1,i_2}^{(0)} \}_{i_1=1,...,N_U, i_2=1,...,N_T}, \{ d_{c,k}(n) \}_{k=1}^{N_U}, \{ d_{s,k}(n) \}_{k=1}^{N_U} \}.
\]

The conditional probability of symbol error for user \( j \), conditioned on \( S \), is given by

\[
P_j^{\text{TH}}(E|S) = P_r\{(\hat{d}_{c,j}(n) \neq d_{c,j}(n)) \cup (\hat{d}_{s,j}(n) \neq d_{s,j}(n))|S\}
= 1 - P_r\{(\hat{d}_{c,j}(n) = d_{c,j}(n)) \cap (\hat{d}_{s,j}(n) = d_{s,j}(n))|S\}
= 1 - P_r\{\hat{d}_{c,j}(n) = d_{c,j}(n)|S\}P_r\{\hat{d}_{s,j}(n) = d_{s,j}(n)|S\},
\]

(4.14)

where the last equality is due to the fact that the real and imaginary parts of the thermal noise component of the test statistic are independent, and we have conditioned over all other remaining random variables.

When \( \hat{d}_{c,j}(n) = d_{c,j}(n) \) then, by (4.12), \( y_j^{(R)}(n) \in [d_{c,j}(n) - 1, d_{c,j}(n) + 1] \). Therefore, the conditional probability of a correct decision for the in-phase component of the desired symbol, conditioned on \( S \), is given by

\[
Pr\{\hat{d}_{c,j}(n) = d_{c,j}(n)|S\}
= Pr\{d_{c,j}(n) - 1 \leq y_j^{(R)}(n) < d_{c,j}(n) + 1|S\}
= Pr\left\{d_{c,j}(n) - 1 \leq m_2\sqrt{M} \left( d_{c,j}(n) + e_j^{(R)}(n) + \tilde{N}_{T,j}^{(R)}(n) \right) < d_{c,j}(n) + 1 \bigg| S \right\}.
\]

(4.15)

where we have substituted for \( y_j^{(R)}(n) \) using (4.9).

Equation (4.15) specifies the conditional probability of the received signal falling in the decision region of all points congruent to \( d_{c,j}(n) \). The modulo operation is a periodic sawtooth function, with period \( 2\sqrt{M} \), taking on values in \([ -\sqrt{M}, \sqrt{M} ) \). It is necessary to take into account that the component of the test statistic due to AWGN,
\( \tilde{N}_{T,j}^{(R)}(n) \), is an unbounded random variable, and that therefore so is the argument of the modulo operation, \( X \triangleq d_{c,j}(n) + e_{j}^{(R)}(n) + \tilde{N}_{T,j}^{(R)}(n) \). Therefore,

\[
m_{2\sqrt{M}}(X) = X - 2\sqrt{Mi},
\]  

(4.16)

when \( X \) falls in the \( i^{th} \) period of the modulo operation given by

\[
\left[ 2\sqrt{Mi} - \sqrt{M}, 2\sqrt{Mi} + \sqrt{M} \right],
\]  

(4.17)

for each \( i \in \mathbb{Z} \).

We can write the region in (4.15) as the union of the events corresponding to the distinct points, each of which fall in a different period, that give the same value of the modulo operation, giving

\[
\left\{ d_{c,j}(n) - 1 \leq m_{2\sqrt{M}}(X) < d_{c,j}(n) + 1 \right\}
\]

\[
= \bigcup_{i = -\infty}^{\infty} \left\{ 2\sqrt{Mi} + d_{c,j}(n) - 1 \leq X < 2\sqrt{Mi} + d_{c,j}(n) + 1 \right\}
\]

\[
= \bigcup_{i = -\infty}^{\infty} \left\{ 2\sqrt{Mi} + d_{c,j}(n) - 1 \leq d_{c,j}(n) + e_{j}^{(R)}(n) + \tilde{N}_{T,j}^{(R)}(n) < 2\sqrt{Mi} + d_{c,j}(n) + 1 \right\}
\]

\[
= \bigcup_{i = -\infty}^{\infty} \left\{ 2\sqrt{Mi} - 1 \leq e_{j}^{(R)}(n) + \tilde{N}_{T,j}^{(R)}(n) < 2\sqrt{Mi} + 1 \right\}.
\]  

(4.18)

Since the probability of the union of disjoint events is equal to the sum of the probability of each event, plugging (4.18) into (4.15) gives

\[
Pr\{ \hat{d}_{c,j}(n) = d_{c,j}(n) | S \}
\]

\[
= \sum_{i = -\infty}^{\infty} Pr \left\{ 2\sqrt{Mi} - 1 \leq e_{j}^{(R)}(n) + \tilde{N}_{T,j}^{(R)}(n) < 2\sqrt{Mi} + 1 | S \right\}.
\]  

(4.19)

Similarly, it is possible to show that the conditional probability of a correct decision for the quadrature component of the desired symbol, conditioned on \( S \), is given by

\[
Pr\{ \hat{d}_{s,j}(n) = d_{s,j}(n) | S \}
\]

\[
= \sum_{i = -\infty}^{\infty} Pr \left\{ 2\sqrt{Mi} - 1 \leq e_{j}^{(R)}(n) + \tilde{N}_{T,j}^{(I)}(n) < 2\sqrt{Mi} + 1 | S \right\}.
\]  

(4.20)
Substituting (4.19) and (4.20) in (4.14) gives

\[ P_j^{(TH)}(E|S) = 1 - \Pr\{d_{c,j}(n) = d_{c,j}(n)|S\} \]

\[ = 1 - \left( \sum_{i=-\infty}^{\infty} \Pr\left\{ 2\sqrt{M}i - 1 \leq e_j^{(R)}(n) + \tilde{N}_{T,j}^{(R)}(n) < 2\sqrt{M}i + 1 | S \right\} \right) \cdot \left( \sum_{i=-\infty}^{\infty} \Pr\left\{ 2\sqrt{M}i - 1 \leq e_j^{(R)}(n) + \tilde{N}_{T,j}^{(R)}(n) < 2\sqrt{M}i + 1 | S \right\} \right). \]

To evaluate the unconditional probability of symbol error, we need the joint density of the conditioning random variables. Since these are independent, we average the conditional probability of symbol error over the product of their densities. Then the unconditional probability of symbol error for user \( j \) can be written as

\[ P_j^{(TH)}(E) = \left( \frac{1}{\sqrt{M}} \right)^{2N_U} \prod_{i=-\infty}^{\infty} \int_{-\infty}^{\infty} \sum_{d_{c,1} \in A} \cdots \sum_{d_{c,N_U} \in A} \sum_{d_{s,1} \in A} \cdots \sum_{d_{s,N_T} \in A} \cdot P_j^{(TH)}\left( E \left| \{h_{i_1,i_2} = a_{i_1,i_2}\}_{i_1=1,...,N_U,i_2=1,...,N_T}, \right. \right. \]

\[ \{\Delta h_{i_1,i_2} = \Delta a_{i_1,i_2}\}_{i_1=1,...,N_U,i_2=1,...,N_T}, \]

\[ \{d_{c,k}(n) = \tilde{d}_{c,k}\}_{k=1}^{N_U}, \{d_{s,k}(n) = \tilde{d}_{s,k}\}_{k=1}^{N_T} \]

\[ f_{h_{1,1}}(a_{1,1}) \cdots f_{h_{N_U,N_T}}(a_{N_U,N_T}) \cdot f_{\Delta h_{1,1}}(\Delta a_{1,1}) \cdots f_{\Delta h_{N_U,N_T}}(\Delta a_{N_U,N_T}) \cdot da_{1,1} \cdots da_{N_U,N_T} d\Delta a_{1,1} \cdots d\Delta a_{N_U,N_T}, \]

where the densities \( f_{h_{i,j}} \) and \( f_{\Delta h_{i,j}} \) are circularly symmetric complex Gaussian with zero-mean and variance \( \sigma_h^2 \) and \( \sigma_e^2 \), respectively.
Substituting (4.21) in (4.22), we have

\[
P_j^{(TH)}(E) = \left( \frac{1}{\sqrt{M}} \right)^{2N_U} \sum_{i=-\infty}^{\infty} \sum_{d_{c,k} \in A} \sum_{d_{s,NU} \in A} \sum_{\tilde{d}_{s,NU} \in A} \sum_{\tilde{d}_{c,NU} \in A} \sum_{\tilde{d}_{s,NU} \in A} \left[ \Phi \left( \frac{2\sqrt{M}i - 1 - \tilde{e}_j^{(R)}(n)}{\sigma} \right) - \Phi \left( \frac{2\sqrt{M}i - 1 - \tilde{e}_j^{(I)}(n)}{\sigma} \right) \right] \]

\[
= \left( \frac{1}{\sqrt{M}} \right)^{2N_U} \sum_{i=-\infty}^{\infty} \sum_{d_{c,k} \in A} \sum_{d_{s,NU} \in A} \sum_{\tilde{d}_{s,NU} \in A} \sum_{\tilde{d}_{c,NU} \in A} \sum_{\tilde{d}_{s,NU} \in A} \left[ \Phi \left( \frac{2\sqrt{M}i + 1 - \tilde{e}_j^{(I)}(n)}{\sigma} \right) - \Phi \left( \frac{2\sqrt{M}i - 1 - \tilde{e}_j^{(I)}(n)}{\sigma} \right) \right] \]

where \( \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt \), \( \sigma^2 \) is defined in (4.11), and \( \tilde{e}_j^{(R)}(n) \) and \( \tilde{e}_j^{(I)}(n) \) are the values of \( e_j^{(R)}(n) \) and \( e_j^{(I)}(n) \), respectively, with the substitution

\[
\{ h_{i_1,i_2} = a_{i_1,i_2} \}_{i_1=1,...,N_U, i_2=1,...,N_T}, \{ \Delta h_{i_1,i_2} = \Delta a_{i_1,i_2} \}_{i_1=1,...,N_U, i_2=1,...,N_T},
\]

\[
\{ d_{c,k}(n) = \tilde{d}_{c,k} \}_{k=1}^{N_U}, \text{ and } \{ d_{s,k}(n) = \tilde{d}_{s,k} \}_{k=1}^{N_U}. \text{ The first equality in (4.23) is due to the fact that both } \tilde{N}_{T,j}^{(R)}(n) \text{ and } \tilde{N}_{T,j}^{(I)}(n) \text{ are independent of } S. \text{ This is because we assume that the thermal noise in the downlink transmission is independent of the thermal noise which is present when making the estimate of the channel in the uplink transmission.}

Let \( \gamma_b \equiv E^{(b)}_av / N_0 \), where \( E^{(b)}_av \) is the average transmitted energy-per-bit. Then, using (4.8), the fact that \( E^{(s)}_av = (\log_2 M) E^{(b)}_av \), \( \sigma^2 = N_0 / p^2 \), and substituting for \( p \), the
unconditional probability of symbol error given in (4.23) can be expressed as

\[ P_j^{(TH)}(E) = \left( \frac{1}{\sqrt{M}} \right)^{2N_U} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \sum_{\hat{d}_{c,1} \in A} \cdots \sum_{\hat{d}_{c,N_U} \in A} \sum_{\hat{d}_{s,1} \in A} \cdots \sum_{\hat{d}_{s,N_U} \in A} \]

\[
\times \left( 1 - \sum_{i=-\infty}^{\infty} \left\{ \Phi \left( \left( 2\sqrt{M}i + 1 - \hat{e}_{j}^{(R)}(n) \right) \left( \frac{\log_2 M}{\gamma_{b}} \right)^{1/2} \right) \right\} \right)
\times f_{h_{1,1}}(a_{1,1}) \cdots f_{h_{N_U,N_T}}(a_{N_U,N_T}) \cdot f_{\Delta h_{1,1}}(\Delta a_{1,1}) \cdots f_{\Delta h_{N_U,N_T}}(\Delta a_{N_U,N_T})
\times d\Delta a_{1,1} \cdots d\Delta a_{N_U,N_T}.
\]

Note that the \{\hat{r}_{i,i}\} and \hat{Q} can be written in terms of the elements of \( H \) and \( \Delta H \), by using the Gram-Schmidt orthogonalization procedure.

For notational convenience, set \( G \triangleq (\hat{H})^H \). By the Gram-Schmidt orthogonalization procedure, let

\[ u_1 = g_1, \quad u_n = g_n - \sum_{j=1}^{n-1} (e_j^H g_n) e_j, \quad n = 2, \ldots, N_U, \] (4.24)

where

\[ e_n = \frac{u_n}{|u_n|}, \quad n = 1, \ldots, N_U, \]

\( (e_j^H g_n) e_j \) is the projection of \( g_n \) onto \( e_j \), and \( |u_n| = \sqrt{\sum_{i=1}^{N_T} |u_{i,n}|^2} \) is the \( l^2 \)-norm of vector \( u_n \). Alternatively, (4.24) can be written as

\[ g_1 = u_1, \quad g_n = u_n + \sum_{j=1}^{n-1} (e_j^H g_n) e_j, \quad n = 2, \ldots, N_U. \] (4.25)
Using the fact that $u_n = e_n |u_n|$, (4.25) can be written in matrix form as

$$G = [g_1, \ldots, g_{NU}]$$

$$= [e_1, \ldots, e_{NU}] Q \triangleq \hat{R}$$

Therefore, the diagonal elements of the upper-right triangular matrix $R$, are given by

$$r_{n,n} = |u_n|$$

$$= \left[ \left( g_n - \sum_{j=1}^{n-1} (e_j^H g_n) e_j \right)^H \left( g_n - \sum_{j=1}^{n-1} (e_j^H g_n) e_j \right) \right]^{1/2}$$

$$= \left[ |g_n|^2 - 2 \sum_{j=1}^{n-1} |e_j^H g_n|^2 + \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} (e_i^H g_n) (e_j^H g_n) (e_i^H e_j) \right]^{1/2}$$

$$= \left[ |g_n|^2 - 2 \sum_{j=1}^{n-1} |e_j^H g_n|^2 + \sum_{j=1}^{n-1} |e_j^H g_n|^2 \right]^{1/2}$$

$$= \left[ |g_n|^2 - \sum_{j=1}^{n-1} |e_j^H g_n|^2 \right]^{1/2}$$

for $n = 1, \ldots, NU$, where we have used the fact that the $\{e_j\}$ are orthonormal.

### 4.4 Numerical Results

The unconditional probability of error is evaluated by numerical integration using the variance reducing technique of importance sampling [56]. Since the region of integration is large, and the integrand is small over most of the multidimensional space, then approximating the distribution and sampling accordingly gives more accurate results for a given number of samples than does blind sampling.

Figure 4.2 illustrates the probability of symbol error as a function of the SNR-per-bit with $M = 2$, $N_T = N_U = 2$ and $\sigma_e^2 = 0, 0.001, 0.01, $ and $0.1$. The analytical results obtained by numerical integration coincide with the simulation results. The simulation
Figure 4.2: The probability of symbol error as a function of the SNR-per-bit. Here, $N_T = N_U = 2$ and the variance of the estimation error takes on the values 0, 0.001, 0.01, and 0.1.

results are obtained by averaging the probability of symbol error over 100,000 realizations of the channel. As $\sigma_e^2$ increases, performance degrades significantly compared to the case when the channel is known ($\sigma_e^2 = 0$).

### 4.5 Summary

We derived the effects of channel estimation errors on the performance of a system employing Tomlinson-Harashima precoding and the QR decomposition, operating over MISO frequency-flat fading channels with decentralized receivers. Simulation and numerical results demonstrated the accuracy of the analysis, which quantified the penalty associated with Tomlinson-Harashima precoding and imperfect channel estimation.

### 4.6 Acknowledgements

This chapter is, in part, a reprint of material that is in preparation for submission under the title “The Effects of Channel Estimation Errors on a Nonlinear Precoder for MISO Channels with Decentralized Receivers,” P. Amihood, E. Masry, L. B. Milstein, and J. G. Proakis. The dissertation author was the primary researcher and author of
this publication.
Summary and Possible Extensions

In this dissertation, we analyzed the performance of interference suppression techniques for multiple antenna systems. Due to the interference caused by the superposition of the signals transmitted from the multiple antennas and to the intersymbol interference in the presence of frequency selective fading, both spatial and temporal equalization are needed for interference suppression.

In Chapter 2, we considered a MIMO-SU channel, where the receive antennas are co-located and the system benefits from receive diversity. We analyzed the performance of a multicode direct sequence spread spectrum system operating over a frequency selective fading MIMO channel. The spatial and temporal interference is suppressed by employing a different spreading sequence on each transmit antenna and by RAKE reception, respectively. We obtained an expression for the asymptotic variance and established the asymptotic normality of the multi-antenna interference for sufficiently large processing gain. This was used to derive an expression for the probability of bit error which we evaluated numerically.

We showed that a system with randomly spaced multipath arrivals yields better performance compared to a system with equally spaced multipath arrivals. Also, for a fixed total data rate, we demonstrated the advantage of decreasing the number of transmit antennas while increasing the number of codes, and for a fixed total diversity order, we demonstrated the advantage of decreasing the number of RAKE taps while increasing the number of receive antennas.
In Chapter 3, we considered a MIMO-BC channel, where the users are geographically separated. We assumed that, due to complexity and power constraints, the receivers cannot perform spatial and temporal equalization to suppress the interference. Instead, the transmitter performs spatial and temporal equalization by precoding. We derived the probability of symbol error for a system employing Tomlinson-Harashima precoding and operating over SISO channels with ISI. Using the QR decomposition, we extended this result to MISO frequency selective fading channels with decentralized receivers. We obtained the statistics of the QR decomposition, which enabled us to find an expression for the unconditional probability of symbol error. Furthermore, we investigated the effect of the optimal ordering of the decentralized receivers, obtaining a closed-form expression for the desired probability density function when two transmit antennas and two receivers are employed. We compared the performance of the ordered QR decomposition with the ZF and MMSE linear precoders. Simulation and numerical results demonstrated the accuracy of the analysis, which quantified the penalty associated with Tomlinson-Harashima precoding.

The performance of precoding in MIMO-BC channels is dependent on the accuracy and availability of CSI at the transmitter. In Chapter 4, we derived the effects of channel estimation errors on the performance of a system employing Tomlinson-Harashima precoding and the QR decomposition, operating over MISO frequency-flat fading channels with decentralized receivers. This is the same system as in Chapter 3, except that we did not consider ordering of the decentralized receivers. Simulation and numerical results demonstrated the accuracy of the analysis, which quantified the penalty associated imperfect channel estimation.

There are various possible extensions of the work in this dissertation. In Chapter 2, we derived asymptotic normality in a spread spectrum system based on first principles. The methods and tools that we used can possibly be applied to other similar problems. For example, suppose there are a limited number of dominant out-of-cell interferers in a cellular system, where the interferers each use different spreading sequences. While it may not be accurate to model the interference as asymptotically normal since there may only be a small number of out-of-cell interferers, it may be the case that the interference is asymptotically normal for sufficiently large processing gain.

Another possible extension of Chapter 2 is how to model the self-interference.
While we would like to apply the same techniques that we used to prove asymptotic normality of the multi-antenna interference, we cannot assume the spreading sequence is random, since we know the spreading sequence and we use it for despreading. One possibility is to assume the spreading sequences are known and to design them appropriately. For example, it is possible to select certain families of spreading sequences that have favorable partial cross-correlation properties [68,69]. However, the processing gain required to achieve favorable bounds on the partial cross-correlation may be too large. Therefore, while the conditional variance may converge for very large processing gain, the rate of convergence of the central limit theorem may be too slow.

An immediate extension of Chapters 3 and 4 includes a comprehensive analysis and comparison of linear and nonlinear precoders, with and without channel estimation. Furthermore, it is important to understand the performance of precoding compared to equalizing at the receiver. Whether it is advantageous for the transmitter to perform equalization as opposed to the receiver is a question that can be posed by considering the costs associated with the receiver feeding back an estimate of the channel to the transmitter. On the other hand, the benefits of equalizing at the receiver can be considered in light of the costs associated with using imperfectly detected symbols in order to perform successive interference cancellation.
Appendix A

Proof of Theorem 2.1

Step A. Establishing equivalent asymptotic normality

In order to show $Y_{A,L}^{(i)}$ is asymptotically normal, it is possible to simplify the analysis by ignoring appropriate terms. The simplified random variable we show to be asymptotically normal is only a function of the current symbol. We make use of the following result.

Proposition A.1

For general random sequences $X_L, U_L$, suppose $X_L$ converges in distribution to $N(0, \sigma^2)$, written $X_L \overset{D}{\to} N(0, \sigma^2)$, and $E|U_L| \to 0$ as $L \to \infty$. Then $X_L + U_L \overset{D}{\to} N(0, \sigma^2)$ as $L \to \infty$.

Proof. We need to show $E \left[ e^{jw(X_L + U_L)} \right] \to e^{-w/2\sigma^2}$ as $L \to \infty$. Since

$$|E \left[ e^{jw(X_L + U_L)} \right] - e^{-w/2\sigma^2}| \leq \left| E \left[ e^{jw(X_L + U_L)} - e^{jwX_L} \right] \right| + \left| E \left[ e^{jwX_L} - e^{-w/2\sigma^2} \right] \right|,$$

and $E \left[ e^{jwX_L} \right] \to e^{-w/2\sigma^2}$, it remains to show $\left| E \left[ e^{jw(X_L + U_L)} - e^{jwX_L} \right] \right| \to 0$, as $L \to \infty$. Note that

$$\left| E \left[ e^{jw(X_L + U_L)} - e^{jwX_L} \right] \right| = \left| E \left[ e^{jwX_L} (e^{jwU_L} - 1) \right] \right|$$

$$\leq E \left| e^{jwU_L} - 1 \right|$$

$$\leq |w|E|U_L|$$

$$\to 0, \quad \forall w,$$

since $E|U_L| \to 0$ as $L \to \infty$, and, for any real $x$, $\left| e^{ix} - 1 \right| \leq x$. □
Write
\[ Z_L^{(i)}(m) = \frac{1}{L} \sum_{n=0}^{L-1} b(n)p_i(n+m)q_i(n + m) \]
\[ \triangleq \frac{1}{L} \sum_{n=0}^{L-1} Z_L^{(i)}(m,n), \quad m = -(K-1), \ldots, K-1. \]

Breaking apart the summation over \( n \), we have from (2.5),
\[ Y_{A,L}^{(i)} = \frac{1}{L} \sum_{m=-(K-1)}^{K-1} \alpha_m \sum_{n=K-1-m}^{L-(K-1)-1-m} Z_L^{(i)}(m,n) \]
\[ \triangleq X_L^{(i)} \]
\[ + \frac{1}{L} \sum_{m=-(K-1)}^{K-1} \alpha_m \left[ \sum_{n=0}^{(K-1)-1-m} Z_L^{(i)}(m,n) + \sum_{n=L-(K-1)-1-m}^{L-1} Z_L^{(i)}(m,n) \right], \]
\[ \triangleq U_L^{(i)} \]

Since \( |Z_L^{(i)}(m,n)| \leq 1 \), then \( E |\alpha_m Z_L^{(i)}(m,n)| \leq |\alpha_m| \) and
\[ E \left| U_L^{(i)} \right| \leq \frac{1}{L} \sum_{m=-(K-1)}^{K-1} 2(K-1-m)|\alpha_m| \to 0, \]
as \( L \to \infty \). Therefore, by Proposition A.1, it is sufficient to show \( X_L^{(i)} \) is asymptotically normal.

Substituting for \( Z_L^{(i)}(m,n) \), \( X_L^{(i)} \) can be written as
\[ X_L^{(i)} = \frac{1}{L} \sum_{m=-(K-1)}^{K-1} \alpha_m \sum_{n=K-1-m}^{L-(K-1)-1-m} b(n)q_i(n + m)p_i(n + m) \]
\[ = \frac{1}{L} \sum_{n=L_1}^{L_2} W_n^{(i)} p_i(n), \]
where \( L_1 \triangleq K-1, \quad L_2 \triangleq L-(K-1)-1 \), and
\[ W_n^{(i)} \triangleq \sum_{m=-(K-1)}^{K-1} \alpha_m b(n - m)q_i(n). \]

Note \( X_L^{(i)} \) is a function only of the current symbols, \( \{d_{i,l}(0)\} \).

**Step B. Asymptotic variance**
To compute the variance of $X^{(i)}_L$, conditioned on $\{p_1(n)\}$ and $\{d_{i,l}(0)\}$, consider

$$
\sigma^2_X(L) \triangleq \text{Var}\left[ X^{(i)}_L \Big| \{p_1(n)\}, \{d_{i,l}(0)\} \right]
$$

$$
= \frac{1}{L^2} \sum_{n=L_1}^{L_2} \left( W^{(i)}_n \right)^2 \text{Var}[p_1(n)]
$$

$$
= \frac{1}{L^2} \sum_{n=L_1}^{L_2} \sum_{m_1=-(K-1)}^{K-1} \sum_{m_2=-(K-1)}^{K-1} \alpha_{m_1} \alpha_{m_2} b(n-m_1) b(n-m_2) q^2_i(n)
$$

$$
= \frac{1}{L^2} \sum_{m_1=-(K-1)}^{K-1} \alpha_{m_1}^2 \left[ (L - 2(K-1)) N_c + \sum_{n=L_1}^{L_2} \sum_{l_1=0}^{N_c} \sum_{l_2=0}^{N_c} s_{i,l_1}(n) s_{i,l_2}(n) d_{i,l_1}(0) d_{i,l_2}(0) \right]
$$

$$
+ \frac{1}{L^2} \sum_{n=L_1}^{L_2} \sum_{m_1=-(K-1)}^{K-1} \sum_{m_2=-(K-1)}^{K-1} \sum_{m_1 \neq m_2} \alpha_{m_1} \alpha_{m_2} b(n-m_1) b(n-m_2) q^2_i(n).
$$

Let

$$
B_1 \triangleq N_c \sum_{m_1=-(K-1)}^{K-1} \alpha_{m_1}^2,
$$

and

$$
Q^{(i)}_L \triangleq - \frac{2(K-1) N_c}{L} \sum_{m_1=-(K-1)}^{K-1} \alpha_{m_1}^2 + \frac{1}{L} \sum_{m_1=-(K-1)}^{K-1} \alpha_{m_1}^2 \sum_{n=L_1}^{L_2} \sum_{l_1=0}^{N_c} \sum_{l_2=0}^{N_c} s_{i,l_1}(n) s_{i,l_2}(n) d_{i,l_1}(0) d_{i,l_2}(0)
$$

$$
+ \frac{1}{L} \sum_{n=L_1}^{L_2} \sum_{m_1=-(K-1)}^{K-1} \sum_{m_2=-(K-1)}^{K-1} \sum_{m_1 \neq m_2} \alpha_{m_1} \alpha_{m_2} b(n-m_1) b(n-m_2) q^2_i(n)
$$

$$
= \sum_{j=1}^{3} Q^{(i)}_{L,j}.
$$

Then we can write

$$
L \sigma^2_X(L) = B_1 + Q^{(i)}_L.
$$
Note $Q_{L,1}^{(i)} = \text{const}/L$ while we have yet to determine the behavior of $Q_{L,j}^{(i)}$, $j = 2, 3$, as a function of $L$. Note that $Q_{L,j}^{(i)}$, $j = 2, 3$, are functions of $\{p_1(n)\}$ and $\{d_{i,l}(0)\}$. We show that as $L \to \infty$, these two terms converge to zero. This implies that the conditional variance is asymptotically a constant which is independent of the conditioning random variables. This greatly simplifies the computation of the probability of error since we do not need to average over the conditioning random variables. Moreover, this establishes that the rate of convergence of the conditional variance is $1/L$, which in turn specifies the rate of convergence of the central limit theorem.

The following result specifies the rate of decay of the conditional variance $\sigma_X^2(L)$.

**Proposition A.2**

$$L\sigma_X^2(L) = L \text{Var} \left[ X_L^{(i)} \mid \{p_1(n)\}, \{d_{i,l}(0)\} \right] \to B_1 \text{ almost surely as } L \to \infty.$$ 

**Proof.** We first prove $E \left[ \left| Q_L^{(i)} \right|^4 \right] \leq \text{const}/L^2$ in Appendix B. By Markov’s inequality [70],

$$P\left[ \left| Q_L^{(i)} \right| > \varepsilon \right] \leq \frac{E \left| Q_L^{(i)} \right|^4}{\varepsilon^4} \leq \frac{\text{const}}{L^2 \varepsilon^4}.$$

Take $\varepsilon_L = \frac{1}{L^\nu}$ for some $\nu < \frac{1}{4}$. Then

$$\sum_{L=1}^{\infty} P\left[ \left| Q_L^{(i)} \right| > \varepsilon_L \right] \leq \text{const} \sum_{L=1}^{\infty} \frac{1}{L^{2-4\nu}} < \infty.$$

It follows by the Borel-Cantelli lemma [70] that

$$Q_L^{(i)} = O(\varepsilon_L) = O\left( \frac{1}{L^\nu} \right) \text{ almost surely.}$$

Therefore, the conditional asymptotic variance converges to a constant as a function of $L$, independent of $\{p_1(n)\}$ and $\{d_{i,l}(0)\}$. Furthermore, a valid central limit theorem can be established with rate $O(1/L)$.

**Step C. Asymptotic normality**

We have already shown in Proposition A.2 that

$$L\sigma_X^2(L) \to B_1,$$
with probability one as $L \to \infty$, where $B_1 = N_c \sum_{m=-(K-1)}^{K-1} \alpha_m^2$.

Now consider the conditional characteristic function of $X^{(i)}_L / \sigma_X(L)$,

$$
\psi_L^{(i)}(\omega) = E \left[ \exp \left( j \omega X^{(i)}_L / \sigma_X(L) \right) \mid \{p_1(n)\}, \{d_{i,l}(0)\} \right]
$$

$$
= E \left[ \exp \left( j \omega / \sigma_X(L) \sum_{n=L}^{L_2} W_n^{(i)} p_i(n) \right) \mid \{p_1(n)\}, \{d_{i,l}(0)\} \right]
$$

$$
= \prod_{n=L_1}^{L_2} E \left[ \exp \left( j \omega W_n^{(i)} / \sigma_X(L) \right) / \sigma_X(L) \right) \mid \{p_1(n)\}, \{d_{i,l}(0)\} \right],
$$

since the $\{p_i(n)\}$ are i.i.d. Evaluating the expectation gives

$$
\psi_L^{(i)}(\omega) = \prod_{n=L_1}^{L_2} \cos \left( \frac{\omega W_n^{(i)}}{L \sigma_X(L)} \right).
$$

Letting $u = \omega / (L \sigma_X^2(L))^{\frac{1}{2}}$ and taking the natural logarithm, we need to show, for each fixed $u$,

$$
\sum_{n=L_1}^{L_2} \ln \left[ \cos \left( \frac{u W_n^{(i)}}{\sqrt{L}} \right) \right] \to -\frac{1}{2} u^2 B_1 \text{ as } L \to \infty. \quad (A.1)
$$

We know that $|W_n^{(i)}| \leq N_c \sum_{m=-(K-1)}^{K-1} |\alpha_m| \triangleq c$. This implies $|W_n^{(i)}| \leq c / \sqrt{L}$ uniformly in $n$, which can be made small for large $L$. It is then seen from (A.1) that the argument in the cosine can be assumed small for large $L$.

We need an expansion of $\ln[\cos x]$ for small $x$. The expansions of the cosine and logarithm functions can be bounded by recognizing that they are both alternating series. We have

$$
\cos x = 1 - \frac{x^2}{2!} + E_1,
$$

where $|E_1| \leq \frac{|x|^4}{4}$ for, say, $|x| \leq \frac{\pi}{4}$, and

$$
\ln x = x - 1 + E_2,
$$

where $|E_2| \leq \frac{1}{2} (x-1)^2$, for $|x-1| < 1$.

Hence

$$
\ln[\cos x] = (\cos x - 1) + E_2,
$$
where $|E_2| \leq \frac{1}{2} (\cos x - 1)^2 = \frac{1}{2} (2 \sin^2 \frac{x}{2})^2$. Thus

$$\ln[\cos x] = -\frac{x^2}{2!} + E,$$

where $|E| \leq |E_1| + |E_2| = \frac{|x|^4}{4!} + 2 \sin^4 \frac{x}{2}$, for $|x| \leq \frac{\pi}{4}$.

From (A.1) we then have

$$\sum_{n=L_1}^{L_2} \ln \left[ \cos \left( \frac{u W_n^{(i)}}{\sqrt{L}} \right) \right] = - \sum_{n=L_1}^{L_2} \frac{u^2 \left( W_n^{(i)} \right)^2}{2L} + \tilde{E}$$

$$= -\frac{u^2}{2} \left( L \sigma_X^2(L) \right) + \tilde{E},$$

where

$$|\tilde{E}| \leq \sum_{n=L_1}^{L_2} \left[ \frac{1}{4!} \left( \frac{u W_n^{(i)}}{\sqrt{L}} \right)^4 + 2 \sin^4 \left( \frac{u W_n^{(i)}}{2\sqrt{L}} \right) \right].$$

We show that $\tilde{E} \to 0$ as $L \to \infty$. Using the fact that $|\sin x| \leq |x|$, and $|W_n^{(i)}| \leq c$ uniformly in $n$, gives

$$|\tilde{E}| \leq \frac{(L-2(K-1))}{4!} \left( \frac{uc}{L} \right)^4 + 2(L-2(K-1)) \left( \frac{uc}{2\sqrt{L}} \right)^4$$

$$\leq \frac{u^4 c^4}{4! L} + \frac{u^4 c^4}{8L}$$

$$= O \left( \frac{1}{L} \right).$$

Thus

$$\sum_{n=L_1}^{L_2} \ln \left[ \cos \left( \frac{u W_n^{(i)}}{\sqrt{L}} \right) \right] = -\frac{u^2}{2} \left( L \sigma_X^2(L) \right) + O \left( \frac{1}{L} \right).$$

By Proposition A.2, $L \sigma_X^2(L) \to B_1$ with probability one as $L \to \infty$ and the result follows.

As an alternative proof to establish asymptotic normality, we show that the $r^{th}$ order cumulant of $\sqrt{L}X_L^{(i)}$, conditioned on $\{p_1(n)\}$ and $\{d_{i,0}(0)\}$, tends to zero as $L \to \infty$ for every $r \geq 3$. This implies that $\sqrt{L}X_L^{(i)}$ is conditionally asymptotically normal [71].
We have
\[
\sum_r \left( \sqrt{L} X_L^{(i)} \right) = \frac{1}{L^{r/2}} \sum_{n_1=L_1}^{L_2} \cdots \sum_{n_r=L_1}^{L_2} s_{1,1}(n_1) \cdots s_{1,1}(n_r) p_1(n_1) \cdots p_1(n_r) \sum_{l_1=1}^{N_c} \cdots \sum_{l_r=1}^{N_c} d_{i,l_1}(0) \cdots d_{i,l_r}(0) \cdot \sum_{m_1=-(K-1)}^{K-1} \cdots \sum_{m_r=-(K-1)}^{K-1} \alpha_{m_1} \cdots \alpha_{m_r} \cdot s_{i,l_1}(n_1 + m_1) \cdots s_{i,l_r}(n_r + m_r) \cdot \sum_r \left( p_1(n_1 + m_1), \ldots, p_1(n_r + m_r) \right),
\]
where we exchange the cumulant with the summations over \(n_1, \ldots, n_r, l_1, \ldots, l_r, \) and \(m_1, \ldots, m_r, \) and take the constants outside the cumulant [71]. Noting that the magnitudes of \(s_{i,l}(n), p_1(n), \) and \(d_{i,l}(0)\) are bounded by 1, and using
\[
\alpha_{\text{max}} \triangleq m=-(K-1), \ldots, K-1|\alpha_m|,
\]
the magnitude of the cumulant is bounded by
\[
\left| \sum_r \left( \sqrt{L} X_L^{(i)} \right) \right| \leq \frac{(\alpha_{\text{max}} N_c)^r}{L^{r/2}} \cdot \frac{L_2}{L^{r/2}} \sum_{n_1=L_1}^{L_2} \cdots \sum_{n_r=L_1}^{L_2} \sum_{m_1=-(K-1)}^{K-1} \cdots \sum_{m_r=-(K-1)}^{K-1} \left| \sum_r \left( p_1(n_1 + m_1), \ldots, p_1(n_r + m_r) \right) \right|.
\]
Since \(\{p_1(n)\}\) is an independent sequence in \(n,\) and the cumulant is zero if any subgroup of \(\{p_1(n_1 + m_1), \ldots, p_1(n_r + m_r)\}\) is independent of the remaining group [71], then the only non-zero contribution occurs when \(n_1 + m_1 = \ldots = n_r + m_r.\) Consequently,
\[
\left| \sum_r \left( \sqrt{L} X_L^{(i)} \right) \right| \leq \frac{(\alpha_{\text{max}} N_c)^r}{L^{r/2}} \sum_{m_1=-(K-1)}^{K-1} \cdots \sum_{m_r=-(K-1)}^{K-1} \sum_{n_1=L_1}^{L_2} \cdots \sum_{n_r=L_1}^{L_2} \left| \sum_r \left( p_1(n_1 + m_1), \ldots, p_1(n_r + m_r) \right) \right| \triangleq c_{\text{cum}}((2K-1)\alpha_{\text{max}} N_c)^r \frac{L_2 - L_1 + 1}{L^{r/2}} = A(r) \frac{(L - 2(K-1))}{L^{r/2}} \to 0 \text{ as } L \to \infty,
\]
for \(r \geq 3, \) where (a) follows since the cumulant is the same finite constant for any \(n_1\) and \(m_1,\) and
\[
A(r) \triangleq c_{\text{cum}}((2K-1)\alpha_{\text{max}} N_c)^r.
\]
This proves that $\sqrt{L}X_L^{(i)}$, or equivalently $\sqrt{L}Y_{A,L}^{(i)}$, conditioned on $\{p_1(n)\}$ and $\{d_{i,t}(0)\}$, is conditionally asymptotically normal with zero mean and variance $B_1$, as $L \to \infty$. ■

A.1 Acknowledgements

This appendix is, in part, a reprint of material that has been accepted for publication as “Performance Analysis of High Data Rate MIMO Systems in Frequency Selective Fading Channels,” P. Amihood, E. Masry, L. B. Milstein, and J. G. Proakis, IEEE Transactions on Information Theory, August 2007. The dissertation author was the primary researcher and author of these publications.
Appendix B

Bounding of Moments for Proof of Theorem 2.1

We prove $E \left| Q^{(i)}_L \right|^4 = O(1/L^2)$. Since $Q^{(i)}_{L,1} = \text{const}/L$, we are concerned with the behavior of $Q^{(i)}_{L,j}$, $j = 2, 3$, as a function of $L$. By Minkowski’s inequality [70],

$$\left( E \left| \sum_{j=2}^{3} Q^{(i)}_{L,j} \right|^4 \right)^{1/4} \leq \sum_{j=2}^{3} \left( E \left| Q^{(i)}_{L,j} \right|^4 \right)^{1/4}.$$ 

It is possible to show $E \left| Q^{(i)}_{L,2} \right|^4 = O(1/L^4)$, using the fact that the $\{s_{i,l}(n)\}$ are orthogonal from $n = 0, \ldots, L - 1$, for different $l$.

It remains to show $E \left| Q^{(i)}_{L,3} \right|^4 = O(1/L^2)$. We have

$$E \left[ \left| Q^{(i)}_{L,3} \right|^4 \right] = \frac{1}{L^4} \sum_{m_1=-(K-1)}^{K-1} \sum_{m'_1=-(K-1)}^{K-1} \cdots \sum_{m_4=-(K-1)}^{K-1} \sum_{m'_4=-(K-1)}^{K-1} \alpha_{m_1} \alpha_{m'_1} \cdots \alpha_{m_4} \alpha_{m'_4}$$

$$\cdot \sum_{n_1=L_1}^{L_2} \sum_{n_4=L_1}^{L_2} s_{1,1}(n_1 - m_1)s_{1,1}(n_1 - m'_1) \cdots s_{1,1}(n_4 - m_4)s_{1,1}(n_4 - m'_4) \cdot E \left[ q_i^2(n_1) \cdots q_i^2(n_4) \right].$$

We need to evaluate

$$E \left[ p_1(n_1 - m_1)p_1(n_1 - m'_1) \cdots p_1(n_4 - m_4)p_1(n_4 - m'_4) \right] = \left( B.1 \right)$$

and

$$E \left[ p_1(n_1 - m_1)p_1(n_1 - m'_1) \cdots p_1(n_4 - m_4)p_1(n_4 - m'_4) \right] = \left( B.2 \right)$$

92
We want to show that (B.2) collapses the summations in (B.1), over \( n_1, n_2, n_3, \) and \( n_4, \) to at most two summations.

The expectation in (B.2) is non-zero when the arguments of the variables \( p_1(\cdot) \) are equal in even combinations. For example, if four of the arguments are equal, then it may be possible for the remaining four arguments to be equal but distinct from the first four. In this case, the expectation in (B.2) breaks down into the product of two expectations with four variables in each. We will denote this grouping as 4/4. The following groupings are possible: 8, 6/2, 4/4, 4/2/2, 2/2/2/2.

Note

\[
E[p_1(n_j - m_j)p_1(n_j - m'_j)] = 0, \quad \text{(B.3)}
\]

since \( m_j \neq m'_j. \) Therefore, by (B.3), the 8 and 6/2 groupings are identically zero.

Again, by (B.3), any grouping that requires taking the expectation of the product of four variables is identically zero unless the arguments of the four variables have different \( n_i, \ i = 1, \ldots, 4. \) This applies to the 4/4 and 4/2/2 groupings. Suppose the arguments of the four variables under the expectation do not have distinct \( n_i. \) Then one possible combination is

\[
E[p_1(n_i - m_i)p_1(n_i - m'_i)p_1(n_j - m_j)p_1(n_k - m_k)] = 0,
\]

by (B.3). Consequently, the non-zero contributions of the 4/4 and 4/2/2 groupings occur when the arguments of the four variables under the expectation have distinct \( n_i. \) This collapses the summations over \( n_i, \ i = 1, \ldots, 4, \) in (B.1), to one summation, and it is easy to show that all such contributions are of order \( O(1/L^3). \)

It remains to show the contribution of any 2/2/2/2 grouping is of order \( O(1/L^2). \) By (B.3), the variables \( p_1(n_i - m_i) \) and \( p_1(n_i - m'_i) \), for any \( i = 1, 2, 3, 4, \) must fall under separate expectations. There are two possible cases:

(i) \( E[p_1(n_i - m_i)p_1(n_j - m_j)]E[p_1(n_i - m'_i)p_1(n_j - m'_j)] \)
\[
\cdot E[p_1(n_k - m_k)p_1(n_l - m_l)]E[p_1(n_k - m'_k)p_1(n_l - m'_l)] \quad \text{or}
\]
(ii) \( E[p_1(n_i - m_i)p_1(n_j - m_j)]E[p_1(n_i - m'_i)p_1(n_k - m_k)] \)
\[
\cdot E[p_1(n_l - m_l)p_1(n_j - m'_j)]E[p_1(n_l - m'_l)p_1(n_k - m'_k)],
\]
where \( i \neq j \neq k \neq l. \) Note we ignore all permutations where we swap \( p_1(n_v - m_v) \) and \( p_1(n_v - m'_v) \), \( v = i, j, k, \) or \( l, \) since they yield a contribution of the same order.
An example of case (i) is

\[ E[p_1(n_1 - m_1)p_1(n_2 - m_2)]E[p_1(n_1 - m'_1)p_1(n_2 - m'_2)] \]
\[ \cdot E[p_1(n_3 - m_3)p_1(n_4 - m_4)]E[p_1(n_3 - m'_3)p_1(n_4 - m'_4)] \]
\[ = \delta(n_1 - m_1, n_2 - m_2)\delta(n_1 - m'_1, n_2 - m'_2)\delta(n_3 - m_3, n_4 - m_4)\delta(n_3 - m'_3, n_4 - m'_4). \]

It is enough to consider this example since the other permutations will result in the same behaviour in \( L \), with the only difference being a change in parameters. Let \( I_1 \) be this contribution to \( E \left[ Q_{L,3}^{(i)} \right]^4 \).

Using the fact that \( |s_{i,t}(n)| \leq 1 \) and \( E[ q_{i_1}^2(n_1) \cdots q_{i_4}^2(n_4) ] \leq N_c \), the magnitude of \( I_1 \) is bounded by

\[
|I_1| \leq \frac{N_c}{L^4} \sum_{m_1=-(K-1)}^{K-1} \sum_{m_1 \neq m'_1}^{K-1} \sum_{m_4=-(K-1)}^{K-1} \sum_{m_4 \neq m'_4}^{K-1} |\alpha_{m_1} \alpha_{m'_1} \cdots \alpha_{m_4} \alpha_{m'_4}| \\
\cdot \frac{L_2}{L_1} \cdots \frac{L_2}{L_1} \\
\cdot \delta(n_1 - m_1, n_2 - m_2)\delta(n_1 - m'_1, n_2 - m'_2)\delta(n_3 - m_3, n_4 - m_4)\delta(n_3 - m'_3, n_4 - m'_4) \\
\leq \frac{N_c(2(L - 2(K - 1))^2)}{L^4} \sum_{m_1=-(K-1)}^{K-1} \sum_{m_1 \neq m'_1}^{K-1} \sum_{m_4=-(K-1)}^{K-1} \sum_{m_4 \neq m'_4}^{K-1} \\
\cdot |\alpha_{m_1} \alpha_{m'_1} \cdots \alpha_{m_4} \alpha_{m'_4}| \delta(m_1 - m'_1, m_2 - m'_2)\delta(m_3 - m'_3, m_4 - m'_4) \\
= O \left( \frac{1}{L^2} \right),
\]

where we bound the summations over \( n_1, \ldots, n_4 \), and substitute the values of \( L_1 \) and \( L_2 \) in terms of \( L \).

In case (ii), since \( n_j = n_i - m_i + m_j \) and \( n_k = n_i - m'_i + m_k \), \( j \neq k \), at most two summations over \( n_i \), \( i = 1, 2, 3, 4 \), remain. Consequently, the contribution of case (ii) is also \( O(1/L^2) \).

Therefore, \( E \left[ Q_{L,3}^{(i)} \right]^4 = O(1/L^2) \). Finally, \( E \left[ Q_L^{(i)} \right]^4 = O(1/L^2) \).  

\( \blacksquare \)
B.1 Acknowledgements

This appendix is, in part, a reprint of material that has been accepted for publication as “Performance Analysis of High Data Rate MIMO Systems in Frequency Selective Fading Channels,” P. Amihood, E. Masry, L. B. Milstein, and J. G. Proakis, *IEEE Transactions on Information Theory*, August 2007. The dissertation author was the primary researcher and author of these publications.
Appendix C

Proof of Theorem 2.2

Step A. Establishing equivalent asymptotic normality

Write

\[
Z_{W,L}^{(j)}(k) = \frac{1}{L} \sum_{n=0}^{L-1} b(n) w_j(n + k), \quad k = 0, \ldots, K - 1.
\]

Breaking apart the summation over \(n\),

\[
Y_{W,L}^{(j)} = \frac{1}{L} \sum_{k=0}^{K-1} \alpha_k \sum_{n=K-1-k}^{L-1-(K-1)-k} Z_{W,L}^{(j)}(k, n) \quad \triangleq X_{W,L}^{(j)}
\]

\[
+ \frac{1}{L} \sum_{k=0}^{K-1} \alpha_k \left[ \sum_{n=0}^{(K-1)-k-1} Z_{W,L}^{(j)}(k, n) + \sum_{n=L-(K-1)-k}^{L-1} Z_{W,L}^{(j)}(k, n) \right] \quad \triangleq U_{W,L}^{(j)}.
\]

Since \(|Z_{W,L}^{(j)}(k, n)| \leq 1\), then \(E \left| \alpha_k Z_{W,L}^{(j)}(k, n) \right| \leq |\alpha_k|\) and

\[
E \left| U_{W,L}^{(j)} \right| \leq \frac{1}{L} \sum_{k=0}^{K-1} 2(K - 1 - k)|\alpha_k| \to 0,
\]

as \(L \to \infty\). Therefore, by Proposition A.1, it is sufficient to show \(X_{W,L}^{(j)}\) is asymptotically normal.

Step B. Asymptotic variance
Substituting for \( Z_{W,L}^{(j)}(k,n) \), \( X_{W,L}^{(j)} \) can be written as

\[
X_{W,L}^{(j)} = \frac{1}{L} \sum_{n=L_1}^{L_2} \left[ \sum_{k=0}^{K-1} \alpha_k b(n-k) \right] w_j(n),
\]

where \( L_1 \triangleq K - 1 \) and \( L_2 \triangleq L - (K - 1) - 1 \).

Since the \( \{w_j(n)\} \) are independent for different \( n \), the variance of \( X_{W,L}^{(j)} \), conditioned on \( \{p_1(n)\} \), is given by

\[
\text{Var}_L \left[ X_{W,L}^{(j)} \bigg| \{p_1(n)\} \right] = \frac{\sigma_w^2}{L^2} \sum_{n=L_1}^{L_2} \sum_{k_1=0}^{K-1} \sum_{k_2=0}^{K-1} \alpha_{k_1} \alpha_{k_2} b(n-k_1)b(n-k_2)
\]

\[
= \frac{\sigma_w^2}{L^2} \left[ \sum_{k_1=0}^{K-1} \sum_{k_2=0}^{K-1} \alpha_{k_1} \alpha_{k_2} b(n-k_1)b(n-k_2) \right]
\]

\[
= \frac{\sigma_w^2}{L^2} \left[ (L - 2(K - 1)) \sum_{k_1=0}^{K-1} \alpha_{k_1}^2 + \sum_{n=L_1}^{L_2} \sum_{k_1=0}^{K-1} \sum_{k_2=0}^{K-1} \alpha_{k_1} \alpha_{k_2} b(n-k_1)b(n-k_2) \right]
\]

Let

\[
B_2 \triangleq \frac{\sigma_w^2}{L^2} \sum_{k_1=0}^{K-1} \alpha_{k_1}^2, \quad \text{and}
\]

\[
Q_L \triangleq \frac{-2(K - 1)\sigma_w^2}{L} \sum_{k_1=0}^{K-1} \alpha_{k_1}^2 + \sum_{n=L_1}^{L_2} \sum_{k_1=0}^{K-1} \sum_{k_2=0}^{K-1} \alpha_{k_1} \alpha_{k_2} b(n-k_1)b(n-k_2)
\]

\[
= \sum_{i=1}^{2} Q_{L}^{(i)}
\]

Then we can write

\[
L \text{Var}_L \left[ X_{W,L}^{(j)} \bigg| \{p_1(n)\} \right] = B_2 + Q_L.
\]

**Proposition C.1**

\[
L \text{Var}_L \left[ X_{W,L}^{(j)} \bigg| \{p_1(n)\} \right] \to B_2 \text{ almost surely as } L \to \infty.
\]

**Proof.** The proof is the same as that of Proposition A.2 except that it is necessary to show \( E \left[ |Q_L|^4 \right] \leq \text{const}/L^2 \), which is proved in Appendix D. \( \blacksquare \)
Step C. Asymptotic normality

Since \( X_{W,L}^{(j)} \) is a linear combination of independent normal random variables, as seen in (C.1), then \( X_{W,L}^{(j)} \), conditioned on \( \{p_1(n)\} \), is conditionally asymptotically normal with variance \( B_2 \), as \( L \to \infty \). This proves that \( \sqrt{L}Y_{W,L}^{(j)} \), conditioned on \( \{p_1(n)\} \), is conditionally asymptotically normal with zero mean and variance \( B_2 \), as \( L \to \infty \). ■

C.1 Acknowledgements

This appendix is, in part, a reprint of material that has been accepted for publication as “Performance Analysis of High Data Rate MIMO Systems in Frequency Selective Fading Channels,” P. Amihood, E. Masry, L. B. Milstein, and J. G. Proakis, IEEE Transactions on Information Theory, August 2007. The dissertation author was the primary researcher and author of these publications.
Appendix D

Bounding of Moments for Proof of Theorem 2.2

We prove $E|Q_L|^4 = O(1/L^2)$. Note $Q_L^{(1)} = \text{const}/L$ and

$$E \left[ |Q_L^{(2)}|^4 \right] = \frac{\sigma^8}{L^4} \sum_{k_1=0}^{K-1} \sum_{k'_1=0}^{K-1} \cdots \sum_{k_4=0}^{K-1} \sum_{k'_4=0}^{K-1} \alpha_{k_1} \alpha_{k'_1} \cdots \alpha_{k_4} \alpha_{k'_4} \sum_{n_1=L_1}^{L_2} \cdots \sum_{n_4=L_1}^{L_2} \cdot s_{1,1}(n_1 - k_1)s_{1,1}(n_1 - k'_1) \cdots s_{1,1}(n_4 - k_4)s_{1,1}(n_4 - k'_4)$$

$$\cdot E \left[ p_1(n_1 - k_1)p_1(n_1 - k'_1) \cdots p_1(n_4 - k_4)p_1(n_4 - k'_4) \right].$$

Since the expectation is the same as (B.2), we have $E \left[ |Q_L|^4 \right] = O(1/L^2).$ ■

D.1 Acknowledgements

This appendix is, in part, a reprint of material that has been accepted for publication as “Performance Analysis of High Data Rate MIMO Systems in Frequency Selective Fading Channels,” P. Amihood, E. Masry, L. B. Milstein, and J. G. Proakis, IEEE Transactions on Information Theory, August 2007. The dissertation author was the primary researcher and author of these publications.
Appendix E

Asymptotic Statistics of the Multi-Antenna Interference with Randomly Spaced Multipath Arrivals

Recall the channel is now described by

\[ h_{i,j}(t) = \sum_{k=0}^{K-1} g_{i,j,k} \delta(t - kT_{ch} - \tau_{i,j}), \]

where \( \tau_{i,j} \) is the delay between transmit antenna \( i \) and receive antenna \( j \). The \( \{\tau_{i,j}\} \) are i.i.d. in \( i \) and \( j \), and are each uniformly distributed random variables in \([0, T_{ch}]\).

The complex low-pass-equivalent received signal at the \( j^{th} \) antenna is

\[ \tilde{y}_j(t) = \sum_{i=1}^{N_t} \tilde{x}_i(t) * h_{i,j}(t) + \tilde{w}_j(t), \]

where \( \tilde{w}_j(t) \) is the AWGN process.

In designing the chip matched filter, we need to take into account the differential delay from the desired transmit antenna (\( i = 1 \)) to receive antenna \( j \). Chip matched
filtering gives

\[ y_j(n) = \frac{1}{\sqrt{T_{ch}}} \int_{nT_{ch} + \tau_{1,j}}^{(n+1)T_{ch} + \tau_{1,j}} \tilde{y}_j(t) \, dt \]

\[ = \frac{1}{\sqrt{T_{ch}}} \int_{nT_{ch} + \tau_{1,j}}^{(n+1)T_{ch} + \tau_{1,j}} \sum_{i=1}^{N_t} \sum_{k=0}^{K-1} g_{i,j,k} \tilde{p}_i(t - kT_{ch} - \tau_{i,j}) \]

\[ \cdot \sum_{l=1}^{N_c} \tilde{s}_{i,l}(t - kT_{ch} - \tau_{i,j}) \tilde{d}_{i,l}(t - kT_{ch} - \tau_{i,j}) \, dt + w_j(n) \]

\[ = \sum_{i=1}^{N_t} \sum_{k=0}^{K-1} g_{i,j,k} D_{i,j,k}(n) + w_j(n), \]

where

\[ D_{i,j,k}(n) \]

\[ \triangleq \frac{1}{\sqrt{T_{ch}}} \int_{nT_{ch} + \tau_{1,j}}^{(n+1)T_{ch} + \tau_{1,j}} \tilde{p}_i(t - kT_{ch} - \tau_{i,j}) \sum_{l=1}^{N_c} \tilde{s}_{i,l}(t - kT_{ch} - \tau_{i,j}) \tilde{d}_{i,l}(t - kT_{ch} - \tau_{i,j}) \, dt. \]

Since we transmit rectangular pulses, it is easy to evaluate the integral. Figure E.1 shows the regions of integration.
Figure E.1: Regions of integration for matched filtering in the case of randomly spaced multipath arrivals.
Evaluating the integral gives

\[ D_{i,j,k}(n) = \begin{cases} 
(\tau_{i,j} - \tau_{1,j}) & \left( \frac{1}{\sqrt{T_{ch}}} p_i(n-k-1) \sum_{l=1}^{N_c} s_{i,l}(n-k-1) \right) \\
+ (T_{ch} - (\tau_{i,j} - \tau_{1,j})) & \left( \frac{1}{\sqrt{T_{ch}}} p_i(n-k) \sum_{l=1}^{N_c} s_{i,l}(n-k) \right) \\
& \frac{1}{\sqrt{T_{ch}}} p_i(n-k+1) \sum_{l=1}^{N_c} s_{i,l}(n-k+1) \right) 
\end{cases} 
\]

if \( \tau_{i,j} > \tau_{1,j} \)

\[ \begin{align*}
& (\tau_{i,j} - \tau_{1,j}) \left( \frac{1}{T_{ch}} p_i(n-k-1) \sum_{l=1}^{N_c} s_{i,l}(n-k-1) \right) \\
& + (1 - (\tau_{i,j} - \tau_{1,j})) \left( \frac{1}{T_{ch}} p_i(n-k) \sum_{l=1}^{N_c} s_{i,l}(n-k) \right) \\
& + (\tau_{i,j} - \tau_{1,j}) \left( \frac{1}{T_{ch}} p_i(n-k+1) \sum_{l=1}^{N_c} s_{i,l}(n-k+1) \right) 
\end{align*} \]

\[ = \begin{cases} 
\frac{(\tau_{i,j} - \tau_{1,j})}{T_{ch}} & (p_i(n-k-1) \sum_{l=1}^{N_c} s_{i,l}(n-k-1) \right) \\
+ (1 - (\tau_{i,j} - \tau_{1,j})) & (p_i(n-k) \sum_{l=1}^{N_c} s_{i,l}(n-k) \right) \\
+ (\tau_{i,j} - \tau_{1,j}) & (p_i(n-k+1) \sum_{l=1}^{N_c} s_{i,l}(n-k+1) \right) 
\end{cases} \]

\[ \tau_{i,j} > \tau_{1,j} \]

\[ \begin{align*}
& (\tau_{i,j} - \tau_{1,j}) \left( \frac{1}{T_{ch}} p_i(n-k-1) \sum_{l=1}^{N_c} s_{i,l}(n-k-1) \right) \\
& + (1 - (\tau_{i,j} - \tau_{1,j})) \left( \frac{1}{T_{ch}} p_i(n-k) \sum_{l=1}^{N_c} s_{i,l}(n-k) \right) \\
& + (\tau_{i,j} - \tau_{1,j}) \left( \frac{1}{T_{ch}} p_i(n-k+1) \sum_{l=1}^{N_c} s_{i,l}(n-k+1) \right) 
\end{align*} \]

\[ \tau_{i,j} < \tau_{1,j} \]

where

\[ D_{i,k}(n) \triangleq p_i(n-k) \sum_{l=1}^{N_c} s_{i,l}(n-k) d_{i,l} \left( \left\lfloor \frac{n-k}{L} \right\rfloor \right) = p_i(n-k) q_i(n-k), \]

\[ c_{i,j}^{(0)} \triangleq \frac{\left| \tau_{i,j} - \tau_{1,j} \right|}{T_{ch}}, \]

\[ c_{i,j}^{(1)} \triangleq 1 - \frac{\left| \tau_{i,j} - \tau_{1,j} \right|}{T_{ch}}, \]

At the output of the RAKE and despreader, the contribution due to the multi-antenna interference is \( Z_{A,L} \). This can be written as

\[ Z_{A,L} = \frac{1}{L} \sum_{n=0}^{L-1} s_1(n) \sum_{j=1}^{N_r} p_1(n) \sum_{k_1=0}^{K-1} g_{1,j,k_1} y_j(n + k) \]

\[ = \frac{1}{L} \sum_{n=0}^{L-1} s_1(n) \sum_{j=1}^{N_r} p_1(n) \sum_{k_1=0}^{K-1} g_{1,j,k_1} \sum_{i=2}^{N_t} \sum_{k_2=0}^{K-1} g_{i,j,k_2} D_{i,j,k_2-k_1}(n). \]
Rearranging terms gives

\[ Z_{A,L} = \sum_{j=1}^{N_r} \sum_{i=2}^{N_t} \sum_{k_1=0}^{K-1} K-1 \sum_{j=1}^{N_r} \sum_{i=2}^{N_t} \sum_{k_1=0}^{K-1} g_{1,j,k_1}^* g_{i,j,k_2} \frac{1}{L} \sum_{n=0}^{L-1} s_{1,1}(n) p_1(n) D_{i,j,k_2-k_1}(n) \]

= \frac{1}{L} \sum_{n=0}^{L-1} b(n) D_{i,j,k_2-k_1}(n),

where we set \( b(n) \triangleq s_{1,1}(n) p_1(n) \). Note

\[
\sum_{n=0}^{L-1} b(n) \tilde{D}_{i,j}(n) = \sum_{n=0}^{L-2} b(n+1) \tilde{D}_{i,j}(n) + (b(0) \tilde{D}_{i,j}(n) + b(0) \tilde{D}_{i,j}(n)) + b(L-1) \tilde{D}_{i,j}(n) - b(n) \tilde{D}_{i,j}(n)
\]

\[
\sum_{n=0}^{L-1} b(n) \tilde{D}_{i,j}(n+1) = \sum_{n=0}^{L-2} b(n) \tilde{D}_{i,j}(n) + b(L-2) \tilde{D}_{i,j}(n) + b(L-1) \tilde{D}_{i,j}(n) + b(L-1) \tilde{D}_{i,j}(n+1)
\]

Then

\[
\sum_{n=0}^{L-1} b(n) D_{i,j,k}(n)
\]

\[
= \begin{cases} \sum_{n=0}^{L-2} b(n) \tilde{D}_{i,j}(n) + I_{i,j,k}^{(0)} & \text{if } \tau_{i,j} > \tau_{1,j} \\ \sum_{n=0}^{L-2} b(n) \tilde{D}_{i,j}(n) + I_{i,j,k}^{(0)} & \text{if } \tau_{i,j} < \tau_{1,j} \\ \sum_{n=0}^{L-2} b(n) \tilde{D}_{i,j}(n) + I_{i,j,k}^{(0)} & \text{if } \tau_{i,j} = \tau_{1,j} \\ \sum_{n=0}^{L-2} b(n) \tilde{D}_{i,j}(n) + I_{i,j,k}^{(0)} & \text{if } \tau_{i,j} < \tau_{1,j} \end{cases}
\]

\[
= \sum_{n=0}^{L-2} b(n) \tilde{D}_{i,j}(n) + I_{i,j,k}^{(0)} + I_{i,j,k}^{(1)}(\tau_{i,j})
\]

where

\[
1_A(x) = \begin{cases} 0 & \text{if } x \notin A \\ 1 & \text{if } x \in A \end{cases}
\]

is the indicator function of \( A \),

\[
b_{j}^{(0)}(n) \triangleq c_{i,j}^{(0)} b(n+1) + c_{i,j}^{(1)} b(n)
\]

\[
b_{j}^{(1)}(n) \triangleq c_{i,j}^{(1)} b(n) + c_{i,j}^{(0)} b(n-1)
\]
and

\[ I_{i,j,k}^{(0)} \triangleq c_{i,j}^{(0)}(b(0)\tilde{D}_{i,k}(-1) + b(1)\tilde{D}_{i,k}(0)) + c_{i,j}^{(1)}(b(0)\tilde{D}_{i,k}(0) + b(L-1)\tilde{D}_{i,k}(L-1)) \]

\[ I_{i,j,k}^{(1)} \triangleq c_{i,j}^{(1)}(b(0)\tilde{D}_{i,k}(0) + b(L-1)\tilde{D}_{i,k}(L-1)) \]

\[ + c_{i,j}^{(0)}(b(L-2)\tilde{D}_{i,k}(L-1) + b(L-1)\tilde{D}_{i,k}(L)). \]

Then we can write

\[
Z_{A,L} = \sum_{j=1}^{N_r} \sum_{i=2}^{N_t} \sum_{k_1=0}^{K-1} \sum_{k_2=0}^{K-1} g_{i,j,k_1} g_{i,j,k_2} 1_L \left[ \sum_{n=1}^{L-2} b_j^{(1)(\tau_{1,j})(\tau_{i,j})}(n)\tilde{D}_{i,k_2-k_1}(n) + I_{i,j,k_2-k_1}^{(1)(\tau_{1,j})(\tau_{i,j})} \right] \\
= \sum_{j=1}^{N_r} \sum_{i=2}^{N_t} \left[ \sum_{k_1=0}^{K-1} g_{i,j,k_1} g_{i,j,k_1} 1_L \left[ \sum_{n=1}^{L-2} b_j^{(1)(\tau_{1,j})(\tau_{i,j})}(n)\tilde{D}_{i,0}(n) + I_{i,j,0}^{(1)(\tau_{1,j})(\tau_{i,j})} \right] \right] \\
+ \sum_{k_1=0}^{K-1} \sum_{k_2=0}^{K-1} g_{i,j,k_1} g_{i,j,k_2} 1_L \left[ \sum_{n=1}^{L-2} b_j^{(1)(\tau_{1,j})(\tau_{i,j})}(n)\tilde{D}_{i,k_2-k_1}(n) + I_{i,j,k_2-k_1}^{(1)(\tau_{1,j})(\tau_{i,j})} \right] \\
\triangleq Z_{L}^{(i,j)}(0) + Z_{L}^{(i,j)}(k_1-k_2) \]

The \( \{Z_{L}^{(i,j)}(m)\} \) are again independent for different \( i \). Previously, for fixed \( i \) and different \( j \), the \( \{Z_{L}^{(i,j)}(m)\} \) were identical. Now, for fixed \( i \), and different \( j \), the \( \{Z_{L}^{(i,j)}(m)\} \) are not identical due to the delays. It suffices to establish the joint asymptotic normality \( \{Z_{L}^{(i,j)}(m)\} \), for fixed \( i \).

As before, we use the Cramér device

\[ Y_{A,L}^{(i)} = \sum_{j=1}^{N_r} \beta_j \sum_{m=-(K-1)}^{K-1} \alpha_m Z_{L}^{(i,j)}(m). \]

Now, we can write \( Y_{A,L}^{(i)} = X_L^{(i)} + U_L^{(i)} \), where \( U_L^{(i)} \) contains the \( \{I_{i,-m}^{(1)(\tau_{1,j})(\tau_{i,j})}\} \), and we can easily show \( E[U_L^{(i)}] \to 0 \) as \( L \to \infty \). Then we only need to consider \( X_L^{(i)} \).

Computing the variance of \( X_L^{(i)} \), we have

\[
\text{Var} \left[ X_L^{(i)} \bigg| \{p_1(n)\}, \{d_{i,l}(0)\}, \{\tau_{i,j}\} \right] = \frac{1}{L^2} \sum_{n=L_1}^{L_2} \left( W_n^{(i)} \right)^2 \text{Var}[p_i(n)]
\]

where

\[ W_n^{(i)} \triangleq \sum_{j=1}^{N_r} \beta_j \sum_{m=-(K-1)}^{K-1} \alpha_m b_j^{(1)(\tau_{1,j})(\tau_{i,j})}(n-m)q_i(n). \]
Then

\[
\text{Var}\left[ X_L^{(i)} \mid \{p_1(n), \{d_{i,t} (0)\}, \{\tau_{i,j}\} \right] \\
= \frac{1}{L^2} \sum_{n=L_1}^{L_2} \sum_{j_1=1}^{N_r} \sum_{j_2=1}^{N_r} \beta_{j_1} \beta_{j_2} \sum_{m_1=-(K-1)}^{K-1} \sum_{m_2=-(K-1)}^{K-1} \alpha_{m_1} \alpha_{m_2} \\
\cdot b_{j_1}^{(1)}(n-m_1) b_{j_2}^{(1)}(n-m_2) q_i^2(n) \\
\triangleq \sum_{j_1=1}^{N_r} \sum_{j_2=1}^{N_r} \beta_{j_1} \beta_{j_2} F_L^{(i)}(j_1, j_2),
\]

where again

\[
b_j^{(0)}(n) \triangleq c_{i,j}(0) b(n+1) + c_{i,j}(1) b(n)
\]

and

\[
b_j^{(1)}(n) \triangleq c_{i,j}(1) b(n) + c_{i,j}(0) b(n-1).
\]

**Case 1:** \(1_{[0, \tau_{1,j_1}]}(\tau_{i,j_1}) = 1_{[0, \tau_{1,j_2}]}(\tau_{i,j_2}) = 0

\[
F_L^{(i)}(j_1, j_2) \\
= \frac{1}{L^2} \sum_{m_1=-(K-1)}^{K-1} \alpha_{m_1}^2 \left[ c_{i,j_1}^{(0)} c_{i,j_2}^{(0)} + c_{i,j_1}^{(1)} c_{i,j_2}^{(1)} \right] \sum_{n=L_1}^{L_2} q_i^2(n) \\
+ \left[ c_{i,j_1}^{(0)} c_{i,j_2}^{(1)} + c_{i,j_1}^{(1)} c_{i,j_2}^{(0)} \right] \sum_{n=L_1}^{L_2} b(n+1-m_1) b(n-m_1) q_i^2(n) \\
+ \frac{1}{L^2} \sum_{n=L_1}^{L_2} \sum_{j_1=1}^{N_r} \sum_{j_2=1}^{N_r} \beta_{j_1} \beta_{j_2} \sum_{m_1=-(K-1)}^{K-1} \sum_{m_2=-(K-1)}^{K-1} \alpha_{m_1} \alpha_{m_2} \\
\cdot b_{j_1}^{(0)}(n-m_1) b_{j_2}^{(0)}(n-m_2) q_i^2(n).
\]

Note

\[
\sum_{n=L_1}^{L_2} q_i^2(n) = (L_2 - L_1 + 1) N_c + \sum_{n=L_1}^{L_2} \sum_{l_1=0}^{N_c} \sum_{l_2=0}^{N_c} s_{i,l_1}(n) s_{i,l_2}(n) d_{i,l_1}(0) d_{i,l_2}(0).
\]

Then let

\[
LF_L^{(i)}(j_1, j_2) = B_1^{(i)}(j_1, j_2) + Q_L^{(i)}(j_1, j_2),
\]
where

\[
B_{\text{e}}^{(i)}(j_1, j_2) \triangleq N_c \sum_{m_1 = -(K-1)}^{K-1} \alpha_{m_1}^2 \left[ c_{i,j_1}^{(0)} c_{i,j_2}^{(0)} + c_{i,j_1}^{(1)} c_{i,j_2}^{(1)} \right],
\]

\[
Q_L^{(i)}(j_1, j_2) = \sum_{k=1}^{4} Q_{L,k}^{(i)}(j_1, j_2),
\]

\[
Q_{L,1}^{(i)}(j_1, j_2) = -\frac{\text{const} \cdot N_c}{L} \sum_{m_1 = -(K-1)}^{K-1} \alpha_{m_1}^2 \left[ c_{i,j_1}^{(0)} c_{i,j_2}^{(0)} + c_{i,j_1}^{(1)} c_{i,j_2}^{(1)} \right],
\]

\[
Q_{L,2}^{(i)}(j_1, j_2) = \frac{1}{L} \sum_{m_1 = -(K-1)}^{K-1} \alpha_{m_1}^2 \left[ c_{i,j_1}^{(0)} c_{i,j_2}^{(0)} + c_{i,j_1}^{(1)} c_{i,j_2}^{(1)} \right] \cdot \sum_{n=L_1}^{L_2} \sum_{l_1=0}^{N_c} \sum_{l_2=0}^{N_c} s_{i,l_1}(n) s_{i,l_2}(n) d_{i,l_1}(0) d_{i,l_2}(0),
\]

\[
Q_{L,3}^{(i)}(j_1, j_2) = \frac{1}{L} \sum_{m_1 = -(K-1)}^{K-1} \alpha_{m_1}^2 \left[ c_{i,j_1}^{(0)} c_{i,j_2}^{(1)} + c_{i,j_1}^{(1)} c_{i,j_2}^{(0)} \right] \sum_{n=L_1}^{L_2} b(n+1-m_1) b(n-m_1) q_i^2(n),
\]

and

\[
Q_{L,4}^{(i)}(j_1, j_2) = \frac{1}{L} \sum_{n=L_1}^{L_2} \sum_{m_1 = -(K-1)}^{K-1} \sum_{m_2 = -(K-1)}^{K-1} \alpha_{m_1} \alpha_{m_2} b_{j_1}^{(0)}(n-m_1) b_{j_2}^{(0)}(n-m_2) q_i^2(n).
\]

**Case 2:** \( 1_{[0, \tau_{1,j_1}]}(\tau_{i,j_1}) = 1_{[0, \tau_{1,j_2}]}(\tau_{i,j_2}) = 1 \)

\[
F_L^{(i)}(j_1, j_2) = \frac{1}{L^2} \sum_{m_1 = -(K-1)}^{K-1} \alpha_{m_1}^2 \left[ c_{i,j_1}^{(0)} c_{i,j_2}^{(0)} + c_{i,j_1}^{(1)} c_{i,j_2}^{(1)} \right] \sum_{n=L_1}^{L_2} q_i^2(n)
\]

\[
+ \left[ c_{i,j_1}^{(0)} c_{i,j_2}^{(1)} + c_{i,j_1}^{(1)} c_{i,j_2}^{(0)} \right] \sum_{n=L_1}^{L_2} b(n+1-m_1) b(n-m_1) q_i^2(n)
\]

\[
+ \frac{1}{L^2} \sum_{n=L_1}^{L_2} \sum_{m_1 = -(K-1)}^{K-1} \sum_{m_2 = -(K-1)}^{K-1} \alpha_{m_1} \alpha_{m_2} b_{j_1}^{(1)}(n-m_1) b_{j_2}^{(1)}(n-m_2) q_i^2(n).
\]

Here, \( B_{\text{e}}^{(i)}(j_1, j_2) \) is the same as in Case 1, but \( Q_L^{(i)}(j_1, j_2) \) is slightly different. Since \( Q_L^{(i)}(j_1, j_2) \) is functionally the same, it can be argued it is sufficient to bound Case 1.

**Case 3:** \( 1_{[0, \tau_{1,j_1}]}(\tau_{i,j_1}) = 0 \) and \( 1_{[0, \tau_{1,j_2}]}(\tau_{i,j_2}) = 1 \)
\[ F_L^{(i)}(j_1, j_2) = \frac{1}{L^2} \sum_{m_1 = -(K-1)}^{K-1} \alpha_{m_1}^2 \left[ \binom{c}{i,j_1}^{(1)} \binom{c}{i,j_2}^{(1)} \right] \sum_{n=L_1}^{L_2} q_i^2(n) \]

\[ + \left[ \binom{c}{i,j_1}^{(0)} \binom{c}{i,j_2}^{(0)} \right] \sum_{n=L_1}^{L_2} b(n + 1 - m_1)b(n - 1 - m_1)q_i^2(n) \]

\[ + \left[ \binom{c}{i,j_1}^{(0)} \binom{c}{i,j_2}^{(1)} \right] \sum_{n=L_1}^{L_2} b(n + 1 - m_1)b(n - m_1)q_i^2(n) \]

\[ + \left[ \binom{c}{i,j_1}^{(1)} \binom{c}{i,j_2}^{(0)} \right] \sum_{n=L_1}^{L_2} b(n - 1 - m_1)b(n - m_1)q_i^2(n) \]

\[ + \frac{1}{L^2} \sum_{n=L_1}^{L_2} \sum_{m_1 = -(K-1)}^{K-1} \sum_{m_2 = -(K-1)}^{K-1} \alpha_{m_1} \alpha_{m_2} b_{j_1}^{(0)}(n - m_1)b_{j_2}^{(1)}(n - m_2)q_i^2(n). \]

Let

\[ B_1^{(i)}(j_1, j_2) \triangleq N_c \sum_{m_1 = -(K-1)}^{K-1} \alpha_{m_1}^2 \left[ \binom{c}{i,j_1}^{(1)} \binom{c}{i,j_2}^{(1)} \right] \]

and

\[ Q_L^{(i)}(j_1, j_2) = \sum_{k=1}^{4} Q_{L,k}^{(i)}(j_1, j_2) \]

where

\[ Q_{L,1}^{(i)}(j_1, j_2) = -\frac{\text{const} \cdot N_c}{L} \sum_{m_1 = -(K-1)}^{K-1} \alpha_{m_1}^2 \left[ \binom{c}{i,j_1}^{(1)} \binom{c}{i,j_2}^{(1)} \right], \]

\[ Q_{L,2}^{(i)}(j_1, j_2) = \frac{1}{L} \sum_{m_1 = -(K-1)}^{K-1} \alpha_{m_1}^2 \left[ \binom{c}{i,j_1}^{(1)} \binom{c}{i,j_2}^{(1)} \right] \sum_{n=L_1}^{L_2} \sum_{n_1=0}^{N_c} \sum_{n_{ij}=0}^{N_c} s_{i,j_1}(n)s_{i,j_2}(n)d_{i,j_1}(0)d_{i,j_2}(0), \]

\[ Q_{L,3}^{(i)}(j_1, j_2) = \frac{1}{L} \sum_{m_1 = -(K-1)}^{K-1} \alpha_{m_1}^2 \sum_{n=L_1}^{L_2} \left[ \binom{c}{i,j_1}^{(0)} \binom{c}{i,j_2}^{(0)} \right] b(n + 1 - m_1)b(n - 1 - m_1)q_i^2(n) \]

\[ + \left[ \binom{c}{i,j_1}^{(0)} \binom{c}{i,j_2}^{(1)} \right] b(n + 1 - m_1)b(n - m_1)q_i^2(n) \]

\[ + \left[ \binom{c}{i,j_1}^{(1)} \binom{c}{i,j_2}^{(0)} \right] b(n - 1 - m_1)b(n - m_1)q_i^2(n). \]

and

\[ Q_{L,4}^{(i)}(j_1, j_2) = \frac{1}{L} \sum_{n=L_1}^{L_2} \sum_{m_1 = -(K-1)}^{K-1} \sum_{m_2 = -(K-1)}^{K-1} \alpha_{m_1} \alpha_{m_2} b_{j_1}^{(0)}(n - m_1)b_{j_2}^{(1)}(n - m_2)q_i^2(n). \]
Here, $B^{(i)}_{1}(j_1, j_2)$ is different. The expression is expected to be different since the ‘previous’ chip and the ‘future’ chip are independent. Again, $Q^{(i)}_{L}(j_1, j_2)$ is similar to Case 1 and it is possible to argue it is bounded as desired.

**Case 4:**

The expression is expected to be different since the ‘previous’ chip and the ‘future’ chip are independent. Again, $Q^{(i)}_{L}(j_1, j_2)$ is similar to Case 1 and it is possible to argue it is bounded as desired.

Suppose we prove $E \left[ \left| Q^{(i)}_{L}(j_1, j_2) \right|^4 \right] \leq \text{const}/L^2$. Then

$$L \text{Var} \left[ X^{(i)}_{L} \left\{ {p_1}(n) \right\}, \left\{ d_{i,l}(0) \right\}, \{ \tau_{i,j} \} \right] \to \sum_{j_1=1}^{N_r} \sum_{j_2=1}^{N_r} \beta_{j_1} \beta_{j_2} B^{(i)}_{1}(j_1, j_2) \triangleq B^{(i)}_{1}$$

almost surely as $L \to \infty$.

We can prove $E \left[ \left| Q^{(i)}_{L}(j_1, j_2) \right|^4 \right] \leq \text{const}/L^2$ as in Appendix A except that we must consider more terms and apply appropriate change of variables.

The proof showing asymptotic normality via the characteristic function approach is exactly the same as before, except that we argue

$$|W^{(i)}_{n}| \leq N_c \sum_{j=1}^{N_r} |\beta_j| \sum_{m=-K}^{K-1} |\alpha_m| \left| b^{(1, \tau_{i,j})}_{j}(n-m) \right|$$

$$\leq N_c \sum_{j=1}^{N_r} |\beta_j| \sum_{m=-K}^{K-1} |\alpha_m| \left( |c^{(0)}_{i,j}| + |c^{(1)}_{i,j}| \right)$$

$$= N_c \sum_{j=1}^{N_r} |\beta_j| \sum_{m=-K}^{K-1} |\alpha_m| \left( 1 + 2 \frac{\tau_{i,j}}{T_{ch}} \right)$$

$$\triangleq c,$$

and we continue the rest of the proof as before.

This proves that $\sqrt{L} X^{(i)}_{L}$, or equivalently $\sqrt{L} Y^{(i)}_{A,L}$, conditioned on $\{ p_1(n) \}$, $\{ d_{i,l}(0) \}$, and $\{ \tau_{i,j} \}$, is conditionally asymptotically normal with zero mean and variance $B^{(i)}_{1}$, as $L \to \infty$.

Recall the $\left\{ Z^{(i,j)}_{L}(m) \right\}$ are independent in $i$ for each fixed $L$. As a corollary of the above result, the random variables $\left\{ \sqrt{L} Z^{(i,j)}_{L}(m) \right\}_{j=1}^{N_r}$, for fixed $i$, conditioned on $\{ p_1(n) \}$, $\{ d_{i,l}(0) \}$, and $\{ \tau_{i,j} \}$, are conditionally asymptotically normal with zero means...
and covariance

\[
\text{Cov} \left[ \sqrt{L} Z_{L}^{(i,j_1)}(m_1), \sqrt{L} Z_{L}^{(i,j_2)}(m_2) \bigg| \{p_1(n)\}, \{d_{i,l}(0)\}, \{\tau_{i,j}\} \right]
\]

\[
= N_c \delta(m_1, m_2) \left( \frac{\tau_{i,j_1}}{T_{ch}} \frac{\tau_{i,j_2}}{T_{ch}} + \left( 1 - \frac{\tau_{i,j_1}}{T_{ch}} \right) \left( 1 - \frac{\tau_{i,j_2}}{T_{ch}} \right) \right) \quad \text{if } 1[0, \tau_{i,j_1}] = 1[0, \tau_{i,j_2}]
\]

\[
\triangleq N_c \delta(m_1, m_2) K_i(j_1, j_2),
\]

as \( L \to \infty \). Simplifying,

\[
K_i(j_1, j_2) \triangleq \begin{cases} 
\frac{\tau_{i,j_1}}{T_{ch}} \frac{\tau_{i,j_2}}{T_{ch}} + \left( 1 - \frac{\tau_{i,j_1}}{T_{ch}} \right) \left( 1 - \frac{\tau_{i,j_2}}{T_{ch}} \right) & \tau_{1,j_1} > \tau_{i,j_1} \text{ and } \tau_{1,j_2} > \tau_{i,j_2} \text{ or } \\
\left( 1 - \frac{\tau_{i,j_1}}{T_{ch}} \right) \left( 1 - \frac{\tau_{i,j_2}}{T_{ch}} \right) & \tau_{1,j_1} < \tau_{i,j_1} \text{ and } \tau_{1,j_2} < \tau_{i,j_2} \text{ otherwise}
\end{cases}
\]

### E.1 Acknowledgements

This appendix is, in part, a reprint of material that has been accepted for publication as “Performance Analysis of High Data Rate MIMO Systems in Frequency Selective Fading Channels,” P. Amighoud, E. Masry, L. B. Milstein, and J. G. Proakis, *IEEE Transactions on Information Theory*, August 2007. The dissertation author was the primary researcher and author of these publications.
Appendix F

Joint density of the \( \left\{ \left( r_{ij}^{(i*)} \right)^2 \right\} \) when \( N_T = N_U = 2 \)

Setting \( N_T = N_U = 2 \), writing (3.32) in terms of the elements of the Wishart matrix, and substituting in (3.31), gives

\[
F_{X_1, X_2}(x_1, x_2) = 2 Pr \left\{ \left\{ w_{1,1} \leq x_1, \frac{(w_{1,1}w_{2,2} - |w_{1,2}|^2)}{w_{1,1}} \leq x_2 \right\} \cap S_1 \right\},
\]

where \( X_1 \triangleq \left( r_{1,1}^{(i*)} \right)^2 \), \( X_2 \triangleq \left( r_{2,2}^{(i*)} \right)^2 \), \( F_{X_1, X_2}(x_1, x_2) = Pr\{X_1 < x_1, X_2 < x_2\} \), and

\[
S_1 \triangleq \left\{ w_{1,1}^{-1} + w_{1,1}(w_{1,1}w_{2,2} - |w_{1,2}|^2)^{-1} < w_{2,2}^{-1} + w_{2,2}(w_{1,1}w_{2,2} - |w_{1,2}|^2)^{-1} \right\}.
\]

The joint density of \( w_{1,1}, w_{2,2}, w_{1,2}^{(R)} \), and \( w_{1,2}^{(I)} \) is given by

\[
f_{w_{1,1}, w_{2,2}, w_{1,2}^{(R)}, w_{1,2}^{(I)}}(a_{1,1}, a_{2,2}, a_{1,2}^{(R)}, a_{1,2}^{(I)}) = \frac{1}{\pi^2 \mu^8} e^{-\mu(a_{1,1}+a_{2,2})}, \quad a_{1,1} > 0, a_{2,2} > 0,
\]

\[
a_{1,1}a_{2,2} - \left( a_{1,2}^{(R)} \right)^2 - \left( a_{1,2}^{(I)} \right)^2 > 0.
\]

This is obtained by setting \( N_T = N_U = 2 \) in (3.33). Then we can write

\[
F_{X_1, X_2}(x_1, x_2) = 2 \iiint_S du_1 du_2 du_3 du_4 f_{w_{1,1}, w_{2,2}, w_{1,2}^{(R)}, w_{1,2}^{(I)}}(u_1, u_2, u_3, u_4), \quad \text{(F.2)}
\]

where \( S \triangleq S^{(1)} \cap S^{(2)} \),

\[
S^{(1)} \triangleq \left\{ u_1^{-1} + u_1(u_1u_2 - u_3^2 - u_4^2)^{-1} < u_2^{-1} + u_2(u_1u_2 - u_3^2 - u_4^2)^{-1} \right\}.
\]
and

\[ S^{(2)} \triangleq \left\{ u_1 \leq x_1, \ u_1 (u_1 u_2 - u_3^2 - u_4^2)^{-1} \leq x_2 \right\}. \]

Simplifying the region \( S^{(1)} \), we have

\[ S^{(1)} = \begin{cases} \frac{u_1 - u_2}{u_1 u_2 - u_3^2 - u_4^2} < \frac{u_1 - u_2}{u_1 u_2} \\ \{0 > u_3^2 + u_4^2\} \text{ if } u_1 - u_2 > 0 \\ \{0 < u_3^2 + u_4^2\} \text{ if } u_1 - u_2 < 0. \end{cases} \]

Note that \( u_1 u_2 - u_3^2 - u_4^2 > 0 \) is always true, due to (F.1). Since \( \{0 > u_3^2 + u_4^2\} \) is the empty set and \( \{0 < u_3^2 + u_4^2\} \) is always true, we have \( S^{(1)} = \{u_1 < u_2\} \), and \( S \) simplifies to \( S = \left\{ u_1 \leq x_1, \ u_1 (u_1 u_2 - u_3^2 - u_4^2)^{-1} \leq x_2, \ u_1 < u_2 \right\}. \)

Substituting \( S \) and (F.1) in (F.2), we have

\[
F_{X_1, X_2}(x_1, x_2) = 8 \int_0^{x_1} \int_0^\infty \int_0^\infty \int_0^\infty \frac{1}{\pi \sigma_h \sigma_b} e^{-\frac{1}{\sigma_h^2}(u_1+u_2)} \\
\cdot I \left( u_1 (u_1 u_2 - u_3^2 - u_4^2)^{-1} \leq x_2 \right) I \left( u_1 u_2 - u_3^2 - u_4^2 > 0 \right) \\
\cdot du_3 du_4 du_1 du_2,
\]

where we have used the fact that the integrand is an even function of \( u_3 \) and \( u_4 \), \( I(B) = 1 \) if \( B \) is true, and \( I(B) = 0 \) otherwise. By simplifying the region where \( I \left( u_1 u_2 - u_3^2 - u_4^2 > 0 \right) \) and \( I \left( u_1 (u_1 u_2 - u_3^2 - u_4^2)^{-1} \leq x_2 \right) \) are both non-zero, it is possible to show that their product is given by

\[
I \left( u_1 (u_1 u_2 - u_3^2 - u_4^2)^{-1} \leq x_2 \right) I \left( u_1 u_2 - u_3^2 - u_4^2 > 0 \right)
= I \left( (u_1 (u_2 - x_2) - u_3^2)^{1/2} \leq u_4 < (u_1 u_2 - u_3^2)^{1/2} \right)
\cdot I \left( u_3 < (u_1 u_2 - x_2)^{1/2} \right) I \left( u_2 > x_2 \right)
+ I \left( u_4 < (u_1 u_2 - u_3^2)^{1/2} \right) I \left( (u_1 (u_2 - x_2))^{1/2} < u_3 < (u_1 u_2)^{1/2} \right) I \left( u_2 > x_2 \right)
+ I \left( u_4 < (u_1 u_2 - u_3^2)^{1/2} \right) I \left( u_3 < (u_1 u_2)^{1/2} \right) I \left( u_2 < x_2 \right).
\]
Then (F.3) can be written as

\[ F_{X_1, X_2}(x_1, x_2) = 8 \]

\[
\cdot \left[ \int_0^{x_1} \int_0^{\infty} \int_0^{(u_1(u_2-x_2))^{1/2}} \int_0^{(u_1u_2-u_2^2)^{1/2}} \frac{1}{\pi \sigma_h^2} e^{-\frac{1}{\sigma_h}(u_1+u_2)} du_4 du_3 du_2 du_1 \\
+ \int_0^{x_1} \int_{\max(u_1,x_2)}^{\infty} \int_0^{(u_1u_2)^{1/2}} \int_0^{(u_1u_2-u_2^2)^{1/2}} \frac{1}{\pi \sigma_h^2} e^{-\frac{1}{\sigma_h}(u_1+u_2)} du_4 du_3 du_2 du_1 \\
+ \int_0^{\min(x_1,x_2)} \int_{u_1}^{x_2} \int_0^{(u_1u_2)^{1/2}} \int_0^{(u_1u_2-u_2^2)^{1/2}} \frac{1}{\pi \sigma_h^2} e^{-\frac{1}{\sigma_h}(u_1+u_2)} du_4 du_3 du_2 du_1 \right].
\] (F.4)

Evaluating the integral over \( u_4 \), and noting that

\[
\int_0^\alpha (\alpha^2 - u_3^2)^{1/2} du_3 = \alpha^2 \int_0^1 (1-u^2)^{1/2} du = \alpha^2 \pi/4,
\]

it is possible to show that (F.4) is equal to

\[
F_{X_1, X_2}(x_1, x_2) = \frac{2x_2}{\sigma_h^4} \int_0^{x_1} u_1 e^{-\frac{u_1}{\sigma_h^2}} e^{-\frac{1}{\sigma_h}\max(u_1,x_2)} du_1 \\
+ \frac{2}{\sigma_h^8} \int_0^{\min(x_1,x_2)} u_1 e^{-\frac{u_1}{\sigma_h^2}} \int_{u_1}^{x_2} u_2 e^{-\frac{u_2}{\sigma_h^2}} du_2 du_1. 
\] (F.5)

The partial derivative of (F.5), with respect to \( x_1 \) and \( x_2 \), can be computed separately for \( x_1 > x_2 \) and \( x_1 < x_2 \). Combining the result, for any \( x_1 > 0 \) and \( x_2 > 0 \), we have

\[
\frac{\partial}{\partial x_2} \frac{\partial}{\partial x_1} F_{X_1, X_2}(x_1, x_2) = \frac{2x_1}{\sigma_h^4} e^{-\frac{1}{\sigma_h}(x_1+\max(x_1,x_2))}, 
\] (F.6)

which is the desired joint density when \( N_T = N_U = 2 \).

### F.1 Acknowledgements

This appendix is, in part, a reprint of material published as “Performance Analysis of a Pre-BLAST-DFE Technique for MISO Channels with Decentralized Receivers,” P. Amihood, E. Masry, L. B. Milstein, and J. G. Proakis, IEEE Transactions on Communications, vol. 55, pp. 1385-1396, July 2007. The dissertation author was the primary researcher and author of these publications.
Bibliography


