

This is to certify that the
dissertation entitled

Advanced Approaches in Information Transmission and
Access Control for Wireless Communication Networks

presented by

HUAHUI WANG

has been accepted towards fulfillment
of the requirements for the

Ph.D. degree in Electrical Engineering



Major Professor's Signature

11/28/2006

Date

MSU is an Affirmative Action/Equal Opportunity Institution

LIBRARY
Michigan State
University

PLACE IN RETURN BOX to remove this checkout from your record.
TO AVOID FINES return on or before date due.
MAY BE RECALLED with earlier due date if requested.

DATE DUE	DATE DUE	DATE DUE

**ADVANCED APPROACHES IN
INFORMATION TRANSMISSION AND
ACCESS CONTROL FOR WIRELESS
COMMUNICATION NETWORKS**

By

Huahui Wang

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Electrical and Computer Engineering

2006

ABSTRACT

ADVANCED APPROACHES IN INFORMATION TRANSMISSION AND ACCESS CONTROL FOR WIRELESS COMMUNICATION NETWORKS

By

Huahui Wang

Over the last two decades, wireless communication has seen tremendous growth and has been demonstrated as a robust voice and data transport mechanism. New wireless communication methods and services are enthusiastically adopted by people throughout the world. Driven by the ever increasing demand on high speed wireless multimedia services, development of highly reliable and more efficient wireless communication networks has become the ultimate goal of the research community. In this dissertation, we focus on efficient information transmission and medium access control (MAC) in wireless networks. More specifically, we aim to improve the reliability and efficiency of the wireless networks through the following research thrusts:

First, we investigate the channel tracking scheme in time-varying environments. Accurate channel estimation is essential in ensuring reliable information transmission. However, both time and frequency dispersions in mobile wireless channels cause significant challenges in channel estimation. Conventionally, for time-varying channels, pilot (training) signals are periodically transmitted to achieve accurate channel estimation. Such a scheme is not spectrally efficient due to the considerable overhead signals. In this dissertation, we propose a semi-blind approach to efficiently estimating fast fading channels and jointly detecting transmitted signals. The proposed scheme is shown to have high spectral efficiency and performance reliability. Additional efforts are devoted to studying the extreme case when there are unpredictable abrupt changes in the channel. We propose algorithms to detect such abrupt changes and

suppress the possible error propagation. The proposed algorithms are demonstrated to be effective through simulations.

Next, we explore the MAC protocol design by taking into account the physical (PHY) layer channel capability. In conventional medium access control protocol designs, the physical layer is simply characterized using a binary collision model. Although the model provides a tractable path for network performance analysis, it fails to reflect the physical layer channel capability. Cross-layer medium access control protocol design, which exploits the physical layer signal processing capabilities for MAC performance improvement, has attracted considerable research attention. In this dissertation, taking a mutually interactive perspective, we propose to design an MAC protocol, named hybrid ALOHA, which is in favor of the physical layer, and the improved physical layer, in turn, improves the MAC performance in terms of throughput, stability and delay behavior. Both theoretical and simulation results show that significant performance improvement can be achieved by hybrid ALOHA, in comparison with traditional ALOHA.

Finally, from a mixed analog-digital perspective, we investigate a system with analog inputs and propose a source-aware information transmission scheme to minimize the average input-output distortion. After sampling and source coding, the digital bits could have different levels of significance, i.e., some bits could be more important than others. Such non-uniformity in source coding then calls for non-uniform information transmission to improve system performance. An unequal error protection scheme is proposed to minimize the average distortion. Simulation results demonstrate its effectiveness. Furthermore, a joint quantization-constellation design is investigated under the criterion of distortion minimization. The proposed method generalizes the concept of constellation design from the perspective of joint source-channel coding. The simplicity and power efficiency of the proposed schemes make them particularly attractive for systems with tight power constraints, such as wireless sensor networks and space communications.

Copyright © by
Huahui Wang
2006

Dedicated to my family

ACKNOWLEDGMENTS

I would like to take this opportunity to express my gratitude to my advisor, Dr. Tongtong Li, for her constant support, guidance and encouragement throughout the years. She makes her best endeavors to help me in every issue, giving me advice not only in doing great research but also in personal growth.

I want to thank Dr. Hassan K. Khalil and Dr. Fathi M. Salem from Department of Electrical and Computer Engineering, and Dr. Jonathan I. Hall from Department of Mathematics for serving as my committee. I am deeply indebted to them for their kind support, either in classes or in all thoughtful correspondences. I also want to thank Dr. Jian Ren for valuable discussions and the help in various technical issues.

I am grateful to a lot of friends who make my life so enjoyable at Michigan State University. Thanks go to my lab-mates Weiguo Liang, Qi Ling and Leonard Lightfoot, for the valuable discussions on research and for the advice on the daily life in the United States. Thanks also go to Jun Yang, Feng Wei, Xiaoyong Li, Xiangrong Guo, Liyong Wu, Yun Liu, Yingying Shi, Qionghua Qiang, and Keke Yang for being my tennis partners and for all the fun we have together.

Last, but not the least, I would like to thank my parents, my sister, and my brother-in-law for their love and ceaseless support. Special thanks go to my lovely little niece, Junlin, who always talked to me over the phone using her sweetest voice, "It is nearly noon here. That means midnight in your place. Do you want to go to bed now?" It is really a great comfort to receive such concerns from a 6-year old girl.

TABLE OF CONTENTS

LIST OF FIGURES	x
1 Introduction	1
1.1 Challenges, Motivations and Proposed Research Directions	1
1.1.1 Major Challenges in Wireless Communications	1
1.1.2 Information Transmission and Access Control in Wireless Com- munications	3
1.2 Relationship with Existing Work	5
1.2.1 Channel Estimation and Signal Detection in Time-Varying En- vironment	5
1.2.2 PHY-Aware MAC Protocol Design	6
1.2.3 Source-aware Non-uniform Transmission for Minimum Distortion	7
1.3 Overview of the Dissertation	8
2 Channel Tracking and Signal Detection over Time-Varying Chan- nels	11
2.1 Introduction	11
2.2 System Model	13
2.2.1 Transmitter Structure	13
2.2.2 Channel Model	15
2.2.3 Received Signal	16
2.3 Joint Channel Estimation and Signal Detection	18
2.3.1 Initial Channel Estimation	18
2.3.2 Joint Channel Estimation and Signal Detection	20
2.4 Detection and Tracking of Abrupt Channel Changes	21
2.4.1 Detection of Abrupt Channel Changes	21
2.4.2 Extraction and Tracking of the Abrupt Channel	24
2.5 Simulation Results	26
2.5.1 Performance under the Jakes' Model with No Abrupt Changes	28
2.5.2 Performance under the Jakes' Model with Abrupt Changes . .	30
2.6 Summary	35
3 Mutually Interactive Random Access Control for Wireless Networks	36
3.1 Introduction	37
3.2 The Hybrid ALOHA Protocol	41
3.2.1 System Model	41
3.2.2 Hybrid ALOHA	42

3.3	Throughput Calculation	43
3.4	Stability of Hybrid ALOHA	48
3.4.1	Stability Region for $N = 2$	48
3.4.2	Sufficient Condition of Stability for $N > 2$	53
3.5	Special Case: Stability Inner Bound for $N = 3$	56
3.5.1	Inner Bound Characterization	56
3.5.2	Analysis of the Kernel $K(x, y)$	60
3.5.3	Reduction to the Riemann-Hilbert Problem	62
3.5.4	Solution to the Riemann-Hilbert Problem (3.48)	63
3.6	Delay Performance for $N = 2$	66
3.6.1	Exact Delays for the General Case	66
3.6.2	Delay Bounds for the Symmetric Case	67
3.7	An Illustrative Example	70
3.7.1	System Set-up	70
3.7.2	Evaluation of Channel Estimation Error	72
3.7.3	Probabilities of Successful Signal Receptions	73
3.7.4	Throughput of Hybrid ALOHA with Imperfect Channel Estimation	75
3.7.5	Stability Regions and Delay Bounds with Imperfect Channel Estimation	77
3.8	Summary	81
4	Source-Aware Non-Uniform Transmission for Minimum Distortion	82
4.1	Introduction	82
4.2	Problem Formulation	84
4.3	Joint Source Index Assignment And Constellation Code Design	86
4.4	Source-Aware Non-Uniform Transmission Design	90
4.4.1	Non-Uniform Transmission Based on Gray-Code Constellations	90
4.4.2	Source-Aware Constellation Design	96
4.5	Summary	99
5	Conclusions and Future Work	100
5.1	Conclusions	100
5.2	Future Work	101
	APPENDICES	103
A	Alternative Proof of Proposition 3.1 for $M = 2$.	104
B	Proof of Lemma 3.1.	108
C	Proof of Proposition 3.2.	110

D Proof of Proposition 3.4.	112
E Formulation of Functional Equation (3.27)	114
F Proof of Lemma 3.2	116
G Proof of Lemma 3.3	118
H Proof of Proposition 3.7.	119
BIBLIOGRAPHY	123

LIST OF FIGURES

2.1	Transmitter structure in an uplink MC-CDMA system.	14
2.2	BER performance under the Jake's model with no abrupt changes, $K = 8$, $P = 16$	27
2.3	MSE performances of the proposed channel estimation at different SNR levels.	29
2.4	Impact of timing errors on the BER performance, $SNR = 25dB$: <i>independent offsets</i> means that users have independent timing errors uniformly distributed within interval $[0, 0.25]\mu s$, <i>common offsets</i> means that all users experience the same timing shift.	30
2.5	Accumulated channel deviation ϕ_k when user k ($k = 1, 2, 3, 4$) is identified to be subjected to abrupt channel changes, while user 1 is the true user whose channel undergoes abrupt changes and other users' channels are "smooth". $K = 8$, $P = 16$	32
2.6	Channel estimates (only real part of the channel is shown) with the proposed error propagation control: $SNR = 20dB$, $K = 8$, $P = 16$, user 1 undergoes abrupt channel changes. Only results of 4 out of 8 users are presented.	33
2.7	BER performances of the system over channels with and without abrupt changes. Here " <i>perfect CSI</i> " means that the channel state information is perfectly known at the receiver, including information on abrupt changes.	34
3.1	Mutual interaction of MAC layer and PHY layer to achieve performance optimization.	39
3.2	Illustration of the hybrid ALOHA slot structure for $M = 2$, in the case that two users ($N = 2$) transmit their training sequences at non-overlapping pilot sub-slots.	43
3.3	Throughput comparison between hybrid ALOHA and traditional slotted ALOHA.	47
3.4	Stability regions for different protocols in the two-user case.	50
3.5	Delay bounds for the system with different parameters.	69
3.6	Throughput comparison of different schemes under the specific system settings, $SNR = 10dB$, $\beta = 6dB$, $\bar{\sigma}_{e1}^2 = 0.01$, $\bar{\sigma}_{e2}^2 = 0.06$	76
3.7	Numerical values of τ achieving equilibrium points and τ_{opt} maximizing the throughput of hybrid ALOHA	78

3.8	Comparison of stability regions of different schemes, $SNR = 10dB$, $\beta = 6dB$, $\sigma_{e1}^2 = 0.01$, $\sigma_{e2}^2 = 0.06$	79
3.9	Comparison of delays of different schemes, $SNR = 10dB$, $\beta = 6dB$, $\bar{\sigma}_{e1}^2 = 0.01$, $\bar{\sigma}_{e2}^2 = 0.06$, $p = 0.75$	80
4.1	System Model	85
4.2	Mapping of quantized values to the 16-QAM constellation.	88
4.3	Symmetric 16-AM constellation.	89
4.4	(a) Asymmetric 16-QAM, $d_1 > d_2$; (b) Symmetric 16-QAM with Gray codes.	91
4.5	MSE for different transmission schemes: Uncoded System.	94
4.6	MSE for different transmission schemes: Coded System ($\lambda = d_2/d_1$).	95
4.7	Example 3: (a) Source distribution (b) Normalized MSE under different SNR levels, uniform 4-AM versus the proposed source-aware 4-AM.	98

“Images in this dissertation are presented in color”

CHAPTER 1

Introduction

1.1 Challenges, Motivations and Proposed Research Directions

Over the last two decades, wireless communication has seen tremendous growth and has been demonstrated as a robust voice and data transport mechanism. New wireless communication methods and services are enthusiastically adopted by people throughout the world. Driven by the ever increasing demand on high speed wireless multimedia services, development of highly reliable and more efficient wireless communication networks has become the ultimate goal of the research community.

1.1.1 Major Challenges in Wireless Communications

Compared with its wirelined counterpart, the major challenges in wireless communication system design are caused by the unpredictable distortion and interference in the wireless environment, and by the limitation of the precious spectrum which is always in shortage.

Information Transmission over Time-Varying Wireless Channels The fundamental limitations on the performance of wireless communications are placed by the unpredictable distortion and interference over wireless channels. Unlike wirelined channels which are generally stationary and predictable, wireless channels are time-variant and totally random. Due to high user mobility and the fact that there is no physical boundary in wireless environment, wireless channels are influenced by many

physical factors including mobile radio propagation, the relative motion between the transmitter and the receiver, motions of surrounding objects and diversity in the signal transmission (e.g., time/frequency/space diversity). Effective time-varying channel tracking and efficient transmission schemes play the key roles in increasing the information capacity over wireless channels.

Spectrum Efficiency in Multiple Access Environment Spectrum is precious in wireless communications. In wirelined communication, new and additional bandwidth requirement can often be granted by adding another line between the transmitter and receiver. In wireless environment, however, different services and different users have to share the total available bandwidth. Along with advances in radio frequency (RF) technology, the available spectrum has become much larger than ever before, but the demand on wireless services has also increased enormously. Take the mobile phone as an example. Mobile communication was commercialized in late 1970s and early 1980s. At that time, the mobile phones relied on analog techniques, only voice service was available and few people had access to the mobile network. This is referred to as the first generation wireless communications. In early 1990s, digital communications gradually replaced their analog counterparts and became the dominant transmission technique in wireless communications. In addition to pure voice service, data services, such as short messages and FAX, became available and many more users could be accommodated by the wireless network. This is referred to as the 2nd generation wireless communications. Today, while we are moving to the 3G (3rd generation) era, the number of worldwide wireless subscribers is approaching two billion. High speed multimedia wireless services, such as wireless internet access, wireless electronic commerce, are coming into reality. Spectrum, therefore, has become the most valuable resource. Spectrally efficient wireless system design in multiple access environment has become the impetus in wireless communication area.

1.1.2 Information Transmission and Access Control in Wireless Communications

In wireless communications, users first access the wireless network through the random access channel to build up a handshake with the mobile switch center through one or more base stations, and then start the pay-load information transmission. In this dissertation, rooted in physical layer (PHY) transmitter-receiver design, we will focus on efficient information transmission and PHY-aware medium access control (MAC) in wireless networks. More specifically, we plan to improve the reliability and efficiency of wireless networks through the following research thrusts:

Time-Varying Channel Tracking Accurate channel estimation is essential in ensuring reliable information transmission. In wireless communication, the time-varying wireless channel has both time dispersion (frequency selective fading caused by multipath fading) and frequency dispersion (Doppler spread caused by relative motion between the transmitter and receiver). This raises significant challenges in channel estimation. On the one hand, pilot signals (training bits or symbols) have to be inserted frequently to achieve accurate channel estimation. On the other hand, we need to minimize the overhead bits for the sake of spectral efficiency.

Conventionally, channel estimation in wireless communications is performed through the following procedure. First, divide the time axis into very short time slots. Secondly, approximate the time-varying channel using a time invariant channel in each short slot. Finally, carry out channel estimation by inserting pilot signal in each slot. This method is acceptable in slow-varying environment, but would fail to deliver satisfying performance when fast fading is present. To track the time-varying wireless channels closely, we go back to the analytical characterization of doubly selective channels. Mathematically, a time-varying channel is characterized by $h(t, \tau)$, which denotes the channel response at time instant t to the unit impulse at $t - \tau$. Therefore, instead of carrying out one-dimensional channel estimation as in the conventional method, we put it into a two-dimensional framework and perform channel estimation in the time-frequency domain. Semi-blind approaches are exploited to

minimize the overhead bits and therefore increasing the spectral efficiency.

PHY-Aware MAC Design Conventionally, random access control protocols are designed disjointly by characterizing the underlying physical layer using a binary collision model. Briefly speaking, this means that when two users attempt to access the random channel simultaneously, a collision occurs, and we then say that both are failed; a transmission is successful if and only if one single user transmits.

More recently, researchers started to realize that although it provides a tractable path for network performance analysis, the collision model fails to reflect the physical layer channel capability. Cross-layer medium access control protocol design, which exploits the physical layer signal processing capabilities for MAC performance improvement, has attracted considerable research attention. In this dissertation, taking a mutually interactive perspective, we propose to design an MAC protocol which is in favor of the PHY layer, and the improved physical layer, in turn, improves the MAC performance in terms of throughput, stability as well as delay behavior.

Source-Aware Information Transmission The meeting ground for analog and digital signals is the A/D conversion. In today's most communication systems, after sampling and source coding, all the bits are treated uniformly. According to the information theory, when entropy coding is applied, then all the bits carry an equal amount of information. However, as the *a priori* statistics of the input signals are generally unavailable, non-entropy coding is widely used in practical systems. This implies that some bits are more important than others. This universal existence of non-uniformity in source coding therefore calls for non-uniform information transmission.

In the literature, without taking quantization optimization into consideration, bit-level unequal error protection schemes have been proposed to provide more protection to the most significant bits. Our research on non-uniform information transmission is motivated by the observation that: when the inputs are analog, the ultimate goal of the communication system is to minimize the average input-output distortion, and a communication system that minimizes the bit-error-rate does not necessarily

minimize the average distortion. Taking a joint source-channel coding perspective, we propose to minimize the average distortion through joint quantization design and constellation index assignment.

1.2 Relationship with Existing Work

In this section, we discuss the relationship between the existing and the proposed work.

1.2.1 Channel Estimation and Signal Detection in Time-Varying Environment

In dynamic time-varying environments, traditional channel estimation techniques for time-invariant models are no longer applicable. To track time-varying channels, a widely used approach is to transmit pilot signals periodically such that the receiver can update the channel estimates timely. Such a method results in considerable overhead bits in fast fading channels and costs excess bandwidth.

To improve spectral efficiency, joint channel estimation and signal detection approaches have been proposed to reduce the overhead bits [FC91, KV94, RSAA97, Moh98, AGR98, VD98, WP99]. The main idea is to use the estimated symbols as pseudo-pilot signals for channel estimation, and then improve the subsequent signal detection based on the refined channel estimates. For time-varying channels, joint channel/signal estimation based on Kalman filtering has been representative [Ilt90, IF91, KI02]. As is well known, Kalman filtering results in the optimal minimum variance estimator for channels characterized by the first-order autoregressive (AR) model [AM79], provided the model parameters are known to the receiver. In most of the existing work utilizing Kalman filtering [Ilt90, IF91, KI02, VT95, RPT95, WP98], time-varying channels are modeled as AR processes with known parameters, and no efforts are taken to estimate them. To bridge this modeling gap, [TGZ96] derives a method for estimating the model parameters. In more general cases when chan-

nel model parameters are too complicated to obtain, adaptive receivers have been developed in [KSP03, WP98].

In this dissertation, an alternative joint channel estimation and signal detection scheme is proposed in time-varying environment. The scheme is advantageous in that it has very high bandwidth efficiency while not requiring any model fitting process to obtain channel model parameters. In addition, algorithms are proposed to detect and track abrupt channel changes and thus are capable of suppressing possible error propagations.

1.2.2 PHY-Aware MAC Protocol Design

Regarding cross-layer design, there has been a line of work focusing on performance improvement from joint MAC-PHY perspective. The prevalent design is to exploit the PHY information, e.g., the channel state information (CSI), in random access to achieve optimal system performance in terms of throughput and delays, etc. The pioneering work was carried out by Knopp and Humblet in [KH95], where the CSI is incorporated in random access for ALOHA systems. It was shown that, under the average power constraint, maximizing the sum-rate leads to scheduling the best user to transmit. In [TS99], decentralized use of the CSI was investigated by Telatar and Shamai. It was shown in [TH98] that multiuser diversity is induced when exploiting the CSI, and the maximum sum-rate could be improved with the increase of the number of users. In [QB01], Qin and Berry used the CSI to control the transmission probability in ALOHA. The effect of multiuser diversity on the throughput was analyzed as well.

The schemes discussed above adopted the well-known ideal collision model: when only one user transmits, the packet arrives at the receiver error free. The model is over-simplified in terms that it overlooks the channel effects such as fading and noise, and it also ignores the possibility of successful packets reception in the presence of simultaneous transmissions. A more realistic model, called multipacket reception (MPR) model, was proposed by Ghez et al. in [GVS88]. The proposed MPR model is *symmetric* in the sense of indistinguishable users in the network, while it offers the

generalization that when there are simultaneous transmissions, the reception can be described by conditional probabilities. A more general *asymmetrical* MPR model was proposed in [NMT05], and the performances in terms of system stability and delay behavior for ALOHA networks were investigated.

These cross-layer designs share a common feature that they take into account the PHY-MAC interaction yet in a “one way” fashion. That is, only the impact of the PHY layer information on the MAC layer design is considered, and no effort has been taken to study how the MAC protocol design would influence the PHY layer information transmission. In this dissertation, the MPR model is adopted to analyze the system performance of a proposed ALOHA protocol. The study takes a mutually interactive MAC-PHY perspective to design an MAC protocol that is in favor of the PHY layer information transmission, while the improved PHY layer can in turn improve the overall system performance. Details of the analysis are referred to Chapter 3.

1.2.3 Source-aware Non-uniform Transmission for Minimum Distortion

In [ZG93], with no specifications on the modulation schemes, Zeger proposed a locally optimal index assignment (source coding) solution for minimum input-output distortion, known as pseudo-Gray coding, by using iterative binary switching. Pseudo-Gray coding provided an effective approach in reducing the average distortion of a vector quantized system by rearranging the positions of code vectors in a given codebook. Various channel-optimized quantizers, multistage vector quantizers have been studied in [FV87, PFM93, WF94].

On the other hand, assuming that source index assignment has been done separately, and with the observation that the bits coming out of the source encoder are generally non-uniform (i.e. have different levels of significance), Masnick, Wolf [MW67] and Cover [Cov72] introduced non-uniform modulation (index mapping) schemes, known as unequal error protection codes, for which the more important bits have lower

error rate than other bits. Stemmed from [MW67, Cov72], unequal error protection schemes through both symmetric and asymmetric constellation designs have been further developed in [ROUV93], [MZFLI00a] and [MZFLI00b]. It should be pointed out that the resulted constellation codeword design in [ROUV93, MZFLI00a, MZFLI00b] may no longer be Gray codes.

In this dissertation, taking a mixed analog-digital perspective, we consider average input-output distortion minimization through joint optimization of source index assignment and modulation design. Based on the fact that Gray code ensures minimum bit error rate (BER) when the channel error probability is sufficiently small, a source-aware information transmission approach is proposed by exploiting the non-uniformity in Gray-coded constellations to achieve unequal error protection.

1.3 Overview of the Dissertation

In the dissertation, we address the following specific questions:

- In time-varying environments, how to design efficient channel estimation and signal detection techniques to cope with channel dynamics?
- From the joint MAC-PHY perspective, how can the cross-layer design improve the wireless network performance in terms of throughput, stability and delay?
- Considering data non-uniformity from source coding, how to employ unequal error protection to minimize average input-output distortion?

The dissertation is structured as follows.

Chapter 2 presents a joint channel estimation and multiuser detection method over time-varying channels for multicarrier code division multiple access (MC-CDMA) systems. The general Jakes' model is adopted without any assumption about the channel model parameters. Only one pilot symbol is required at the beginning of each data frame for initial channel estimation. Based on the fact that the channel coefficients of two successive symbols are highly correlated, subsequent channel tracking

and signal detection are carried out iteratively. An effective error control approach is proposed to track abrupt channel changes and extract the user whose channel suffers from such changes. With this process, error propagation can be successfully suppressed and all users channel estimates can be updated accurately.

Chapter 3 goes beyond the PHY layer design to improve overall system performance by considering cross-layer medium access control (MAC) design in wireless networks. Taking a mutually interactive MAC-PHY perspective, we aim to design an MAC protocol that is in favor of the PHY layer information transmission, and the improved PHY, in turn, can improve the MAC performance. Motivated by the fact that as long as good channel estimation can be achieved, advanced signal processing does allow effective signal separation given that the multiuser interference is limited to a certain degree, we propose a novel MAC protocol, named hybrid ALOHA, which allows conditional collision-free channel estimation and makes it possible for simultaneous multiuser transmission. More specifically, short idle sections are introduced into the structure of the traditional ALOHA slot, so that different users could transmit their training sequences in non-overlapping training slots, whereby collision-free channel estimation could be achieved. Relying on the general multipacket reception (MPR) model, in this chapter, quantitative analysis is conducted for the proposed hybrid ALOHA protocol in terms of throughput, stability as well as delay behavior. Both theoretical derivation and simulation examples demonstrate that significant performance improvement can be achieved by hybrid ALOHA, in comparison with the traditional ALOHA protocol.

Chapter 4 considers average input-output distortion minimization through joint optimization of source index assignment and modulation design. First, the general optimization criterion is derived. Secondly, source-aware non-uniform information transmission is proposed by utilizing the inherent non-uniformity in Gray coded constellations, both symmetric and asymmetric. It is found that when channel coding is involved, the proposed scheme delivers significantly better performance in comparison with existing unequal error protection schemes with non-Gray codewords. Further, joint quantization-constellation design is considered under the criterion of distortion

minimization. The proposed method generalizes the concept of constellation design from the perspective of joint source-channel coding. The simplicity and power efficiency of the proposed scheme make it particularly attractive for systems with tight power constraints, such as wireless sensor networks and space communications.

Chapter 5 discusses conclusions and possible extensions of the work covered in this dissertation.

CHAPTER 2

Channel Tracking and Signal Detection over Time-Varying Channels

In this chapter, a joint channel estimation and multiuser detection method is presented for uplink MC-CDMA systems in time-varying environment. Unlike the conventional channel tracking schemes which generally need to insert pilot bits in every OFDM symbol, in the proposed approach, only one pilot symbol is required at the beginning of each data frame. Based on the fact that the channel coefficients of two successive symbols are highly correlated, subsequent channel tracking and signal detection are carried out iteratively. A major contribution of the proposed approach is that it enables robust error propagation control, and is capable of detecting and tracking abrupt channel changes effectively. Simulation examples demonstrate the robustness of the proposed algorithms.

2.1 Introduction

Multicarrier code division multiple access (MC-CDMA) [YLF93, HP97], which amalgamates the advantages of both orthogonal frequency division multiplexing (OFDM) and CDMA, has been identified as a major technique for high speed wireless communications. However, good performance for MC-CDMA is only guaranteed with accurate *channel state information* (CSI) and effective demodulation. As wireless communication at higher frequency bands is becoming possible, relative motion be-

tween the mobile, the base station and the surrounding objects introduces significant frequency dispersions. Traditional channel estimation techniques for time-invariant models are no longer applicable.

To track the time-varying channels, a widely used approach is to insert pilot bits (also referred to as “pilot-carriers” in literature) in each OFDM symbol, and to estimate the channel coefficients utilizing filtering in the frequency domain, see [Cho00, CMS01] for example. In [HKR97] and [KH97], a two-dimensional (2-D) pilot grid is designed to insert pilot bits in both the frequency dimension and the time dimension, based on the 2-D sampling theorem. These pilot-aided techniques yield good performance in fast fading scenarios but result in a considerable amount of overhead bits.

To improve spectral efficiency, joint channel estimation and signal detection approaches have been proposed to reduce the overhead bits [FC91, KV94, RSAA97, Moh98, AGR98, VD98, WP99]. The main idea is to use the estimated symbols as pseudo-pilot signals for channel estimation, and then improve the subsequent signal detection based on the refined channel estimates. For time-varying channels, joint channel/signal estimation based on Kalman filtering has been representative [Ilt90, IF91, KI02]. As is well known, Kalman filtering results in the optimal minimum variance estimator for channels characterized by the first-order autoregressive (AR) model [AM79], provided the model parameters are known to the receiver. In most of the existing works utilizing Kalman filtering [Ilt90, IF91, KI02, VT95, RPT95, WP98], time-varying channels are modeled as AR processes with known parameters, and no efforts are taken to estimate them. To bridge this modeling gap, [TGZ96] derives a method for estimating the model parameters. In more general cases when channel model parameters are unknown, adaptive receivers exploiting least mean square (LMS) and recursive least square (RLS) algorithms are developed for synchronous MC-CDMA systems [KSP03, WP98]. Moreover, in [SMM04], channel estimation is performed using LMS for a quasi-synchronous MC-CDMA system, with the assumptions that the channel is invariant within the training period, and the number of the training blocks are greater than or equal to the number of active users.

In this chapter, assuming a practical quasi-synchronous MC-CDMA framework, we propose a novel joint channel estimation and multiuser detection scheme over fast-fading channels. Instead of assuming a known AR model, we use the general Jakes' model [Jak74], making no assumptions on the knowledge of the channel model parameters. Targeting on higher spectral efficiency, the proposed approach requires only one pilot symbol for initial channel estimation. A coarse estimate of the subsequent symbol is obtained utilizing the CSI of the previous symbol, and then refined through an iterative process. The reliability of such a detection method is based on the fact that, although the channel may vary fast and become uncorrelated over some period, the channel coefficients for two *successive* multicarrier (MC) symbols are highly correlated. Multiple access interference (MAI) can then be reconstructed and canceled out from the composite signal using parallel interference cancelation (PIC) [VA91], hence improve the accuracy of signal detection.

Under circumstances when the channel undergoes unpredictable abrupt changes, joint channel estimation and signal detection schemes are generally subjected to significant error propagation. The major contribution of the proposed scheme is that, while achieving higher spectral efficiency, it is capable of detecting and tracking abrupt channel changes, and thus can ensure robust error propagation control.

The rest of the chapter is organized as follows. Section 2.2 describes the MC-CDMA system as well as the time-varying channel model. Section 2.3 introduces the proposed algorithm for joint channel estimation and signal detection. Section 2.4 is focused on error propagation control for channels subjected to abrupt changes. Simulation results are presented in Section 2.5 and we conclude in Section 2.6.

2.2 System Model

2.2.1 Transmitter Structure

Consider an uplink MC-CDMA system with K users. The transmitter structure of user k is illustrated in Figure 2.1. The input binary stream is first mapped to BPSK

or QPSK symbols and then grouped into J -symbol blocks. The i th block for user k is denoted as $\mathbf{d}_{k,i}$. Each block of data is spread into $N = PJ$ chips using the user's specific pseudo-random spreading codes $\{c_{k,i}(n), n = 0, 1, \dots, N - 1\}$. Here, P is the spreading factor.

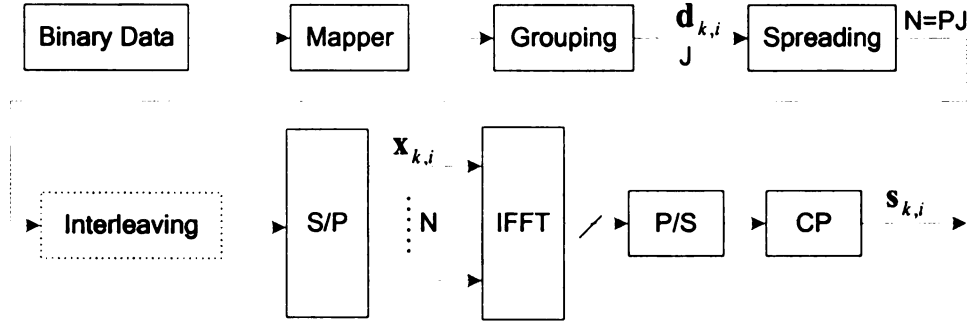


Figure 2.1. Transmitter structure in an uplink MC-CDMA system.

Before the spread signals are modulated on different OFDM subcarriers, a $P \times J$ block interleaver is employed to maximize the diversity and hence reduce the risk that all these chips suffer severe fades at the same time. At the receiver, a corresponding $J \times P$ deinterleaver is employed to reshape the signals. Interleaving plays an important role in improving the system performance, while for simplicity of notation, they are omitted in the subsequent description. In the simulations, however, their effects are taken into account.

The spread/interleaved signals modulated on OFDM subcarriers are saved in a vector $\mathbf{x}_{k,i} = [x_{k,i}(0), \dots, x_{k,i}(N - 1)]^T$. The output of the IFFT block is parallel-to-serial converted, and appended with cyclic prefix (CP) at the beginning of the modulated symbol. To prevent intersymbol interference (ISI), the length of the CP is assumed to be equal to the maximum channel order L . Following the notations in [WG00], the extended OFDM symbol, which is referred to as an MC symbol, can

be described by a transformation given by

$$\mathbf{s}_{k,i} = \mathbf{T}_{cp} \mathbf{F} \mathbf{x}_{k,i}, \quad (2.1)$$

where $\mathbf{T}_{cp} = [\mathbf{I}_{cp}^T, \mathbf{I}_N^T]^T$, $\mathbf{I}_{cp}$ represents the last L rows of an $N \times N$ identity matrix $\mathbf{I}_N$, and $\mathbf{F}$ is the IFFT matrix defined as

$$\mathbf{F} = \frac{1}{\sqrt{N}} \begin{pmatrix} W_N^{00} & \dots & W_N^{0(N-1)} \\ \vdots & \ddots & \vdots \\ W_N^{(N-1)0} & \dots & W_N^{(N-1)(N-1)} \end{pmatrix},$$

with $W_N^{nk} = e^{j2\pi nk/N}$.

2.2.2 Channel Model

The baseband equivalent channel is generally characterized using a multipath time-varying response [Ste92, LJS98],

$$h(t, \tau) = \sum_l \gamma_l(t) \delta(\tau - \tau_l(t)), \quad (2.2)$$

where $\delta(x)$ denotes the Dirac delta function, $\tau_l(t)$ and $\gamma_l(t)$ are the delay and the complex amplitude of the l th path at time instant t , respectively. The channel's frequency response with respect to τ at time t is given by

$$H(t, f) = \sum_l \gamma_l(t) e^{-j2\pi f \tau_l(t)}. \quad (2.3)$$

$\gamma_l(t)$'s are usually considered to be wide-sense stationary (WSS) and independent for different paths. Hence the time-varying correlation function of $H(t, f)$, defined as $R_H(\Delta_t, \Delta_f) = E[H(t + \Delta_t, f + \Delta_f) H^*(t, f)]$, can be decomposed as the product of a time-domain correlation function $R_t(\Delta_t)$ and a frequency-domain correlation function $R_f(\Delta_f)$ [LJS98]:

$$R_H(\Delta_t, \Delta_f) = R_t(\Delta_t) R_f(\Delta_f). \quad (2.4)$$

For an MC-CDMA system with symbol duration T and tone spacing Δ_f , the correlation function for different MC symbols and different tones can be written as

$$R_H(iT, n\Delta_f) = R_t(iT)R_f(n\Delta_f). \quad (2.5)$$

For the Jakes' model [Jak74], the time correlation function can be expressed as

$$R_t(iT) = J_0(2\pi f_d iT), \quad (2.6)$$

where $J_0(x)$ is the zeroth-order Bessel function of the first kind and f_d is the maximum doppler shift caused by frequency dispersion.

In practical systems, unpredictable, abrupt channel changes may occur due to sudden changes in the transmission environment. These changes can no longer be fully characterized using a WSS model. In this chapter, channels with or without abrupt changes are both considered. Abrupt channels are generated by imposing sudden, random changes to channels obtained from the Jakes' model.

2.2.3 Received Signal

Sampling the received signal at $t = mT/N_s$ (chip rate), where $N_s = N + L$, we obtain

$$r(m) = \sum_{k=1}^K r_k(m - \Delta_k) + w(m), \quad (2.7)$$

where $r_k(m - \Delta_k)$ is the received signal for user k , which is the convolution of the transmitted signals and the corresponding channel coefficients, $w(m)$ is the additive white Gaussian noise (AWGN) with zero mean and double-sided power spectral density of $N_0/2$. Δ_k denotes the k th user's propagation delay. As in [SMM04], we consider a quasi-synchronous system where Δ_k is uniformly distributed within the CP. The quasi-synchronization can be realized by providing each user a global positioning system (GPS) based receiver [IM96]. At this stage, for notation simplicity, we assume that Δ_k is known at the receiver. However, simulation result with timing

error will be presented in Section 2.5.

After the CP removal, the i th symbol input to the k th user's FFT demodulator, which is synchronized with user k , is saved in vector $\mathbf{z}_{k,i} = [r(iN_s + L + \Delta_k), \dots, r((i+1)N_s + \Delta_k - 1)]^T$. Separating the desired signal and the MAI components, $\mathbf{z}_{k,i}$ can be expressed as

$$\mathbf{z}_{k,i} = \mathbf{z}_{k,i}^{(D)} + \sum_{j=1, j \neq k}^K \mathbf{z}_{j,i}^{(I)} + \mathbf{w}_{k,i}, \quad i = 0, 1, \dots \quad (2.8)$$

where $\mathbf{z}_{k,i}^{(D)} = [r_k(iN_s + L), \dots, r_k((i+1)N_s - 1)]^T$ is the desired signal for user k , $\mathbf{z}_{j,i}^{(I)} = [r_j(iN_s + L + \Delta_k - \Delta_j), \dots, r_j((i+1)N_s + \Delta_k - \Delta_j - 1)]^T$ represents the interference from user j , and $\mathbf{w}_{k,i}$ is the AWGN noise. After the FFT demodulation, the k th user's signal can be written as

$$\mathbf{y}_{k,i} = \mathbf{F}^\dagger \mathbf{z}_{k,i} = \mathbf{X}_{k,i} \mathbf{H}_{k,i} + \sum_{j=1, j \neq k}^K \mathbf{I}_{j,i} + \mathbf{v}_{k,i}, \quad (2.9)$$

where the superscript $\dagger$ denotes Hermitian transpose, $\mathbf{X}_{k,i} = \text{diag}(\mathbf{x}_{k,i})$ is the desired signal, $\mathbf{v}_{k,i} = \mathbf{F}^\dagger \mathbf{w}_{k,i}$ is the noise term, $\mathbf{I}_{j,i} = \mathbf{F}^\dagger \mathbf{z}_{j,i}^{(I)}$ is the interference from user j . $\mathbf{H}_{k,i} = [H_{k,i}(0), \dots, H_{k,i}(N-1)]^T$ represents the frequency domain channel coefficients of the N subcarriers. If the channel order is L , the corresponding time domain coefficient $\mathbf{h}_{k,i} = [h_{k,i}(0), \dots, h_{k,i}(L)]^T$ is related to $\mathbf{H}_{k,i}$ by

$$\mathbf{h}_{k,i} = [\mathbf{F}]_{L+1} \mathbf{H}_{k,i}, \quad (2.10)$$

where $[\mathbf{F}]_{L+1}$ means that only the first $L+1$ rows of the matrix $\mathbf{F}$ is retained.

2.3 Joint Channel Estimation and Signal Detection

In this section, the proposed approach for joint channel estimation and multiuser detection is presented. Instead of setting aside pilot carriers in each MC symbol, only one pilot symbol is needed at the beginning of each data frame. For each symbol, PIC is exploited to suppress MAI and improve system performance. The proposed algorithm is outlined below:

1. At the beginning of each data frame of user k ($k = 1, \dots, K$), one pilot symbol $\mathbf{x}_{k,0}$ is transmitted to obtain an initial channel estimate $\hat{\mathbf{h}}_{k,0}$;
2. For the i th ($i > 0$) symbol, based on the channel information estimated from the previous symbol, i.e., $\hat{\mathbf{h}}_{k,i-1}$, a rough coherent detection is performed to obtain a tentative symbol estimate $\tilde{\mathbf{x}}_{k,i}$;
3. With the tentative signal $\tilde{\mathbf{x}}_{k,i}$ and the channel information $\hat{\mathbf{h}}_{k,i-1}$, ($k = 1, \dots, K$), MAI is reconstructed and canceled out from the composite signal; a more accurate channel estimate for the i th symbol, $\hat{\mathbf{h}}_{k,i}$, is obtained and then utilized to get a more accurate estimate of the i th symbol, denoted as $\hat{\mathbf{x}}_{k,i}$;
4. Step 3 can be performed iteratively to achieve better performance by feeding $\hat{\mathbf{x}}_{k,i}$ back to further refine the channel estimation and signal detection;
5. $\hat{\mathbf{h}}_{k,i}$ is used as the channel information for the next symbol and Step 2 to Step 4 are repeated to process each symbol recursively in the rest of the data frame.

Details of the algorithm are presented in the following subsections.

2.3.1 Initial Channel Estimation

The initial channel estimation is performed in an OFDMA fashion as below. Suppose the maximum channel order is L , then M ($L < M < N$) subcarriers are chosen to

transmit pilot bits for each user. For simplicity, assume $N/M = Q$ is an integer. The pilot symbol for user k is then given by

$$x_{k,0}(n) = \begin{cases} \pm 1, & n \in \Theta_k, \\ 0, & \text{elsewhere,} \end{cases} \quad (2.11)$$

where ± 1 means the pilot bit is randomly chosen from $\{+1, -1\}$, Θ_k defines a set of uniformly Q -spaced subcarriers for user k ($1 \leq k \leq K$), given by $\Theta_k = \{0 \leq n \leq N-1 \mid n = (k - \lfloor \frac{k-1}{Q} \rfloor Q - 1) + m \cdot Q, m = 0, \dots, M-1\}$, where $\lfloor x \rfloor$ is the maximum integer less than or equal to x . If the number of total active users K is not greater than Q , then $\Theta_k \cap \Theta_l = \emptyset, \forall 1 \leq k, l \leq K, k \neq l$, that is, the users transmit their pilot bits in non-overlapping subcarrier groups. In this case, MAI is mitigated at the training stage, resulting in accurate initial channel estimation.

Remark 2.1 *Note that increasing M improves the performance of channel estimation in the single-user case. However, it reduces the value of Q at the same time. In the case when $K > Q$, more than one user need to transmit pilots on the same subcarriers. The resulting MAI degrades the channel estimation quality. To accommodate more users and perform MAI-free initial channel estimation, one can increase the value of N , yet it requires more bandwidth. An alternative solution is to allocate more MC symbols for initial channel estimation and distribute users' pilots in a time-frequency grid, which mimics a TDMA-FDMA scheme.*

Denoting $y_{k,0}(n)$ as the received signal on subcarrier n , the least-squares (LS) estimate of the channel frequency response on subcarrier n is given by

$$\tilde{H}_{k,0}(n) = y_{k,0}(n)/x_{k,0}(n), \quad n \in \Theta_k. \quad (2.12)$$

Stack the M elements $\tilde{H}_{k,0}(n), n \in \Theta_k$ into one vector $\tilde{\mathbf{H}}_{\Theta_k} = [\tilde{H}_{k,0}(n_1), \dots, \tilde{H}_{k,0}(n_M)]^T$, where $n_i \in \Theta_k$ for $i = 1, \dots, M$. The channel estimate in the time domain $\hat{\mathbf{h}}_{k,0} = [\hat{h}_{k,0}(0), \dots, \hat{h}_{k,0}(L)]^T$ can be obtained by

$$\hat{\mathbf{h}}_{k,0} = [\mathbf{F}]_{L+1, \Theta_k} \tilde{\mathbf{H}}_{\Theta_k}, \quad (2.13)$$

where $[\mathbf{F}]_{L+1, \Theta_k}$ is part of the IFFT matrix with the first $L+1$ rows and the columns with indices belonging to set Θ_k retained.

Augmenting the channel coefficients $\mathbf{h}_{k,0}$ to be $\mathbf{h}_{k,0}^A = [\mathbf{h}_{k,0}^T, \mathbf{0}_{N-L-1}^T]^T$, where $\mathbf{0}_{N-L-1}$ is an all-zero vector of dimension $(N-L-1) \times 1$, the channel response of user k in the frequency domain, $\hat{\mathbf{H}}_{k,0} = [\hat{H}_{k,0}(0), \dots, \hat{H}_{k,0}(N-1)]^T$, is then given by

$$\hat{\mathbf{H}}_{k,0} = \mathbf{F}^\dagger \mathbf{h}_{k,0}^A. \quad (2.14)$$

2.3.2 Joint Channel Estimation and Signal Detection

Relying on high correlation between the channel coefficients of two successive symbols, we use $\hat{\mathbf{h}}_{k,i-1}$ as the coarse CSI for the detection of the i th symbol. When the tentative decision $\tilde{\mathbf{d}}_{k,i}$ is obtained based on $\hat{\mathbf{h}}_{k,i-1}$, MAI is reconstructed and PIC is performed to cancel the interference. The “cleaner” signal for user k is then given by

$$\tilde{z}_{k,i}(n) = r(iN_s + \Delta_k + n) - \sum_{j=1, j \neq k}^K \hat{r}_j(n + \Delta_k - \Delta_j), \quad 0 \leq n \leq N_s - 1, \quad (2.15)$$

where the reconstructed signal for user j , $\hat{r}_j(n + \Delta_k - \Delta_j)$, is a function of $\tilde{\mathbf{d}}_{j,i}$ and $\hat{\mathbf{h}}_{j,i-1}$. It involves a convolution operation which requires $(L+1)$ complex multiplications and L complex additions. Hence the computational complexity of the MAI cancelation in (2.15) is $O[KN_s(L+1)]$, which is slightly larger than that of the matched filter (MF) bank, $O[2KN]$.

After the FFT demodulation, the output is given by

$$\tilde{\mathbf{y}}_{k,i} = \mathbf{X}_{k,i} \mathbf{H}_{k,i} + \sum_{j=1, j \neq k}^K \tilde{\mathbf{I}}_{j,i} + \mathbf{v}_{k,i}. \quad (2.16)$$

The major difference between (2.16) and (2.9) is that MAI has now been reduced to residual interference, and hence more accurate channel estimation can be achieved.

Estimate of $\mathbf{H}_{k,i}$ is then given by

$$\hat{\mathbf{H}}_{k,i} = \tilde{\mathbf{X}}^{-1} \tilde{\mathbf{y}}_{k,i}, \quad (2.17)$$

where $\tilde{\mathbf{X}} = \text{diag}(\tilde{\mathbf{x}}_{k,i})$ with $\tilde{\mathbf{x}}_{k,i}$ originating from the tentative decision $\tilde{\mathbf{d}}_{k,i}$. Due to the diagonal matrix of $\tilde{\mathbf{X}}$, the channel estimation procedure only requires N complex divisions.

With the updated channel coefficients, more accurate MAI can be reconstructed and canceled out. An even “cleaner” signal can be obtained and only single user detection (SUD) is required to obtain $\hat{\mathbf{d}}_{k,i}$, which is the final estimate of the original data $\mathbf{d}_{k,i}$. As discussed above, this PIC refinement procedure takes roughly twice as many arithmetic operations as that of the MF bank.

2.4 Detection and Tracking of Abrupt Channel Changes

The proposed scheme in the previous works well under the Jakes’ model, as will be demonstrated in the simulation examples. However, when channels experience unpredictable abrupt changes, severe error propagation could occur, resulting in significant performance degradation. To achieve tight error propagation control, in this section, simple but robust algorithms are proposed for effective detection and tracking of abrupt channel changes.

2.4.1 Detection of Abrupt Channel Changes

We assume channels are uncorrelated and no channel suffers abrupt changes simultaneously with any other channels. Without loss of generality, the time domain channel coefficients for user l at time i , $\mathbf{h}_{l,i} = [h_{l,i}(0), \dots, h_{l,i}(L)]^T$, can be expressed as

$$\mathbf{h}_{l,i} = \mathbf{h}_{l,i-1} + \Delta \mathbf{h}_{l,i}, \quad (2.18)$$

where $\Delta \mathbf{h}_{l,i}$ is the deviation from the channel coefficient $\mathbf{h}_{l,i-1}$. If we assume user v is *the* user whose channel undergoes abrupt changes at time i , then

$$\|\Delta \mathbf{h}_{v,i}\| \gg \|\Delta \mathbf{h}_{l,i}\|, \text{ for } l \neq v, \quad (2.19)$$

where $\|\Delta \mathbf{h}_{l,i}\| = \sqrt{\sum_{k=0}^L |\Delta h_{l,i}(k)|^2}$. In the extreme case when the channels are invariant, $\|\Delta \mathbf{h}_{l,i}\| = 0$.

The channel response of user l in the frequency domain is then given by

$$\begin{aligned} \mathbf{H}_{l,i} &= \mathbf{F}^\dagger \mathbf{h}_{l,i}^A \\ &= \mathbf{H}_{l,i-1} + \Delta \mathbf{H}_{l,i}, \quad 1 \leq l \leq K, \end{aligned} \quad (2.20)$$

where $\Delta \mathbf{H}_{l,i}$ is the frequency response of $\Delta \mathbf{h}_{l,i}^A = [\Delta \mathbf{h}_{l,i}^T, \mathbf{0}_{N-L-1}^T]^T$. According to Parseval's Theorem, it is also true that $\|\Delta \mathbf{H}_{v,i}\| \gg \|\Delta \mathbf{H}_{l,i}\|$, for $l \neq v$.

At the receiver end, if it is noiseless, after passing the FFT demodulator, the overall signal at subcarrier n is given by

$$r_i(n) = \sum_{l=1}^K H_{l,i}(n) x_{l,i}(n), \quad (2.21)$$

where $H_{l,i}(n)$ and $x_{l,i}(n)$ denote the channel coefficient and the transmitted signal on subcarrier n , respectively.

Following the PIC procedure, after the cancelation of the reconstructed MAI, the “cleaner” signal for user l , $y_{l,i}(n)$, is given by

$$y_{l,i}(n) = r_i(n) - \sum_{j \neq l} \hat{H}_{j,i-1}(n) \hat{x}_{j,i}(n), \quad (2.22)$$

where $\hat{H}_{j,i-1}(n)$ and $\hat{x}_{j,i}(n)$ are the estimates of $H_{j,i-1}(n)$ and $x_{j,i}(n)$, respectively. If the estimates are error-free, i.e., $\hat{x}_{j,i}(n) = x_{j,i}(n)$ and $\hat{H}_{j,i-1}(n) = H_{j,i-1}(n)$, then

$$y_{l,i}(n) = H_{l,i}(n) x_{l,i}(n) + \sum_{j \neq l} \Delta H_{j,i}(n) x_{j,i}(n). \quad (2.23)$$

The estimate of the channel coefficient at time i is then given by

$$\begin{aligned}\hat{H}_{l,i}(n) &= \frac{y_{l,i}(n)}{x_{l,i}(n)} \\ &= H_{l,i}(n) + \sum_{j \neq l} \Delta H_{j,i}(n) \frac{x_{j,i}(n)}{x_{l,i}(n)}.\end{aligned}\quad (2.24)$$

If $l = v$, then $\Delta H_{j,i}(n) \approx 0$ for any $j \neq l$ since there are no abrupt changes in these channels by assumption. As a result, the second term on the right-hand side of (2.24) is approximately 0, and we obtain

$$\hat{H}_{v,i}(n) \approx H_{v,i}(n) = H_{v,i-1}(n) + \Delta H_{v,i}(n). \quad (2.25)$$

If $l \neq v$,

$$\begin{aligned}\hat{H}_{l,i}(n) &= H_{l,i}(n) + \Delta H_{v,i}(n) \frac{x_{v,i}(n)}{x_{l,i}(n)} + \sum_{j \neq l,v} \Delta H_{j,i}(n) \frac{x_{j,i}(n)}{x_{l,i}(n)} \\ &\approx H_{l,i}(n) + \Delta H_{v,i}(n) \frac{x_{v,i}(n)}{x_{l,i}(n)}.\end{aligned}\quad (2.26)$$

Since we utilize constant envelope modulation such as BPSK or QPSK, (2.26) can be rewritten as

$$\begin{aligned}\hat{H}_{l,i}(n) &\approx H_{l,i}(n) + e^{j\theta_{lv}(n)} \Delta H_{v,i}(n) \\ &\approx H_{l,i-1}(n) + e^{j\theta_{lv}(n)} \Delta H_{v,i}(n), \quad l \neq v,\end{aligned}\quad (2.27)$$

where $\theta_{lv}(n)$ is the phase difference between $x_{l,i}(n)$ and $x_{v,i}(n)$. From (2.25) and (2.27), it is clear that for any l ($1 \leq l \leq K$),

$$|\hat{H}_{l,i}(n) - H_{l,i-1}(n)| \approx |\Delta H_{v,i}(n)|. \quad (2.28)$$

That is, when one user's channel is subjected to a sudden change, the channel estimates of all other users present a similar abrupt deviation. This finding leads to the following algorithm.

Proposition 2.1 An abrupt channel change is considered to have occurred at time instant i , if any one of the following K inequalities holds:

$$\frac{1}{N} \sum_{n=0}^{N-1} |\hat{H}_{l,i}(n) - \hat{H}_{l,i-1}(n)|^2 > \lambda, \quad 1 \leq l \leq K, \quad (2.29)$$

where λ is a “toleration threshold” which satisfies

$$\max_{l \neq v} \left\{ \frac{1}{N} \sum_{n=0}^{N-1} |\Delta H_{l,i}(n)|^2 \right\} < \lambda < \frac{1}{N} \sum_{n=0}^{N-1} |\Delta H_{v,i}(n)|^2. \quad (2.30)$$

Justified by Parseval’s Theorem, in practice, the algorithm can be carried out in the time domain. The “toleration threshold” λ is obtained through a recursive process. Assuming the abrupt change occurs at time i , it can be detected if any one of the following inequalities holds:

$$\frac{1}{L+1} \sum_{j=0}^L |\hat{h}_{l,i}(j) - \hat{h}_{l,i-1}(j)|^2 > \lambda(i), \quad 1 \leq l \leq K, \quad (2.31)$$

where $\lambda(i)$ is recursively obtained by

$$\lambda(i) = \max_l \left\{ \frac{1}{L+1} \sum_{j=0}^L |\hat{h}_{l,i-1}(j) - \hat{h}_{l,i-2}(j)|^2 \right\} + \epsilon. \quad (2.32)$$

When the channels are “smooth”, $|\hat{h}_{l,i-1}(j) - \hat{h}_{l,i-2}(j)|^2$ does not deviate very much for different i ’s. ϵ is a small empirical value larger than such deviation.

2.4.2 Extraction and Tracking of the Abrupt Channel

When the receiver detects the abrupt change at time i , it should identify which user’s channel is subjected to the change. From (2.10) and (2.25), the channel estimate for

user v at time i , $\hat{\mathbf{h}}_{v,i} = [\hat{h}_{v,i}(0), \dots, \hat{h}_{v,i}(L)]^T$, is given by

$$\begin{aligned}\hat{\mathbf{h}}_{v,i} &= [\mathbf{F}]_{L+1} \hat{\mathbf{H}}_{v,i} \\ &\approx \mathbf{h}_{v,i} = \mathbf{h}_{v,i-1} + \Delta \mathbf{h}_{v,i}.\end{aligned}\quad (2.33)$$

Similarly, from (2.10) and (2.27) the channel estimate for user l ($l \neq v$) is given by

$$\begin{aligned}\hat{\mathbf{h}}_{l,i} &= [\mathbf{F}]_{L+1} \hat{\mathbf{H}}_{l,i} \\ &= \mathbf{h}_{l,i-1} + \Delta \tilde{\mathbf{h}}_{l,i},\end{aligned}\quad (2.34)$$

where $\Delta \tilde{\mathbf{h}}_{l,i}$ deviates greatly from $\Delta \mathbf{h}_{l,i}$ due to the phase term $e^{j\theta_{lv}(n)}$ in (2.27).

In (2.33), $\hat{\mathbf{h}}_{v,i}$ approximates $\mathbf{h}_{v,i}$ due to the reason that the reconstructed signals of other users are well approximated. While in (2.34), $\{\hat{\mathbf{h}}_{l,i}, l \neq v\}$ deviate from $\mathbf{h}_{l,i}$ since they are corrupted by the strong interference from user v . It is then natural that we use $\hat{\mathbf{h}}_{v,i}$ to reconstruct user v 's signal and resort to $\{\hat{\mathbf{h}}_{l,i-1}, l \neq v\}$ to reconstruct other users' signals.

If user v has been correctly identified, the “cleaner” signal for user l ($l \neq v$) is then given by

$$\tilde{z}_{l,i}^v(n) = r(iN_s + \Delta_l + n) - \tilde{r}_v(n + \Delta_l - \Delta_v) - \sum_{j \neq l,v} \hat{r}_j(n + \Delta_l - \Delta_j), \quad (2.35)$$

where $\tilde{r}_v(n + \Delta_l - \Delta_v)$ is the reconstructed signal for user v , being a function of $\hat{\mathbf{h}}_{v,i}$ instead of $\hat{\mathbf{h}}_{v,i-1}$. Following the analysis in Section 2.4, it can be seen that after extracting the reconstructed signal of user v , the channel estimate of user l ($l \neq v$), $\tilde{\mathbf{h}}_{l,i}^v$, can be given by

$$\tilde{\mathbf{h}}_{l,i}^v \approx \mathbf{h}_{l,i-1} + \Delta \mathbf{h}_{l,i}. \quad (2.36)$$

Hence, if user v has been identified correctly and its reconstructed signal using $\hat{\mathbf{h}}_{v,i}$ has been extracted, then $\|\tilde{\mathbf{h}}_{l,i}^v - \hat{\mathbf{h}}_{l,i-1}\| \approx \|\Delta \mathbf{h}_{l,i}\|$ is a small value. However, if the abrupt channel is incorrectly identified to be user k ($k \neq v$), the channel estimate of user l ($l \neq v, k$), $\tilde{\mathbf{h}}_{l,i}^k$, deviates from $\mathbf{h}_{l,i-1}$ significantly because, as shown in (2.34),

significant interference is introduced when $\hat{\mathbf{h}}_{k,i}$ is used to reconstruct the signal of user k .

Therefore, the value of $\|\tilde{\mathbf{h}}_{l,i}^k - \hat{\mathbf{h}}_{l,i-1}\|$ ($l \neq k$) is only minimized when the user who experienced abrupt channel changes is correctly identified, i.e., when $k = v$. The algorithm is analytically presented below.

Proposition 2.2 If the abrupt channel change at time instant i has been detected, user v , whose channel undergoes the abrupt change, can be identified through:

$$v = \arg \min_{\Omega_k} \left\{ \sum_{l \neq k} \|\tilde{\mathbf{h}}_{l,i}^k - \hat{\mathbf{h}}_{l,i-1}\| \right\}, \quad (2.37)$$

where $\Omega_k = \{k : 1 \leq k \leq K\}$.

Remark 2.2 An optimal solution to perform joint abrupt channel identification and signal detection is given by

$$(v, \{\tilde{\mathbf{d}}_{m,i}\}) = \arg \min_{(\Omega_k, \Omega_d)} \left\{ \sum_{l \neq k} \|\tilde{\mathbf{h}}_{l,i}^k - \hat{\mathbf{h}}_{l,i-1}\| \right\}, \quad (2.38)$$

where $\Omega_d = \{\mathbf{d}_{m,i} | \mathbf{d}_{m,i} \in \{-1, 1\}^J, 1 \leq m \leq K\}$ for BPSK, and $\Omega_d = \{\mathbf{d}_{m,i} | \mathbf{d}_{m,i} \in \{\pm 1 \pm j\}^J, 1 \leq m \leq K\}$ for QPSK. The complexity of this optimal algorithm grows exponentially with the number JK , making it only applicable to some small systems. However, suboptimal solutions could be found when antenna arrays are employed at the base station. When the antennas are employed to minimize the channel correlation between each other, it is then possible to obtain reliable $\{\tilde{\mathbf{d}}_{m,i}, 1 \leq m \leq K\}$ through the channels over which no abrupt changes occur.

2.5 Simulation Results

In this section, performance of the proposed algorithm is illustrated through simulation examples. In the simulations, QPSK is adopted as the modulation scheme. The entire channel bandwidth is 8MHz and divided into $N = 128$ subchannels, which im-

plies that the informative part of an MC symbol has a duration of $16\mu s$. All channels are fast-fading, and are characterized using an 8-ray multipath model with exponential power delay profiles. The extended MC symbol has a CP of $1\mu s$ to eliminate the ISI. The system is quasi-synchronous with the initial transmission delays being uniformly distributed within $[0, 1]\mu s$. We assume that the carrier frequency is $2GHz$, and the maximum Doppler shift is $200Hz$, which corresponds to a vehicle speed of $110km/h$. For diversity, two receive antennas are employed at the base station. They are assumed to be far apart enough such that the channels are uncorrelated with each other. The spreading factor is $P = 16$ and the system is half loaded with the number of active users being $K = 8$.

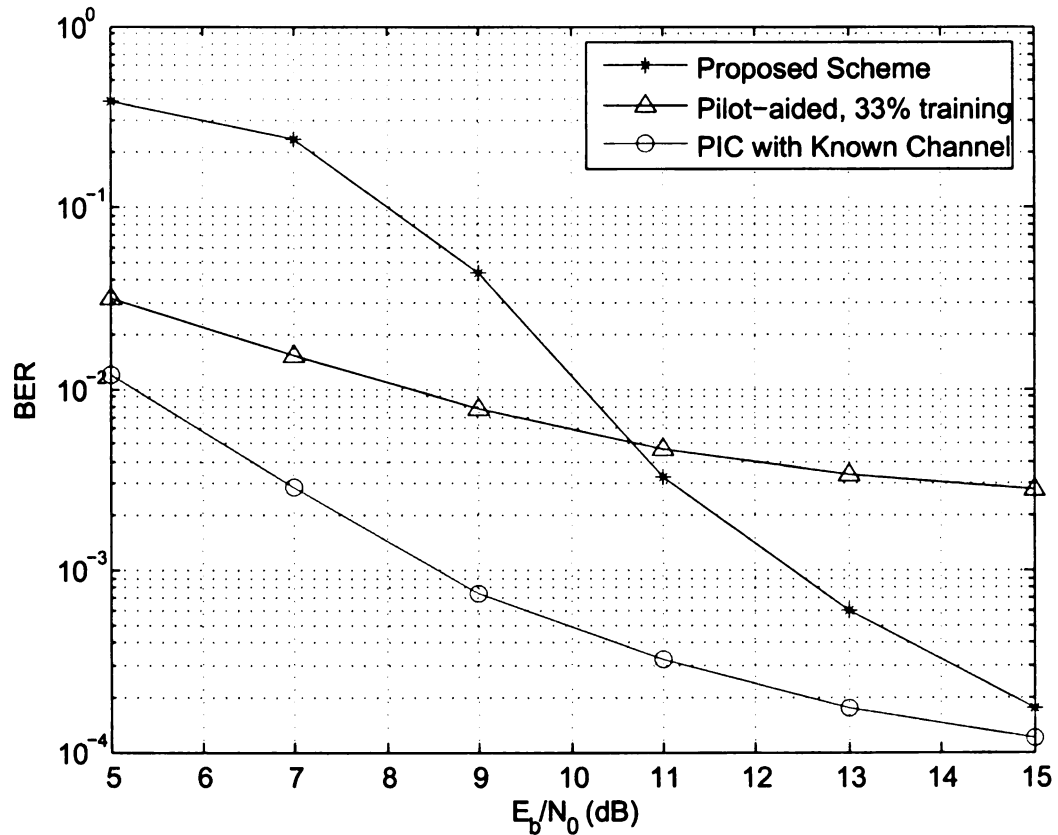


Figure 2.2. BER performance under the Jake's model with no abrupt changes, $K = 8$, $P = 16$.

2.5.1 Performance under the Jakes' Model with No Abrupt Changes

Bit error rate (BER) performance of the system under the Jake's model is illustrated in Figure 2.2, where no abrupt change is introduced. In this simulation example, $M = P = 16$ subcarriers are used for transmitting each user's pilot bits, $K = 8$ users transmit their pilot bits on different subcarriers following the pattern described in Section 2.3.1. Only one iteration is performed for channel update. It is found through simulations that one iteration is good enough to provide near-optimum channel update, and additional iterations can barely improve the system performance. In Figure 2.2, the circled line corresponds to the method using PIC with perfect CSI knowledge, which serves as the benchmark for BER comparison. It is demonstrated that the proposed scheme approaches the ideal case when SNR increases.

Performance of the pilot-aided scheme is illustrated in Figure 2.2 as well. In the fast fading environment, the pilot-aided schemes require that the pilots be inserted frequently to guarantee good performance. In this comparison, the training signals occupy 33% of the total data length and the linear interpolation technique [CEPB02] is exploited for channel interpolation. It is shown that the proposed scheme delivers a significantly better performance than the pilot-aided schemes when the SNR is larger than 12 dB. The underlying reason is that the proposed channel update in high SNR regions becomes more accurate and thus improves the overall system performance. This argument is further justified by the results shown in Figure 2.3, where the mean-squared error (MSE) performances of the proposed channel estimation at different SNR levels are plotted. It is shown that for the SNR at 20dB, the channel estimation MSE is much smaller than that of the initial estimates, while for the SNR at 12dB, the MSE is roughly at the same level with the initial estimates.

Figure 2.4 shows the BER performances of the proposed scheme for systems with imperfect time synchronization. Two scenarios are considered: (i) All users have the same timing shift, referred to as "*common offset*"; (ii) The users have independent timing errors uniformly distributed within a certain interval, referred to as "*inde-*

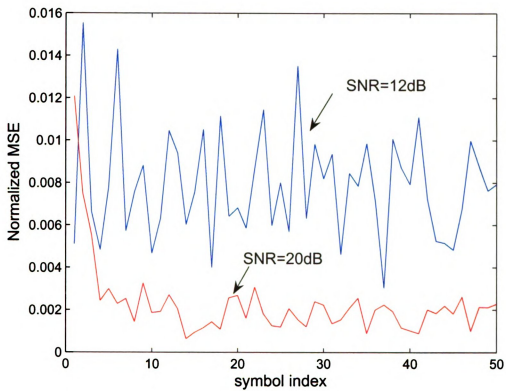


Figure 2.3. MSE performances of the proposed channel estimation at different SNR levels.

pendent offsets". It is observed that the more active users in the system, the more sensitive the system is to timing errors. As can be seen from Figure 2.4, the half loaded system works well when the timing error is insignificant. However, as expected, when timing jitters get worse, performance degradation is no longer negligible.

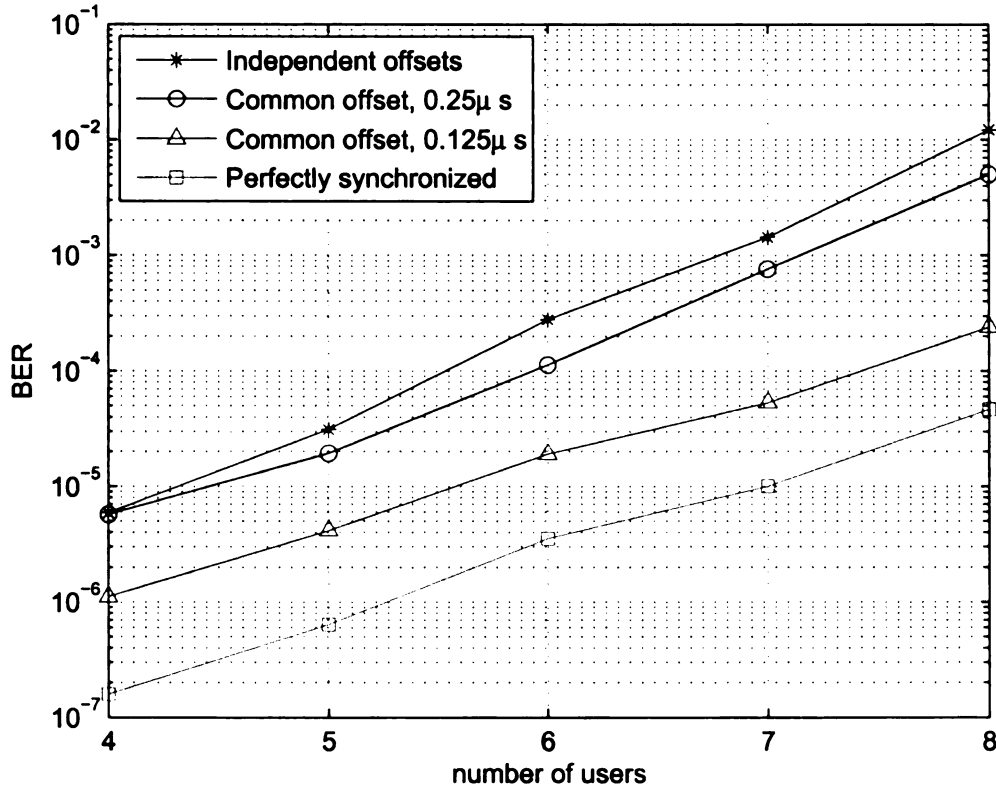


Figure 2.4. Impact of timing errors on the BER performance, $SNR = 25dB$: *independent offsets* means that users have independent timing errors uniformly distributed within interval $[0, 0.25]\mu s$, *common offsets* means that all users experience the same timing shift.

2.5.2 Performance under the Jakes' Model with Abrupt Changes

In the case when abrupt changes occur, significant performance degradation could happen due to error propagation. Therefore, special efforts need to be taken to track

these unpredictable changes. In the simulation example, we introduce the abrupt change based on the following model:

$$\mathbf{g}_{v,i} = \tilde{\alpha} \mathbf{h}_{v,i} + \mathbf{w}_{v,i}, \quad (2.39)$$

where $\mathbf{h}_{v,i}$ is user v 's channel response at time instant i . It is generated from the Jakes' model and is normalized with $E[\|\mathbf{h}_{v,i}\|^2] = 1$. $\mathbf{g}_{v,i}$ is the new channel response with the abrupt change introduced by $\tilde{\alpha}$ and $\mathbf{w}_{v,i}$, where $\tilde{\alpha} = \text{diag}(\alpha)$. α and $\mathbf{w}_{v,i}$ are two $(L+1) \times 1$ complex Gaussian vector with zero mean and covariance matrices of $\sigma_\alpha^2 \mathbf{I}_{L+1}$ and $\sigma_w^2 \mathbf{I}_{L+1}$, respectively.

In the simulations, we assume that the channels are uncorrelated. User 1 undergoes abrupt channel changes at the first receiver antenna, while the channel at the second antenna is "smooth". Hence the detected signal at the second antenna can assist the channel update procedure at the first antenna. Specifically, the parameters for the abrupt channel model introduced above are $\sigma_\alpha^2 = 0.06$ and $\sigma_w^2 = 0.5$. The channels of the other users are assumed to be "smooth". Under the above settings, the abrupt channel can be successfully detected with the value of ϵ in (2.32) empirically set to be 0.2.

Suppose the system is half loaded with $K = 8$. By varying k from 1 to K in (2.37), we have different evaluation of the term $\phi_k = \sum_{l \neq k} \|\tilde{\mathbf{h}}_{l,i}^k - \hat{\mathbf{h}}_{l,i-1}\|$, which corresponds to the *accumulated channel deviation* when user k is identified to be subjected to abrupt channel changes. It is shown in Figure 2.5 that ϕ_k is minimized when $k = v = 1$, i.e., when the right user has been identified. Only results of 4 out of 8 users are presented, and the rest of the "smooth" users are omitted for clearer illustration.

In the simulation example, it has been assumed that a burst of abrupt changes appear in user 1's channel. As can be seen from Figure 2.6, accuracy of channel estimation is greatly improved after the abrupt channel is correctly identified. It is shown in Figure 2.7 that with effective abrupt channel tracking, disastrous error propagation is avoided, and the signal detection performance (BER) in systems with

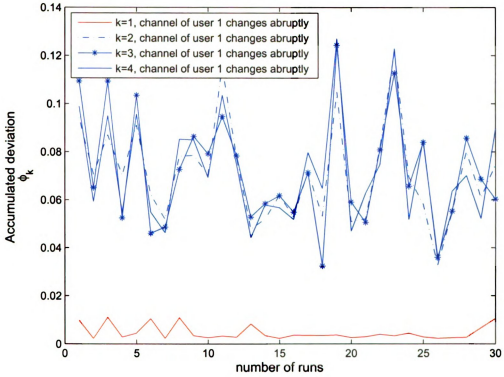


Figure 2.5. Accumulated channel deviation ϕ_k when user k ($k = 1, 2, 3, 4$) is identified to be subjected to abrupt channel changes, while user 1 is the true user whose channel undergoes abrupt changes and other users' channels are "smooth". $K = 8$, $P = 16$.

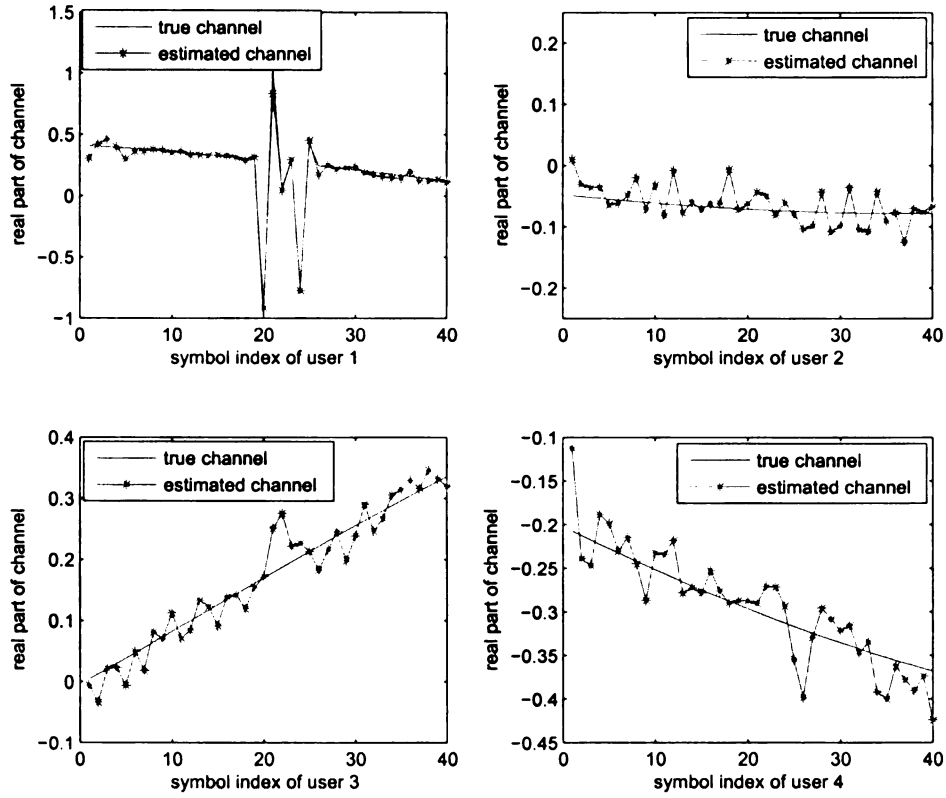


Figure 2.6. Channel estimates (only real part of the channel is shown) with the proposed error propagation control: $SNR = 20dB$, $K = 8$, $P = 16$, user 1 undergoes abrupt channel changes. Only results of 4 out of 8 users are presented.

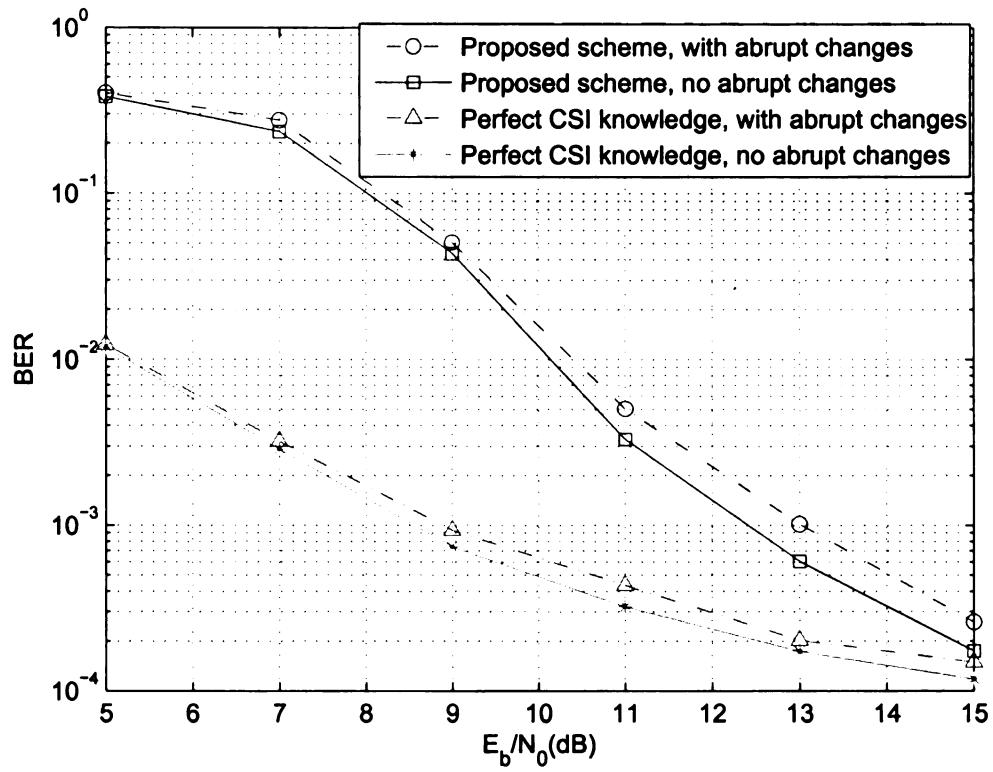


Figure 2.7. BER performances of the system over channels with and without abrupt changes. Here “*perfect CSI*” means that the channel state information is perfectly known at the receiver, including information on abrupt changes.

abrupt changes is comparable with that of systems with no abrupt changes.

2.6 Summary

In this chapter, a recursive approach is proposed for joint channel estimation and signal detection of quasi-synchronous MC-CDMA systems over time-varying channels. Only one pilot symbol is needed for initial channel estimation. Subsequent signal detection and channel update are carried out by exploiting the channel information from the previous symbol. The proposed scheme hence possesses much higher spectral efficiency than conventional pilot-aided methods. Furthermore, for channels subjected to unpredictably sudden changes, effective abrupt channel tracking algorithms have been proposed to ensure tight error propagation control. Simulation results demonstrated the effectiveness and robustness of the proposed algorithms.

CHAPTER 3

Mutually Interactive Random Access Control for Wireless Networks

This chapter considers cross-layer medium access control (MAC) protocol design in wireless networks. Taking a mutually interactive MAC-PHY perspective, we aim to design an MAC protocol that is in favor of the PHY layer information transmission, and the improved PHY, in turn, can improve the MAC performance. Motivated by the fact that as long as good channel estimation can be achieved, advanced signal processing does allow effective signal separation given that the multiuser interference is limited to a certain degree, we propose a novel MAC protocol, named hybrid ALOHA, which allows conditional collision-free channel estimation and simultaneous multiuser transmission. More specifically, short idle sections are introduced into the structure of the traditional ALOHA slot, so that it is possible for different users to transmit their training sequences in non-overlapping training slots, whereby collision-free channel estimation could be achieved. Relying on the general multipacket reception (MPR) model, quantitative analysis is conducted for the proposed hybrid ALOHA protocol in terms of throughput, stability as well as delay behavior. Significant performance improvement can be achieved in comparison with the traditional ALOHA protocol based either on the collision model or the MPR model.

3.1 Introduction

Over the past decades, studies on medium access control (MAC) have largely relied on the “collision” model: a packet is successfully received if and only if there are no concurrent transmissions [Rob72, Gal85, EH98]. MAC protocols such as traditional ALOHA [Abr70, Rob72], the tree algorithms [Cap79], the window random-access algorithm [PPK89] and a number of adaptive MAC schemes [HG81, Hlu82, KY78], have all been developed by assuming that the physical (PHY) layer is characterized as an idealized collision model [ZT03b]. Although it is convenient for analysis, the collision model fails to represent all the characteristics of the PHY layer. As pointed out by Tong *et al.* in [TNV04], the model is optimistic as it ignores the channel effects such as fading and noise on reception, whereas the model is also pessimistic since it overlooks the possibility that the packets may be successfully decoded even in the presence of simultaneous multiuser transmissions.

Going beyond the collision model and assuming a multiuser PHY layer, Ghez, Verdú and Schwartz proposed the multipacket reception (MPR) model in [GVS88, GVS89], where the reception of packets is characterized by conditional probabilities instead of deterministic failures or successes. The MPR model provides a more realistic interface between the PHY layer and the MAC layer, and propels the research on cross-layer MAC design. In [Met76, RSG87, HKS89, Zor95], the impact of MPR on the performance of slotted ALOHA is studied by investigating the capture model, which, being a subclass of the MPR, assumes that *at most* one user’s packet can be successfully decoded when multiple users transmit at the same time. More examples on protocol design for networks in the presence of capture can be found in [LPK89, YD00, AD01, SC85]. Based on the general MPR model beyond the capture channel, in [ZT03b] and [ZT03a], a multi-queue service room (MQSR) protocol and a simpler dynamic queue protocol have been developed to adaptively grant channel access to users, respectively.

One limitation with the MPR model of Ghez et al. is that it assumes a symmetric model with indistinguishable users. To differentiate users in a multimedia

networks where each user may have his own rate, in [NMT05], Naware, Mergen and Tong introduced a more general, asymmetric MPR model as the interfaces between the MAC layer and the PHY layer. Based on the generalized MPR model, their study [NMT05] on the stability and delay of slotted ALOHA leads to an interesting result: *if the PHY layer is reasonably reliable, then slotted ALOHA with transmission probability 1 is optimal in the sense that the stability region coincides with the MAC capacity region. Hence, there is no need for sophisticated MAC protocols when the optimality is achieved.* Here, the *MAC capacity region* is defined as the maximum throughput achievable by any MAC protocols without considering queue stability [GT02, MT05, NMT05]. The *stability region* of a particular MAC protocol is defined to be the set of arrival rates for which there exists retransmission probabilities that make the probability of buffer overflow arbitrarily small [Szp94, LE99, NMT05]. Generally, the stability region is contained in the capacity region. The result in [NMT05] implies that the burden at the MAC layer can be reduced significantly if we can ensure a reasonably good PHY layer.

As is well known, being an important characteristic that directly influences the quality of the PHY layer, the channel state information (CSI) plays a critical role in signal detection and estimation. The more accurate the CSI estimation is, the more likely that the signal can be recovered at the receiver, and hence the stronger the MPR capability. There has been a line of work studying the effect of CSI on resource allocation, e.g., [AT05, TH98, SV01, TE93]. Assumptions are usually made that either transmitters or receivers can track the channel state, and the knowledge of such information is employed to improve the network capacity. In other words, existing works have been focused on improving the MAC performance by exploiting the PHY layer transmission capability. Here, we turn to examine the effect of the MAC protocol design on the PHY layer performance, and study how the mutual MAC-PHY interaction will influence the overall system capacity. More specifically, as illustrated in Figure 3.1, *taking a mutually interactive MAC-PHY perspective, we aim to design an MAC protocol that is in favor of the PHY layer information transmission, and the improved PHY layer, in turn, can improve the performance in*

the sense of throughput, stability and delay.

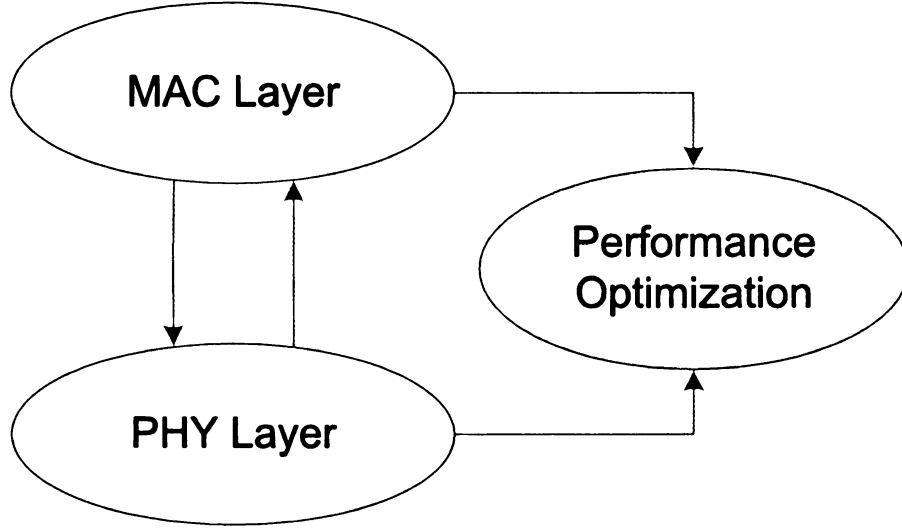


Figure 3.1. Mutual interaction of MAC layer and PHY layer to achieve performance optimization.

In this chapter, we propose a *hybrid ALOHA* random access protocol which makes collision-free channel estimation possible and allows simultaneous multiuser data transmission, based on the fact that as long as good channel estimation can be achieved, advanced signal processing does allow effective signal separation given that the multiuser interference is limited to a certain degree. In the hybrid ALOHA protocol, idle sections are inserted to each slot to make it possible for different users to transmit their training sequences at nonoverlapping sub-slots, therefore allowing collision-free channel estimation.

The hybrid ALOHA protocol is evaluated in terms of throughput, stability and delay behaviors. The major results are outlined as follows:

- **Throughput** Hybrid ALOHA allows conditional simultaneous transmission. Depending on whether the training sequences collide or not, the quality of CSI estimation varies, resulting in varying MPR capabilities. Relying on the existence of collision-free channel estimation, the increased MPR capability compensates for the bandwidth cost of the idle sections in the protocol and leads

to significant improvement in the system throughput.

- **Stability region** We characterize the stability region for the two-user system and derive the sufficient condition for the stability of the general finite-user system. Stability inner bound for the special three-user system is also characterized, which turns out to be solving a Riemann-Hilbert Problem. Study in the two-user case shows that when the MPR capability reaches a certain critical level, hybrid ALOHA has a convex stability region identical to that of a TDMA scheme, which possesses a frame guard time of the same length with the idle section of hybrid ALOHA. In the cases when the MPR capabilities are beyond the critical level, hybrid ALOHA outperforms TDMA by containing the stability region of the latter as a subset. The convexity of the stability region shows the optimality of the protocol in terms of its coincidence with the capacity region.
- **Delay behavior** Based on the general MPR model, we study the delay performance for the two-user system. The “exact” delays for the asymmetric system, and the upper and lower bounds of the average delay for the symmetric case are derived. It is shown that in the circumstances when the MPR channels are sufficiently strong, the difference between the upper and lower bounds becomes trivial and the actual delay can be roughly determined. When the MPR model is reduced to the capture model, the two bounds merge and we obtain the exact delay, which coincides with the result of [NMT05].

The rest of the chapter is organized as follows. In Section 3.2 the system model and the proposed hybrid ALOHA protocol are presented. In Section 3.3, the throughput of the system is analyzed. In Section 3.4 the stability region is derived for the two-user case. In addition, the sufficient condition for stability is derived for the general finite-user system. Section 3.5 studies exclusively the stability inner bound of the specific three-user system. The solution is found to be that of a Riemann-Hilbert boundary value problem. The delay behavior is studied in Section 3.6. To provide in-depth insights into the problem, an illustrative example is presented in Section 3.7. We conclude in Section 3.8.

3.2 The Hybrid ALOHA Protocol

3.2.1 System Model

Consider a wireless network with a set $\mathcal{N}$ of users, $\mathcal{N} = \{1, 2, \dots, N\}$, communicating with a common access point. Each user is equipped with an infinite buffer for storing arriving and backlogged packets. The packet arrival processes are assumed to be independent from user to user. The channel is slotted in time, with slot period larger than the packet length. When the buffer of the i th ($i \in \mathcal{N}$) user is nonempty, he tosses a coin to determine whether to transmit in the current slot or not. The probability of transmission of the i th user at each slot is assumed to be p_i . Packets are assumed to be of equal size for all users and composed of two parts: the first part is the *training sequence* for channel estimation, and the second part is the *information data*. The length of the training sequence is typically much smaller than that of the information data. The arrivals of the i th user are assumed to be independent and identically distributed (i.i.d) Bernoulli random variables from slot to slot, with the average number of arrivals being λ_i *packets per slot*. If this arrival rate is measured in *packets per unit time*, we then use Λ_i to denote it. Throughout the paper, we use the slot length of traditional ALOHA as the reference time unit and denote it as 1. Note that the slot length is protocol-dependent and may not always be 1.

We adopt the general MPR model in [NMT05] where the multiuser PHY layer is characterized by a set of conditional probabilities. For any subset $\mathcal{S} \subseteq \mathcal{N}$ of users transmitting in a slot, the *marginal* probability of successfully receiving packets from users in $\mathcal{R} \subseteq \mathcal{S}$, given that users in $\mathcal{S}$ transmit, is defined as

$$q_{\mathcal{R}|\mathcal{S}} = \sum_{\mathcal{U}: \mathcal{R} \subseteq \mathcal{U} \subseteq \mathcal{S}} q_{\mathcal{U},\mathcal{S}}, \quad (3.1)$$

where $q_{\mathcal{U},\mathcal{S}}$ is the conditional probability of reception defined as

$$q_{\mathcal{U},\mathcal{S}} = \Pr\{\text{only packets from } \mathcal{U} \text{ are successfully received} \mid \text{users in } \mathcal{S} \text{ transmit}\}, \mathcal{U} \subseteq \mathcal{S}. \quad (3.2)$$

In the two-user case, for example, $\mathcal{N} = \{1, 2\}$, for $i = 1, 2$,

$$\begin{aligned} q_{i,\{i\}} &= Pr\{\text{user } i \text{ is successful} \mid \text{only user } i \text{ transmits}\}, \\ q_{i,\{1,2\}} &= Pr\{\text{user } i \text{ is successful} \mid \text{both users transmit}\}, \\ q_{\{1,2\},\{1,2\}} &= Pr\{\text{both users are successful} \mid \text{both users transmit}\} \end{aligned} \quad (3.3)$$

and the marginal probabilities

$$q_{i|\{i\}} = q_{i,\{i\}}, \quad q_{i|\{1,2\}} = q_{i,\{1,2\}} + q_{\{1,2\},\{1,2\}}. \quad (3.4)$$

Clearly, $q_{\mathcal{R}|\mathcal{S}}$ depends on the quality of the PHY layer.

Assume that at the end of each slot, the receiver gives an instantaneous feedback of all the packets that were successfully received to all the users. The users remove successful packets from their buffers while unsuccessful packets are retained. Let N_i^t denote the queue length of the i th user at the beginning of time slot t , the queue evolution function for the i th ($i \in \{1, 2, \dots, N\}$) queue is given by [Szp94, NMT05]

$$N_i^{t+1} = [N_i^t - Y_i^t]^+ + \beta_i^t, \quad (3.5)$$

where β_i^t is the number of arrivals during the t th slot to the i th user with $E(\beta_i^t) = \lambda_i < \infty$, Y_i^t is the Bernoulli random variable denoting the number of departures from queue i in time slot t , and $[x]^+ = \max(0, x)$.

3.2.2 Hybrid ALOHA

The proposed hybrid ALOHA protocol aims at improving MPR capability by allowing conditional collision-free channel estimation and simultaneous transmission. In hybrid ALOHA, each slot contains sub-slots including training sections, data sections and idle sections. Idle sections are introduced to make it possible for different users to transmit their training sequences at non-overlapping sub-slots, whereby collision-free channel estimation could be achieved. We assume that the physical layer can accommodate

M users, that is, given reasonably accurate channel estimation, the user's packet can be successfully decoded if and only if there are no more than M simultaneous users. In hybrid ALOHA, $M - 1$ idle sections are inserted to each traditional ALOHA slot, such that collision-free channel estimation is made possible for each user. Note that in the proposed hybrid ALOHA, each slot has multiple pilot sub-slots, and each user can randomly select a pilot sub-slot to transmit his training sequence.

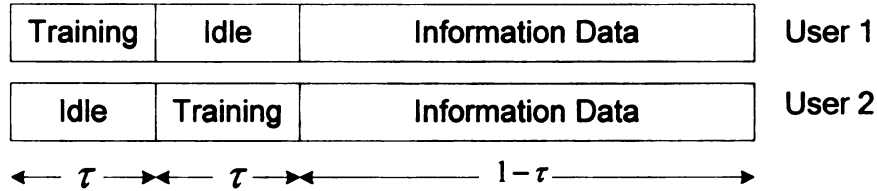


Figure 3.2. Illustration of the hybrid ALOHA slot structure for $M = 2$, in the case that two users ($N = 2$) transmit their training sequences at non-overlapping pilot sub-slots.

Figure 3.2 illustrates the slot structure of hybrid ALOHA in the case of $M = 2$, i.e., successful transmission is possible only when there are no more than two simultaneous users. Each slot has $M + 1 = 3$ sub-slots. The preceding two sub-slots, each having a length of τ , are the “pilot sub-slots” reserved for training sequences. *When a user is involved in a transmission, we assume that the selection of the pilot sub-slots is of equal probability.* The information data is always transmitted in the last sub-slot referred to as the “data sub-slot”. We assume that the length of the data sub-slot is $1 - \tau$ with $\tau \ll 1$. Recall that the length of the traditional ALOHA slot is used as the reference time unit, denoted as 1, which consists of a training sub-slot of length τ and a data sub-slot of length $1 - \tau$.

3.3 Throughput Calculation

As a figure of merit, the throughput of the system, which is a function of τ and denoted here as $\nu(\tau)$, is calculated below. The throughput is determined by finding

the average traffic successfully getting through the channel, i.e., $E[T_t]$, where $E[\cdot]$ is the expectation operator and T_t denotes the number of packets successfully getting through the channel during slot t . Let Λ be the overall arrival rate to the system, in the unit of *number of packets per unit time*. The average traffic per slot is $R(\tau) = \Lambda(M\tau + 1 - \tau)$. Let A_S be the event that users in $\mathcal{S}$ transmit, and $A_S^{\mathcal{R}}$ the event that only users in $\mathcal{R}$ succeed given that users in $\mathcal{S}$ transmit. The throughput for a general MPR model is given by

$$\begin{aligned}
\nu(\tau) &= E[T_t] \\
&= \sum_{\mathcal{S} \subseteq \mathcal{N}} E[T_t | A_S] \cdot Pr\{A_S\} \\
&= \sum_{\mathcal{S} \subseteq \mathcal{N}} \sum_{\mathcal{R} \subseteq \mathcal{S}} E[T_t | A_S^{\mathcal{R}}] \cdot Pr\{A_S^{\mathcal{R}}\} \cdot Pr\{A_S\} \\
&= \sum_{\mathcal{S} \subseteq \mathcal{N}} \sum_{\mathcal{R} \subseteq \mathcal{S}} R(\tau) \frac{|\mathcal{R}|}{|\mathcal{S}|} \cdot q_{\mathcal{R}, \mathcal{S}} \cdot Pr\{A_S\}, \tag{3.6}
\end{aligned}$$

Here $|\mathcal{R}|$ and $|\mathcal{S}|$ denote the number of elements in sets $\mathcal{R}$ and $\mathcal{S}$ respectively, and it follows from equation (3.2) $Pr\{A_S^{\mathcal{R}}\} = q_{\mathcal{R}, \mathcal{S}}$.

For the symmetric MPR model where users are indistinguishable, a more tractable form can be obtained. Denote A_K as the event that K users transmit in one slot and denote A_K^i as the event that exactly i out of K users succeed in transmission, then

$$\begin{aligned}
\nu(\tau) &= E[T_t] \\
&= \sum_{K=1}^M E[T_t | A_K] \cdot Pr\{A_K\} \\
&= R(\tau) \sum_{K=1}^M \sum_{i=1}^K \frac{i}{K} Pr\{A_K^i | A_K\} Pr\{A_K\}, \tag{3.7}
\end{aligned}$$

where $Pr\{A_K^i | A_K\}$ is determined by the MPR capability of the system, whereas $Pr\{A_K\}$ relies on the packet arrival model. We have assumed that the packet arrivals for the users follow the Bernoulli processes, and the overall arrival rate is Λ . If the number of users in the system is sufficiently large, the binomial distribution can

be closely approximated using the Poisson distribution [Fis63]. The closed-form expression of $Pr\{A_K\}$ can be obtained by assuming that the packets arrival follows the Poisson distribution with parameter $R(\tau) = \Lambda(M\tau + 1 - \tau)$. The result is summarized in the following proposition.

Proposition 3.1 *Under the symmetric MPR model, as well as the Poisson approximation with parameter $R(\tau) = \Lambda(M\tau + 1 - \tau)$, the throughput $\nu(\tau)$ of hybrid ALOHA, measured in packets per hybrid ALOHA slot, is given by*

$$\nu(\tau) = R(\tau) \sum_{K=1}^M \sum_{i=1}^K \frac{i}{K} Pr\{A_K^i | A_K\} \frac{R(\tau)^{K-1} e^{-R(\tau)}}{(K-1)!}. \quad (3.8)$$

To further simplify the expression of $\nu(\tau)$, we assume that users colliding in pilot sub-slots fail in transmission, whereas those who have collision-free channel estimation survive, given that there are no more than M users transmitting simultaneously in one slot. This simplified model is called the **Simplistic Assumption**. Evaluation of the throughput under this assumption is carried out in the following. When $M = 1$, Eq. (3.8) results in $\nu(\tau) = R(\tau)e^{-R(\tau)}$, which corresponds to the traditional slotted ALOHA system. For $M = 2$, assuming users are symmetric, it is then easy to know that $Pr\{A_1^1 | A_1\} = 1$, $Pr\{A_2^1 | A_2\} = 0$ and $Pr\{A_2^2 | A_2\} = 1/2$. Thus,

$$\begin{aligned} \nu(\tau) &= R(\tau) \left[e^{-R(\tau)} Pr\{A_1^1 | A_1\} + R(\tau) e^{-R(\tau)} \left(\frac{1}{2} Pr\{A_2^1 | A_2\} + Pr\{A_2^2 | A_2\} \right) \right] \\ &= [R(\tau)^2/2 + R(\tau)] e^{-R(\tau)}. \end{aligned} \quad (3.9)$$

Similarly, when $M = 3$, we have $\nu(\tau) = R(\tau)e^{-R(\tau)} [1 + \frac{2}{3}R(\tau) + \frac{2}{9}R(\tau)^2]$.

An alternative proof for the $M = 2$ case is provided in Appendix A, where we calculate the throughput of the system based on the Bernoulli arrivals model. The result coincides with that in (3.9) and demonstrates the validity of the Poisson approximation in Proposition 3.1.

Remark 3.1 *Clearly, the throughput of the system is a function of τ . Increasing τ allows a longer training sequence, hereby resulting in better channel estimation.*

More specifically, as will be shown in Section 3.7, the mean-squared error of channel estimation is inversely proportional to τ . Hence, increasing τ will result in improved MPR capability, which is desirable to improve the overall system performance. On the other hand, however, a larger τ also implies longer idle periods, leading to spectral inefficiency. Therefore, a tradeoff should be made on the design of τ for different applications. Please refer to Section 3.7 for a specific example.

Figure 3.3 presents the throughput comparison of the schemes for $M = 1, 2, 3$, respectively. Here we choose $\tau = 0.05$. The reference time unit is 1, which is the length of one slot in the case of $M = 1$, corresponding to the traditional ALOHA scheme. Note that in Proposition 1, the throughput is measured in the unit of “packets” per “hybrid ALOHA slot”, with slot length $[1 + (M - 1)\tau]$. For throughput comparison with traditional ALOHA, we need to normalize the throughput of hybrid ALOHA by the slot length. In other words, the throughput is normalized to “number of packets per unit-time”, which is $\nu(\tau)/[1 + (M - 1)\tau]$. It is shown that when $M = 2$, the proposed scheme has a 46% gain over traditional ALOHA ($M = 1$) in terms of maximum throughput, and the gain doubles for the $M = 3$ case. That is, the throughput tends to increase as M increases. At the point where the maximum throughput is achieved, the proposed protocol has an expected number of attempting transmissions approximately 1.4 and 1.95 for $M = 2$ and $M = 3$, respectively, much larger than that of traditional slotted ALOHA which takes on the value 1. This can be intuitively understood as: the larger the value of M , the stronger the MPR capability. As will be seen in Section 3.7, due to the small length of τ , the cost of extra bandwidth can be compensated for by the improved MPR capability.

Without loss of generality, in what follows, the stability regions and the delay performance of the hybrid ALOHA system are derived for the specific case of $M = 2$, that is, at most two simultaneous users can transmit in one slot in order for successful reception.

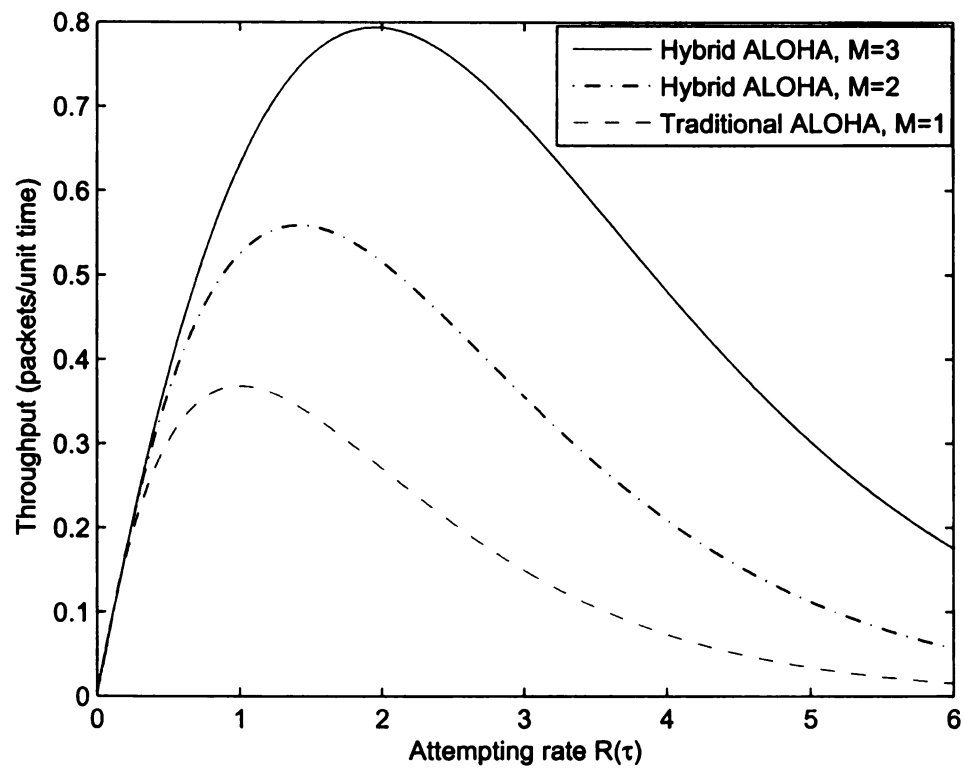


Figure 3.3. Throughput comparison between hybrid ALOHA and traditional slotted ALOHA.

3.4 Stability of Hybrid ALOHA

We study the stability performance of hybrid ALOHA in this section.

Definition 3.1 [Szp94] *A multidimensional stochastic process $\mathbf{N}^t = (N_1^t, \dots, N_N^t)$ is said to be stable if for $\mathbf{x} \in \mathbb{N}^N$ the following holds*

$$\lim_{t \rightarrow \infty} \Pr\{\mathbf{N}^t < \mathbf{x}\} = F(\mathbf{x}) \quad \text{and} \quad \lim_{\mathbf{x} \rightarrow \infty} F(\mathbf{x}) = 1, \quad (3.10)$$

where $F(\mathbf{x})$ is the limiting distribution function.

For an N -user slotted ALOHA system, the **stability region** ($\mathfrak{S}_{ALOHA}$) is defined as the set of arrival rate $\boldsymbol{\lambda} = [\lambda_1, \lambda_2, \dots, \lambda_N]$ for which there exists a transmission probability vector $\mathbf{p} = [p_1, p_2, \dots, p_N]$ such that the queues in the system are stable. In literature, studies on the stability conditions of ALOHA systems were mainly based on the collision model, see [TM79, RE88, Ana91, Szp94, LE99], for example. In [NMT05], the stability analysis of slotted ALOHA based on the general MPR model is investigated. In this section, under the *Simplistic Assumption*, we first analyze the stability region of hybrid ALOHA for the two-user case and then derive the sufficient condition for stability of hybrid ALOHA in the case of $N > 2$.

3.4.1 Stability Region for $N = 2$

When $N = 2$, suppose the arrival rates for the two users are λ_1 and λ_2 (packets per slot), and their transmission probabilities are p_1 and p_2 , respectively. We have the following result.

Lemma 3.1 *Under the Simplistic Assumption, for a fixed transmission probability vector $\mathbf{p} = [p_1, p_2]$, the stability region of hybrid ALOHA is given by*

$$\lambda_1 \leq p_1 - \frac{p_1 \lambda_2}{2 - p_1}, \quad \text{for } \lambda_2 \leq p_2 - p_1 p_2 / 2, \quad (3.11)$$

and

$$\lambda_2 \leq p_2 - \frac{p_2 \lambda_1}{2 - p_2}, \quad \text{for } \lambda_1 \leq p_1 - p_1 p_2 / 2, \quad (3.12)$$

where λ_1 and λ_2 are the arrival rates for the two users in the unit of packets per slot.

Proof: Please referred to Appendix B.

Let $\mathfrak{S}_{H-ALOHA}(\mathbf{p})$ be the stability region of hybrid ALOHA for a fixed $\mathbf{p}$, then the overall stability region of hybrid ALOHA can be characterized as

$$\mathfrak{S}_{H-ALOHA} = \bigcup_{\mathbf{p} \in [0,1]^N} \mathfrak{S}_{H-ALOHA}(\mathbf{p}). \quad (3.13)$$

By taking the closure of all the regions as $\mathbf{p}$ varies over $[0,1]^2$, we can obtain the stability region of the proposed scheme. In [NMT05], an alternative approach is used to find the stability region by solving a corresponding constrained optimization problem. The approach is adopted here to obtain the following result.

Proposition 3.2 *Under the Simplistic Assumption, the stability region of hybrid ALOHA coincides with that of TDMA and is characterized by the region below*

$$\begin{aligned} \mathfrak{S}_{H-ALOHA} = \{(\lambda_1, \lambda_2) : (\lambda_1, \lambda_2) \geq (0, 0), (\lambda_1, \lambda_2) \text{ lies below the line} \\ \lambda_1 + \lambda_2 = 1, 0 \leq \lambda_1 \leq 1\}, \end{aligned} \quad (3.14)$$

where λ_1 and λ_2 are the arrival rates for the two users in the unit of packets per slot.

Proof: Please refer to Appendix C.

Recall that in this chapter, the length of the traditional ALOHA slot is assumed to be 1. Due to insertion of the idle sections, the slot length in hybrid ALOHA is larger than 1. For fair comparison, let Λ_i denote the number of arrivals per unit time when the arrival rate for user i is λ_i packets per hybrid ALOHA slot, that is, $\lambda_i = \Lambda_i[1 + (M - 1)\tau]$. Figure 3.4 portrays the stability regions of four different protocols in the two-user case. The stability region for hybrid ALOHA is bounded by the curve $\Lambda_1 + \Lambda_2 = 1/(1 + \tau)$. In other words, the maximum arrival rate for each user is limited by $1/(1 + \tau)$. This can be understood in the extreme case when there is only one user transmitting. In this circumstance the maximum stable arrival rate is

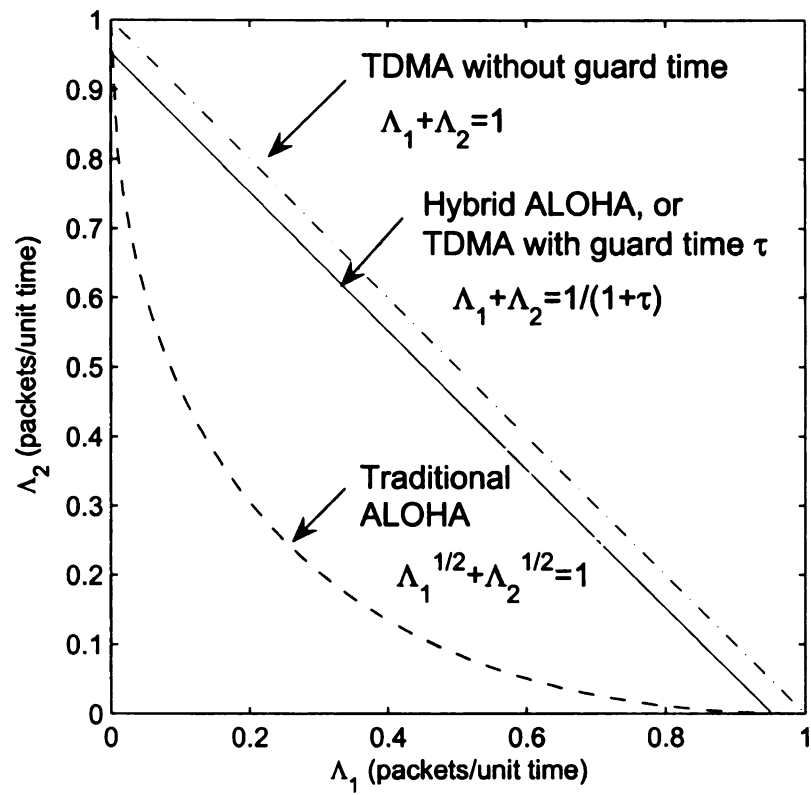


Figure 3.4. Stability regions for different protocols in the two-user case.

one packet per slot, corresponding to $1/(1 + \tau)$ packets per unit time. It can be seen that the stability region of hybrid ALOHA is identical to that of the TDMA protocol where a guard time of τ is present, and contains the stability region of traditional ALOHA as a subset except in the extreme case when only one user transmits.

A heuristic explanation to the stability coincidence of the hybrid ALOHA protocol and the TDMA protocol with guard times is that for hybrid ALOHA, under the *Simplistic Assumption*, the probability for both two simultaneous users to be successful is $1/2$, since the chance that their training sequences are transmitted in two different pilot sub-slots is 50%. Such time multiplexing feature produces the same effect as TDMA in terms of stability.

The *capacity region* of a system is defined as follows.

Definition 3.2 [NMT05] *Capacity region is the closure of the set of all achievable rates, where a rate $\lambda = (\lambda_1, \dots, \lambda_N)$ is called achievable if there exists a scheduling policy (S^*) with delivery rate greater than λ , i.e.,*

$$\lim_{T \rightarrow \infty} \inf \frac{1}{T} \sum_{t=0}^{T-1} \mathbf{1}(i \in \mathcal{R}_{S^*}(t)) \geq \lambda_i \quad \forall i \text{ a.s.} \quad (3.15)$$

where $\mathcal{R}_S(t)$ is the set of successful transmissions when the set of users $\mathcal{S}_S(t)$ transmit in slot t under scheduling policy S . Capacity region ($\mathfrak{C}$) is the closure of the set of all achievable rates.

If we define $\mathfrak{S}$ as the union of the stability regions over all MAC protocols, then $\mathfrak{S}$ provides the set of stable *arrival* rates with all MAC protocols. Generally, $\mathfrak{S} \subseteq \mathfrak{C}$ [VNT04], and then we have the relation: $\mathfrak{S}_{H_ALOHA} \subseteq \mathfrak{S} \subseteq \mathfrak{C}$. However, it has been proved in [NMT05] that when the MPR capacity of an MAC protocol reaches a critical level such that its stability region becomes convex, then the stability region coincides with the capacity region. As pointed out in [TNV04], the capacity defined here mimics the Shannon capacity limit, but is bounded by straight lines since successful transmission is measured in the unit of packets instead of bits.

Remark 3.2 *The result in Proposition 3.2 states that under the Simplistic As-*

sumption, the MPR capability of hybrid ALOHA reaches a critical level that makes the stability region $\mathfrak{S}_{H_ALOHA}$ convex. Following the above discussion, we have $\mathfrak{S}_{H_ALOHA} = \mathfrak{S} = \mathfrak{C}$. This implies that under the Simplistic Assumption, hybrid ALOHA is optimal in the sense that it can stabilize all rates that can be stabilized by any centralized or decentralized MAC protocols. The optimality is achieved with the transmission probabilities $\mathbf{p} = [1, 1]$, that is, both users always transmit packets if their queues are not empty, and no transmission control is required. This can be seen by letting $\mathbf{p} = [1, 1]$ in Lemma 3.1, where the resultant stability region coincides with the result given in Proposition 3.2.

Remark 3.3 Beyond the Simplistic Assumption, in the case when the collision of the training sequences does not necessarily lead to reception failure(s), stronger MPR capability could be achieved. The convexity of the stability region remains and hybrid ALOHA can outperform the TDMA scheme. More details can be found in Section 3.7.

For the general reception model, the stability of slotted ALOHA is derived in [NMT05] using the marginal probabilities of success. Similar derivations apply to the proposed hybrid ALOHA, resulting in the same result as in [NMT05], which is stated below.

Proposition 3.3 If $Q_1 = q_{1|\{1\}} - q_{1|\{1,2\}} \geq 0$ and $Q_2 = q_{2|\{2\}} - q_{2|\{1,2\}} \geq 0$, then the stability region of hybrid ALOHA for the general reception model is given by $\mathfrak{S}_1 \cap \mathfrak{S}_2$ where

$$\begin{aligned} \mathfrak{S}_1 &= \{(\lambda_1, \lambda_2) : (\lambda_1, \lambda_2) \geq (0, 0), (\lambda_1, \lambda_2) \text{ lies below the curve} \\ &\quad \lambda_2 = f(\lambda_1; q_{1|\{1\}}, q_{2|\{2\}}, Q_1, Q_2)\}, \end{aligned} \quad (3.16)$$

and

$$\begin{aligned} \mathfrak{S}_2 &= \{(\lambda_1, \lambda_2) : (\lambda_1, \lambda_2) \geq (0, 0), (\lambda_1, \lambda_2) \text{ lies below the curve} \\ &\quad \lambda_1 = f(\lambda_2; q_{2|\{2\}}, q_{1|\{1\}}, Q_2, Q_1)\}, \end{aligned} \quad (3.17)$$

where

$$f(\lambda; \alpha, \beta, \gamma, \delta) = \begin{cases} \beta - \frac{\lambda\delta}{\alpha - \gamma}, & \lambda \in \mathcal{I}_1, \\ \frac{(\sqrt{\alpha\beta} - \sqrt{\lambda\delta})^2}{\gamma}, & \lambda \in \mathcal{I}_2. \end{cases} \quad (3.18)$$

with

$$\mathcal{I}_1 = [0, \frac{\beta(\alpha - \gamma)^2}{\alpha\delta}] \text{ and } \mathcal{I}_2 = [\frac{\beta(\alpha - \gamma)^2}{\alpha\delta}, \frac{\alpha\beta}{\delta}]. \quad (3.19)$$

If either Q_1 or Q_2 equals zero, then we assume $\frac{1}{0} = \infty$ and the result still holds.

Proof: Please refer to [NMT05].

3.4.2 Sufficient Condition of Stability for $N > 2$

In the case of $N > 2$, stability regions are difficult to obtain for general ALOHA systems even under the collision model. However, in literature, stability conditions for a *fixed* transmission probability vector $\mathbf{p}$ have been developed under various models. Szpankowski [Szp94] derived a sufficient and necessary condition for stability of the ALOHA system with the collision model. In [NMT05], Naware, Mergen and Tong provides the sufficient condition for stability of the ALOHA system under the generalized MPR model. The stochastic dominance¹ [RE88] and the Loynes Theorem [Loy62] were the keys to these derivations, which, associated with the *Simplistic Assumption*, are adopted here to derive the sufficient condition for stability of hybrid ALOHA.

Under the *Simplistic Assumption*, it is easy to verify that the departure process Y_i^t in (3.5) can be written as

$$\begin{aligned} Y_i^t = & \sum_{j=1}^2 [1 - \chi(R_{i,\bar{j}}^t)] \{ R_{i,j}^t \cdot (1 - \sum_{k \in \mathcal{N} \setminus \{j\}} R_{k,j}^t \chi(N_k^t))^+ \\ & \cdot [(2 - \sum_{k \in \mathcal{N} \setminus \{j\}} R_{k,\bar{j}}^t \chi(N_k^t))^+ - (1 - \sum_{k \in \mathcal{N} \setminus \{j\}} R_{k,\bar{j}}^t \chi(N_k^t))^+] \}, \end{aligned} \quad (3.20)$$

where j is the index of the pilot sub-slot and $\bar{j} = \{1, 2\} \setminus \{j\}$; $p_{i,j}$ is the transmission

¹A real random variable X is said to stochastically dominate a real random variable Y if $\forall z \in \mathbb{R}$, $\Pr\{X > z\} \geq \Pr\{Y > z\}$. This dominance is denoted by $X \geq_{st} Y$.

probability of the training sequence at sub-slot j for user i , hence the transmission probability of user i in one slot is given by $p_i = \sum_{j=1}^2 p_{i,j}$. $R_{i,j}^t$ is the i.i.d Bernoulli($p_{i,j}$) random variable for $1 \leq i \leq N$, indicating the “coin tossing” outcomes that determine the transmission attempts. For hybrid ALOHA, since one user cannot transmit in both pilot sub-slots, $R_{i,\bar{j}}^t = 1 - R_{i,j}^t$. The function $\chi(k)$ is defined as

$$\chi(k) = \begin{cases} 0, & k = 0, \\ 1, & k > 0. \end{cases}$$

We construct a modified system as follows. Let $\mathcal{P} = \{\mathcal{S}, \mathcal{U}\}$ be a partition of $\mathcal{N}$ such that users in $\mathcal{S} \neq \mathcal{N}$ work exactly in the same manner as in the original system, while users in $\mathcal{U}$ *persistently* transmit *dummy* packets even if their queues are empty. Users in $\mathcal{U}$ are called *persistent* and those in $\mathcal{S}$ *nonpersistent*. Such modified system is denoted by $\Theta^{\mathcal{P}}$. Let $\mathbf{N}_{\mathcal{P}}^t = (\mathbf{N}_{\mathcal{S}}^t, \mathbf{N}_{\mathcal{U}}^t)$ denote the queue lengths in $\Theta^{\mathcal{P}}$ and it can be proved that $\mathbf{N}_{\mathcal{P}}^t$ stochastically dominates $\mathbf{N}^t$ of the original system provided that the initial conditions are identical [TM79, Szp86].

By the construction above, the process $\mathbf{N}_{\mathcal{S}}^t$ is an $|\mathcal{S}|$ -dimensional Markov chain that mimics the behavior of the original system. Note that the system consisting of users in $\mathcal{S}$ forms a smaller copy of the original system with modified reception probabilities. Generally, for any $\mathcal{S}' \subseteq \mathcal{S}$ and $i \in \mathcal{S}'$, the modified reception probabilities for the smaller system consisting of the stand-alone nonpersistent set $\mathcal{S}$ are given by [NMT05]

$$q_{i|\mathcal{S}'}^{\mathcal{S}} = \sum_{T \subseteq \mathcal{U}} \left(\prod_{j \in T} p_j \prod_{k \in \mathcal{U} \setminus T} \bar{p}_k \right) q_{i|\mathcal{S}' \cup T}^{\mathcal{N}}. \quad (3.21)$$

Induction arguments can then be applied to establish the stability conditions. We further assume that the Markov chain $\mathbf{N}_{\mathcal{S}}^t$ is stationary and ergodic. We denote the stationary version as $\bar{\mathbf{N}}_{\mathcal{S}}^0$ to indicate that the process starts from the stationary distribution.

Let $Y_i^t(\mathcal{P})$ be the departure process from the i th queue in the dominant system

$\Theta^{\mathcal{P}}$, then we have

$$\begin{aligned}
Y_i^t(\mathcal{P}) = & \sum_{j=1}^2 [1 - \chi(R_{i,\bar{j}}^t)] R_{i,j}^t \cdot \mathbf{1} \left[\sum_{k \in \mathcal{S} \setminus \{i\}} R_{k,j}^t \chi(N_k^t) + \sum_{k \in \mathcal{U} \setminus \{i\}} R_{k,j}^t = 0 \right] \\
& \cdot \mathbf{1} \left[\sum_{k \in \mathcal{S} \setminus \{i\}} R_{k,\bar{j}}^t \chi(N_k^t) + \sum_{k \in \mathcal{U} \setminus \{i\}} R_{k,\bar{j}}^t \leq 1 \right]. \tag{3.22}
\end{aligned}$$

Given that $Y_i^t(\mathcal{P})$ is stationary, we denote $P_{succ}^i(\mathcal{P}) = E[Y_i^t(\mathcal{P})]$ the probability of a successful transmission from the i th user in system $\Theta^{\mathcal{P}}$, which is given by

$$\begin{aligned}
P_{succ}^i(\mathcal{P}) = & p_i \prod_{k \in \mathcal{U} \setminus \{i\}} (1 - p_k) \sum_{\mathbf{z}_{\mathcal{S}} \in \{0,1\}^{|\mathcal{S}|}} \left(Pr\{\chi(\bar{\mathbf{N}}_{\mathcal{S}}^0) = \mathbf{z}_{\mathcal{S}}\} \prod_{k \in \mathcal{S} \setminus \{i\}} (1 - p_k)^{z_k} \right) \\
& + \left(\sum_{j=1}^2 p_{i,j} \sum_{k \in \mathcal{U} \setminus \{i\}} [p_{k,\bar{j}} \prod_{k' \in \mathcal{U} \setminus \{i,k\}} (1 - p_{k'})] \right) \cdot \\
& \cdot \sum_{\mathbf{z}_{\mathcal{S}} \in \{0,1\}^{|\mathcal{S}|}} \left(Pr\{\chi(\bar{\mathbf{N}}_{\mathcal{S}}^0) = \mathbf{z}_{\mathcal{S}}\} \prod_{k \in \mathcal{S} \setminus \{i\}} (1 - p_k)^{z_k} \right) \\
& + \prod_{k \in \mathcal{U} \setminus \{i\}} (1 - p_k) \left(\sum_{j=1}^2 p_{i,j} \sum_{\mathbf{z}_{\mathcal{S}} \in \{0,1\}^{|\mathcal{S}|}} Pr\{\chi(\bar{\mathbf{N}}_{\mathcal{S}}^0) = \mathbf{z}_{\mathcal{S}}\} \cdot \right. \\
& \cdot \left. \left[\sum_{k \in \mathcal{S} \setminus \{i\}} \mathbf{1}(z_k = 1) p_{k,\bar{j}} \prod_{k' \in \mathcal{S} \setminus \{i,k\}} (1 - p_{k'})^{z_{k'}} \right] \right), \tag{3.23}
\end{aligned}$$

where $\chi(\bar{\mathbf{N}}_{\mathcal{S}}^0) = (\chi(\bar{N}_{i_1}^0), \dots, \chi(\bar{N}_{i_{|\mathcal{S}|}}^0))$ with $i_k \in \mathcal{S}$ for all $k = 1, \dots, |\mathcal{S}|$. The first term of the right-hand side of (3.23) represents the probability of successful transmission of user i when no one else transmits. The second and third terms represent the probabilities of successful transmission of user i when there is another user involved in transmission. The second term corresponds to the case when the extra user belongs to $\mathcal{U}$, and the third term corresponds to the case when the extra user belongs to $\mathcal{S}$.

Under the partition $\mathcal{P} = (\mathcal{S}, \mathcal{U})$, let $\mathfrak{S}_{\mathcal{N}}$ and $\mathfrak{S}_{\mathcal{S}}$ denote the stability regions for the whole system $\Theta^{\mathcal{P}}$ and the system consisting of nonpersistent users $\mathcal{S}$, respectively.

Define a region

$$\mathfrak{S}_{\mathcal{N}} = \bigcup_{\mathcal{P}} \{\lambda_{\mathcal{N}} = \{\lambda_1, \dots, \lambda_N\} : \lambda_k < P_{succ}^k(\mathcal{P}) \ \forall k \in \mathcal{U} \text{ and } \lambda_{\mathcal{S}} \in \mathfrak{S}_{\mathcal{S}}\}, \quad (3.24)$$

then we have the following result.

Proposition 3.4 *For a fixed transmission probability vector $\mathbf{p}$, under the Simplistic Assumption, if $\lambda_{\mathcal{N}} \in \mathfrak{S}_{\mathcal{N}}$, then the hybrid ALOHA system is stable, i.e., $\mathfrak{S}_{\mathcal{N}} \subseteq \mathfrak{S}_{H-ALOHA}(\mathbf{p})$.*

Proof: Please refer to Appendix D.

In the proof, mathematical induction is used to reduce the dimensionality of the problem and then the Loynes Theorem is applied to the isolated general queue. For partition $\mathcal{P} = \{\mathcal{S}, \mathcal{U}\}$, since $|\mathcal{S}| < N$, by induction $\lambda_{\mathcal{S}} \in \mathfrak{S}_{\mathcal{S}}$ is sufficient for stability of the system $\mathcal{S}$. This makes the departure process in $\mathcal{U}$ stationary and ergodic. Moreover, by the Loynes Theorem, queues in $\mathcal{U}$ are stable with $\lambda_k < P_{succ}^k(\mathcal{P}), \forall k \in \mathcal{U}$. The conditions are then sufficient for $\Theta^{\mathcal{P}}$ to be stable. With the property of the stochastic dominance of $\Theta^{\mathcal{P}}$ over the original system, the conditions are also sufficient for hybrid ALOHA to be stable.

3.5 Special Case: Stability Inner Bound for $N = 3$

3.5.1 Inner Bound Characterization

Recall that $\mathcal{P} = \{\mathcal{S}, \mathcal{U}\}$ is a partition of $\mathcal{N}$. Consider the three partitions $\mathcal{P}_i = \{\mathcal{S}, \mathcal{U}\} = \{\mathcal{N}_i, \{i\}\}$, where $\mathcal{N}_i = \mathcal{N} \setminus \{i\}$, $i = 1, 2, 3$. Define the other three partitions $\bar{\mathcal{P}}_i = \{\mathcal{S}, \mathcal{U}\} = \{\{i\}, \mathcal{N}_i\}$. From (3.23) it is easy to verify that $P_{succ}^i(\bar{\mathcal{P}}_i) < P_{succ}^i(\mathcal{P}_i)$ for any i . Hence it follows from Proposition 3.4 that the stability inner bound $\mathcal{R}$ is the sum of three regions $\mathcal{R}_i$ corresponding to $\mathcal{P}_i$, $i = 1, 2, 3$, respectively. In what follows, we inspect in details the region $\mathcal{R}_3$ corresponding to $\mathcal{P}_3 = \{\{1, 2\}, \{3\}\}$. The other two regions $\mathcal{R}_1$ and $\mathcal{R}_2$ can be easily obtained through the similar procedures.

Let N_i^t and β_i^t denote respectively the queue length and the number of arriving packets of user i ($i = 1, 2$) in time slot t . Let $F_3(x, y)$ be the moment generating function of the joint arrival process for user 1 and 2. Thus, for $|x| \leq 1, |y| \leq 1, t \in \mathbb{N}$

$$F_3(x, y) = E(x^{\beta_1^t} y^{\beta_2^t}) = (x\lambda_1 + \bar{\lambda}_1)(y\lambda_2 + \bar{\lambda}_2), \quad (3.25)$$

where $\bar{\lambda}_i = 1 - \lambda_i$, for $i = 1, 2$.

The investigated system model implies that (N_1^t, N_2^t) is an irreducible, aperiodic Markov chain. Hence the stability of the system is equivalent to existence of a unique stationary (limiting) distribution. Let $G_3(x, y)$ be the moment generating function of the joint stationary queue process, *viz.*,

$$G_3(x, y) = \lim_{t \rightarrow \infty} E(x^{N_1^t} y^{N_2^t}). \quad (3.26)$$

In Appendix E we derive the following functional equation:

$$K(x, y)G_3(x, y) = a(x, y)G_3(0, y) + b(x, y)G_3(x, 0) + c(x, y)G_3(0, 0), \quad (3.27)$$

with

$$K(x, y) = \frac{1}{(x\lambda_1 + \bar{\lambda}_1)(y\lambda_2 + \bar{\lambda}_2)} - \frac{p_1\bar{p}_2(1 - p_3/2)}{x} - \frac{p_2\bar{p}_1(1 - p_3/2)}{y} - \frac{p_1p_2\bar{p}_3}{2xy} - [1 - (1 - p_3/2)(p_1\bar{p}_2 + \bar{p}_1p_2) - p_1p_2\bar{p}_3/2], \quad (3.28)$$

$$a(x, y) = \frac{p_1p_2(1 - p_3/2)}{y} + \frac{p_1\bar{p}_2(1 - p_3/2)}{x} - \frac{p_1p_2\bar{p}_3}{2xy} + p_1(1 - p_3/2)(1 - 2p_2) + p_1p_2\bar{p}_3/2, \quad (3.29)$$

$$b(x, y) = \frac{p_1p_2(1 - p_3/2)}{x} + \frac{p_2\bar{p}_1(1 - p_3/2)}{y} - \frac{p_1p_2\bar{p}_3}{2xy} + p_2(1 - p_3/2)(1 - 2p_1) + p_1p_2\bar{p}_3/2, \quad (3.30)$$

$$c(x, y) = \frac{p_1 p_2 \bar{p}_3}{2xy} - \frac{p_1 p_2 (1 - p_3/2)}{y} - \frac{p_1 p_2 (1 - p_3/2)}{x} + \frac{p_1 p_2 (3 - p_3)}{2}, \quad (3.31)$$

In (3.27), taking $y = 1$, divided by $x - 1$ and then taking $x = 1$, we have

$$\begin{aligned} -\lambda_1 + \frac{1}{2}p_1[2 - (p_2 + p_3)] = \\ \frac{1}{2}p_1[2 - (p_2 + p_3)]G_3(0, 1) - \frac{p_1 p_2}{2}G_3(1, 0) + \frac{p_1 p_2}{2}G_3(0, 0). \end{aligned} \quad (3.32)$$

Exchanging the roles of x and y and repeating the above operations, we have

$$\begin{aligned} -\lambda_2 + \frac{1}{2}p_2[2 - (p_1 + p_3)] = \\ -\frac{p_1 p_2}{2}G_3(0, 1) + \frac{1}{2}p_2[2 - (p_1 + p_3)]G_3(1, 0) + \frac{p_1 p_2}{2}G_3(0, 0). \end{aligned} \quad (3.33)$$

Note that in deriving these equations we exploit the facts that $F_3(1, 1) = 1$ and $\lambda_i = E(\beta_i^t)$ for $i = 1, 2$.

Define $P_3(z_1, z_2) \doteq Pr\{\chi(\bar{N}_1) = z_1, \chi(\bar{N}_2) = z_2\}$ with user 3 being the persistent one. Then from (E.4)-(E.7) we have

$$\begin{aligned} P_3(0, 0) &= G_3(0, 0); \\ P_3(1, 0) &= G_3(1, 0) - G_3(0, 0); \\ P_3(0, 1) &= G_3(0, 1) - G_3(0, 0); \\ P_3(1, 1) &= 1 + G_3(0, 0) - G_3(0, 1) - G_3(1, 0). \end{aligned}$$

Using these relations and (3.23) and assuming $p_{i,j} = p_i/2$ for $i = 1, 2, 3, j = 1, 2$, we

can easily have the following results.

$$\begin{aligned}
P_{succ}^1(\mathcal{P}_3) &= p_1 \bar{p}_3 \left[1 - (1 - G_3(1, 0)p_2) \right] + (p_{1,1}p_{3,2} + p_{1,2}p_{3,1}) \left[1 - (1 - G_3(1, 0))p_2 \right] \\
&\quad + \bar{p}_3 \left[(p_{1,1}p_{2,2} + p_{1,2}p_{2,1})(P_3(0, 1) + P_3(1, 0)) \right] \\
&= p_1 \left[1 - \frac{1}{2}p_3 - \frac{1}{2}p_2(1 - G_3(1, 0)) \right] \\
P_{succ}^2(\mathcal{P}_3) &= p_2 \bar{p}_3 \left[1 - (1 - G_3(0, 1)p_1) \right] + (p_{2,2}p_{3,1} + p_{2,1}p_{3,2}) \left[1 - (1 - G_3(0, 1))p_1 \right] \\
&\quad + \bar{p}_3 \left[(p_{2,1}p_{1,2} + p_{2,2}p_{1,1})(P_3(0, 1) + P_3(1, 0)) \right] \\
&= p_2 \left[1 - \frac{1}{2}p_3 - \frac{1}{2}p_1(1 - G_3(0, 1)) \right] \\
P_{succ}^3(\mathcal{P}_3) &= p_3 \left[P_3(0, 0) + (1 - \frac{1}{2}p_1)P_3(1, 0) + (1 - \frac{1}{2}p_2)P_3(0, 1) \right. \\
&\quad \left. + (1 - \frac{1}{2}p_1 - \frac{1}{2}p_2)P_3(1, 1) \right] \\
&= p_3 \left[1 - \frac{1}{2}p_1(1 - G_3(0, 1)) - \frac{1}{2}p_2(1 - G_3(1, 0)) \right] \tag{3.34}
\end{aligned}$$

Stability region $\mathcal{R}_3$ can then be characterized by $\mathcal{R}_3 = \{\lambda_i < P_{succ}^i(\mathcal{P}_3), i = 1, 2, 3\}$. The other two regions $\mathcal{R}_1$ and $\mathcal{R}_2$ can be similarly calculated and the following result is then readily in form.

Proposition 3.5 *For $N = 3$, hybrid ALOHA is stable in the region $\mathcal{R} = \bigcup_{k=1}^3 \mathcal{R}_k$, with*

$$\mathcal{R}_k = \{\lambda_i < P_{succ}^i(\mathcal{P}_k), i = 1, 2, 3\}, \tag{3.35}$$

where $P_{succ}^i(\mathcal{P}_k)$ for $k = 3, i = 1, 2, 3$, are shown in (3.34), and it can be calculated in the same manner for $k = 1, 2, i = 1, 2, 3$.

It is seen from (3.34) that the probability of success $P_{succ}^i(\mathcal{P}_k)$ depends explicitly on $G_k(1, 0)$ and $G_k(0, 1)$. These functions are generally nonlinear functions of the input rates. It will be shown in the sequel that explicit expression of these functions can be found by solving a general Riemann-Hilbert problem. The analysis follows the procedures given in [FI79, Nai85].

3.5.2 Analysis of the Kernel $K(x, y)$

The analysis of the kernel $K(x, y)$ is the key to solving the functional equation (3.27). For description simplicity, in what follows we omit the subscript of the function $G_3(x, y)$. Since $G(x, y)$ is analytic in $|x| < 1, |y| < 1$ and continuous in $|x| \leq 1, |y| \leq 1$, the right-hand side of the equation (3.27) must vanish whenever the “kernel” $K(x, y)$ vanishes for $|x| \leq 1, |y| \leq 1$.

Solving for x the equation $K(x, y) = 0$, we shall have a root $x = f_x(y)$ satisfying

$$A(y)x^2 + B(y)x + C(y) = 0, \quad (3.36)$$

where, by defining $\Delta = 1 - (1 - p_3/2)(p_1\bar{p}_2 + \bar{p}_1p_2) - p_1p_2\bar{p}_3/2$, we have

$$A(y) = \lambda_1[\bar{p}_1p_2(1 - p_3/2)(\lambda_2 + \frac{\bar{\lambda}_2}{y}) + \Delta(y\lambda_2 + \bar{\lambda}_2)], \quad (3.37)$$

$$B(y) = [\lambda_1p_1\bar{p}_2(1 - p_3/2) + \bar{\lambda}_1\Delta](y\lambda_2 + \bar{\lambda}_2) + [\bar{\lambda}_1\bar{p}_1p_2(1 - p_3/2) + \lambda_1\frac{p_1p_2\bar{p}_3}{2}](\lambda_2 + \frac{\bar{\lambda}_2}{y}) - 1, \quad (3.38)$$

$$C(y) = \bar{\lambda}_1p_1\bar{p}_2(1 - p_3/2)(y\lambda_2 + \bar{\lambda}_2) + \bar{\lambda}_1\frac{p_1p_2\bar{p}_3}{2}(\lambda_2 + \frac{\bar{\lambda}_2}{y}). \quad (3.39)$$

Therefore,

$$f_x(y) = \frac{-B(y) \pm \sqrt{D(y)}}{2A(y)}, \quad (3.40)$$

with $D(y) = B(y)^2 - 4A(y)C(y)$. Defining $t(y, \phi) = -B(y) - 2\cos(\phi)\sqrt{A(y)C(y)}$, we have the following lemma.

Lemma 3.2 *For $\phi \in [0, 2\pi]$, the equation $t(y, \phi) = 0$ has exactly two real roots $y = r_1(\phi)$ and $y = r_2(\phi)$ which satisfy $0 < r_1(\phi) < 1 < r_2(\phi)$.*

Proof: Please refer to Appendix F.

Since $D(y) = t(y, 0)t(y, \pi)$, it is readily shown that $y_1 = r_1(\pi), y_2 = r_1(0), y_3 = r_2(0), y_4 = r_2(\pi)$ are the four zeros of $D(y)$ (branch points of $f_x(y)$) satisfying $0 < y_1 < y_2 < 1 < y_3 < y_4$. This result is evidently valid for the branch points x_1, x_2, x_3, x_4 of the function $f_y(x)$. The following lemma then holds.

Lemma 3.3 *The equation $K(x, y) = 0$ has one root $x = k(y)$ which is an analytic algebraic function of y in the whole complex plane cut along $[y_1, y_2] \cup [y_3, y_4]$. Moreover, $|k(y)| \leq 1$ for $|y| = 1$.*

Similar propositions apply to $f_y(x)$: i.e., there exists $y = h(x)$ such that $K(x, h(x)) = 0$ with $|h(x)| \leq 1$ for $|x| = 1$.

Proof: Please refer to Appendix G.

Defining

$$\rho(\phi) \doteq \left(\frac{\bar{\lambda}_1[p_1\bar{p}_2(1 - p_3/2)r_1(\phi) + p_1p_2\bar{p}_3/2]}{\lambda_1[\Delta r_1(\phi) + \bar{p}_1p_2(1 - p_3/2)]} \right)^{1/2}, \quad (3.41)$$

we then have the following result.

Lemma 3.4 $k(r_1(\phi)) = \rho(\phi)e^{i\phi}$ for $\phi \in [0, 2\pi]$.

Proof: From Lemma 3.3, $f_x(y) = k(y)$ is the algebraic branch of $K(x, y) = 0$ such that $|x(y)| \leq 1$ for $|y| = 1$. Denote $k_c(y)$ as the other root of equation $K(x, y) = 0$. It is shown that for $y \in \mathbb{C} \setminus [y_1, y_2] \cup [y_3, y_4]$, the minus and plus signs in (3.40) correspond to $k(y)$ and $k_c(y)$, respectively (compute $k(1)$ and $k_c(1)$).

Observe that $y = r_1(\phi)$ sweeps twice the cut $[y_1, y_2]$ as ϕ traverses the interval $[0, 2\pi]$, then $k(r_1(\phi))$ and $k_c(r_1(\phi))$ are two conjugate complex numbers satisfying

$$k(r_1(\phi))k_c(r_1(\phi)) = \frac{C(r_1(\phi))}{A(r_1(\phi))} = |\rho(\phi)|^2. \quad (3.42)$$

From the definition of the algebraic branch, it is shown that $k(r_1(\phi)) = \rho(\phi)e^{i\phi}, \phi \in [0, 2\pi]$. $\square$

The image of the cut $[y_1, y_2]$ under the mapping $x = k(y)$ is then denoted as $L_x \doteq \{x \in \mathbb{C} : x = \rho(\phi)e^{i\phi}, \phi \in [0, 2\pi]\}$, which is a smooth closed contour enclosing 0.

3.5.3 Reduction to the Riemann-Hilbert Problem

It will be shown in this subsection that the problem of finding out the expressions of function $G(x, 0)$ and $G(0, y)$ reduces to solving a general Riemann-Hilbert boundary value problem.

Riemann-Hilbert Problem Let L^+ be a finite or infinite region, bounded by a smooth contour L . It is required to find a function $\Phi(z)$, holomorphic in L^+ and continuous in $L^+ \cup L$, satisfying the boundary condition

$$\Re[U(z)\Phi(z)] = V(z) \text{ on } L, \quad (3.43)$$

where $U(z), V(z)$ are continuous functions given on L .

The formulation of the boundary value problem is illustrated below.

For pairs (x, y) satisfying $K(x, y) = 0$ and $|x| \leq 1, |y| \leq 1$, we should have

$$a(x, y)G(0, y) + b(x, y)G(x, 0) + c(x, y)G(0, 0) = 0. \quad (3.44)$$

Defining $D = \{y \in \mathbb{C} : |y| \leq 1, |k(y)| \leq 1\}$ and $D_c = \{y \in \mathbb{C} : |y| \leq 1, |k(y)| > 1\}$, we have for $y \in D$

$$a(k(y), y)G(0, y) + b(k(y), y)G(k(y), 0) = -c(k(y), y)G(0, 0). \quad (3.45)$$

$G(0, y)$ and $G(x, 0)$ are analytic in $D \setminus [y_1, y_2]$. When $|y| \leq 1$, $k(y)$ is in the region containing the curve L_x . Then (3.45) can be used to continue $G(x, 0)$ as a meromorphic function up to L_x . The eventual poles of $G(x, 0)$ are the zeros of $b(k(y), y)$ for $y \in D_c$.

Since for $y \in [y_1, y_2]$, the power series expansion of $G(0, y)$ has positive coefficients, it is easily seen that $\Re\{iG(0, y)\} = 0$ for $y \in [y_1, y_2]$. Hence it follows that

$$\Re\left\{\frac{ib(x, h(x))}{a(x, h(x))}G(x, 0)\right\} = \Re\left\{\frac{-ic(x, h(x))}{a(x, h(x))}G(0, 0)\right\}, \text{ for } x \in L_x. \quad (3.46)$$

Suppose x_a is the eventual zero of $a(x, h(x))$ on L_x . We reformulate (3.46) into

$$\Re\left\{\frac{ib(x, h(x))(x - x_a)^\alpha}{a(x, h(x))}G(x, 0)\right\} = \Re\left\{\frac{-ic(x, h(x))(x - x_a)^\alpha}{a(x, h(x))}G(0, 0)\right\} \quad (3.47)$$

where $\alpha = 0$ when the zero x_a does not exist and $\alpha = 1$ otherwise.

Similarly, denote x_b as an eventual root of $b(x, h(x)) = 0$ on L_x . Let

$$U(x) = \frac{b(x, h(x))(x - x_a)^\alpha}{a(x, h(x))(x - x_b)^\gamma},$$

where $\gamma = 0$ when the root does not exist and $\gamma = 1$ otherwise. Furthermore, let

$$V(x) = \Re\left\{\frac{-ic(x, h(x))(x - x_a)^\alpha}{a(x, h(x))}G(0, 0)\right\}.$$

The problem then reduces to finding a function $\tilde{G}(x)$ which is analytic in L_x^+ , continuous in $L_x \cup L_x^+$ and satisfies

$$\Re\{iU(x)\tilde{G}(x)\} = V(x). \quad (3.48)$$

Consequently, from (3.47), we have

$$G(x, 0) = \frac{\tilde{G}(x)}{(x - x_b)^\gamma}. \quad (3.49)$$

This is typically a Riemann-Hilbert boundary value problem. The function $G(0, y)$ can be obtained in the same manner.

3.5.4 Solution to the Riemann-Hilbert Problem (3.48)

The solution of (3.48) can be directly obtained as in [Mus53, pp. 99-107] whenever the contour L_x is a unit circle. The problem is more complicated when L_x is arbitrary. Fortunately, Riemann's mapping theorem guarantees the existence of a conformal mapping which maps L_x conformally onto the unit circle. Such a mapping can be fulfilled at the aid of the Theodorsen's procedure [CB83, pp. 70-73], which performs

the inverse mapping from the unit circle to L_x and is stated in the following lemma.

Lemma 3.5 [CB83] *The conformal mapping f_0 of the unit circle $z = e^{it}$, $t \in [0, 2\pi]$ onto the curve $L_x = \{x : x = \rho(\phi)e^{i\phi}, \phi \in [0, 2\pi]\}$ is defined as*

$$f_0(e^{it}) = \rho(\phi(t))e^{i\phi(t)}, \quad (3.50)$$

where f_0 is assumed to be normalized by $f_0(0) = 0$ and $f_0'(0) > 0$. For $|z| < 1$, $f_0(z)$ is uniquely determined by

$$f_0(z) = z \exp\left[\frac{1}{2\pi} \int_0^{2\pi} \log \rho(\phi(t)) \frac{e^{it} + z}{e^{it} - z} dt\right], \text{ for } |z| < 1, \quad (3.51)$$

where $\phi(t)$ satisfies:

$$\phi(t) = t - \frac{1}{2\pi} \int_0^{2\pi} \log \rho(\phi(\omega)) \cot \frac{1}{2}(\omega - t) d\omega, \quad 0 \leq t \leq 2\pi. \quad (3.52)$$

This is Theodorsen's integral equation for $\phi(t)$; it is a nonlinear, singular integral equation.

The details of solving the equations in Lemma 3.5 can be found in [CB83] and will not be discussed in this chapter. We denote $f(z)$ as the inverse of $f_0(z)$. Using Lemma 3.5 and the methods in [CB83, pp. 68-69] and [Mus53, pp. 99-107], we can find the solution to the Riemann-Hilbert problem (3.48) as below.

Define the index of the non-homogeneous Riemann-Hilbert boundary value problem (3.48) as

$$\kappa = \frac{-1}{\pi} \arg[U(x)]_{x \in L_x}, \quad (3.53)$$

where $\arg[U(x)]_{x \in L_x}$ is the variation of the argument of the function $U(x)$ when x moves along the contour L_x in the positive direction.

We present the solution to (3.48) for $\kappa \geq 0$ as below. The function $\tilde{G}(x) = H(f(x))$ where

$$H(z) = \frac{z^\kappa F(z)}{2\pi} \int_{|t|=1} \frac{t^{-\kappa} V(t) dt}{U(t) F^+(t) t} - \left\{ \int_{|t|=1} \frac{V(t) dt}{U(t) F^+(t) (t-z)} + \right. \\ \left. + z^\kappa \int_{|t|=1} \frac{t^{-\kappa} V(t) dt}{U(t) F^+(t) (t-z)} \right\} \frac{F(z)}{2\pi}. \quad (3.54)$$

In the above equation,

$$F(z) = C \cdot e^{\Gamma(z)}, \quad (3.55)$$

where C is a nonzero constant and

$$\Gamma(z) = \frac{1}{2\pi i} \int_{|t|=1} \frac{\log[t^{-\kappa} J(t)] dt}{t-z}, \quad (3.56)$$

with

$$J(t) = \frac{\overline{iU(t)}}{iU(t)}, \quad (3.57)$$

and

$$U(t) = \frac{b(f_0(t), h(f_0(t)))(f_0(t) - x_a)^\alpha}{a(f_0(t), h(f_0(t)))(f_0(t) - x_b)^\gamma}. \quad (3.58)$$

$$V(t) = \Re \left\{ \frac{-ic(f_0(t), h(f_0(t)))(f_0(t) - x_a)^\alpha}{a(f_0(t), h(f_0(t)))} G(0, 0) \right\}. \quad (3.59)$$

For $t_0 \in L$, $F^+(t_0)$ is defined to be the limit when t approaches t_0 along any path, which remains, however, on the left of L , i.e.,

$$F^+(t_0) = \lim_{t \rightarrow t_0, t \in L^+} F(t). \quad (3.60)$$

Applying the Plemelj-Sokhotski formulas [CB83, pp. 32], it is shown that in (3.54)

$$F^+(t_0) = \frac{1}{2} F(t_0) + \frac{1}{2\pi i} \int_{|t|=1} \frac{F(t)}{t - t_0} dt. \quad (3.61)$$

Finally, $G(x, 0)$ is obtained through (3.49).

On the other hand, $G(0, y)$ can be computed through the similar procedures, and consequently the stability bound in Proposition 3.4 is explicitly determined.

3.6 Delay Performance for $N = 2$

We consider to study the delay performance of hybrid ALOHA in this section. Based on the collision model, the “exact” delays in the two-user slotted ALOHA systems with symmetric and asymmetric arrival rates and transmission probabilities are calculated in [SS83] and [Nai85], respectively. In [NMT05], the average delay is computed for the capture model in a symmetrical two-user system. It is hard to get the exact delay for $N > 2$, but bounds on the average delay have been obtained in [TM79, Szp86, EZ87] for the collision model, and in [SS00] for the capture model.

3.6.1 Exact Delays for the General Case

In this section, we derive the “exact” delay for the general *asymmetric* $N = 2$ hybrid ALOHA system.

It can be verified that in the $N = 2$ case, we still have the functional equation (3.27), with the exceptions that probability p_3 should be set to 0 and that the subscript of function $G_3(x, y)$ should be dropped. As discussed in the previous section, through this functional equation and by means of solving the reduced Riemann-Hilbert problem, we can obtain the explicit expressions of $G(x, 0)$ and $G(0, y)$. Consequently, the derivatives $\frac{dG(x, 0)}{dx}|_{x=1} \doteq G'_1$ and $\frac{dG(0, y)}{dy}|_{y=1} \doteq G'_2$ can be computed, and the following proposition holds:

Proposition 3.6 *For the $N = 2$ hybrid ALOHA system, under the Simplistic Assumption, when the system is stable with $\lambda_i < p_i - p_i p_{\bar{i}}/2$, where $\bar{i} = \{1, 2\} \setminus \{i\}$ and $i = 1, 2$, the exact delays D_i , $i = 1, 2$, are given by*

$$D_i = \frac{2\lambda_i(1 - \lambda_i) - p_i p_{\bar{i}} G'_i}{2\lambda_i(p_i - p_i p_{\bar{i}}/2 - \lambda_i)}, \quad i = 1, 2. \quad (3.62)$$

Proof: It can be easily known from (3.34) that if $\lambda_i < p_i - p_i p_i^*/2$, the system is stable. The queue length for either user can then be calculated as below.

Take $y = 1$ in the functional equation (3.27). Differentiating both sides of the equation twice in the variable x , and making $x = 1$, we have user 1's queue length $N_1 \doteq \frac{dG(x,1)}{dx}|_{x=1}$ as below.

$$N_1 = \frac{1}{2(p_1 - p_1 p_2/2 - \lambda_1)} \left\{ p_1 p_2 [G(1,0) - G(0,0) - G'_1] + p_1(2 - 3p_2)G(0,1) - 2\lambda_1^2 + p_1(2 - p_2) \right\}. \quad (3.63)$$

In (3.27), taking $y = 1$, divided by $x - 1$ and making $x = 1$, we have

$$p_1(2 - 3p_2)G(0,1) + p_1 p_2 [G(1,0) - G(0,0)] = 2\lambda_1 - p_1(2 - p_2). \quad (3.64)$$

Therefore, from (3.63) and (3.64),

$$N_1 = \frac{2\lambda_1(1 - \lambda_1) - p_1 p_2 G'_1}{2(p_1 - p_1 p_2/2 - \lambda_1)} \quad (3.65)$$

User 2's queue length $N_2 \doteq \frac{dG(1,y)}{dy}|_{y=1}$ can be similarly obtained.

The exact delay D_i of the i th user ($i = 1, 2$) can then be obtained by applying Little's theorem, i.e.,

$$D_i = N_i / \lambda_i, \quad i=1,2. \quad (3.66)$$

This completes the proof of Proposition 3.6. $\square$

3.6.2 Delay Bounds for the Symmetric Case

The exact delays can be obtained yet they are hard to calculate. For numerical results, we derive the upper and lower bounds for the symmetric two-user system based on the MPR model. We assume that each user has an infinite buffer to store arriving and backlogged packets. The arrivals to the i th user are i.i.d. Bernoulli with mean λ_i in every slot, and the arrivals are independent across users.

Let $\lambda_1 = \lambda_2 = \lambda$ be the arrival rate measured in packet per hybrid ALOHA slot and p the transmission probability of both users. If we denote $q_{1|\{1\}} = q_{2|\{2\}} = a$, $q_{1,\{1,2\}} = q_{2,\{1,2\}} = b$ and $q_{\{1,2\}|\{1,2\}} = c$, the bounds of the delay are given by the following proposition.

Proposition 3.7 *If the system is stable, i.e., $\lambda < pa + p^2(b + c - a)$, the lower and upper bounds of the average delay for either user are given by*

$$D_{Lower} = \frac{1}{a} \left[\frac{a(1 - \lambda) + p(b + c - a)(1 - \lambda/2)}{pa + p^2(b + c - a) - \lambda} \right], \quad (3.67)$$

and

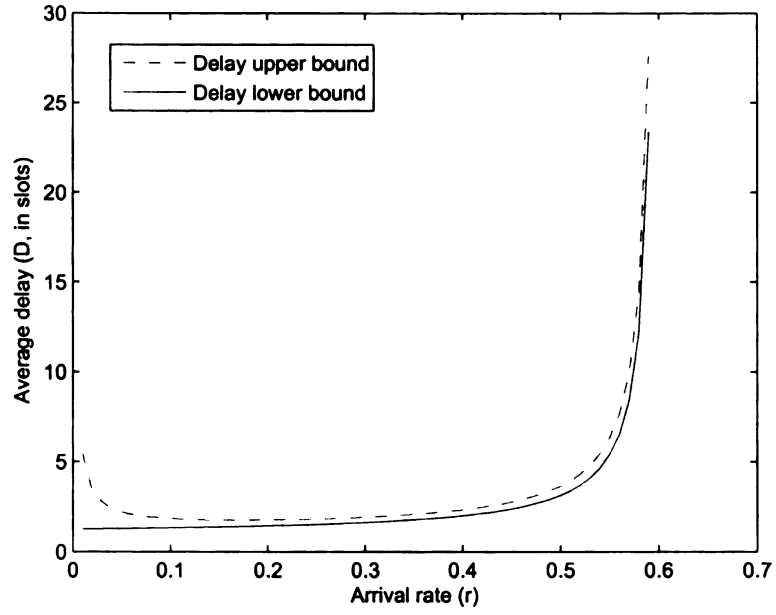
$$D_{Upper} = D_{Lower} + \frac{p^3 c [a - (b + c)]}{2ar[pa + p^2(b + c - a) - \lambda]}. \quad (3.68)$$

Proof: Please refer to Appendix H.

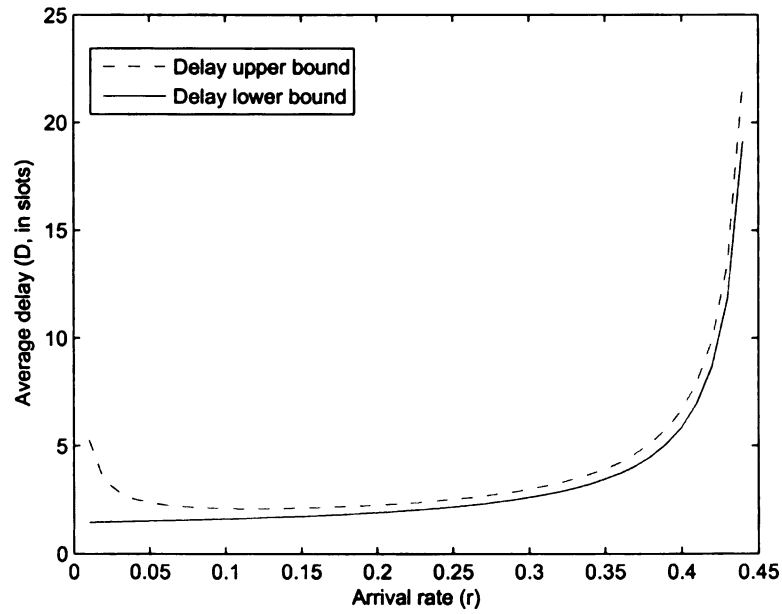
When the arrival rate λ is low (e.g. $\lambda < 0.1$), the given upper bound is quite loose, as illustrated in Figure 3.5. However, if the MPR capability of the system is strong enough, for small λ 's we can approximate the desired average delay with the lower bound. Intuitively, if the arrival rate λ is small and the MPR capability is strong enough to handle all the packet deliveries, then the probability that the two queues are empty at the steady state tends to be 1, which, known from Appendix H, is equivalent to $\lim_{t \rightarrow \infty} E(1[N_1^t = 0, N_2^t = 0]) = 1$. Therefore, $\lim_{t \rightarrow \infty} E(1[N_1^t > 0, N_2^t > 0]) = 0$, and the analysis in Appendix B shows that the two bounds merge and become the exact average delay.

For moderate λ 's, if the MPR capability is relatively strong with $b + c$ getting close to a , or if the transmission probability p is relatively small, the two bounds get very close to each other, as shown in Figure 3.5(a) and Figure 3.5(b). In these cases, the actual delay of the system can be roughly determined, e.g., by taking the average of the two bounds. Note that when the MPR model reduces to the capture model, $q_{\{1,2\}|\{1,2\}} = c = 0$, the two delay bounds merge, and we obtain the “exact” delay which coincides with the result in [NMT05].

Corollary 3.1 *For a symmetric two-user hybrid ALOHA system with arrival rates*



(a) $p = 1$, $a = 0.8$, $b = 0.4$, $c = 0.2$.



(b) $p = 0.7$, $a = 1$, $b = 0.3$, $c = 0.2$.

Figure 3.5. Delay bounds for the system with different parameters.

$\lambda_1 = \lambda_2 = \lambda$ and transmission probabilities $p_1 = p_2 = p$, under the *Simplistic Assumption*, when the system is stable with $\lambda < p - p^2/2$, the lower and upper bounds of the average delay are given by

$$D_{Lower} = \frac{2(1 - \lambda) - p(1 - \lambda/2)}{2(p - p^2/2 - \lambda)}, \quad (3.69)$$

$$D_{Upper} = D_{Lower} + \frac{p^3}{8\lambda[p - p^2/2 - \lambda]}. \quad (3.70)$$

3.7 An Illustrative Example

In this section, we provide an example to illustrate the performance of the proposed protocol. The example is no longer limited by the *Simplistic Assumption*. The channel estimation in the presence of noise is considered, and the possibility that two training sequences collide in the same pilot sub-slot is also taken into account. The throughput, stability and delay performances of hybrid ALOHA are evaluated and compared with that of traditional ALOHA based on the collision model as well as the MPR model.

3.7.1 System Set-up

Consider a two-user system. Both users communicate with the base station which employs an P -element linear antenna array. The signals of the two users arrive at the array with spatial angles $\theta = [\theta_1, \theta_2]$ with respect to the array normal. Suppose the array elements are uniformly spaced with a distance of one wavelength. The array response matrix $\mathbf{V}(\theta)$ is then given by

$$\mathbf{V}(\theta) = \begin{pmatrix} 1, & e^{j2\pi \cos \theta_1}, & \dots, & e^{j2\pi(P-1) \cos \theta_1} \\ 1, & e^{j2\pi \cos \theta_2}, & \dots, & e^{j2\pi(P-1) \cos \theta_2} \end{pmatrix}^T.$$

The two users access the system based on hybrid ALOHA and the time slots are assumed to be synchronized. The received signal at the base station $\mathbf{y}$ is given by

$$\mathbf{y} = \mathbf{V}(\theta)\mathbf{H}\mathbf{s} + \mathbf{n}, \quad (3.71)$$

where $\mathbf{H} = \text{diag}[h_1, h_2]$ is the diagonal matrix containing the channel fading parameters for the two users. For simplicity we assume flat fading with each $h_i (i = 1, 2)$ being a zero mean complex Gaussian random variable of unit variance, i.e., $E[h_i] = 0$ and $E[|h_i|^2] = 1$; $\mathbf{s} = [s_1, s_2]^T$, where $s_i = 1$ or -1 denotes the symbol transmitted by user i in a particular symbol interval (no intersymbol interference is considered in this example). $\mathbf{n}$ is the additive white Gaussian noise with zero mean and variance $\sigma_n^2 \mathbf{I}_P$, and the noise is assumed to be independent of the information sequence.

We assume that the channel fading is independent for two users. Furthermore, the fading is i.i.d from slot to slot yet remains unchanged during each slot. At the base station, the pilot symbol assisted channel estimate $\hat{h}_i$ differs from the actual channel h_i by an independent complex Gaussian random variable Δh_i , i.e.,

$$\hat{h}_i = h_i + \Delta h_i, \quad i = 1, 2, \quad (3.72)$$

where Δh_i are i.i.d with zero mean and variance σ_e^2 from slot to slot. Thus $\hat{h}_i' s$ are complex Gaussian random variables with zero mean and variance $\hat{\sigma}^2 = 1 + \sigma_e^2$. The actual channel h_i can be written in terms of $\hat{h}_i$ as [GL03]

$$h_i = \rho \hat{h}_i + \tilde{h}_i, \quad (3.73)$$

where $\rho = 1/(1 + \sigma_e^2)$, $\tilde{h}_i' s$ are i.i.d complex Gaussian random variables with zero mean and variance $\tilde{\sigma}^2 = \sigma_e^2/(1 + \sigma_e^2)$, and $E[\tilde{h}_i \hat{h}_j^*] = 0$ for $i, j = 1, 2$.

Considering coherent reception, signal processing at the receiver produces the output

$$\mathbf{z} = \hat{\mathbf{H}}^* \mathbf{W} \mathbf{y} = \hat{\mathbf{H}}^* \mathbf{W} \mathbf{V}(\theta) \mathbf{H} \mathbf{s} + \hat{\mathbf{H}}^* \mathbf{W} \mathbf{n}, \quad (3.74)$$

where $*$ denotes the complex conjugate operator, $\hat{\mathbf{H}} = \text{diag}[\hat{h}_1, \hat{h}_2]$ contains the estimated channel coefficients for the two users, $\mathbf{W}$ represents the beamforming weight matrix. The i th row of $\mathbf{W}$, $\mathbf{w}_i$, represents the weight vector for the i th user.

If we denote the i th column of $\mathbf{V}(\theta)$ as $\mathbf{v}_i$, (3.74) can be rewritten as

$$\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} \mathbf{w}_1 \mathbf{v}_1 h_1 \hat{h}_1^* & \mathbf{w}_1 \mathbf{v}_2 h_2 \hat{h}_1^* \\ \mathbf{w}_2 \mathbf{v}_1 h_1 \hat{h}_2^* & \mathbf{w}_2 \mathbf{v}_2 h_2 \hat{h}_2^* \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} + \begin{bmatrix} \hat{h}_1^* \mathbf{w}_1 \mathbf{n} \\ \hat{h}_2^* \mathbf{w}_2 \mathbf{n} \end{bmatrix}. \quad (3.75)$$

Assuming that the data source is i.i.d., and without loss of generality considering the case when $s_1 = 1$ (results for $s_1 = -1$ can be similarly obtained), from (3.73) and (3.75), it is easy to see

$$E[z_1 | \hat{h}_1] = \mathbf{w}_1 \mathbf{v}_1 \rho |\hat{h}_1|^2, \quad (3.76)$$

and

$$\text{Var}[z_1 | \hat{h}_1] = |\hat{h}_1|^2 (|\mathbf{w}_1 \mathbf{v}_1|^2 \tilde{\sigma}^2 + |\mathbf{w}_1 \mathbf{v}_2|^2 + \mathbf{w}_1 \mathbf{w}_1^H \sigma_n^2), \quad (3.77)$$

Given the estimate of the channel, the signal to interference plus noise ratio (SINR) is then given by

$$E[\text{SINR} | \hat{h}_1] = \frac{\rho^2 |\hat{h}_1|^2}{\tilde{\sigma}^2 + \frac{|\mathbf{w}_1 \mathbf{v}_2|^2}{|\mathbf{w}_1 \mathbf{v}_1|^2} + \frac{\mathbf{w}_1 \mathbf{w}_1^H}{|\mathbf{w}_1 \mathbf{v}_1|^2} \sigma_n^2}. \quad (3.78)$$

If the interference from user 2 is not present, then only the noise accounts, and we have

$$E[\text{SNR} | \hat{h}_1] = \frac{\rho^2 |\hat{h}_1|^2}{\tilde{\sigma}^2 + \frac{\mathbf{w}_1 \mathbf{w}_1^H}{|\mathbf{w}_1 \mathbf{v}_1|^2} \sigma_n^2}. \quad (3.79)$$

3.7.2 Evaluation of Channel Estimation Error

Suppose user 1 transmits a pilot symbol s_1 in one of the pilot sub-slots, and it is contaminated by the interference s_2 from user 2 and the noise. The received composite signal is given by

$$\mathbf{y} = \mathbf{v}_1 h_1 s_1 + \mathbf{v}_2 h_2 s_2 + \mathbf{n}. \quad (3.80)$$

Regardless of the existence of user 2, the “brute force” estimation of h_1 is given by

$$\hat{h}_1 = \frac{\mathbf{v}_1^H \mathbf{y}}{|\mathbf{v}_1^H \mathbf{v}_1| s_1} = h_1 + \frac{\mathbf{v}_1^H \mathbf{v}_2 h_2 s_2}{|\mathbf{v}_1^H \mathbf{v}_1| s_1} + \frac{\mathbf{v}_1^H \mathbf{n}}{\mathbf{v}_1^H \mathbf{v}_1 s_1}. \quad (3.81)$$

The mean-squared error of the estimation is

$$\sigma_e^2 = E[|h_1 - \hat{h}_1|^2] = \frac{|\mathbf{v}_1^H \mathbf{v}_2|^2}{|\mathbf{v}_1^H \mathbf{v}_1|^2} + \frac{\sigma_n^2}{|\mathbf{v}_1^H \mathbf{v}_1|}. \quad (3.82)$$

If $L(\tau)$ pilot symbols are transmitted, where $L(\tau)$ is proportional to τ , then by averaging, the mean-squared error of the channel estimation is reduced by a factor of $1/L(\tau)$, i.e.,

$$\bar{\sigma}_e^2(\tau) = \frac{1}{L(\tau)} \left(\frac{|\mathbf{v}_1^H \mathbf{v}_2|^2}{|\mathbf{v}_1^H \mathbf{v}_1|^2} + \frac{\sigma_n^2}{|\mathbf{v}_1^H \mathbf{v}_1|} \right). \quad (3.83)$$

In the absence of user 2, the channel estimation error is only introduced by the noise, hence

$$\bar{\sigma}_e^2(\tau) = \frac{1}{L(\tau)} \frac{\sigma_n^2}{|\mathbf{v}_1^H \mathbf{v}_1|}. \quad (3.84)$$

3.7.3 Probabilities of Successful Signal Receptions

With the derivations above, we can then adopt the SINR threshold model [ZR94] for packet reception. Under this model, user 1's packet is assumed to be successfully received if $E[\text{SINR}|\hat{h}_1] > \beta$, where β is the predetermined threshold required by system QoS.

When only user 1 transmits, we have

$$q_1|\{1\} = \text{Pr}\left\{ \frac{\rho_1^2 |\hat{h}_1|^2}{\tilde{\sigma}_1^2 + \frac{\mathbf{w}_1 \mathbf{w}_1^H}{|\mathbf{w}_1 \mathbf{v}_1|^2} \sigma_n^2} > \beta \right\}, \quad (3.85)$$

where $\rho_1^2 = 1/(1 + \bar{\sigma}_{e1}^2)$, $\tilde{\sigma}_1^2 = \bar{\sigma}_{e1}^2/(1 + \bar{\sigma}_{e1}^2)$, and $\bar{\sigma}_{e1}^2$ is the variance of the collision-free channel estimation error given by (3.84).

Since $|\hat{h}_1|^2 \sim \exp(1/\hat{\sigma}^2)$, the above probability is given by

$$\begin{aligned} q_{1|\{1\}} &= \frac{1}{\hat{\sigma}^2} \int_{\beta(\hat{\sigma}^2 + \frac{\mathbf{w}_1 \mathbf{w}_1^H}{|\mathbf{w}_1 \mathbf{v}_1|^2} \sigma_n^2)/\rho^2}^{\infty} e^{-\frac{x}{\hat{\sigma}^2}} dx \\ &= \exp\{-\beta[\sigma_{e1}^2 + (1 + \sigma_{e1}^2) \frac{\mathbf{w}_1 \mathbf{w}_1^H}{|\mathbf{w}_1 \mathbf{v}_1|^2} \sigma_n^2]\}. \end{aligned} \quad (3.86)$$

Similarly, the probability of user 1's packet being successfully received when two users stagger the training sequences in two pilot sub-slots, $\tilde{q}_{1|\{1,2\}}$, is given by

$$\begin{aligned} \tilde{q}_{1|\{1,2\}} &= Pr\left\{\frac{\rho_1^2 |\hat{h}_1|^2}{\hat{\sigma}_1^2 + \frac{|\mathbf{w}_1 \mathbf{v}_2|^2}{|\mathbf{w}_1 \mathbf{v}_1|^2} + \frac{\mathbf{w}_1 \mathbf{w}_1^H}{|\mathbf{w}_1 \mathbf{v}_1|^2} \sigma_n^2} > \beta\right\} \\ &= \exp\{-\beta[\sigma_{e1}^2 + (1 + \sigma_{e1}^2) \left(\frac{|\mathbf{w}_1 \mathbf{v}_2|^2}{|\mathbf{w}_1 \mathbf{v}_1|^2} + \frac{\mathbf{w}_1 \mathbf{w}_1^H}{|\mathbf{w}_1 \mathbf{v}_1|^2} \sigma_n^2\right)]\}. \end{aligned} \quad (3.87)$$

Consider the more practical case beyond *the Simplistic Assumption*, there is a probability that both users succeed even if their training sequences collide in one pilot sub-slot. Let $\tilde{\tilde{q}}_{1|\{1,2\}}$ correspond to the probability of user 1's packet being successfully received when two users' training sequences collide. Equation (3.87) applies to obtain the value of $\tilde{\tilde{q}}_{1|\{1,2\}}$. Nevertheless, the collision-free channel estimation error σ_{e1}^2 should be replaced by the estimation error with collision, σ_{e2}^2 , which is calculated using (3.83).

In Section 3.3, we have assumed that the user transmits the training sequence in either pilot sub-slot with equal probability. Hence, given that two users transmit packets concurrently in one time slot, the probability that user 1 is successfully received is

$$q_{1|\{1,2\}} = \frac{1}{2} \tilde{q}_{1|\{1,2\}} + \frac{1}{2} \tilde{\tilde{q}}_{1|\{1,2\}}. \quad (3.88)$$

Similarly, under the assumption that two users are symmetric, the probability that both packets are correctly received when two users transmit in one time slot,

$q_{\{1,2\}|\{1,2\}}$, is given by

$$\begin{aligned} q_{\{1,2\}|\{1,2\}} = & \frac{1}{2} \exp\{-2\beta[\sigma_{e1}^2 + (1 + \sigma_{e1}^2)(\frac{|\mathbf{w}_1 \mathbf{v}_2|^2}{|\mathbf{w}_1 \mathbf{v}_1|^2} + \frac{\mathbf{w}_1 \mathbf{w}_1^H}{|\mathbf{w}_1 \mathbf{v}_1|^2} \sigma_n^2)]\} + \\ & \frac{1}{2} \exp\{-2\beta[\sigma_{e2}^2 + (1 + \sigma_{e2}^2)(\frac{|\mathbf{w}_1 \mathbf{v}_2|^2}{|\mathbf{w}_1 \mathbf{v}_1|^2} + \frac{\mathbf{w}_1 \mathbf{w}_1^H}{|\mathbf{w}_1 \mathbf{v}_1|^2} \sigma_n^2)]\}. \end{aligned} \quad (3.89)$$

Other probabilities can be obtained through similar calculations, and we can now obtain the throughput, stability region and the delay bounds of the proposed protocol.

In the sequel, we assume that at the base station, the linear antenna array has $P = 5$ elements, and the impinging signals are relatively close with $\theta = [54^\circ, 64^\circ]$. The front-end processing exploits matched filter with the beamforming weight matrix $\mathbf{W} = \mathbf{V}^H(\theta)$. Furthermore, we assume the packets are of small size containing 40 symbols per packet, with a training length of 2 symbols corresponding to $\tau = 0.05$.

3.7.4 Throughput of Hybrid ALOHA with Imperfect Channel Estimation

Under these settings, Figure 3.6 presents the overall throughput of hybrid ALOHA with $M = 2$, as well as the throughput of traditional ALOHA based on both the collision model and the MPR model. Note that in the scheme of traditional ALOHA based on MPR model, simultaneous packet receptions are possible: successful signal decoding can be achieved as long as the SINR of the user's signal is beyond the threshold required by QoS. As shown in (3.85)-(3.87), the probabilities of successful transmissions are related to τ and are dependent on the noise power, interference level, and channel estimation accuracy. Following Proposition 3.1, the normalized throughput of hybrid ALOHA, $\nu_{H_ALOHA}(\tau)$, and that of hybrid ALOHA based on the MPR model, $\nu_{ALOHA}(\tau)$, can be evaluated through these probabilities. The results presented in Figure 3.6 show that hybrid ALOHA outperforms traditional ALOHA due to the improved MPR capability.

As we have mentioned in **Remark 1**, the training length τ should be designed according to specific system requirements. In this example, the optimal τ maximizing

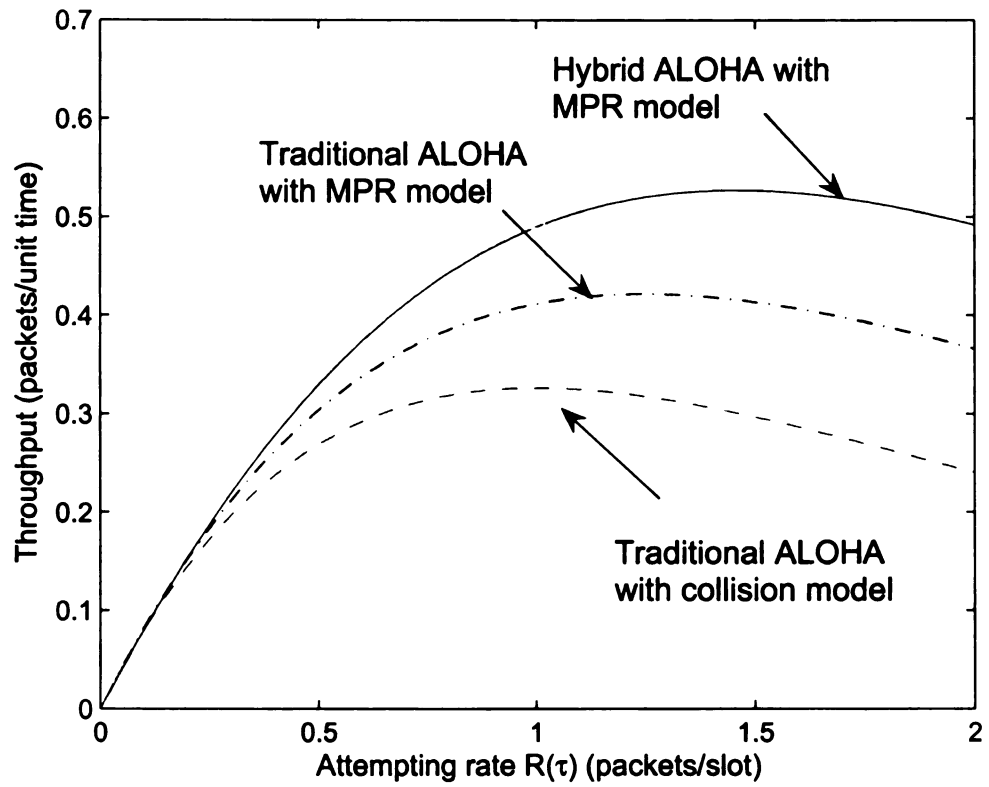


Figure 3.6. Throughput comparison of different schemes under the specific system settings, $SNR = 10dB$, $\beta = 6dB$, $\bar{\sigma}_{e1}^2 = 0.01$, $\bar{\sigma}_{e2}^2 = 0.06$.

the throughput of hybrid ALOHA $\nu_{H_ALOHA}(\tau)$ can be obtained as:

$$\tau_{opt} = \arg \max_{\tau \in \Omega} \{\nu_{H_ALOHA}(\tau)\}, \quad (3.90)$$

where $\Omega = \{\tau \mid 0 < \tau < 1, L(\tau) \in \mathbb{N}\}$, $L(\tau)$ is the number of training symbols defined in Section 3.7.2. The general closed-form expression of τ_{opt} is difficult to find. However, the numerical result can be obtained through computer programming. In Figure 3.7, the dashed line (red color) represents the optimal τ against the arrival rates of each (symmetric) user under the system settings mentioned above. In addition, an “equilibrium” curve (defined as the value of τ such that $\nu_{H_ALOHA}(\tau) = \nu_{ALOHA}(\tau)$) is also presented in Figure 3.7, below which the throughput of hybrid ALOHA is always higher than that of traditional ALOHA with the same training length τ . Beyond the equilibrium curve, the increment of τ would degrade the spectral efficiency compared to that of the traditional ALOHA. Note that the staircase-like shapes in both curves reflect the constraint that the training sequence should contain integer number of symbols, i.e., $L(\tau) \in \mathbb{N}$.

In Figure 3.7, it is as expected that when arrival rates are very low, equilibrium values do not exist. This is because the insertion of idle sections under the circumstances of rare collisions makes hybrid ALOHA inferior to traditional ALOHA.

3.7.5 Stability Regions and Delay Bounds with Imperfect Channel Estimation

Figure 3.8 shows the stability regions for hybrid ALOHA under the aforementioned system settings. The arrival rates are in the unit of packets per unit time. For moderate arrival rates, hybrid ALOHA is shown to have the best performance due to the MPR improvement. On the other hand, in the cases when the arrival rate of one user is at the low end while that of the other user is at the high end, the stability region for hybrid ALOHA is contained in that of traditional ALOHA, and this degradation reflects the impact of idle sections. Refer to Figure 3.4, it can be seen that within certain ranges, the stability region of hybrid ALOHA contains that

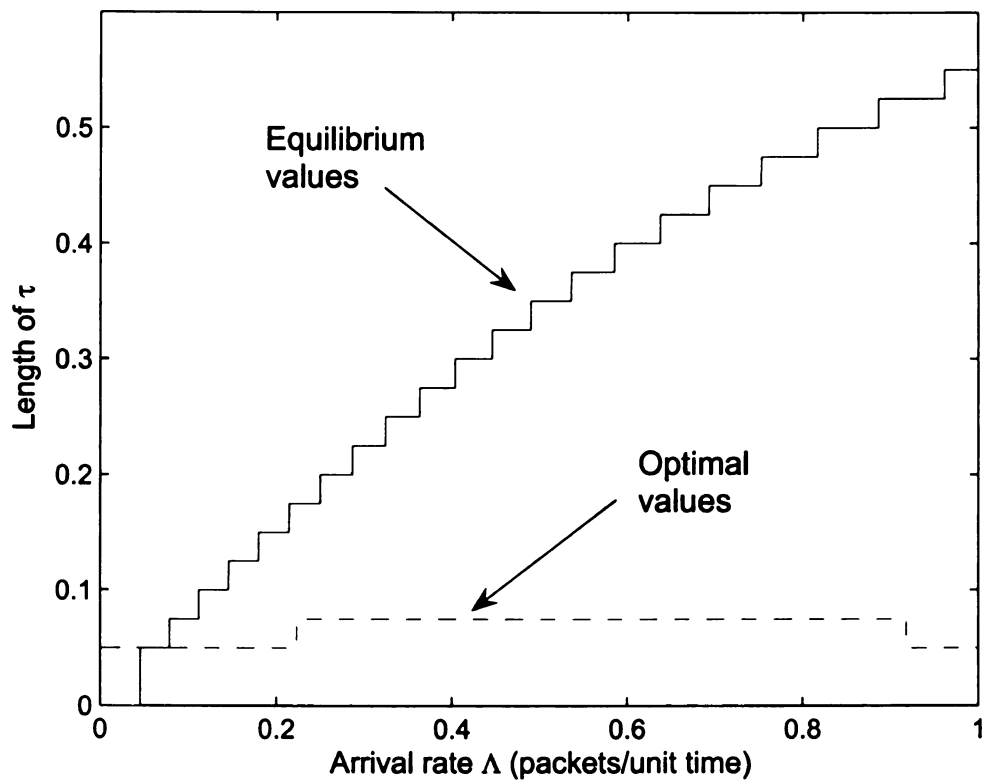


Figure 3.7. Numerical values of τ achieving equilibrium points and τ_{opt} maximizing the throughput of hybrid ALOHA .

of TDMA as well, even for the scheme without any guard time.

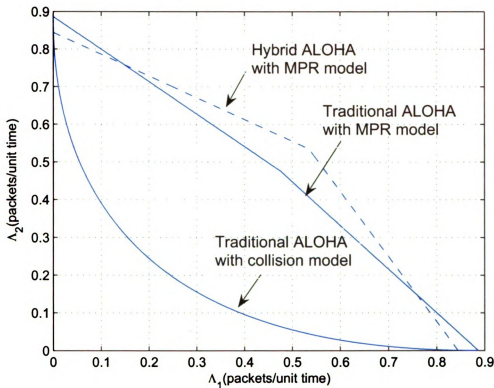


Figure 3.8. Comparison of stability regions of different schemes, $SNR = 10dB$, $\beta = 6dB$, $\sigma_{e1}^2 = 0.01$, $\sigma_{e2}^2 = 0.06$.

Figure 3.9 illustrates the delays (bounds) of three different schemes as a function of the arrival rate for fixed transmission probability $p = 0.75$. Hybrid ALOHA results in the smallest delay among the three schemes for moderate rates, while the trend is reversed when the rates are small. The gains over traditional ALOHA with the MPR model seem relatively minor from the figure. One of the reasons is that the system contains only a small number of users ($N = 2$), which limits the possible improvement range. Much larger gap could be observed in systems with more users, which deserves further investigation.

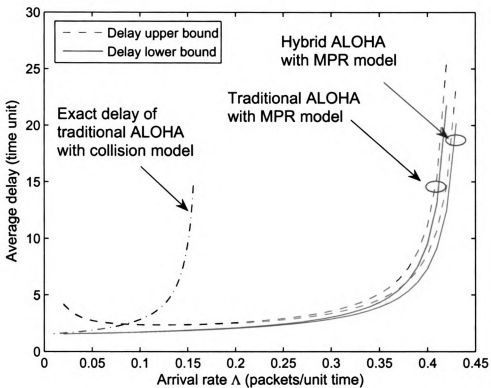


Figure 3.9. Comparison of delays of different schemes, $SNR = 10dB$, $\beta = 6dB$, $\bar{\sigma}_{e1}^2 = 0.01$, $\bar{\sigma}_{e2}^2 = 0.06$, $p = 0.75$.

3.8 Summary

In this chapter, taking a cross-layer (MAC and PHY) perspective, we propose a novel MAC protocol named hybrid ALOHA for wireless networks. Idle sections are introduced into the traditional ALOHA slot to make it possible for collision-free channel estimation in multiple access environment. Improved channel estimation increases the MPR capability and results in better system performances in terms of throughput, stability and delay. The major results can be briefly summarized as follows. **First**, we simplify the analysis by adopting *the Simplistic Assumption*, in which a transmission is successful if and only if collision-free channel estimation can be achieved, given that the noise and the interference is limited within a certain degree. Under the simplified model, the throughput of the system is obtained. The performance demonstrates a monotonicity with the increase of the MPR capability. **Secondly**, we characterize the stability region of hybrid ALOHA in the two-user case, and derive the sufficient condition for system stability in the more general finite-user case. Characterization of the stability inner bound for the special $N = 3$ case is present. It is shown in the two-user case that, under *the Simplistic Assumption*, the hybrid ALOHA system reaches a critical MPR capability such that it has an identical stability region with that of the TDMA scheme. The characterization of the stability inner bound for the $N = 3$ case can be reduced to solving a general Riemann-Hilbert problem. **Thirdly**, we study the delay performance for the two-user system under the general MPR model. The “exact” delays for the asymmetric system and the upper and lower bounds of the average delay for the symmetric case are derived. It is shown that when the MPR channel reduces to the capture channel, the two bounds merge and we get the “exact” delay of the system which coincides with the result in [NMT05]. **Finally**, beyond the *Simplistic Assumption*, the effectiveness of the hybrid ALOHA protocol is demonstrated through a practical example using array signal processing.

CHAPTER 4

Source-Aware Non-Uniform Transmission for Minimum Distortion

This chapter considers average input-output distortion minimization through joint optimization of source index assignment and modulation design. First, the general optimization criterion is derived. Secondly, source-aware non-uniform information transmission is proposed by utilizing the inherent non-uniformity in Gray coded constellations, both symmetric and asymmetric. It is found that when channel coding is involved, the proposed scheme delivers significantly better performance in comparison with existing unequal error protection schemes with non-Gray codewords. Furthermore, joint quantization-constellation design is investigated under the criterion of distortion minimization. The proposed method generalizes the concept of constellation design from the perspective of joint source-channel coding. The simplicity and power efficiency of the proposed schemes make them particularly attractive for systems with tight power constraints, such as wireless sensor networks and space communications.

4.1 Introduction

Given a source with rate R bits/second and a channel with capacity C bits/second, Shannon's well known channel capacity theorem says that if $R < C$, then there exists a combination of source and channel coders such that the source can be communicated

over the channel with fidelity arbitrarily close to perfect. This theorem essentially implies that the source coding and channel coding are fundamentally *separable* without loss of performance for the overall system.

Following the “*separation principle*”, in most modern communication systems, source coding and channel coding are treated independently. In other words, source representation is designed disjointly from information transmission. After A/D conversion of analog source signals, the bit streams are then *uniformly* encoded and mapped to symbols prior to transmission, uniform bit-error-rate (BER) has been serving as one of the most commonly used performance measures. However, for system with analog input, the ultimate goal of the communication system is to minimize the overall input-output distortion. While BER plays the dominant role in distortion minimization, a communication system that minimizes the BER does not necessarily minimize the overall input-output distortion. Therefore, novel system design that can achieve simultaneous BER and distortion minimization is highly desirable.

In [ZG93], with no specifications on the modulation schemes, Zeger proposed a locally optimal index assignment (source coding) solution for minimum input-output distortion, known as pseudo-Gray coding, by using iterative binary switching. Pseudo-Gray coding provided an effective approach in reducing the average distortion of a vector quantized system by rearranging the positions of code vectors in a given codebook. Various channel-optimized quantizers, multistage vector quantizers have been studied in [FV87, PFM93, WF94]. On the other hand, assuming source index assignment has been done separately, and with the observation that the bits come out of the source encoder are generally non-uniform (i.e. have different levels of significance), Masnick, Wolf [MW67] and Cover [Cov72] introduced non-uniform modulation (index mapping) schemes, known as unequal error protection codes, for which the more important bits have lower error rate than other bits. Stemmed from [MW67, Cov72], unequal error protection through both symmetric and asymmetric constellations have been further developed in [ROUV93], [MZFLI00a] and [MZFLI00b]. It should be pointed out that the resulted constellation codeword design in [ROUV93, MZFLI00a, MZFLI00b] may no longer be Gray codes.

In this chapter, taking a mixed analog-digital perspective, we consider average input-output distortion minimization through joint optimization of source index assignment and modulation design. *First*, we derive the general optimization criterion, and obtain that the necessary condition for simultaneous minimization of BER and distortion is that the quantization codebook and the modulation constellation should have the same dimensionality. *Secondly*, based on the fact that Gray code ensures minimum BER when the channel error probability is sufficiently small, we propose a novel source-aware information transmission approach by exploiting the non-uniformity in Gray-coded constellations. It is shown that when there is no channel coding, the proposed approach has comparable performance with existing modulation schemes in [ROUV93, MZFLI00a, MZFLI00b]. But when channel coding is involved, our design outperforms the existing schemes with big margin. The underlying argument is that after channel coding, BER dominates the distortion minimization as the original geometric structure of the quantization codebook can no longer be maintained. The simplicity and power efficiency of the proposed scheme make it particularly attractive for systems with tight power constraints, such as wireless sensor networks and space communications.

4.2 Problem Formulation

Consider a digital communication system with analog input, as shown in Figure 4.1. Let $\mathbf{x}_k$ be the discrete-time analog input vector resulted from uniform sampling of a continuous signal $x(t)$.

$\mathbf{x}_k$ is first fed into a quantizer Q , which is a mapping of n -dimensional Euclidean space $\mathbf{R}^n$ to a finite set $P \subset \mathbf{R}^n$, given by $Q : \mathbf{R}^n \rightarrow P$, where $P = \{P_0, P_1, \dots, P_{M-1}\}$ is the quantization codebook with $P_i \in \mathbf{R}^n$ for $0 \leq i \leq M-1$. We assume that the size of P is $|P| = M = 2^m$, where $m \geq 0$ is an integer. Let $\mathbf{y}_k = Q(\mathbf{x}_k)$ denote the quantization value of $\mathbf{x}_k$. $\mathbf{y}_k$ is coded into a binary sequence through an index assignment function π , and is then fed into a source-aware digital channel encoder and a modulator, i.e., the most significant bits (MSB) and least sig-

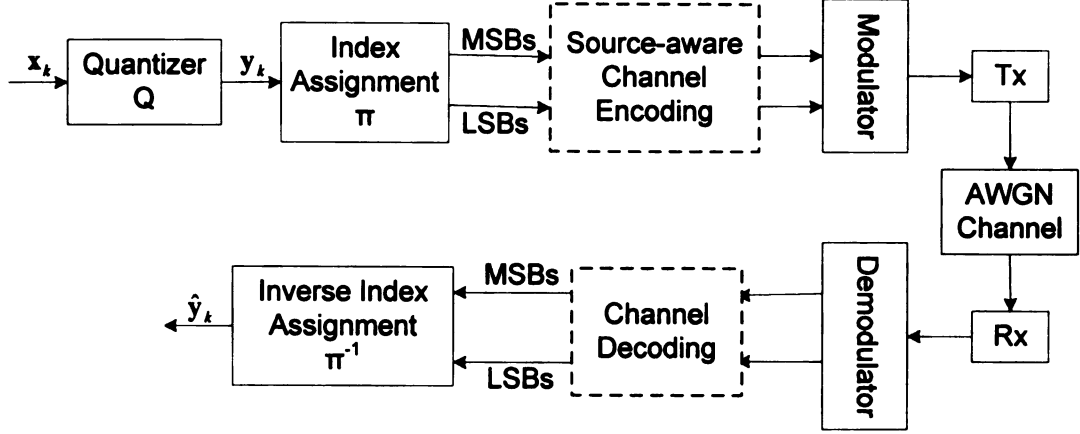


Figure 4.1. System Model

nificant bits (LSB) will be treated distinctly. Let $\hat{\mathbf{y}}_k$ denote the receiver output, which is an estimate of the quantization value $\mathbf{y}_k$, the averaged input-output distortion is then given by

$$D_0 = E\{d(\mathbf{x}_k, \hat{\mathbf{y}}_k)\}, \quad (4.1)$$

where $d : \mathbf{R}^n \times \mathbf{R}^n \rightarrow \mathbf{R}$ is a non-negative function that measures the distance between two vectors in $\mathbf{R}^n$.

Consider the widely used mean-square distortion function $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|^2$. In this case, the optimal quantizer satisfies the well-known nearest neighbor and centroid conditions [ZG93, ZG88], then the overall distortion D_0 can be decomposed into two parts, namely, the distortion due to quantization noise, and the distortion due to channel noise [Mes79], denoted as $\mathbf{n}_q$ and $\mathbf{n}_c$, respectively. That is,

$$\mathbf{x}_k - \hat{\mathbf{y}}_k = \underbrace{(\mathbf{x}_k - \mathbf{y}_k)}_{\mathbf{n}_q} + \underbrace{(\mathbf{y}_k - \hat{\mathbf{y}}_k)}_{\mathbf{n}_c}$$

When the quantizer satisfies the centroid condition, $E\{\mathbf{n}_q\} = 0$. Note that the quantization noise and the channel noise are independent, we have $E\{\mathbf{n}_q \mathbf{n}_c^H\} =$

$E\{\mathbf{n}_c \mathbf{n}_q^H\} = 0$. It then follows that

$$\begin{aligned} D_0 &= E\{\|\mathbf{n}_q\|^2\} + E\{\|\mathbf{n}_c\|^2\} \\ &= E\{\|\mathbf{x}_k - \mathbf{y}_k\|^2\} + E\{\|\mathbf{y}_k - \hat{\mathbf{y}}_k\|^2\}. \end{aligned} \quad (4.2)$$

When the quantizer is optimal, the distortion due to quantization error is minimized. Minimization of D_0 is thus reduced to minimizing the distortion only due to the channel noise

$$D = E\{d(\mathbf{y}_k, \hat{\mathbf{y}}_k)\}. \quad (4.3)$$

In the sequel, we will discuss joint source index assignment and constellation codeword design for minimum distortion, as well as non-uniform transmission based on Gray-coded constellations.

4.3 Joint Source Index Assignment And Constellation Code Design

In this section, we consider to minimize the distortion $D = E\{\|\mathbf{y}_k - \hat{\mathbf{y}}_k\|^2\}$ through joint design of source index assignment and index mapping.

Write $\mathbf{y}_k$ as

$$\mathbf{y}_k = \hat{\mathbf{y}}_k + \mathbf{e}_k \quad (4.4)$$

where $\mathbf{y}_k, \hat{\mathbf{y}}_k \in P = \{P_0, P_1, \dots, P_{M-1}\}$, and $\mathbf{e}_k$ is the estimation error. For $0 \leq i \leq M-1$, define

$$E_i = \{P_i - P_j, 0 \leq j \leq M-1\}. \quad (4.5)$$

It then follows that

$$\begin{aligned}
D &= E\{\|\mathbf{y}_k - \hat{\mathbf{y}}_k\|^2\} \\
&= \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} \|P_i - P_j\|^2 p(\hat{\mathbf{y}}_k = P_j | \mathbf{y}_k = P_i) p(\mathbf{y}_k = P_i) \\
&= \sum_{i=0}^{M-1} p(\mathbf{y}_k = P_i) \sum_{\mathbf{e}_k \in E_i} \|\mathbf{e}_k\|^2 p(\mathbf{e}_k).
\end{aligned} \tag{4.6}$$

Here $p(x)$ denote the probability that x occurs.

For efficient transmission, each quantizer output $\mathbf{y}_k$ is first coded to a binary sequence then mapped to a symbol in a constellation Ω . When the signal-to-noise ratio (SNR) is reasonably high, as it is for most useful communication systems, each transmitted symbol is more likely to be mistaken for one of its neighbors than for far more distant symbols. Therefore, to minimize the distortion D , the optimal index assignment and constellation codeword design should map the neighboring quantization vectors from the quantization codebook P to neighboring symbols in constellation Ω . More specifically, the optimal 1-1 mapping $S : P \rightarrow \Omega$ should satisfy the following condition:

$$\text{(C1) Let } P_i, P_j, \tilde{P}_i, \tilde{P}_j \in P, \text{ then } d(P_i, P_j) \leq d(\tilde{P}_i, \tilde{P}_j) \text{ if and only if } d(S(P_i), S(P_j)) \leq d(S(\tilde{P}_i), S(\tilde{P}_j)).$$

That is, ideally, an isomorphic mapping that reserves the geometric structure should exist between the quantization codebook P and the constellation Ω . *When the quantizer is optimal, and the constellation is Gray coded, condition (C1) ensures the equivalence between minimizing the BER and minimizing the average distortion D .*

Next we look at the necessary conditions for the existence of S that satisfies (C1). First, assume that the size of the constellation $|\Omega| = |P| = M$, and then look at the case when $|\Omega| < |P|$.

We start with systems equipped with scalar quantizers and two-dimensional constellations. Consider a system with a 4-bit uniform scalar quantizer and a 16-QAM

constellation, see Figure 4.2. Since $P = \{P_0, P_1, \dots, P_{15}\} \subset \mathbb{R}$, without loss of generality, assume $P_0 < P_1 < \dots < P_{15}$. As can be seen, each P_i has at most two nearest neighbors, but a symbol in a 16-QAM constellation can have as many as 4 nearest neighbors. It is then impossible to find an $S : P \rightarrow \Omega$ that satisfies (C1).

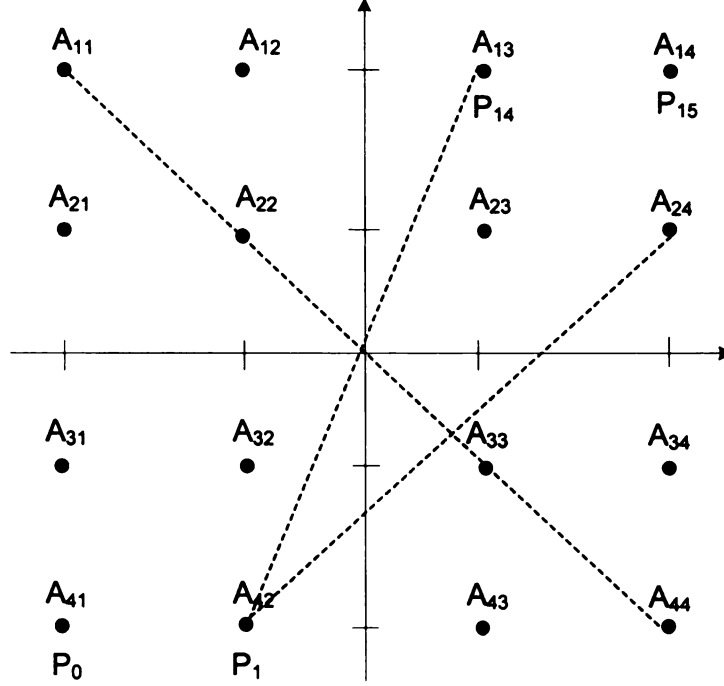


Figure 4.2. Mapping of quantized values to the 16-QAM constellation.

In fact, assuming there is an $S : P \rightarrow \Omega$ that satisfies (C1), then we should have

$$d(S(P_0), S(P_{15})) = \max_{x_1, x_2 \in \Omega} d(x_1, x_2), \quad (4.7)$$

$$d(S(P_0), S(P_1)) = d(S(P_{14}), S(P_{15})) = \min_{x_1, x_2 \in \Omega} d(x_1, x_2), \quad (4.8)$$

$$d(S(P_1), S(P_{14})) \geq \max_{P_i, P_j \in P} \{d(S(P_i), S(P_j))\}, \quad i, j \neq 0, 1, 14, 15. \quad (4.9)$$

Without loss of generality, assume $S(P_0) = A_{41}$, $S(P_{15}) = A_{14}$, as illustrated in Figure 4.2. Now consider the pair P_1 and P_{14} . For (4.8) to be satisfied, P_1 and P_{14} should be mapped to the nearest neighbors of P_0 and P_{15} respectively. Without loss of generality, assume $S(P_1) = A_{42}$. Since $d(A_{42}, A_{13}) > d(A_{42}, A_{24})$, according to

(C1), we should have $S(P_{14}) = A_{13}$. However, this violates (4.9), since $d(A_{42}, A_{13}) < d(A_{11}, A_{44})$ but A_{11}, A_{44} will correspond to points from $\{P_i, i \neq 0, 1, 14, 15\}$. This implies that, to satisfy (4.9), P_1 and P_{14} should be mapped to the pair A_{11} and A_{44} . Clearly, this violates (4.8). Therefore, an S that satisfies (C1) does not exist.

More generally, we have:

Lemma 4.1 *For systems utilizing scalar quantizer with codebook P and a symmetric (two-dimensional) rectangular or square constellation Ω with $|\Omega| = |P| = 2^m$, $m > 1$, there is no 1-1 mapping $S : P \rightarrow \Omega$ that satisfies (C1).*

For the 4-bit scalar quantizer discussed above, instead of 16-QAM, consider the one-dimensional constellation 16-AM with Gray code, see Figure 4.3. Still denoting

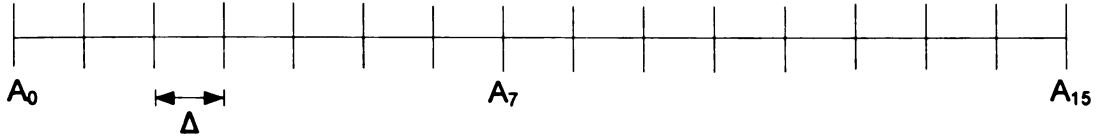


Figure 4.3. Symmetric 16-AM constellation.

the constellation with Ω , clearly the 1-1 mapping $S : P \rightarrow \Omega$ defined by $S(P_i) = A_i$, $0 \leq i \leq 15$, satisfies (C1), and it minimizes the BER and the distortion D simultaneously.

Remark 4.1 *Our discussions in the one-dimensional case imply that: simultaneous minimization of BER and distortion D essentially requires that there exists a 1-1 mapping $S : P \rightarrow \Omega_0$ which satisfies condition (C1), where $\Omega_0 \subseteq \Omega$ is a real subset of Ω or Ω itself. This result provides another demonstration to the well known fact that: the essence to obtaining larger coding gain is to design codes in a subspace of signal space with higher dimensionality, as a larger minimum distance can be obtained with the same signal power. For example, two-dimensional constellation such as QAM would be a natural choice for two-dimensional vector quantizers. For multidimensional vector quantizers, multidimensional constellations would fit best [FW89]. In the case*

when $|\Omega| < |P|$, more than one constellation symbols are needed to represent one quantization value. Again, the multidimensional signal constellation obtained from the Cartesian product Ω^N should be exploited and is currently under investigation.

Due to lack of *a priori* statistical information of the input signal, non-entropy coding is widely used for various sources in practice. That is, in each quantization codeword, some bits are more significant than other bits, and an error in a significant bit will result in large distortion than that in a less significant bit. This universal existence of non-uniformity in source coding calls for non-uniform information transmission, also known as unequal error protection, in which the most significant bits have lower bit-error-rates than other bits. In the following, we consider source-aware non-uniform transmission design along the line of joint source index assignment and constellation design.

4.4 Source-Aware Non-Uniform Transmission Design

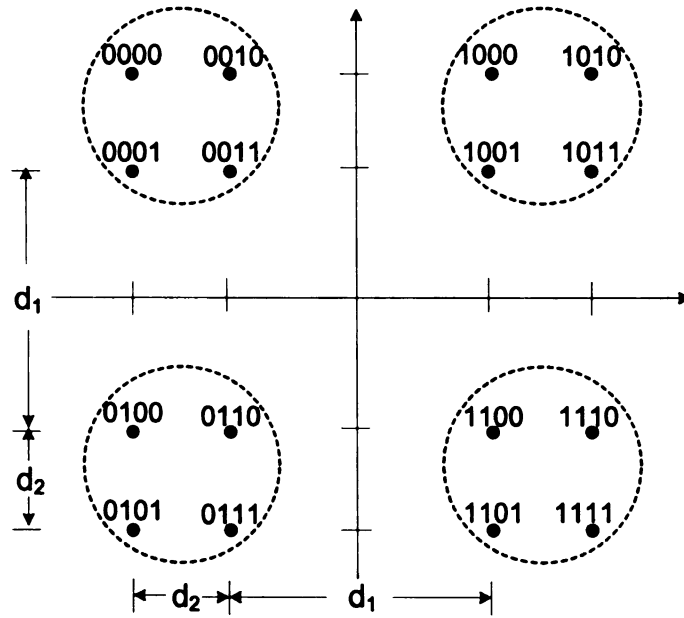
4.4.1 Non-Uniform Transmission Based on Gray-Code Constellations

Non-uniformity exists in most constellations except BPSK and QPSK, either they are asymmetric or even if they are perfectly symmetric.

Asymmetric constellations, see Figure 4.4(a) for example, were originally developed for multiresolution (MR) broadcast in Digital HDTV [Cov72, ROUV93], and further analyzed in [MZFLI00b]. The asymmetric constellations were designed to provide more protection to the more significant bits by grouping bits into clouds leading by the most significant bits, and the minimum distance between the clouds is larger than the minimum distance between symbols within a cloud.

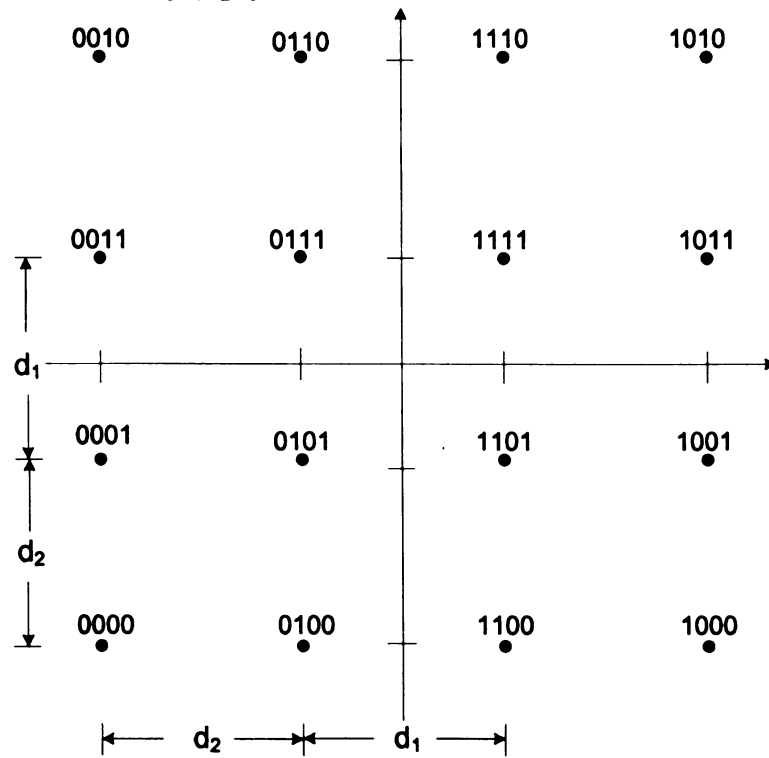
For symmetric constellations, an unequal error protection scheme based on block partitioning is provided in [MZFLI00a], which is a generalization of the Ungerboeck's

16-QAM: $b_0b_1b_2b_3$



(a)

16-QAM: $b_0b_1b_2b_3$



(b)

Figure 4.4. (a) Asymmetric 16-QAM, $d_1 > d_2$; (b) Symmetric 16-QAM with Gray codes.

mapping by set partitioning. With block partitioning, the number of nearest neighbors is minimized for each bit level b_i . It turns out that the resulted codeword design coincide with that of the MR scheme with $d_1 = d_2$. As can be seen, constellations resulted from either block partitioning or the MR scheme may no longer be Gray-coded.

As is well known, Gray codes are developed to minimize the bit-error-rate, in which the nearest neighbors correspond to bit groups that differ by only one position. Here, we revisit the non-uniformity in constellations with Gray codes, and introduce a non-uniform transmission scheme based on Gray-coded constellations. In the following, we illustrate the idea through Gray coded 16-QAM constellation shown in Figure 4.4(b).

In 16-QAM, each codeword has the form $b_0b_1b_2b_3$. If we go through the 16 symbols in Figure 4.4(b), there are altogether 24 nearest neighbor bit changes, among which b_0 and b_2 each changes 4 times, and b_1 and b_3 each changes 8 times. Note that when channel probability error is sufficiently small, a bit error corresponding to each bit location b_i is most likely to occur when the nearest neighbor has a different value in that specific bit location, i.e., among neighboring pairs where a change occurs. Let P_e denote the error probability, then this implies that when SNR is reasonably high, $P_e(b_0) = P_e(b_2) = \frac{1}{2}P_e(b_1) = \frac{1}{2}P_e(b_3)$. Our analysis coincides with the theoretical results in [LLCL99, YH00, CY02].

We propose to minimize the average distortion by exploiting the inherent non-uniformity in Gray-coded constellations, that is, to *map the more significant bits from the source encoder to bit locations with lower error probability in constellations with Gray codes*. For example, consider a 4-bit quantizer and a 16-QAM constellation as in Figure 4.4(b), the two MSBs will be mapped to b_0 and b_2 , while the two LSBs be mapped to b_1 and b_3 .

The proposed approach can be applied to both symmetric and asymmetric constellations. To illustrate the performance, we compare the proposed Gray-code based non-uniform transmission scheme with the block partitioning based approaches [MZFLI00a, MZFLI00b] for both coded and uncoded systems (note that the

MR scheme [ROUV93] is only for asymmetric constellations and coincides with the block partitioning based method in the asymmetric case [MZFLI00a, MZFLI00b]).

Example 1 The source is assumed to be analog with the amplitude uniformly distributed within $[0,100]$, quantized using a 12-bit uniform quantizer. We consider various 16-QAM constellations, both symmetric and asymmetric. First, each 12-bit quantization output $b_0b_1 \cdots b_{11}$ is partitioned into three 4-bit strings: $b_0b_1b_6b_7$, $b_2b_3b_8b_9$, $b_4b_5b_{10}b_{11}$, then mapped to both symmetric and asymmetric 16-QAM constellations based on the block partitioning (BP) scheme or the proposed Gray-code based non-uniform transmission scheme. By random index assignment, we mean that no distinction is made on MSBs and LSBs, and the strings are mapped to the Gray coded constellation based on their original bit arrangements $b_0b_1b_2b_3$, $b_4b_5b_6b_7$, $b_8b_9b_{10}b_{11}$. The result is shown in Figure 4.5.

Example 2 In this example, impact of channel coding is investigated for both systematic and non-systematic coding schemes. Using the same source as in Example 1, a 10-bit uniform quantizer is connected with a source-aware channel encoder, for which the first 4 MSBs are fed to a rate 1/3 convolutional (or Turbo) encoder and the rest 6 bits are fed to a rate 1/2 convolutional(or Turbo) encoder, respectively. The channel coding output are then mapped to 16-QAM constellations non-uniformly based on the block partitioning approach and the proposed mapping scheme. The result is shown in Figure 4.6.

Remark 4.2 *As demonstrated by the simulation results, while the proposed approach has comparable performance with existing unequal error protection methods for uncoded systems (i.e. when there is no channel coding), the Gray-code based non-uniform transmission outperforms the non-Gray-coded methods (i.e., the MR method and the block partitioning based approach) with big margins when channel coding is involved. The underlying arguments are: (i) Channel coding may change the geometric structure of the uncoded symbols; (ii) When SNR is reasonably high, BER of the more significant bits vanishes, and BER of the less significant bits dominates the overall distortion, and hence Gray coded constellations result in much better performance.*

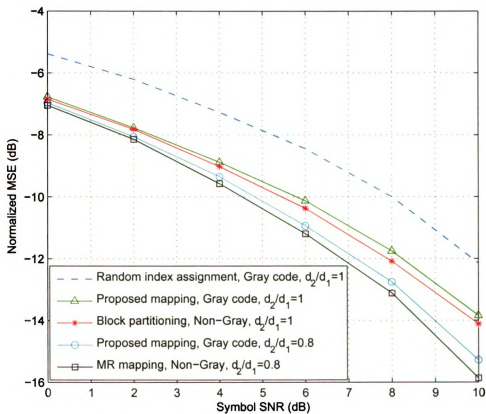


Figure 4.5. MSE for different transmission schemes: Uncoded System.

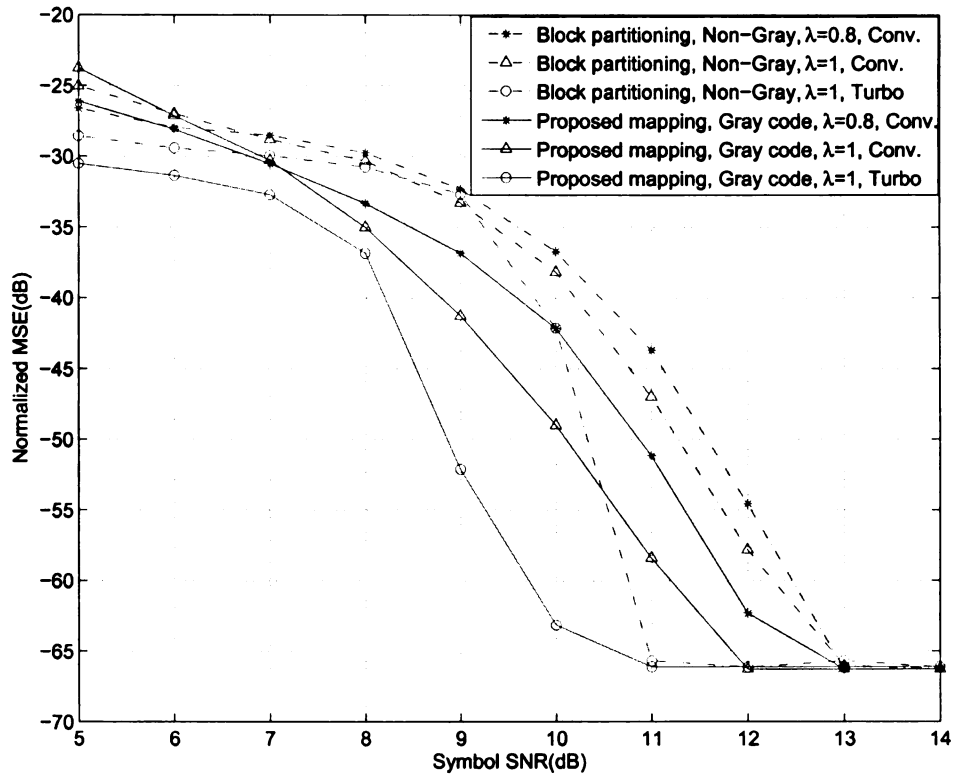


Figure 4.6. MSE for different transmission schemes: Coded System ($\lambda = d_2/d_1$).

4.4.2 Source-Aware Constellation Design

In literature, constellation design has largely been separated from quantizer design. In this section, we consider joint quantizer-constellation design for minimum distortion.

Following our discussions in Section 4.3 and Section 4.4.1, we propose to incorporate the source information reflected in optimal quantizer design into constellation design. Note that the optimal quantizers minimize the average distortion between the original sampled values and the quantization values, when considering memoryless AWGN channels, optimal quantizers can be exploited directly for constellation design. We illustrate this idea through the following example.

Consider a non-uniform scalar quantizer with four possible quantization values. Assuming the quantization code book is $P = \{P_1, \dots, P_4\}$, where $P_i < P_{i+1}$ for $i = 1, 2, 3$ and each P_i occurs with probability $p(P_i) = p_i$ for $i = 1, \dots, 4$. Define $D_{ij} = |P_i - P_j|^2$. Along the lines of Lemma 4.1, we consider the design of a 4-AM constellation $\Omega = \{A_1, \dots, A_4\}$ and assume that the 1-1 map $S : \Omega \rightarrow P$ is designed to satisfy condition (C1). We further assume that the quantizer is optimal, that is, it satisfies the nearest neighbor rule and the centroid criterion. Define $d_{ij} = |P_i - P_j|$ for $i, j = 1, \dots, 4$, $d_i = |A_i - A_{i+1}|$ for $i = 1, 2, 3$, and let D_i be the average distortion corresponding to symbol A_i for $i = 1, \dots, 4$. Consider a memoryless AWGN channel, for which the noise is zero mean and with variance σ^2 , then we have

$$D_1 = d_{12}^2 \left[Q\left(\frac{d_1}{2\sigma}\right) - Q\left(\frac{2d_1 + d_2}{2\sigma}\right) \right] + d_{14}^2 Q\left(\frac{2d_1 + 2d_2 + d_3}{2\sigma}\right) + d_{13}^2 \left[Q\left(\frac{2d_1 + d_2}{2\sigma}\right) - Q\left(\frac{2d_1 + 2d_2 + d_3}{2\sigma}\right) \right]. \quad (4.10)$$

$$D_2 = d_{12}^2 Q\left(\frac{d_1}{2\sigma}\right) + d_{23}^2 \left[Q\left(\frac{d_2}{2\sigma}\right) - Q\left(\frac{2d_2 + d_3}{2\sigma}\right) \right] + d_{24}^2 Q\left(\frac{2d_2 + d_3}{2\sigma}\right). \quad (4.11)$$

$$D_3 = d_{13}^2 Q\left(\frac{d_1 + 2d_2}{2\sigma}\right) + d_{23}^2 \left[Q\left(\frac{d_2}{2\sigma}\right) - Q\left(\frac{2d_2 + d_1}{2\sigma}\right) \right] + d_{34}^2 Q\left(\frac{d_3}{2\sigma}\right). \quad (4.12)$$

$$D_4 = d_{34}^2 \left[Q\left(\frac{d_3}{2\sigma}\right) - Q\left(\frac{d_2 + 2d_3}{2\sigma}\right) \right] + d_{14}^2 Q\left(\frac{d_1 + 2d_2 + 2d_3}{2\sigma}\right) + d_{24}^2 \left[Q\left(\frac{d_2 + 2d_3}{2\sigma}\right) - Q\left(\frac{d_1 + 2d_2 + 2d_3}{2\sigma}\right) \right]. \quad (4.13)$$

where $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-t^2/2} dt$. The overall average distortion can be written as

$$D = \sum_{i=1}^4 p_i D_i. \quad (4.14)$$

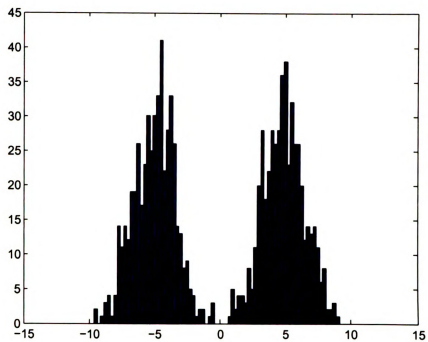
Defining $\gamma_1 = \frac{d_1}{\sigma}$, $\gamma_2 = \frac{d_2}{d_1}$, $\gamma_3 = \frac{d_3}{d_1}$, the problem of optimal constellation design for minimum average distortion is reduced to finding $\gamma_1, \gamma_2, \gamma_3$ such that D is minimized, subjected to a power constraint, that is

$$\min_{\gamma_1, \gamma_2, \gamma_3} D, \quad \text{subjected to } P_s \leq C, \quad (4.15)$$

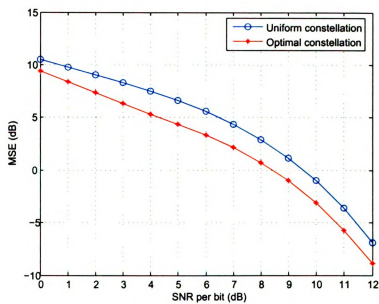
where P_s is the average symbol power and C is a constant. The method used in this example can be extended directly to more general cases. We illustrate the proposed approach through an example.

Example 3 Consider a source consists of two Gaussian distributed random processes centered at ± 5 with variance $\sigma^2 = (\frac{5}{3})^2$, as shown in Figure 4.7(a). Using a 4-level optimal quantizer with codebook $P = \{-6.40, -3.86, 3.69, 6.2\}$, normalized MSE under different SNR levels is shown in Figure 4.7(b) for both uniform 4-AM and the proposed source-aware 4-AM.

From the simulation result, it can be seen that while symmetric constellations are optimal for uniformly distributed sources, asymmetric constellations reduce the average distortion significantly for sources that require non-uniform quantization. Compared with the asymmetric constellations originally designed for multiresolution broadcasting [Cov72, ROUV93], the proposed joint quantizer-constellation design scheme generalizes the concept of non-uniform constellation design from the perspective of joint source-channel coding.



(a)



(b)

Figure 4.7. Example 3: (a) Source distribution (b) Normalized MSE under different SNR levels, uniform 4-AM versus the proposed source-aware 4-AM.

4.5 Summary

In this chapter, we studied joint optimization of source index assignment and modulation design for overall input-output distortion minimizations in communication systems. Taking a joint source-channel coding perspective, distortion minimization was carried out through Gray-code based non-uniform mapping and joint quantizer-constellation design. The power efficiency of the proposed source-aware non-uniform transmission scheme makes it attractive for any digital systems with analog inputs, particularly for systems with tight power constraints such as wireless sensor networks and space communications.

CHAPTER 5

Conclusions and Future Work

5.1 Conclusions

In this dissertation, we explore reliable information transmission and efficient medium access issues in wireless communication networks.

Specifically, in **Chapter 2** we investigate at the physical layer the channel estimation and signal detection problem for time-varying channels, with the objective to improving spectral efficiency with satisfying reliability. We reveal that channel correlation can be exploited in a very simple way to jointly estimate channel response and detect transmitted signals. High spectral efficiency can be achieved with satisfying information transmission reliability, in terms of bit error rate and channel tracking accuracy. Additional efforts are devoted to studying the extreme case when the channel correlation is lost due to some abrupt changes. We propose an algorithm to detect such abrupt changes and suppress the possible error propagation. The proposed algorithms are demonstrated to be effective through simulations.

Next, going beyond the physical layer, we inspect in **Chapter 3** the medium access control problem yet in a fashion of joint MAC-PHY optimization. A novel MAC protocol named hybrid ALOHA is proposed, where the slots are purposely structured in order to improve network performance. Idle sections are introduced to make it possible for collision-free channel estimation even when multiple users transmit simultaneously in one slot. Improved channel estimation increases the MPR capability and results in better system performances in terms of throughput, stability and delay. It is demonstrated that as the length of the idle section is small compared to that of the data section, the bandwidth cost introduced by the idle sections can be

compensated for by the increased MPR capability, and the throughput of the network is greatly increased.

We then characterize in the stability region of hybrid ALOHA in the two-user case and derive the sufficient condition for stability for the general finite-user case. Characterization of the stability inner bound for the special $N = 3$ case is also present, which is reduced to solving a general Riemann-Hilbert problem. Delay performance for the two-user system is also studied under the general MPR model. We obtain the "exact" delays of the asymmetric case, as well as the upper and lower delay bounds for the symmetric system. The two bounds could be fairly close to each other under certain conditions such as a strong MPR capability. When the MPR channel reduces to the capture channel, the two bounds merge and we have the "exact" delay of the system which coincides with the result in [NMT05].

Lastly, in **Chapter 4**, we studied joint optimization of source index assignment and constellation design for overall input-output distortion minimizations in digital wireless systems with analog inputs. It turns out that when systems which minimizes the BER does not necessarily minimize the input-output distortion, non-uniform mapping to constellation with Gray codes plays a dominant role in source-aware non-uniform information transmission. An unequal error protection scheme is proposed and investigated in the uncoded and coding systems. Simulation results demonstrate its effectiveness. Further, joint quantization-constellation design is considered under the criterion of distortion minimization. The proposed method generalizes the concept of constellation design from the perspective of joint source-channel coding. The simplicity and power efficiency are two major features of the proposed schemes, making them particularly attractive for systems with tight power constraints, such as the wireless sensor networks.

5.2 Future Work

We propose the following extensions to the work presented in the dissertation.

Extensions in PHY-Aware MAC Design

- In Chapter 3, we characterize the stability region and the delay performance for the simplest cases. General results for arbitrary users system deserve further investigation.
- We have derived the delays of the two-user system for given transmission probabilities. Future work can be concentrated on finding the optimal transmission probabilities which minimize the delays.
- As the proposed hybrid ALOHA slot is structured in time domain, extension to the frequency domain could lead to interesting results. The problem reduces to the medium access control design under the OFDMA framework and deserves thorough investigation.

Further Directions in Joint Source-Channel Coding Design

- In Chapter 4, performance of the proposed information transmission scheme can be further studied under more complicated channel conditions. Exploitation of space and frequency diversities can be investigated. One further research direction is the analysis of the relationship between source rate, transmit power and average distortion. The specifically interesting topic is how to distribute power among different classes of bits to minimize the average distortion, when the source rate is fixed and the average power is constrained. Information theoretic analysis is expected to be largely involved.

APPENDICES

APPENDIX A

Alternative Proof of Proposition

3.1 for $M = 2$.

In this section, we calculate the throughput of the hybrid ALOHA system for $M = 2$, based on the Bernoulli arrivals model.

According to the system model in Section 3.2, the behavior of hybrid ALOHA can be modeled as a discrete-time Markov chain. Taking the number of backlogged packets at the beginning of each hybrid ALOHA slot as the state, and the beginning of each time slot as the transition time of the states [BG92], the throughput of the system for $M = 2$ based on the Bernoulli arrivals model can be obtained as below.

Suppose there are m users. Let n be the number of backlogged users at the beginning of a given time slot. Assume the users, either backlogged or unbacklogged, transmit the training sequences independently. Let p_b be the probability that the backlogged user transmits the training sequence in either one of the $M = 2$ pilot sub-slots. Let p_u be the probability that each of the $m - n$ unbacklogged users transmits in either pilot sub-slot. Here we have assumed the transmissions in pilot sub-slots are of equal probability, i.e., $P_b = \frac{1}{2}$.

Let $P_u(n, i)$ be the probability that i unbacklogged users transmit packets in a given pilot sub-slot, and $P_b(n, i)$ be the probability that i backlogged users transmit [BG92], then according to the assumptions above, the probabilities are given by

$$P_u(n, i) = \binom{m-n}{i} p_u^i (1 - p_u)^{(m-n-i)}, \quad (\text{A.1})$$

$$P_b(n, i) = \binom{n}{i} p_b^i (1 - p_b)^{(n-i)}. \quad (\text{A.2})$$

Based on the *Simplistic Assumption*, the probability of successful transmission for one user in one of the time slots is given by

$$\begin{aligned}
P_{succ}(1) &= P_u(n, 1)P_u(n, 0)P_b(n, 0)^2 + P_u(n, 0)P_u(n, 1)P_b(n, 0)^2 \\
&\quad + P_u(n, 0)^2P_b(n, 1)P_b(n, 0) + P_u(n, 0)^2P_b(n, 0)P_b(n, 1) \\
&= 2[P_u(n, 1)P_u(n, 0)P_b(n, 0)^2 + P_u(n, 0)^2P_b(n, 1)P_b(n, 0)] \\
&= 2\left[(m - n)p_u(1 - p_u)^{2(m-n)-1}(1 - p_b)^{2n} \right. \\
&\quad \left. + n(1 - p_u)^{2(m-n)}p_b(1 - p_b)^{2n-1}\right], \tag{A.3}
\end{aligned}$$

where $P_u(n, 1)P_u(n, 0)P_b(n, 0)^2$ represents the probability that only one unbacklogged user transmits the training sequence in the first pilot sub-slot, and no one else, either unbacklogged or backlogged, transmits in the same time slot. Other terms in the above equation can be interpreted similarly.

The probability of successful transmission for exactly two users in one of the time slots can be obtained as

$$\begin{aligned}
P_{succ}(2) &= P_u(n, 1)P_u(n, 0)P_b(n, 0)P_b(n, 1) + P_u(n, 0)P_u(n, 1)P_b(n, 1)P_b(n, 0) \\
&\quad + P_u(n, 1)^2P_b(n, 0)^2 + P_u(n, 0)^2P_b(n, 1)^2 \\
&= 2P_u(n, 1)P_u(n, 0)P_b(n, 0)P_b(n, 1) + P_u(n, 1)^2P_b(n, 0)^2 \\
&\quad + P_u(n, 0)^2P_b(n, 1)^2 \\
&= 2(m - n)np_u(1 - p_u)^{2(m-n)-1}p_b(1 - p_b)^{2n-1} \\
&\quad + (m - n)^2p_u^2(1 - p_u)^{2(m-n)-2}(1 - p_b)^{2n} \\
&\quad + (1 - p_u)^{2(m-n)}n^2p_b^2(1 - p_b)^{2n-2}. \tag{A.4}
\end{aligned}$$

The average number of successfully transmitted packets per slot can then be calcu-

lated as

$$\begin{aligned}
E_{succ} &= 1 \cdot P_{succ}(1) + 2 \cdot P_{succ}(2) \\
&= (1 - p_u)^{2(m-n)-2} (1 - p_b)^{2n-2} \left[2(m-n)p_u(1 - p_u)(1 - p_b)^2 \right. \\
&\quad + 2n(1 - p_u)^2 p_b(1 - p_b) + 4(m-n)n p_u p_b(1 - p_u)(1 - p_b) \\
&\quad \left. + 2(m-n)^2 p_u^2 (1 - p_b)^2 + 2n^2 p_b^2 (1 - p_u)^2 \right]. \tag{A.5}
\end{aligned}$$

For small x , the approximation $(1 - x)^{(n-1)} \approx (1 - x)^n \approx e^{-nx}$ holds. By neglecting the higher order terms of the probabilities, E_{succ} can be approximated as

$$\begin{aligned}
E_{succ} &\simeq e^{-[2(m-n)p_u + 2np_b]} \times \\
&\quad [2(m-n)^2 p_u^2 + 2n^2 p_b^2 + 4(m-n)n p_u p_b + 2(m-n)p_u + 2np_b]. \tag{A.6}
\end{aligned}$$

Let $R(n) = 2(m-n)p_u + 2np_b$, then

$$E_{succ} \simeq e^{-R(n)} \left[\frac{1}{2} R(n)^2 + R(n) \right]. \tag{A.7}$$

where $R(n)$ can be interpreted as the expected number of attempting transmissions at state n of the Markov chain. The first term in $R(n)$ is the attempting transmission from unbacklogged users, and the second term corresponds to the attempts from backlogged ones. As can be seen, the result coincides with that of (3.9).

When $m \rightarrow \infty$ and the probabilities of the transmissions are small. The Poisson approximation applies and the result in Proposition 3.1 can be alternatively obtained as below.

Since $2p_u$ is the total probability that unbacklogged users transmit in a given time slot, assuming the packets arrival follows the Poisson process with overall arrival rate of Λ packets per unit time, it is easy to know $2p_u = 1 - e^{-\Lambda(1+\tau)/m}$, where $e^{-\Lambda(1+\tau)/m}$ is the probability that no packets arrive during the previous time slot. Under the infinite-user model, i.e., $m \rightarrow \infty$, using the approximation $e^{-x} \approx 1 - x$,

and assuming p_b is small, $R(n)$ can be approximated as a function of τ : $R(\tau) \approx \Lambda(1 + \tau) + 2np_b \approx \Lambda(1 + \tau)$. The departure rate (i.e., the expected departures per slot) is then given by

$$\nu(\tau) = E_{succ} = [R(\tau)^2/2 + R(\tau)]e^{-R(\tau)}, \quad (\text{A.8})$$

which is identical to (3.9).

APPENDIX B

Proof of Lemma 3.1.

The proof is carried out by using the stochastic dominance argument. Let S denote the original system of queues. To resolve the interactions between the queues, as in [RE88, NMT05], we construct a hypothetical *dominant* system S' , in which one of the queues continues to transmit *dummy* packets when the queue is empty. S' dominates S since the queue sizes in S' are not necessarily smaller than those in S , given that the two systems are identical in terms of arrivals and attempted transmissions.

Suppose user 1 in S' transmits dummy packets when his queue is empty, then user 2 always sees a probability of success given by $u_2 = p_2\bar{p}_1q_{2|\{2\}} + p_1p_2q_{2|\{1,2\}}$. Under the *Simplistic Assumption*, it is easy to have $q_{i|\{i\}} = 1$ and $q_{i|\{1,2\}} = 0.5$, $i = 1, 2$. Hence the probability of success for user 2 is $u_2 = p_2\bar{p}_1 + p_1p_2/2 = p_2 - p_1p_2/2$. The probability of success for user 2 in S' is always lower than that in system S , where the probability of success for user 2 oscillates between $p_2 - p_1p_2/2$ and p_2 , depending on whether queue 1 is empty or not. Suppose the arrival rate for user 2 is λ_2 . It is easy to know that user 2 operates as a discrete-time M/M/1 system and is stable if and only if

$$\lambda_2 \leq p_2 - p_1p_2/2. \quad (\text{B.1})$$

If queue 2 is stable, according to the M/M/1 formula, queue 2 is nonempty with a probability of $\lambda_2/u_2 = \lambda_2/(p_2 - p_1p_2/2)$. And it is empty with probability $1 - \lambda_2/u_2$. User 1 sees a probability of success of $p_1q_{1|1} = p_1$ when queue 2 is empty, and $p_1\bar{p}_2q_{1|1} + p_1p_2q_{1|\{1,2\}} = p_1 - p_1p_2/2$ when queue 2 is nonempty. It is shown in [RE88] that a sufficient condition for the stability of queue 1 is that, the arrival rate of user

1, λ_1 , be less than the average probability of success, which is given by

$$p_1(1 - \lambda_2/u_2) + (p_1 - p_1p_2/2)(\lambda_2/u_2) = p_1 - \frac{p_1\lambda_2}{2 - p_1}, \quad (\text{B.2})$$

i.e., user 1 is stable when

$$\lambda_1 \leq p_1 - \frac{p_1\lambda_2}{2 - p_1}. \quad (\text{B.3})$$

Therefore, the dummy packet system is stable (with the possible exception of boundary points) if and only if

$$\lambda_1 \leq p_1 - \frac{p_1\lambda_2}{2 - p_1}, \quad \lambda_2 \leq p_2 - p_1p_2/2. \quad (\text{B.4})$$

Exchanging the roles of user 1 and user 2 in the dominance system S' , i.e., user 2 transmits dummy packets when the queue is empty, the stability conditions for S' become

$$\lambda_2 \leq p_2 - \frac{p_2\lambda_1}{2 - p_2}, \quad \lambda_1 \leq p_1 - p_1p_2/2. \quad (\text{B.5})$$

By stochastic dominance, when S' is stable, S is also stable. Hence (B.4) and (B.5) are sufficient conditions for S to be stable.

On the other hand, these are also necessary conditions. The proof follows the argument that as long as the queues in the systems are not empty, the dominant system S' and the original system S are “indistinguishable” [RE88]. Therefore, when the conditions to maintain a stable S' are violated, system S also becomes unstable. Consequently, the conditions given in Lemma 3.1 are then necessary for stability as well.

APPENDIX C

Proof of Proposition 3.2.

From Lemma 3.1, the stability region for a fixed transmission probability vector $\mathbf{p}$ is known. To find the union of all the stability regions as $\mathbf{p}$ varies over $[0, 1]^2$, a constrained optimization problem is formed as below. Under the constraints (B.4) and (B.5), we fix one rate, say λ_1 , and maximize the other rate λ_2 as $\mathbf{p}$ varies over $[0, 1]^2$. Hence a stability region $\mathfrak{S}_1$ is obtained. Exchanging the roles of λ_1 and λ_2 , another stability region $\mathfrak{S}_2$ can be obtained. $\mathfrak{S}_1 \cap \mathfrak{S}_2$ produces $\mathfrak{S}_{H_ALOHA}$.

Write the boundary given by (B.4) and (B.5) as

$$\lambda_1 = p_1 - \frac{p_1 \lambda_2}{2 - p_1}, \quad 0 \leq \lambda_2 \leq p_2 - p_1 p_2 / 2. \quad (\text{C.1})$$

$$\lambda_2 = p_2 - \frac{p_2 \lambda_1}{2 - p_2}, \quad 0 \leq \lambda_1 \leq p_1 - p_1 p_2 / 2. \quad (\text{C.2})$$

In (C.2), fixing λ_1 and differentiating λ_2 with respect to p_2 , we have

$$\frac{d\lambda_2}{dp_2} = 1 - \frac{2\lambda_1}{(2 - p_2)^2}. \quad (\text{C.3})$$

Since $0 \leq \lambda_1 \leq p_1(1 - p_2/2)$, then $0 \leq 2\lambda_1/(2 - p_2)^2 \leq p_1/(2 - p_2)$. Under the constraint that $0 \leq p_1, p_2 \leq 1$, it is easy to know $\frac{d\lambda_2}{dp_2} \geq 0$. Hence λ_2 is a monotonic non-decreasing function with respect to p_2 . Since in (C.2), p_2 is maximized at $p_2 = 2(1 - \lambda_1)$ for $0 \leq \lambda_1 \leq 1/2$, and at 1 for $1/2 \leq \lambda_1 \leq 1$, we then have

$$\lambda_{2,max} = 1 - \lambda_1, \quad \text{for } 0 \leq \lambda_1 \leq 1. \quad (\text{C.4})$$

Similarly, the maximization for (C.1) gives

$$\lambda_{1,max} = 1 - \lambda_2, \text{ for } 0 \leq \lambda_2 \leq 1. \quad (C.5)$$

(C.4) and (C.5) represent two overlapping regions, and the stability region for the proposed protocol is then given by

$$\begin{aligned} \mathfrak{S}_{H-ALOHA} = \{(\lambda_1, \lambda_2) : (\lambda_1, \lambda_2) \geq (0, 0), (\lambda_1, \lambda_2) \text{ lies below the line} \\ \lambda_1 + \lambda_2 = 1, 0 \leq \lambda_1 \leq 1\}. \end{aligned} \quad (C.6)$$

APPENDIX D

Proof of Proposition 3.4.

Mathematical induction on $N = |\mathcal{M}|$ is applied to prove the claim. Assuming that one transmits the training sequence in either pilot sub-slot with equal transmission probability, we can easily see that the region given by (3.24) is the stability region for $N = 2$ characterized in Section 3.4. Suppose $\mathcal{M} = \{1, 2\}$ and $\mathcal{P} = \{\mathcal{S}, \mathcal{U}\} = \{\{2\}, \{1\}\}$. If $p_{i,1} = p_{i,2} = \frac{1}{2}p_i$, then from (3.23) we have

$$P_{succ}^2(\mathcal{P}) = p_2(1 - p_1) + (p_{2,1}p_{1,2} + p_{2,2}p_{1,1}) = p_2 - p_1p_2/2, \quad (\text{D.1})$$

and

$$\begin{aligned} P_{succ}^1(\mathcal{P}) &= p_1[Pr\{\chi(\overline{N}_2^0) = 0\} + (1 - p_2)Pr\{\chi(\overline{N}_2^0) = 1\}] \\ &\quad + [p_{1,1}Pr\{\chi(\overline{N}_2^0) = 1\}p_{2,2} + p_{1,2}Pr\{\chi(\overline{N}_2^0) = 1\}p_{2,1}]. \end{aligned} \quad (\text{D.2})$$

Let $u_2 = p_2 - p_1p_2/2$, then for $\lambda_2 < P_{succ}^2(\mathcal{P})$, the probability $Pr\{\overline{N}_2^0 \geq 1\} = \lambda_2/u_2$. We have

$$\begin{aligned} P_{succ}^1(\mathcal{P}) &= p_1[(1 - \frac{\lambda_2}{u_2} + (1 - p_2)\frac{\lambda_2}{u_2}] + \frac{1}{2}p_1p_2\frac{\lambda_2}{u_2} \\ &= p_1 - \frac{p_1\lambda_2}{2 - p_1}. \end{aligned} \quad (\text{D.3})$$

Hence we obtain the stability condition identical to (3.11). Similarly, the condition of (3.12) can be obtained by performing the partition as $\mathcal{P} = \{\mathcal{S}, \mathcal{U}\} = \{\{1\}, \{2\}\}$. The union of the regions under these two partitions characterizes the stability region given in Lemma 3.1. And this proves Proposition 3.4 in the case of $N = 2$.

Now apply induction on N and assume that the claim is true for hybrid ALOHA systems with number of users less than N . We consider the dominant system $\Theta^{\mathcal{P}}$ under the partition $\mathcal{P} = \{\mathcal{S}, \mathcal{U}\}$, and concentrate on proving the stability condition for the system represented by $\bar{\mathbf{N}}_{\mathcal{P}}^t = (\bar{\mathbf{N}}_{\mathcal{S}}^t, \bar{\mathbf{N}}_{\mathcal{U}}^t)$. It is easy to know from the induction arguments that $\bar{\mathbf{N}}_{\mathcal{S}}^t$ is stable for $\lambda_{\mathcal{S}} \in \mathfrak{S}_{\mathcal{S}}$, where $\lambda_{\mathcal{S}} = (\lambda_{i_1}, \dots, \lambda_{i_{|\mathcal{S}|}})$ with $i_k \in \mathcal{S}$ for $1 \leq k \leq |\mathcal{S}|$. Note that in the dominant system $\Theta^{\mathcal{P}}$, the set of users restricted to $\mathcal{S}$ are “decoupled” from those in $\mathcal{U}$, and they appear to be a smaller copy of the original system with modified reception probabilities given by (3.21). Since $|\mathcal{S}| < N$, then by the induction hypothesis, the queues in $\mathcal{S}$ are stable if $\lambda_{\mathcal{S}} \in \mathfrak{S}_{\mathcal{S}}$.

Let $\lambda_{\mathcal{S}} \in \mathfrak{S}_{\mathcal{S}}$, then we know $\bar{\mathbf{N}}_{\mathcal{S}}^t$ is stationary and ergodic. Since the coin tosses and packet reception events are independent from slot to slot, it follows from (3.22) that if the process $\bar{\mathbf{N}}_{\mathcal{S}}^t$ is initialized from its unique stationary distribution, then for any $i \in \mathcal{U}$, the process $Y_i^t(\mathcal{P})$ is also stationary and ergodic. By the Loynes Theorem, it is shown that the persistent queue i is stable in $\Theta^{\mathcal{P}}$ when $\lambda_i < E[Y_i^t(\mathcal{P})] = P_{succ}^i(\mathcal{P})$, where $P_{succ}^i(\mathcal{P})$ is given by (3.23).

Therefore, if $\lambda_{\mathcal{S}} \in \mathfrak{S}_{\mathcal{S}}$ and $\lambda_i < P_{succ}^i(\mathcal{P})$ for all $i \in \mathcal{U}$, each queue in the dominant system $\Theta^{\mathcal{P}}$ is stable. This implies the stability of the process $\bar{\mathbf{N}}_{\mathcal{P}}^t$ from Lemma 5 in [Szp94]. By virtue of the stochastic dominance, the original system is then also stable. Since the partition $\mathcal{P}$ is arbitrary, the final condition for stability is the union over all the partitions.

APPENDIX E

Formulation of Functional Equation (3.27)

Let $D_{10}(t)$ be a binary-valued random variable that takes value 1 if $N_1^t > 0, N_2^t = 0$ and the departure from queue 1 is successful. Similarly, $D_{01}(t)$ is a binary-valued random variable that takes value 1 if $N_2^t > 0, N_1^t = 0$ and the departure from queue 2 is successful. In the situation when both queues are nonempty, the binary-valued variables $D_{11}^i(t)$ for $i = 1, 2$ take value 1 when departure from queue i is successful. The recursive equations for N_i^t are given as

$$N_1^{t+1} = \begin{cases} \beta_1^t, & N_1^t = 0, N_2^t = 0; \\ \beta_1^t + N_1^t - D_{10}(t), & N_1^t > 0, N_2^t = 0; \\ \beta_1^t, & N_1^t = 0, N_2^t > 0; \\ \beta_1^t + N_1^t - D_{11}^1(t), & N_1^t > 0, N_2^t > 0. \end{cases} \quad (\text{E.1})$$

and

$$N_2^{t+1} = \begin{cases} \beta_2^t, & N_1^t = 0, N_2^t = 0; \\ \beta_2^t, & N_1^t > 0, N_2^t = 0; \\ \beta_2^t + N_2^t - D_{10}(t), & N_1^t = 0, N_2^t > 0; \\ \beta_2^t + N_2^t - D_{11}^2(t), & N_1^t > 0, N_2^t > 0. \end{cases} \quad (\text{E.2})$$

If a persistent user 3 exists, i.e., we investigate the partition $\mathcal{P}_3$, then from (E.1) and (E.2) we have

$$\begin{aligned}
E(x^{N_1^{t+1}} y^{N_2^{t+1}}) &= E(x^{\beta_1^t} y^{\beta_2^t}) \{ E(1[N_1^t = 0, N_2^t = 0]) + \\
&\quad + E(x^{N_1^t} 1[N_1^t > 0, N_2^t = 0]) \cdot [\frac{\Delta_1}{x} + 1 - \Delta_1] + \\
&\quad + E(y^{N_2^t} 1[N_1^t = 0, N_2^t > 0]) \cdot [\frac{\Delta_2}{y} + 1 - \Delta_2] + \\
&\quad + E(x^{N_1^t} y^{N_2^t} 1[N_1^t > 0, N_2^t > 0]) \\
&\quad \times [\frac{\Delta_3}{x} + \frac{\Delta_4}{y} + \frac{\Delta_5}{xy} + 1 - \Delta_3 - \Delta_4 - \Delta_5] \}, \quad (E.3)
\end{aligned}$$

where $1[\cdot]$ is the indicator function and

$$\begin{aligned}
\Delta_1 &= p_1 \left(\bar{p}_3 q_{\{1\}, \{1\}} + p_3 q_{\{1\} | \{1,3\}} \right); \\
\Delta_2 &= p_2 \left(\bar{p}_3 q_{\{2\}, \{2\}} + p_3 q_{\{2\} | \{2,3\}} \right); \\
\Delta_3 &= p_1 \left(\bar{p}_2 \bar{p}_3 q_{\{1\}, \{1\}} + p_2 \bar{p}_3 q_{\{1\}, \{1,2\}} + \bar{p}_2 p_3 q_{\{1\} | \{1,3\}} \right); \\
\Delta_4 &= p_2 \left(\bar{p}_1 \bar{p}_3 q_{\{2\}, \{2\}} + p_1 \bar{p}_3 q_{\{2\}, \{1,2\}} + \bar{p}_1 p_3 q_{\{2\} | \{2,3\}} \right); \\
\Delta_5 &= p_1 p_2 \bar{p}_3 q_{\{1,2\}, \{1,2\}}.
\end{aligned}$$

Note that

$$G_3(0, 0) = \lim_{t \rightarrow \infty} E(1[N_1^t = 0, N_2^t = 0]), \quad (E.4)$$

$$G_3(x, 0) - G_3(0, 0) = \lim_{t \rightarrow \infty} E(x^{N_1^t} 1[N_1^t > 0, N_2^t = 0]), \quad (E.5)$$

$$G_3(0, y) - G_3(0, 0) = \lim_{t \rightarrow \infty} E(y^{N_2^t} 1[N_1^t = 0, N_2^t > 0]), \quad (E.6)$$

$$G_3(x, y) + G_3(0, 0) - G_3(x, 0) - G_3(0, y) = \lim_{t \rightarrow \infty} E(x^{N_1^t} y^{N_2^t} 1[N_1^t > 0, N_2^t > 0]). \quad (E.7)$$

Assuming the Simplistic Assumption, from (E.3), it follows that

$$K(x, y) G_3(x, y) = a(x, y) G_3(0, y) + b(x, y) G_3(x, 0) + c(x, y) G_3(0, 0), \quad (E.8)$$

where $K(x, y)$, $a(x, y)$, $b(x, y)$ and $c(x, y)$ are given in (3.28).

APPENDIX F

Proof of Lemma 3.2

When $y \rightarrow 0^+$,

$$\begin{aligned} A(0^+) &\sim [\lambda_1 \bar{p}_1 p_2 (1 - p_3/2)](\lambda_2 + \frac{\bar{\lambda}_2}{y}); \\ B(0^+) &\sim [-\lambda_1 \bar{p}_1 p_2 (1 - p_3/2) - \lambda_1 \frac{p_1 p_2 \bar{p}_3}{2}](\lambda_2 + \frac{\bar{\lambda}_2}{y}); \\ C(0^+) &\sim [\bar{\lambda}_1 \frac{p_1 p_2 \bar{p}_3}{2}](\lambda_2 + \frac{\bar{\lambda}_2}{y}). \end{aligned} \tag{F.1}$$

Hence

$$\begin{aligned} t(0^+, \phi) &= -B(0^+) - 2 \cos \phi \sqrt{A(0^+)C(0^+)} \\ &< -\{[\lambda_1 \bar{p}_1 p_2 (1 - p_3/2) + \lambda_1 \frac{p_1 p_2 \bar{p}_3}{2}] \\ &\quad + 2 \sqrt{[\lambda_1 \bar{p}_1 p_2 (1 - p_3/2)] \cdot [\lambda_1 \frac{p_1 p_2 \bar{p}_3}{2}]}\}(\lambda_2 + \frac{\bar{\lambda}_2}{y}) \\ &= -\infty, \end{aligned} \tag{F.2}$$

that is, $t(0^+, \phi) = -\infty$.

When $y = 1$, then

$$\begin{aligned} A(1) &= \lambda_1 [1 - p_1 \bar{p}_2 (1 - p_3/2) - \frac{p_1 p_2 \bar{p}_3}{2}], \\ B(1) &= (2\lambda_1 - 1)[p_1 \bar{p}_2 (1 - p_3/2) + \frac{p_1 p_2 \bar{p}_3}{2}] - \lambda_1, \\ C(1) &= \bar{\lambda}_1 [p_1 \bar{p}_2 (1 - p_3/2) + \frac{p_1 p_2 \bar{p}_3}{2}]. \end{aligned} \tag{F.3}$$

It can be easily seen that $-B(1) = A(1) + C(1)$, hence $t(1, \phi) = -B(1) -$

$$2 \cos \phi \sqrt{A(1)C(1)} > 0.$$

As $y \rightarrow \infty$,

$$\begin{aligned} A(y) &\sim \lambda_1 \lambda_2 \Delta y, \\ B(y) &\sim [\lambda_1 p_1 \bar{p}_2 (1 - p_3/2) + \bar{\lambda}_1 \Delta] \lambda_2 y, \\ C(y) &\sim \bar{\lambda}_1 \lambda_2 p_1 \bar{p}_2 (1 - p_3/2) y, \end{aligned} \tag{F.4}$$

then

$$\begin{aligned} t(y, \phi) &< -B(y) + 2\sqrt{A(y)C(y)} \\ &= -y\lambda_2 \left\{ \lambda_1 p_1 \bar{p}_2 (1 - p_3/2) + \bar{\lambda}_1 \Delta - 2\sqrt{(\bar{\lambda}_1 \Delta)[\lambda_1 p_1 \bar{p}_2 (1 - p_3/2)]} \right\} \\ &= -y\lambda_2 \left(\sqrt{\lambda_1 p_1 \bar{p}_2 (1 - p_3/2)} - \sqrt{\bar{\lambda}_1 \Delta} \right)^2 \end{aligned} \tag{F.5}$$

Hence, $t(\infty, \phi) < 0$.

Consequently, $t(y, \phi) = 0$ has at least two real roots $r_1(\phi)$ and $r_2(\phi)$ satisfying $0 < r_1(\phi) < 1 < r_2(\phi)$. Since $y^2 t(y, \phi) t(y, \phi + \pi)$ is a polynomial of degree four in the variable y , it can be deduced that $t(y, \phi)$ has exactly two real roots, and this completes the proof.

APPENDIX G

Proof of Lemma 3.3

The first part of the lemma results from the general theory of polynomials of two complex variables.

The second assertion is proved by using Rouché's theorem as below:

$K(x, y)$ can be rewritten as

$$K(x, y) = \frac{xy - xyF(x, y)g(x, y)}{xyF(x, y)}, \quad (\text{G.1})$$

where

$$g(x, y) = \frac{p_1\bar{p}_2(1 - p_3/2)}{x} + \frac{p_2\bar{p}_1(1 - p_3/2)}{y} + \frac{p_1p_2\bar{p}_3}{2xy} + \Delta. \quad (\text{G.2})$$

For $|y| = 1$ and $y \neq 1$, $|x| = 1$,

$$\begin{aligned} |xyF(x, y)g(x, y)| &= |(x\lambda_1 + \bar{\lambda}_1)(y\lambda_2 + \bar{\lambda}_2)[p_1\bar{p}_2(1 - p_3/2)y \\ &\quad + p_2\bar{p}_1(1 - p_3/2)x + \frac{p_1p_2\bar{p}_3}{2} + \Delta xy]| \\ &\leq 1 \\ &= |xy|. \end{aligned} \quad (\text{G.3})$$

Based on Rouché's theorem, this implies that for $|y| = 1$, $y \neq 1$, there exists exactly one x , $|x| < 1$ such that $xy - xyF(x, y)g(x, y) = 0$ and hence $K(x, y) = 0$.

For $y = 1$, $K(x, 1) = 0$ reduces to

$$(x - 1)\left(x - \frac{\bar{\lambda}_1[p_1\bar{p}_2(1 - p_3/2) + p_1p - 2\bar{p}_3/2]}{\lambda[1 - p_1\bar{p}_2(1 - p_3/2) - p_1p - 2\bar{p}_3/2]}\right) = 0. \quad (\text{G.4})$$

When $\lambda_1 < p_1\bar{p}_2(1 - p_3/2) + p_1p_2\bar{p}_3/2$ as implied by (3.34), $x = 1$ is the only root. $\square$

APPENDIX H

Proof of Proposition 3.7.

Let N_i^t denote the queue length of user i ($i = 1, 2$) at time slot t . Let β_i^t denote the number of packets that arrive at the i th user's queue in time slot t . Let $F(x, y)$ be the moment generating function of the joint arrival process. Thus, for $|x| \leq 1$, $|y| \leq 1$, $t \in \mathbb{N}$

$$F(x, y) = E(x^{\beta_1^t} y^{\beta_2^t}) = (x\lambda + \bar{\lambda})(y\lambda + \bar{\lambda}), \quad (\text{H.1})$$

where $\bar{\lambda} = 1 - \lambda$. From the queue evolution equation (3.5), it can be seen that

$$\begin{aligned} E(x^{N_1^{t+1}} y^{N_2^{t+1}}) &= F(x, y)[E(\mathbf{1}[N_1^t = 0, N_2^t = 0]) \\ &\quad + (\frac{pa}{x} + 1 - pa)E(x^{N_1^t} \mathbf{1}[N_1^t > 0, N_2^t = 0]) \\ &\quad + (\frac{pa}{y} + 1 - pa)E(y^{N_2^t} \mathbf{1}[N_1^t = 0, N_2^t > 0]) \\ &\quad + ((pa + p^2(b - a))(\frac{1}{x} + \frac{1}{y}) + \frac{p^2c}{xy} + 1 - 2(pa + p^2(b - a)) - p^2c)) \\ &\quad \times E(x^{N_1^t} y^{N_2^t} \mathbf{1}[N_1^t > 0, N_2^t > 0])], \end{aligned} \quad (\text{H.2})$$

where $\mathbf{1}[\cdot]$ is the indicator function. According to Lemma 1 in [NMT05], if $\lambda < pa + p^2(b + c - a)$, then the system is stable. Since (N_1^t, N_2^t) is an irreducible, aperiodic Markov chain, stability is equivalent to existence of a unique stationary (limiting) distribution. Let $G(x, y)$ be the moment generating function of the joint stationary queue process, *viz.*,

$$G(x, y) = \lim_{t \rightarrow \infty} E(x^{N_1^t} y^{N_2^t}). \quad (\text{H.3})$$

Note that

$$G(0, 0) = \lim_{t \rightarrow \infty} E(\mathbf{1}[N_1^t = 0, N_2^t = 0]), \quad (\text{H.4})$$

$$G(x, 0) - G(0, 0) = \lim_{t \rightarrow \infty} E(x^{N_1^t} \mathbf{1}[N_1^t > 0, N_2^t = 0]), \quad (\text{H.5})$$

$$G(0, y) - G(0, 0) = \lim_{t \rightarrow \infty} E(y^{N_2^t} \mathbf{1}[N_1^t = 0, N_2^t > 0]), \quad (\text{H.6})$$

$$G(x, y) + G(0, 0) - G(x, 0) - G(0, y) = \lim_{t \rightarrow \infty} E(x^{N_1^t} y^{N_2^t} \mathbf{1}[N_1^t > 0, N_2^t > 0]). \quad (\text{H.7})$$

From (H.2), it follows that

$$\begin{aligned} G(x, y) = & F(x, y) \{ C(x, y) [G(x, y) + G(0, 0) - G(x, 0) - G(0, y)] \\ & + B(x, y) [G(0, y) - G(0, 0)] + A(x, y) [G(x, 0) - G(0, 0)] + G(0, 0) \} \end{aligned} \quad (\text{H.8})$$

where

$$\begin{aligned} A(x, y) &= \frac{pa}{x} + 1 - pa, \\ B(x, y) &= \frac{pa}{y} + 1 - pa, \\ C(x, y) &= [pa + p^2(b - a)] \left(\frac{1}{x} + \frac{1}{y} \right) + \frac{p^2c}{xy} + 1 - 2[pa + p^2(b - a)] - p^2c. \end{aligned} \quad (\text{H.9})$$

Using the symmetry property of $G(x, y)$, we have $G(1, 1) = 1$ and $G(1, 0) = G(0, 1)$. In (H.8), let $y = 1$ and take derivative with respect to x at both sides and then let $x = 1$, we find

$$(2\theta - pa)G(1, 0) - (\theta - pa)G(0, 0) = \theta - \lambda, \quad (\text{H.10})$$

where $\theta = pa + p^2(b + c - a)$. Furthermore, (H.8) can be rewritten as

$$\begin{aligned} G(x, y) = & \frac{F(x, y)}{1 - F(x, y)C(x, y)} \times \{ C(x, y) [G(0, 0) - G(x, 0) - G(0, y)] + \\ & B(x, y) [G(0, y) - G(0, 0)] + A(x, y) [G(x, 0) - G(0, 0)] + G(0, 0) \} \end{aligned} \quad (\text{H.11})$$

Let $G_1(x, y) \equiv dG(x, y)/dx$. In (H.11), letting $y = 1$, taking the derivative with respect to x and then letting $x = 1$, we obtain

$$G_1(1, 1) = \frac{p^2(b + c - a)G_1(1, 0) + \lambda\bar{\lambda}}{\theta - \lambda}. \quad (\text{H.12})$$

Based on (H.11), a more tedious calculation will lead to the following result.

$$\begin{aligned} \frac{G(x, x)}{dx} \Big|_{x=1} = & 2\lambda + \frac{(2\theta - pa)G_1(1, 0)}{\theta - \lambda} + \frac{\lambda^2 + 2\lambda - 4\lambda\theta + p^2c}{2(\theta - \lambda)} \\ & - \frac{p^2c[G(1, 0) - \frac{1}{2}G(0, 0)]}{\theta - \lambda}. \end{aligned} \quad (\text{H.13})$$

If we use the fact that

$$\frac{G(x, x)}{dx} \Big|_{x=1} = G_1(1, 1) + G_2(1, 1) = 2G_1(1, 1), \quad (\text{H.14})$$

and associate with the result from (H.12), we can solve for $G_1(1, 0)$ and $G_1(1, 1)$ to obtain

$$G_1(1, 1) = \frac{2\lambda\theta - \lambda^2pa - \lambda^2\theta}{2pa(\theta - \lambda)} - \frac{p^3c(b + c - a)[G(1, 1) + G(0, 0) - G(1, 0) - G(0, 1)]}{2a(\theta - \lambda)}. \quad (\text{H.15})$$

Since $G_1(1, 1)$ is equal to the mean queue length of the users, applying Little's Theorem [Lit61], we have the average delay D as

$$D = \frac{G_1(1, 1)}{\lambda} = \frac{1}{a} \left[\frac{a(1 - \lambda) + p(b + c - a)(1 - \lambda/2)}{pa + p^2(b + c - a) - \lambda} \right] + \phi, \quad (\text{H.16})$$

where

$$\phi = - \frac{p^3c(b + c - a)[G(1, 1) + G(0, 0) - G(1, 0) - G(0, 1)]}{2ar(pa + p^2(b + c - a) - \lambda)}. \quad (\text{H.17})$$

Since $G(1, 1) + G(0, 0) - G(1, 0) - G(0, 1) = \lim_{t \rightarrow \infty} E(1[N_1^t > 0, N_2^t > 0])$ (as (H.7) implies), associated with $b + c - a < 0$ and $\lambda < pa + p^2(b + c - a)$, we then conclude that

$$0 \leq \phi \leq \frac{-p^3c(b + c - a)}{2ar[pa + p^2(b + c - a) - \lambda]}. \quad (\text{H.18})$$

Therefore, the average delay is lower bounded by

$$D_{\text{Lower}} = \frac{1}{a} \left[\frac{a(1 - \lambda) + p(b + c - a)(1 - \lambda/2)}{pa + p^2(b + c - a) - \lambda} \right], \quad (\text{H.19})$$

and upper bounded by

$$D_{Upper} = \frac{1}{a} \left[\frac{a(1-\lambda) + p(b+c-a)(1-\lambda/2)}{pa + p^2(b+c-a) - \lambda} \right] - \frac{p^3 c(b+c-a)}{2ar[pa + p^2(b+c-a) - \lambda]}. \quad (\text{H.20})$$

From the analysis above, we see that the bounds become the exact delay of the system when $c = 0$ or $\lim_{t \rightarrow \infty} E(1[N_1^t > 0, N_2^t > 0]) = 0$.

BIBLIOGRAPHY

BIBLIOGRAPHY

- [Abr70] N. Abramson. The aloha system - another alternative for computer communications. In *Proc. Fall Joint Comput. Conf., AFIPS Conf*, volume 44, pages 281-285, Montvale, NJ, 1970.
- [AD01] D. E. Ayyildiz and H. Delic. Adaptive random-access algorithm with improved delay performance. *Int. J. Commun. Syst.*, 14:531-539, Jun 2001.
- [AGR98] P. D. Alexander, A. J. Grant, and M. C. Reed. Iterative detection in code division multiple access with error control coding. *European Trans. Telecommun.*, 9:419-425, Sept/Oct 1998.
- [AM79] B. D. O. Anderson and J. R. Moore, editors. *Optimal filtering*. Prentice Hall, NJ, 1979.
- [Ana91] V. Anantharam. Stability region of the finite user slotted ALOHA protocol. *IEEE Trans. Inform. Theory*, 37:535-540, May 1991.
- [AT05] S. Adireddy and L. Tong. Exploiting decentralized channel state information for random access. *IEEE Trans. Inform. Theory*, 51:537-561, Feb 2005.
- [BG92] D. Bertsekas and R. Gallager. *Data Networks*. Prentice Hall, Upper Saddle River, NJ, 2nd edition, 1992.
- [Cap79] J. I. Capetanakis. Tree algorithms for packet broadcast channels. *IEEE Trans. Inform. Theory*, 25:505-515, Sept 1979.
- [CB83] J. W. Cohen and O. J. Boxma. *Boundary value problems in queueing system analysis*. North-Holland mathematics studies, 79. North-Holland, New York, Oxford, 1983.
- [CEPB02] S. Coleri, M. Ergen, A. Puri, and A. Bahai. Channel estimation techniques based on pilot arrangement in OFDM systems. *IEEE Trans. Broadcasting*, 48:223-229, 2002.
- [Cho00] J. Choi. Channel estimation for coherent multi-carrier CDMA systems over fast fading channels. In *Proc. IEEE Vehicular Technology Conference (VTC)*, volume 1, pages 400-404, Tokyo, Japan, May 2000.

- [CMS01] J. Cai, J. W. Mark, and X. Shen. Fast fading channel estimation in multicarrier-CDMA systems. In *Proc. IEEE Global Telecommunications Conference (GLOBECOM)*, volume 5, pages 3135–3139, November 2001.
- [Cov72] T. M. Cover. Broadcast channels. *IEEE Trans. Inform. Theory*, IT-18:2–14, Jan 1972.
- [CY02] K. Cho and D. Yoon. On the general BER expression of one- and two-dimensional amplitude modulations. *IEEE Trans. Commun.*, 50:1074–1080, Jul 2002.
- [EH98] A. Ephremides and B. Hajek. Information theory and communication networks: An unconsummated union. *IEEE Trans. Inform. Theory*, 44:2416–2434, Oct 1998.
- [EZ87] A. Ephremides and R. Z. Zhu. Delay analysis of interacting queues with an approximate model. *IEEE Trans. Commun.*, COM-35:194–201, Feb 1987.
- [FC91] M. Feder and J.A. Catipovic. Algorithms for joint channel estimation and data recovery - application to equalization in underwater communications. *IEEE Journal of Oceanic Engineering*, 16:42–55, January 1991.
- [FI79] G. Fayolle and R. Iasnagorodski. Two coupled processes: The reduction to a riemann-hilbert problem. In *Z. Wahrscheinlichkeitstheorie*, volume 47, pages 325–351, 1979.
- [Fis63] M. Fisz. *Probability theory and mathematical statistics*. Wiley publications in statistics. New York, Wiley, 3rd edition, 1963.
- [FV87] N. Farvardin and V. Vaishampayan. Optimal quantizer design for noisy channels: an approach to combined source-channel coding. *IEEE Trans. Inform. Theory*, 35:827–838, Nov 1987.
- [FW89] G. D. Forney and Lee-Fang Wei. Multidimensional constellations - Part I: Introduction, figures of merit, and generalized cross constellations. *IEEE Trans. Inform. Theory*, 7:877–892, Aug 1989.
- [Gal85] R. Gallager. A perspective on multiaccess channels. *IEEE Trans. Inform. Theory*, IT-31:124–142, Mar 1985.
- [GL03] D. Gu and C. Leung. Performance analysis of transmit diversity scheme with imperfect channel estimation. *Elect. Lett.*, 39:420–403, Feb 2003.

- [GT02] M. Grossglauser and D. Tse. Mobility increases the capacity of wireless adhoc networks. *IEEE/ACM Trans. networking*, 10:477–486, Aug 2002.
- [GVS88] S. Ghez, S. Verdú, and S. Schwartz. Stability properties of slotted ALOHA with multipacket reception capability. *IEEE Trans. Automat. Cont.*, 33:640–649, Jul 1988.
- [GVS89] S. Ghez, S. Verdú, and S. Schwartz. Optimal decentralized control in the random access multipacket channel. *IEEE Trans. Automat. Cont.*, 34:1153–1163, Nov 1989.
- [HG81] M. G. Hluchyj and R. G. Gallager. Multiaccess of a slotted channel by finitely many users. In *Proc. Nat. Telecommunications Conf.*, pages D4.2.1–D4.2.7, New Orleans, LA, Aug 1981.
- [HKR97] P. Hoeher, S. Kaiser, and P. Robertson. Two-dimensional pilot-symbol-aided channel estimation by Wiener filtering. In *Proc. IEEE Int. Conf. on Acustics, Speech and Signal Processing (ICASSP)*, volume 3, pages 1845–1848, April 1997.
- [HKS89] I. M. I. Habbab, M. Kavehrad, and C. E. W. Sundberg. ALOHA with capture over slow and fast fading radio channels with coding and diversity. *IEEE J. Select. Areas Commun.*, 7:79–88, Jan 1989.
- [Hlu82] M. G. Hluchyj. Multiple access window protocol: Analysis for large finite populations. In *Proc. IEEE Conf. Decision and Control*, pages 589–595, New York, 1982.
- [HP97] S. Hara and R. Prasad. Overview of Multicarrier CDMA. *IEEE Commun. Mag.*, 35:126–133, December 1997.
- [IF91] R. A. Iltis and A. W. Fuxjaeger. A digital DS spread-spectrum receiver with joint channel and Doppler shift estimation. *IEEE Trans. Commun.*, 39:1255–1267, August 1991.
- [Ilt90] R. A. Iltis. Joint estimation of PN code delay and multipath using the extended Kalman filter. *IEEE Trans. Commun.*, 38:1677–1685, October 1990.
- [IM96] R. A. Iltis and L. Mailaender. Multiuser detection of quasi-synchronous CDMA signals using linear decorrelators. *IEEE Trans. Commun.*, 44:1561–1571, November 1996.
- [Jak74] W. C. Jakes, editor. *Microwave Mobile Communications*. IEEE Piscataway, New York, 1974.

- [KH95] R. Knopp and P. Humblet. Information capacity and power control in single cell multi-user communications. In *Proc. Intl. Conf. Comm.*, pages 331–335, Seattle, WA, Jun 1995.
- [KH97] S. Kaiser and P. Hoeher. Performance of multicarrier CDMA with channel estimation in two dimensions. In *Proc. IEEE int. Symp. on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pages 115–119, 1997.
- [KI02] K. Kim and R. A. Iltis. Joint detection and channel estimation algorithms for QS-CDMA signals over time-varying channels. *IEEE Trans. Commun.*, 50:845–855, May 2002.
- [KSP03] D. N. Kalofonos, M. Stojanovic, and J. Proakis. Performance of adaptive MC-CDMA detectors in rapidly fading Rayleigh channels. *IEEE Trans. Wireless Commun.*, 2:229–239, March 2003.
- [KV94] G. K. Kaleh and R. Vallet. Joint parameter estimation and symbol detection for linear or nonlinear unknown channels. *IEEE Trans. Commun.*, 42:2406–2413, 1994.
- [KY78] L. Kleinrock and Y. Yemini. An optimal adaptive scheme for multiple access broadcast communications. In *Proc. Int. Conf. Communications*, pages 7.2.1–7.2.5, Jun 1978.
- [LE99] W. Luo and A. Ephremides. Stability of N interacting queues in random-access systems. *IEEE Trans. Inform. Theory*, 45:1579–1587, Jul 1999.
- [Lit61] J. D. C. Little. A proof for the queueing formula $L = \lambda W$. *Op. Res.* 9, pages 383–387, 1961.
- [LJS98] Y. Li, L. J. Cimini, Jr., and N. R. Sollenberger. Robust channel estimation for OFDM systems with rapid dispersive fading channels. *IEEE Trans. Commun.*, 46:902–915, July 1998.
- [LLCL99] J. Lu, K. B. Letaief, J. C-I Chuang, and M. L. Liou. M-PSK and M-QAM BER computation using signal-space concepts. *IEEE Trans. Commun.*, 47:181–184, Feb 1999.
- [Loy62] R. M. Loynes. The stability of a queue with non-independent inter-arrival and service times. *Proc. Cambridge Philos. Soc.*, 58:497–520, 1962.
- [LPK89] D. F. Lyons and P. Papantoni-Kazakos. A window random access algorithm for environments with capture. *IEEE Trans. Commun.*, 37:766–770, Jul 1989.

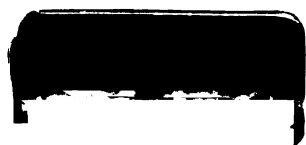
- [Mes79] D. G. Messerschmitt. Accumulation of distortion in tandem communication links. *IEEE Trans. Inform. Theory*, IT-25:692–698, Nov 1979.
- [Met76] J. L. Metzner. On improving utilization in ALOHA networks. *IEEE Trans. Commun.*, COM-24:447–448, Apr 1976.
- [Moh98] M. Moher. An iterative multiuser decoder for near-capacity communications. *IEEE Trans. Commun.*, 46:870–880, July 1998.
- [MT05] G. Mergen and L. Tong. Stability and capacity of regular wireless networks. *IEEE Trans. Inform. Theory*, 51:1938–1953, Jun 2005.
- [Mus53] N. I. Mushkelishvili. *Singular integral equation*. Noordhoff, Groningen, 1953.
- [MW67] B. Masnick and J. Wolf. On linear unequal error protection codes. *IEEE Transactions on Information Theory*, 13:600–607, October 1967.
- [MZFLI00a] R. H. Morelos-Zaragoza, M. P. C. Fossorier, S. Lin, and H. Imai. Multilevel coded modulation for unequal error protection and multistage decoding .i. symmetric constellations. *IEEE Trans. Commun.*, 48:204–213, Feb 2000.
- [MZFLI00b] R. H. Morelos-Zaragoza, M. P. C. Fossorier, S. Lin, and H. Imai. Multilevel coded modulation for unequal error protection and multistage decoding. ii. asymmetric constellations. *IEEE Trans. Commun.*, 48:774–786, May 2000.
- [Nai85] P. Nain. Analysis of a two node ALOHA network with infinite capacity buffers. In *Proc. Int. Sem. Computer Netwkg.*, pages 2.2.1–2.2.16, Tokyo, Japan, 1985.
- [NMT05] V. Naware, G. Mergen, and L. Tong. Stability and delay of finite user slotted ALOHA with multipacket reception. *IEEE Trans. Inform. Theory*, 51:2636–2656, Jul 2005.
- [PFM93] N. Phamdo, N. Farvardin, and T. Moriya. A unified approach to tree-structured and multistage vector quantization for noisy channels. *IEEE Trans. Inform. Theory*, 39:835–850, May 1993.
- [PPK89] M. Paterakis and P. Papantoni-Kazakos. A simple window random-access algorithm with advantageous properties. *IEEE Trans. Inform. Theory*, 35:1124–1130, Sept 1989.

- [QB01] X. Qin and R. Berry. Exploiting multiuser diversity in wireless ALOHA networks. In *Proc. Allerton Conf. Communication, Control and Computing*, pages 793–794, Allerton, IL, Oct 2001.
- [RE88] R. Rao and A. Ephremides. On the stability of interacting queues in a multi-access system. *IEEE Trans. Inform. Theory*, 34:918–930, Sept 1988.
- [Rob72] L. G. Roberts. *Aloha packet system with and without slots and capture*. Stanford Research Institute, Advanced Research Projects Agency, Network Information Center, 1972. ASS Note 8.
- [ROUV93] K. Ramchandran, A. Ortega, K. M. Uz, and M. Vetterli. Multiresolution broadcast for digital HDTV using joint source/channel coding. *IEEE J. Select Areas Commun.*, 11:6–23, Jan 1993.
- [RPT95] R. Raheli, A. Polydoros, and C. Tzou. Per-survivor processing: A general approach to MLSE in uncertain environments. *IEEE Trans. Commun.*, 43:354–364, Feb/Mar/Apr 1995.
- [RSAA97] M. C. Reed, C. B. Schlegel, P. D. Alexander, and J. A. Asenstorfer. Iterative multiuser detection for DS-CDMA with FEC. In *Proc. Int. Symp. Turbo Codes and Related Topics*, pages 162–165, Brest, France, 1997.
- [RSG87] B. Ramamurthi, A. M. Saleh, and D. J. Goodman. Perfect capture ALOHA for local radio communications. *IEEE J. Select. Areas Commun.*, SAC-5:806–813, Jun 1987.
- [SC85] M. Sidi and I. Cidon. Splitting protocols in presence of capture. *IEEE Trans. Inform. Theory*, IT-31:295–301, Mar 1985.
- [SMM04] L. Sanguinetti, M. Morelli, and U. Mengali. Channel estimation and tracking for MC-CDMA signals. *European Trans. Telecommun.*, 15:249–258, May 2004.
- [SS83] M. Sidi and A. Segall. Two interfering queues in packet-radio networks. *IEEE Trans. Commun.*, COM-31:123–129, Jan 1983.
- [SS00] J. Sant and V. Sharma. Performance analysis of a slotted-ALOHA protocol on a capture channel with fading. *Queueing Systems, Theory and Applications*, 34:1–35, 2000.
- [Ste92] R. Steel. *Mobile Radio Communications*. IEEE press, New York, 1992.

- [SV01] S. Shamai and S. Verdu. The impact of frequency-flat fading on the spectral efficiency of CDMA. *IEEE Trans. Inform. Theory*, 47:1302–1327, Sept 2001.
- [Szp86] W. Szpankowski. Bounds for queue lengths in a contention packet broadcast system. *IEEE Trans. Commun.*, pages 1132–1140, Nov 1986.
- [Szp94] W. Szpankowski. Stability conditions for some multiqueue distributed systems: buffered random access systems. *Adv. Appl. Probab.*, 26:498–515, Jun 1994.
- [TE93] L. Tassiulas and A. Ephremides. Dynamic server allocation to parallel queues with randomly varying connectivity. *IEEE Trans. Inform. Theory*, 39:466–478, Mar 1993.
- [TGZ96] M. K. Tsatsanis, G. B. Giannakis, and G. Zhou. Estimation and equalization of fading channels with random coefficients. In *Proc. IEEE Int. Conf. on Acustics, Speech and Signal Processing (ICASSP)*, volume 2, pages 1093–1096, May 1996.
- [TH98] D. N. C. Tse and S. V. Hanly. Multiaccess fading channels. i. polymatroid structure, optimal resource allocation and throughput capacities. *IEEE Trans. Inform. Theory*, 44:2796–2815, Nov 1998.
- [TM79] B. S. Tsybakov and V. A. Mikhailov. Ergodicity of slotted ALOHA system. *Problems of Information Transmission*, 15:73–87, Mar 1979.
- [TNV04] L. Tong, V. Naware, and P. Venkitasubramaniam. Signal processing in random access. *IEEE Signal Processing Magazine*, pages 29–39, Sept 2004.
- [TS99] I. E. Telatar and S. Shamai. Some information theoretic aspects of decentralized power control in multiple access fading channels. In *Proc. Inform. Theory and Networking Workshop*, pages 23–23, Piscataway, NJ, 1999.
- [VA91] M. K. Varanasi and B. Aazhang. Near-optimum detection in synchronous code division multiple access systems. *IEEE Trans. Commun.*, 39:725–736, May 1991.
- [VD98] M. C. Valenti and B. D. Woerner. Iterative multiuser detection for convolutionally coded asynchronous DS-CDMA. In *Proc. IEEE Int. Symp. Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pages 213–217, Boston, MA, Sept 1998.

- [VNT04] G. Mergen V. Naware and L. Tong. Stability and delay of finite user slotted ALOHA with multipacket reception. Tech. Rep. ACSP-TR-08-04-01, Cornell Univ., Ithaca, NY, 2004.
- [VT95] G. M. Vitetta and D. P. Taylor. Maximum likelihood decoding of uncoded and coded PSK signal sequences transmitted over Rayleigh flat-fading channels. *IEEE Trans. Commun.*, 43:2750–2758, November 1995.
- [WF94] M. Wang and T. R. Fischer. Trellis-coded quantization designed for noisy channels. *IEEE Trans. Inform. Theory*, 40:1792–1802, Nov 1994.
- [WG00] Z. Wang and G. B. Giannakis. Wireless multicarrier communications: where Fourier meets Shannon. *IEEE Signal Processing Mag.*, 17:29–48, May 2000.
- [WP98] X. Wang and H. V. Poor. Adaptive joint multiuser detection and channel estimation in multipath fading CDMA channels. *Wireless Networks*, 4:453–470, 1998.
- [WP99] X. Wang and H. V. Poor. Iterative (turbo) soft interference cancellation and decoding for coded CDMA. *IEEE Trans. Commun.*, 47:1046–1061, 1999.
- [YD00] B. Yucel and H. Delic. Mobile radio window random-access algorithm with diversity. *IEEE Trans. Veh. Technol.*, 49:2060–2070, Nov 2000.
- [YH00] L. Yang and L. Hanzo. A recursive algorithm for the error probability evaluation of M-QAM. *IEEE Commun. Lett.*, 4:304–306, Oct 2000.
- [YLF93] N. Yee, J-P. Linnartz, and G. Fettweis. Multi-carrier CDMA in indoor wireless radio networks. In *Proc. IEEE int. Symp. on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pages 109–113, Yokohama, Japan, Sept 1993.
- [ZG88] K. Zeger and A. Gersho. Vector quantization design for memoryless noisy channels. In *Proc. IEEE Int. Conf. Commun.*, pages 1593–1597, Philadelphia, PA, Jun 1988.
- [ZG93] K. Zeger and A. Gersho. Pseudo-gray coding. *IEEE Trans. Commun.*, 38(12):2147–2158, Dec 1993.
- [Zor95] M. Zorzi. Mobile radio slotted ALOHA with capture and diversity. *Wireless Networks*, 1:227–239, May 1995.
- [ZR94] M. Zorzi and R.R.Rao. Capture and retransmission control in mobile radio. *IEEE J. Select. Areas Commun.*, SAC-12:1289–1298, 1994.

- [ZT03a] Q. Zhao and L. Tong. A dynamic queue protocol for multiaccess wireless networks with multipacket reception. *IEEE Trans. Wireless Commun.*, 11:125–137, Feb 2003.
- [ZT03b] Q. Zhao and L. Tong. A multiqueue service room MAC protocol for wireless networks with multipacket reception. *IEEE/ACM Trans. Networking*, 11:125–137, Feb 2003.



MICHIGAN STATE UNIVERSITY LIBRARIES



3 1293 02845 5768