2006-07-07

ARQ Techniques for MIMO Communication Systems

Zhihong Ding
Brigham Young University - Provo

ARQ TECHNIQUES FOR MIMO COMMUNICATION SYSTEMS

by

Zhihong Ding

A dissertation submitted to the faculty of

Brigham Young University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Electrical and Computer Engineering

Brigham Young University

August 2006

BRIGHAM YOUNG UNIVERSITY

GRADUATE COMMITTEE APPROVAL

of a dissertation submitted by

Zhihong Ding

This dissertation has been read by each member of the following graduate committee and by majority vote has been found to be satisfactory.

Date	Michael D. Rice, Chair
Date	Michael A. Jensen
Date	Brian D. Jeffs
Date	Richard L. Frost
Date	A. Lee Swindlehurst

BRIGHAM YOUNG UNIVERSITY

As chair of the candidate's graduate committee, I have read the dissertation of Zhihong Ding in its final form and have found that (1) its format, citations, and bibliographical style are consistent and acceptable and fulfill university and department style requirements; (2) its illustrative materials including figures, tables, and charts are in place; and (3) the final manuscript is satisfactory to the graduate committee and is ready for submission to the university library.

Date

Michael D. Rice
Chair, Graduate Committee

Accepted for the Department

Michael A. Jensen
Graduate Coordinator

Accepted for the College

Alan R. Parkinson
Dean, Ira A. Fulton College of
Engineering and Technology

ABSTRACT

ARQ TECHNIQUES FOR MIMO COMMUNICATION SYSTEMS

Zhihong Ding

Electrical and Computer Engineering

Doctor of Philosophy

Multiple-input multiple-output (MIMO) communication systems employ multiple antennas at the transmitter and the receiver. Multiple antennas provide capacity gain and/or robust performance over single antenna communications. Traditional automatic-repeat-request (ARQ) techniques developed for single-input single-output (SISO) communication systems have to be modified in order to be employed in MIMO communication systems. In this dissertation, we propose and analysis some ARQ techniques for MIMO communication systems.

The basic retransmission protocols of ARQ, stop-and-wait (SW-ARQ), go-back-N (GBN-ARQ), and selective repeat (SR-ARQ), designed for SISO communication systems are generalized for parallel multichannel communication systems. The generalized ARQ protocols seek to improve the channel utilization of multiple parallel channels with different transmission rates and different packet error rates. The generalized ARQ protocols are shown to improve the transmission delay as well.

A type-I hybrid-ARQ error control is used to illustrate the throughput gain of employing ARQ error control into MIMO communication systems. With the channel information known at both the transmitter and the receiver, the MIMO channel

is converted into a set of parallel independent subchannels. The performance of the type-I hybrid-ARQ error control is presented. Simulation results show the throughput gain of using an ARQ scheme in MIMO communication systems.

When the channel state information is unknown to the transmitter, error control codes that span both space and time, so-called space-time coding, are explored in order to obtained spatial diversity. As a consequence, the coding scheme used for ARQ error control has to be designed in order to consider coding across both space and time. In this dissertation, we design a set of retransmission codes for a type-II hybrid-ARQ scheme employing the multidimensional space-time trellis code as the forward error control code. A concept of sup-optimal partitioning of the (super-)constellation is proposed. The hybrid-ARQ error control scheme, consisting of the optimal code for each transmission, outperforms the hybrid-ARQ error control scheme, consisting of the same code for all transmissions.

ACKNOWLEDGMENTS

I sincerely thank my advisor, Dr. Michael Rice, for his patience, encouragement, and guidance over the years. I appreciate him for his constant support. Dr. Rice is alway ready to answer my questions and inspire me whenever needed.

I would also like to thank the other members of my committee: Dr. Richard Lee Frost, Dr. Brian Jeffs, Dr. Michael Jensen, and Dr. A. Lee Swindlehurst for their help. I gratefully acknowledge Dr. Wynn Stirling, Dr. Dah-Jay Lee, and Dr. David Long for their encouragement and instruction. I owe special thanks to Dr. Chris Peel and Dr. Thomas Svantesson for sharing good research ideas and suggestions.

I also appreciate my friends for sharing with me some wonderful times, helping me go through the hard times, and some of them being companions and fellow mothers with me. It would be a long list to mention all of the friends I am indebted to. I gratefully thank all of them.

My husband, my child, my father, and my brother deserve a warm and special acknowledgement for their love and care.

Finally, I would like to dedicate this dissertation to my mother. Her love and encouragement were in the end what made this dissertation possible. I wish she has a peaceful mind in heaven.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Background and Motivation

Automatic-Repeat-Request (ARQ) protocols are an error control technique for data transmission in which the receiver detects transmission errors in a message and automatically requests a retransmission from the transmitter [49]. Most of the ARQ techniques were developed for single-input single-output (SISO) communication systems. When we use an ARQ protocol in a multiple-input multiple-output (MIMO) communication system, some retransmission protocols must be generalized and new error control coding should be designed in order to realize the full potential of MIMO communications.

There are three basic retransmission protocols: stop-and-wait (SW-ARQ), go-back-N (GBN-ARQ), and selective repeat (SR-ARQ) [49, 38]. The retransmission protocols determine how retransmission requests are handled by the transmitter and receiver. The basic stop-and-wait ARQ scheme is illustrated in Figure 1.1. The transmitter sends a packet to the receiver and waits for an acknowledgement. A positive acknowledgement (ACK) from the receiver indicates that the transmitted packet has been successfully received, and the transmitter sends the next packet in the queue. A negative acknowledgement (NAK) from the receiver indicates that the transmitted packet has been detected in error; the transmitter then resends the packet again and waits for an acknowledgement. Since the transmitter is idle while waiting for the acknowledgment, this scheme is inefficient when the round-trip delay is large. If we are willing to allow for some buffering in the transmitter, pipelined ARQ

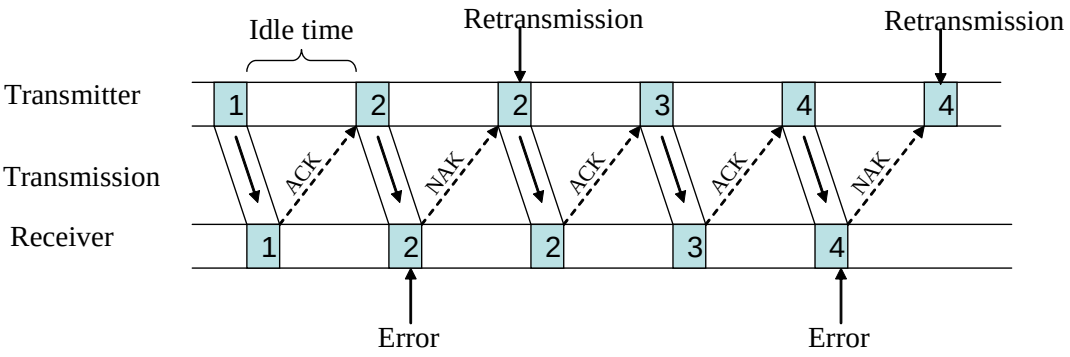


Figure 1.1: Stop-and-Wait ARQ.

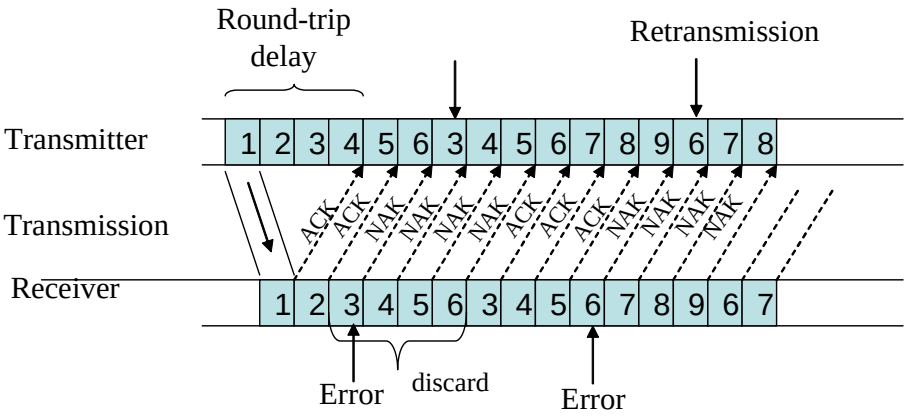


Figure 1.2: Go-Back-N ARQ with $N = 4$.

protocols, such as go-back-N (GBN) or selective repeat (SR), as shown in Figure 1.2 and Figure 1.3, are used.

In GBN-ARQ protocol, the transmitter sends packets in a continuous stream. When the receiver detects an error in a received packet, it sends a retransmission request for that packet and waits for its second copy. All subsequent incoming packets are ignored until the second packet is received. By ignoring the packets that follow a retransmission request, receiver buffering is avoided. In SR-ARQ protocol, buffering is allowed in both the transmitter and the receiver. In this case, the transmitter sends a continuous stream of packets and re-sends only those packets that were

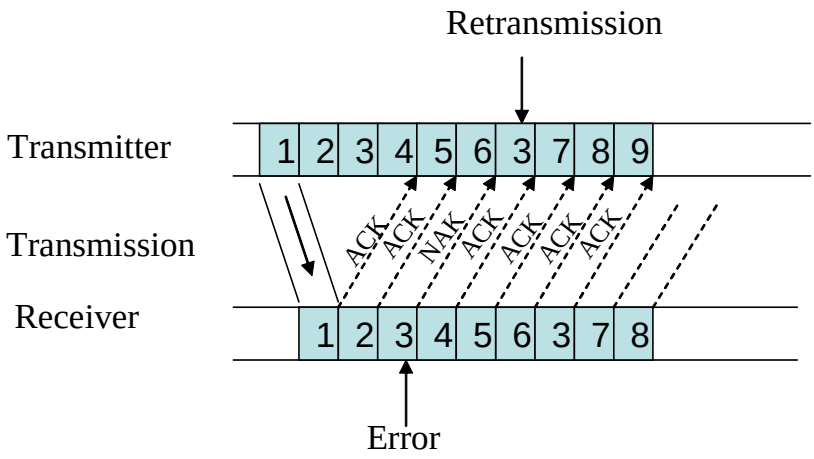


Figure 1.3: Selective-Repeat ARQ.

negatively acknowledged. The SR-ARQ protocol has the highest throughput. But receiver must be able to cope with our-of-order packets.

These three retransmission protocols were originally designed for single channel transmission. In these systems, the transmitter sends one packet at a time over the channel. When a temporally sequential data stream is transmitted over M multiple-parallel channels, a *block* consisting of M packets is sent: one packet over each of the constituent parallel channels. Due to the strong connection between the temporal nature of the data stream and SW and GBN retransmission protocols, these protocols must be generalized in order to realize the full potential of parallel multi-channel communications.

Previous work on the retransmission protocols of multichannel ARQ includes Chang and Yang [5]; Wu, Vassiliadis, and Chung [51]; and Anagnostou and Protonotarios [2]. The performance analysis given in those papers assume that all of the parallel channels were identical (i.e., each has the same transmission rate and packet error probability). A packet to be retransmitted is simply assigned to the next available channel. Which channel is used is unimportant, since all the channels are the same. When the channels are *different*, it is important which channel is used for retransmission. This behavior was first observed by Shacham in 1987 [36] in an

analysis of overall resequencing delay. He noted that proper channel assignment for retransmission could have an effect on throughput performance. Shacham and Shin [37] described and analyzed a modified SR ARQ protocol for use over parallel channels with the same transmission rate but different packet error rates. In this dissertation, we present the generalized ARQ protocols that seek to improve the channel utilization (a generalization of system throughput) when applied multiple parallel channels with different transmission rates and different packet error rates. These generalized ARQ protocols are later shown to improve the transmission delay performance as well.

MIMO communication systems employ multiple antennas at the transmitter and the receiver. A MIMO system takes advantage of the spatial diversity that is obtained by spatially separated antennas in a dense multipath scattering environment. MIMO systems may be implemented in a number of different ways to obtain either a capacity gain or to obtain a diversity gain to combat signal fading. Generally, there are three categories of MIMO techniques. The first type exploits knowledge of the channel at the transmitter to achieve near capacity. The second class uses a layered approach to increase capacity. One popular example of such a system is V-BLAST proposed by Foschini et al. [12], where full spatial diversity is usually not achieved. The third class aims to improve the power efficiency by maximizing spatial diversity. Such techniques include space-time block codes (STBC) [1, 43] and space-time trellis codes (STTC) [44]. The error control coding used in MIMO communication systems may or may not be the same as the SISO communication systems.

When the channel information is known to both the transmitter and the receiver, the spatio-temporal vector-coding (STVC) [31] converts the MIMO channel into a set of parallel independent subchannels. It decomposes the channel coefficient matrix using a singular value decomposition (SVD) and uses these decomposed unitary matrices as pre- and post-filters at the transmitter and the receiver to achieve near capacity as shown in Figure 1.4. Single ARQ and multiple ARQ schemes can be used for STVC MIMO system as shown in Figure 1.4. In the single ARQ scheme (top plot of Figure 1.4), all the transmit antennas share a unique encoder (CRC). In other words, the ARQ is unaware of the presence of MIMO. Multiple ARQ scheme

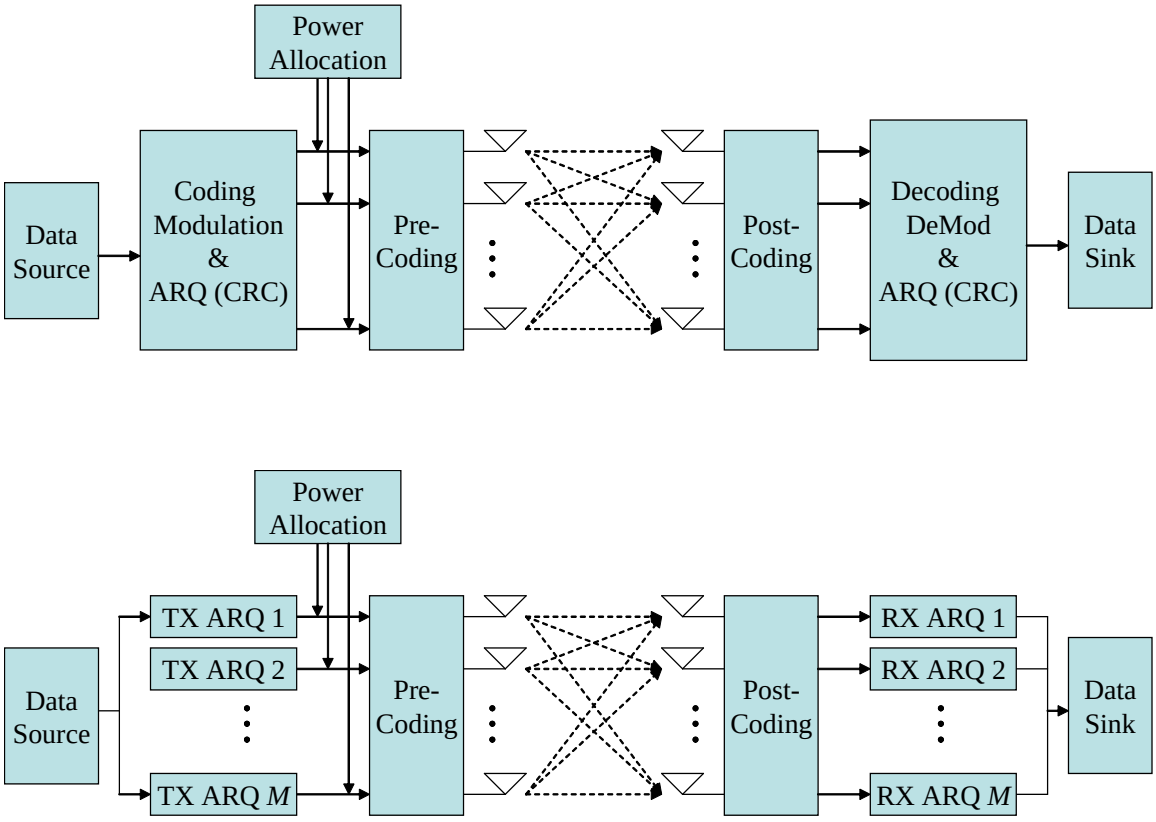


Figure 1.4: MIMO communication with channel state information at the transmitter. (top) Single ARQ scheme. (bottom) Multiple ARQ scheme.

(bottom plot of Figure 1.4) uses one encoder for each transmit antenna. In this case, each subchannel can be treated as a SISO channel and the error correction code used on each subchannel can be the same as the one used in SISO communication systems.

Others have studied packet retransmissions in MIMO systems with channel state information available at the transmitter. Sun and Ding et al. [42, 41] propose linear ARQ precoders in flat-fading MIMO system with the objective of maximizing the mutual information delivered by multiple transmissions of the same packet [42] or minimizing the mean square error between the transmitted data and the joint receiver output [41] . The optimal linear precoders combine the waterfilling power loading and the optimal pairing of singular vectors in the current retransmission with

previous transmissions. Single ARQ scheme is used in [42, 41]. The data stream transmitted over multiple subchannels are treated as a single packet and detected and (re)transmitted all together. Since the substreams emitted from various transmit antennas encounter distinct propagation channels and thus have different error statistics, Zheng et al. [57] have shown that the multiple ARQ scheme results in a throughput improvement compared with single ARQ scheme.

In this dissertation, we show the performance of a type-I HARQ scheme of MIMO communications, where we assume that the channel information is known at both the transmitter and the receiver. Multiple ARQ scheme is considered. A set of multidimensional trellis code modulations (MSTTC) has been used as the error correction code. The throughput gain of using ARQ scheme in MIMO systems has been illustrated.

When the channel state information is unknown to the transmitter, error control codes that span both space and time, so-called space-time coding, are explored in order to obtained spatial diversity (Figure 1.5). Such techniques include space-time trellis codes (STTCs) [44, 13], multidimensional space-time trellis codes (MSTTCs) [18, 19], and space-time block codes (STBCs) [1, 43], all of which are designed for the case that the channel state information is available at the receiver but not at the transmitter; and unitary space-time codes (USTM) [15, 14, 16], which are designed for the case that the channel state information is available at neither the receiver nor the transmitter.

Seok and Lee [22] present a hybrid-ARQ scheme employing different STTCs for each transmission which are optimal on different operating SNR ranges. These codes were found using a computer search. The hybrid-ARQ scheme, consisting of the optimal STTC for each transmission, outperforms the hybrid-ARQ scheme, consisting of the same STTC for all transmissions.

In this dissertation, we consider the hybrid-ARQ scheme employing the MSTTC as the forward error control (FEC) code. Different MSTTC codes are designed for each transmission based on different partition chains of the super-constellation of MSTTCs.

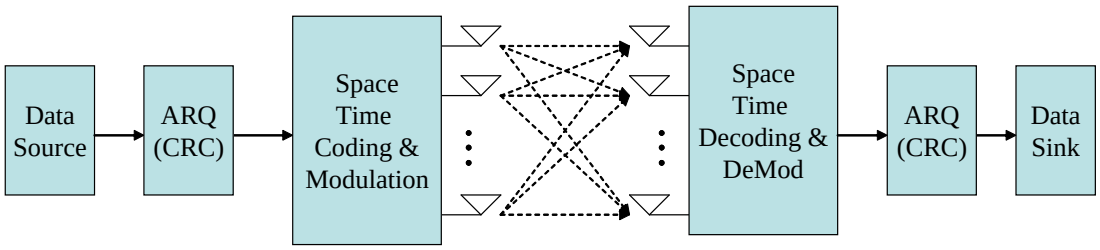


Figure 1.5: MIMO communication system using space time coding.

Some other ARQ schemes combined with MIMO communications have been proposed with channel state information at the receiver without using the standard space time coding explored for FEC MIMO channel. Samra and Ding in [34, 35] proposed a space-time block code using symbol mapping diversity where the bit-to-symbol mapping is adapted for each ARQ retransmission. Onggosanusi et al. [28] introduced methods for combining packet transmissions by using zero-forcing and minimum mean squared error (MMSE) receivers. Nguyen and Ingram [47] investigated hybrid ARQ protocols for systems that use recursive space-time codes and a turbo space-time hybrid ARQ scheme. Koike et al. [24] proposed a hybrid ARQ scheme employing trellis-coded modulation (TCM) reassignment and antenna permutation.

1.2 Contributions

- ARQ error control for parallel multichannel communications

  Generalized ARQ protocols are proposed that seek to improve the channel utilization and transmission delay when applied to parallel multichanel communication systems. Channel utilization and transmission delay of SW, GBN and SR ARQ protocols over parallel multichannels are analyzed. Simulation results show that the generalized ARQ protocols improve both the channel utilization and transmission delay performance of SW and GBN ARQ over parallel multichannel. A conference paper [9] has been published and a journal article [7] accepted based on this work.

- Type-I hybrid-ARQ using MTCM spatio-temporal vector coding for MIMO systems

  In order to show the capacity gain of using ARQ scheme in MIMO systems, we consider a type-I hybrid-ARQ scheme for MIMO communications with the channel state information at both the transmitter and the receiver. Spatio-temporal vector coding [31] has been used to convert the MIMO channel into parallel channels. The performance of the type-I hybrid-ARQ scheme over quasi-static flat fading MIMO channel has been analyzed. The capacity gain of using ARQ scheme in MIMO systems has been illustrated using simulations of a set of multidimensional trellis code modulations. A conference paper [8] has been published based on this work.

- A type-II hybrid-ARQ error control for multidimensional space-time trellis codes in quasi-static flat fading channels

  We present the space-time code design for hybrid-ARQ error control over MIMO channel employing MSTTCs are the FEC codes. The MSTTCs used for retransmission are designed using the sub-optimal partition of the super-constellation of the MSTTCs. Simulation results show that the hybrid-ARQ scheme, consisting of the optimal MSTTC for each transmission, outperforms the hybrid-ARQ scheme, consisting of the same MSTTC for all transmissions. This works is in preparation for submission to *IEEE Transactions on Wireless Communications*.

1.3 Organization

The remainder of this dissertation is organized as follows. Chapter 2 gives the channel utilization expressions of the SW, GBN, and SR ARQ protocols over a communication link consisting of multiple parallel channels with different transmission rates and different packet error rates. Generalized ARQ protocols are proposed to improve the channel utilization when applied to multiple parallel channels. At the end of Chapter 2, we derive the transmission delay of the SW, GBN, and SR ARQ

protocols over a communication link consisting of multiple parallel channels with different transmission rates and different packet error rates. Simulation results show that the generalized ARQ protocols improve the transmission delay of SW and GBN ARQ as well. In Chapter 3, we show the performance improvement of employing a type-I hybrid-ARQ scheme in MIMO systems where we assume that the channel state information is available at both the transmitter and the receiver. Spatio-Temporal Vector coding [31] has been used to convert the MIMO channel into parallel channels. In Chapter 4, we present a hybrid-ARQ scheme employing the multidimensional space-time code (MSTTC) as the FEC code. The retransmission codes are designed based on the sub-optimal partition of the super-constellation of the MSTTC. Finally, we offer conclusions in Chapter 5.

Chapter 2

ARQ Error Control for Parallel Channel Communications

2.1 Introduction

Historically, automatic-repeat-request (ARQ) protocols have been designed assuming temporally sequential communication over a single channel [49, 38]. In these systems, the transmitter sends one packet at a time over the channel. The three basic retransmission protocols are stop-and-wait (SW), go-back-N (GBN), and selective-repeat (SR) [49, 38]. For the stop-and-wait (SW) ARQ protocol, the transmitter waits until it receives an positive acknowledgement (ACK) or negative acknowledgement (NAK) from the receiver before resuming transmission. If the round-trip delay is large enough, then the SW ARQ protocol is inefficient and pipelined ARQ protocols, such as go-back-N (GBN) or selective repeat (SR), are used.

When transmitting a temporally sequential data stream over a single channel communication system, the data are partitioned into packets and transmitted one-by-one over the channel. When a temporally sequential data stream is transmitted over M multiple-parallel channels, a *block* consisting of M packets is sent: one packet over each of the constituent parallel channels. Due to the strong connection between the temporal nature of the data stream and SW and GBN retransmission protocols, these protocols must be generalized in order to realize the full potential of parallel multi-channel communications.

Multiple, parallel channels can be created in the frequency domain by using Orthogonal Frequency-Division Multiplexing (OFDM) or Discrete Multitone (DMT) modulation [4, 33, 46], in the code domain using vector coding [26, 23], or in space

using multiple transmit antennas [31]. In data networks, adjacent nodes may be connected by more than one link [37]. In this case, the multiple links present multiple parallel channels to the transmitter.

In this chapter, the SW, GBN, and SR ARQ protocols are analyzed over a communication link consisting of multiple parallel channels with different transmission rates and different packet error rates. The analysis leads to definitions of generalized ARQ protocols that improve channel utilization (a generalization of system throughput) when applied multiple parallel channels. At the end of the chapter, we show that the generalized ARQ protocols improve the transmission (delivering) delay, as well.

The operation of ARQ in a system that maps sequential data to multiple, parallel channels for transmission is somewhat different than it is for single channel systems. The operation of ARQ error control in a multichannel system is described in Section 2.2. In Section 2.3, it is shown that the generalized ARQ protocols take the form of channel assignment rules for packets to be retransmitted. The channel assignment rules are a function of the transmission rates and packet error rates associated with each channel. Simulation results are presented in Section 2.4 that demonstrate the gains of channel utilizations that can be obtained by using the proper packet-to-channel assignment rules. In Section 2.5, the transmission (deliver) delay of ARQ error control in a multichannel system has been derived. Simulation results demonstrate the reductions in transmission delays that can be obtained by using the proper packet-to-channel assignment rules. Conclusions are summarized in Section 2.6.

2.2 ARQ Error Control in a Multichannel System

The system model is illustrated in Figure 2.1. The communication link between the transmitter and receiver consists of M parallel channels. The m-th channel is characterized by its transmission rate R_m, measured in bits/symbol, and its packet error rate P_m for $m = 1, 2, \ldots, M$. We assume that the signal-to-noise (SNR) of each channel is known. Given a modulation type, it is usually straight-forward to compute the bit error rate or packet error rate for a given instantaneous SNR. All

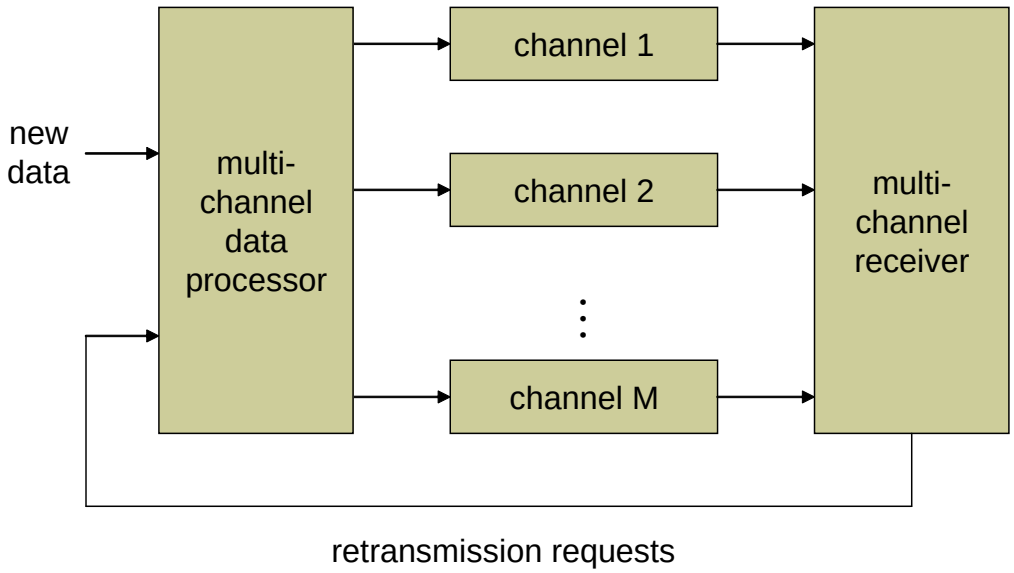


Figure 2.1: An abstraction of a communication system using multiple parallel channels.

channels share a set of *sequence numbers* which are used by the multichannel data processor to make the packet-to-channel assignments. The number of retransmissions is not restricted and the feedback channel is assumed error free.

Given a single input data stream, the multichannel data processor divides the input data stream into packets and assigns a sequence number to each packet. The sequence number preserves the sequential ordering of the packets in the input data stream. As a consequence, when a previously transmitted packet has to be retransmitted, it will have the lowest sequence number in the transmit queue. The multichannel data processor assigns the next M packets in the transmit queue to the M parallel channels. The multichannel receiver generates ACKs and NAKs for each packet on each of the M parallel channels and reassembles the accepted packets into a single data stream. Buffering is assumed available to handle out-of-order packet reception.

Since the channels can operate at different transmission rates, the concept of *channel utilization* is used in place of throughput as the performance measure.

Channel utilization is the average information data rate over the parallel channel measured in bits/symbol. Note that for the single channel system, the channel utilization is the same as the normalized throughput defined in [55] and [56].

Chang and Yang [5], Wu, Vassiliadis, and Chung [51], and Anagnostou and Protonotarios [2] investigated the throughput performance of multichannel ARQ protocols where all of the parallel channels were identical (i.e., each has the same transmission rate and packet error probability). Since all the channels are the same, the throughput and transmission delay performance is not a function of packet-to-channel assignment. Shacham [36] shown that when the channels are different, which channel is used for retransmission will affect the overall resequencing delay of SR ARQ. A modified SR ARQ protocol for used over parallel channels with the same transmission rate but different packet error rates was described and analyzed in 1992 by Shacham and Shin [37].

The case of parallel channels with different transmission characteristics is relevant to modern communication systems. Parallel channels with the same transmission rate, but different packet error rates can occur in an OFDM system experiencing frequency selective fading (e.g., some of the tones are suffering from more severe fading-induced attenuation than others). The parallel channel point of view for DMT/OFDM was exploited in [39, 6] to obtain improved bit allocation and bit loading algorithms. Recent results reported in [29] treated DMT tones as parallel channels that could be "clustered" to produce efficient fractional bit loading algorithms that did not require significant trellis modifications in the receiver. Likewise, parallel channels with the different transmission rates, but the same bit error rate result in MIMO systems using spatio-temporal coding with power allocation assignments obtained using a "water-filling" solution [31].

In the next section, necessary conditions for packet-to-channel assignment rules that improve channel utilization are derived for the SW, GBN, and SR retransmission protocols.

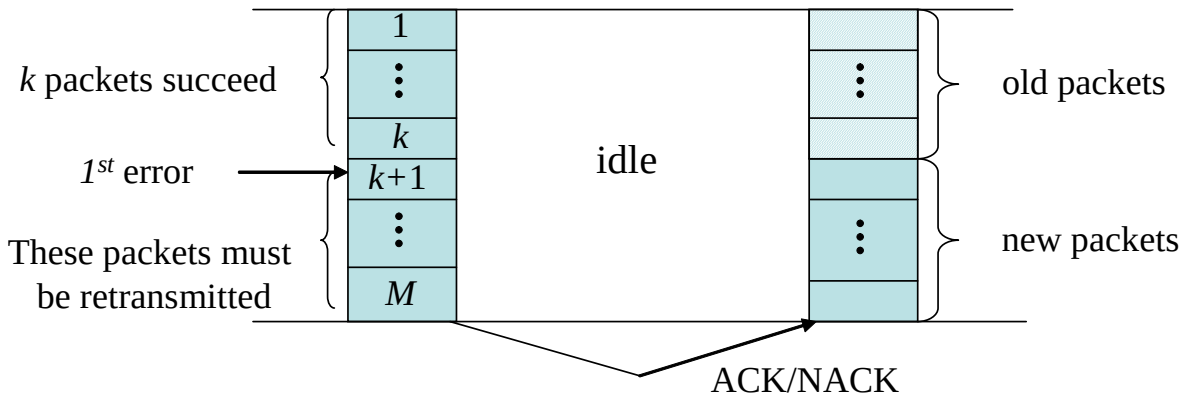


Figure 2.2: Stop and wait ARQ for multiple parallel channel.

2.3 Packet-to-Channel Assignment Rules

A mathematical expression for the channel utilization is derived for each of the three retransmission protocols. This expression is then used as the basis for defining packet-to-channel assignment rules that improve the channel utilization for each case.

2.3.1 Packet-to-Channel Assignment Rule for SW ARQ

In the SW ARQ protocol, the transmitter sends a block of M packets to the receiver and waits for acknowledgement from the receiver before it sends the next block of packets.

The transmitter is idle while waiting for the acknowldegment. Let D be the "idle time" or round-trip delay measured in packet times. Suppose an ACK is received for the packets sent on channels $1, 2, \ldots, k$ and a NAK is received for channel $k+1$. Since no buffering is provided at the receiver, the packets originally transmitted over channels $k+1, k+2, \ldots, M$ have to be retransmitted (See Figure 2.2), no matter they have been succefully transmitted or not. The channel utilization is derived as following.

The probability that the first k packets in the parallel block of M packets is successfully transmitted is

$$P_{\mathrm{S}}(k) = \begin{cases} \prod_{i=1}^{M}(1-P_i) & k = M \\ \prod_{i=1}^{k}(1-P_i)P_{k+1} & k \neq M \end{cases}. \tag{2.1}$$

Thus, the channel utilization can be expressed as

$$\eta = \frac{\sum_{k=1}^{M}\left(\sum_{i=1}^{k} R_i\right) \cdot P_{\mathrm{S}}(k)}{1+D} \tag{2.2}$$

where P_i and R_i are the packet error probability and the transmission rate of the ith channel. Using the matrices

$$\mathbf{R} = \begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_M \end{bmatrix} \qquad \mathbf{P} = \begin{bmatrix} P_{\mathrm{S}}(1) \\ P_{\mathrm{S}}(2) \\ \vdots \\ P_{\mathrm{S}}(M) \end{bmatrix}, \tag{2.3}$$

the channel utilization may be expressed in matrix form as

$$\eta = \frac{1}{1+D}\mathbf{R}^{\mathcal{T}}\mathbf{V}\mathbf{P} \tag{2.4}$$

where $\mathbf{R}^{\mathcal{T}}$ denotes the transpose of $\mathbf{R}$ and $\mathbf{V}$ is an upper triangular matrix consisting of all ones.

The expression (2.4) represents the channel utilization for a particular ordering of channels (as indicated by the position of the channel transmission rates and channel packet error rates in the matrices $\mathbf{R}$ and $\mathbf{P}$, respectively). Now consider a new ordering represented by switching the order of channel index i and channel index $i+1$. In this ordering, the channel transmission rates and packet error rates are

summarized by the matrices

$$\mathbf{R}' = \begin{bmatrix} R_1 & R_2 & \cdots & R_{i+1} & R_i & \cdots & R_M \end{bmatrix}^{\mathcal{T}} \text{ and} \tag{2.5}$$

$$\mathbf{P}' = \begin{bmatrix} P_{\mathrm{S}}(1) \\ P_{\mathrm{S}}(2) \\ \vdots \\ P_{\mathrm{S}}(i-2) \\ (1-P_1)(1-P_2)\cdots(1-P_{i-1})P_{i+1} \\ (1-P_1)(1-P_2)\cdots(1-P_{i-1})(1-P_{i+1})P_i \\ P_{\mathrm{S}}(i+1) \\ \vdots \\ P_{\mathrm{S}}(M) \end{bmatrix}. \tag{2.6}$$

The channel utilization for this ordering of channels is

$$\eta' = \frac{1}{1+D}\mathbf{R}'^{\mathcal{T}}\mathbf{V}\mathbf{P}'. \tag{2.7}$$

The difference between the two channel utilizations is

$$(\eta - \eta')(1+D) = (\mathbf{R}-\mathbf{R}')^{\mathcal{T}}\,\mathbf{V}\mathbf{P} + \mathbf{R}'^{\mathcal{T}}\mathbf{V}\,(\mathbf{P}-\mathbf{P}') \tag{2.8}$$

$$= \begin{bmatrix} 0 & \cdots & 0 & R_i - R_{i+1} & R_{i+1} - R_i & 0 & \cdots & 0 \end{bmatrix} \mathbf{V}\mathbf{P}$$

$$+ \mathbf{R}'^{\mathcal{T}}\mathbf{V} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ (P_i - P_{i+1})\prod_{n=1}^{i-1}(1-P_n) \\ (P_{i+1} - P_i)\prod_{n=1}^{i-1}(1-P_n) \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{2.9}$$

$$= (R_i - R_{i+1}) \prod_{n=1}^{i-1} (1-P_n)(1-P_i)P_{i+1}$$
$$+ \sum_{n=1}^{i-1} R_n \prod_{n=1}^{i-1} (1-P_n)(P_i - P_{i+1})$$
$$- \left(\sum_{n=1}^{i-1} R_n + R_{i+1} \right) \prod_{n=1}^{i-1} (1-P_n)(P_i - P_{i+1})$$

$$= \left[(R_i - R_{i+1})(1-P_i) P_{i+1} - R_{i+1}(P_i - P_{i+1}) \right] \times \prod_{n=1}^{i-1} (1-P_n). \quad (2.10)$$

A necessary condition for the original ordering to be optimal is that

$$\Delta_i = (R_i - R_{i+1})(1-P_i) P_{i+1} - R_{i+1}(P_i - P_{i+1}) > 0 \quad (2.11)$$

for $i = 1, 2, \ldots, M-1$. Five important special cases should be noted.

- All channels have identical transmission rates and packet error rates. In this case, $R_i = R_{i+1}$ and $P_i = P_{i+1}$ for all $i = 1, 2, \cdots, M-1$. Then Δ_i is zero for all i. Channel ordering in the assignment rule does not matter.

- All channels have the same transmission rate but different packet error rates. Let $R_i = R$ be the common transmission rate for $i = 1, 2, \ldots, M$. Then the necessary condition (2.11) becomes

  $$\Delta_i = -R(P_i - P_{i+1}) > 0 \Rightarrow P_i < P_{i+1}. \quad (2.12)$$

  This means the channels should be ordered from lowest packet error rate to highest packet error rate.

- All channels have different transmission rates but the same packet error rates. Let $P = P_i$ be the common packet error rate for $i = 1, 2, \ldots, M$. Then the necessary condition (2.11) becomes

  $$\Delta_i = (R_i - R_{i+1})(1-P)P > 0 \Rightarrow R_i > R_{i+1}. \quad (2.13)$$

  This means the channels should be ordered from highest transmission rate to lowest transmission rate.

- All channels have different transmission rates and different packet error rates but the packet error rate of each channel is proportional to the transmission rate. This case occurs when each channel is designed to have the same bit error rate[1]. Let $P_i = LR_i$ for $i = 1, 2, \cdots, M$. Then the necessary condition (2.11) becomes

$$\begin{aligned}\Delta_i &= (R_i - R_{i+1})(1 - P_i)LP_{i+1} - R_{i+1}L(R_i - R_{i+1}) \\ &= L(R_i - R_{i+1})[(1 - P_i)P_{i+1} - P_{i+1}] \\ &= L(R_i - R_{i+1})[-P_iP_{i+1}] > 0. \Rightarrow R_i < R_{i+1}. \end{aligned} \tag{2.14}$$

  This means the channels should be ordered from lowest transmission rate (lowest packet error rate) to highest transmission rate (highest packet error rate). We will show later that this case can be applied into MIMO communication system when the channel state information is available at both the transmitter and the receiver.

- All channels have different transmission rates and different packet error rates but all packet error rates satisfy $P_i \ll 1$ for $i = 1, 2, \cdots, M$. In other words, the packet error rates of the channels are relatively small. In this case, the necessary condition (2.11) implies that

$$\begin{aligned}(R_i - R_{i+1})\, P_{i+1} &> \frac{R_{i+1}\,(P_i - P_{i+1})}{1 - P_i} \\ &\approx R_{i+1}\,(P_i - P_{i+1}) \qquad (2.15) \\ \Rightarrow \frac{R_i}{P_i} &> \frac{R_{i+1}}{P_{i+1}}. \qquad (2.16)\end{aligned}$$

  The interpretation of this result is that the channels should be ordered (in descending order) based on the ratio of transmission rate to packet error rate. Note that the second and third special cases are special cases of this scenario.

[1]To see that this is so, let P_b be the common bit error rate for each channel and suppose that the packet length, $\mathcal{L}$ (measured in symbols), is also the same for each channel. The number of bits transmitted in a length-$\mathcal{L}$ packet over channel m with rate R_m is $L_{b,m} = \mathcal{L}R_m$. The packet error rate may be expressed as $P_m = 1 - (1 - P_b)^{L_{b,m}} \approx L_{b,m}P_b$. Substituting we obtain $P_m \approx \mathcal{L}P_bR_m$ which may be expressed as LR_m using $L = \mathcal{L}P_b$.

In all cases, the ordering is based on a quantitative measure of the channel quality. A channel with a higher transmission rate, or lower packet error rate (or both) has a higher quality. The reason the higher quality channels should be ordered first lies in the details of how the SW ARQ protocol assigns sequentially available packets in the transmit queue to the parallel channels. If a transmission in the first channel fails, then the packets sent in channels 2 and higher must also be retransmitted. This must be the case since SW ARQ does not provide any buffering at the receiver for reordering packets received out of order.

2.3.2 Packet-to-Channel Assignment Rule for GBN ARQ

In the GBN ARQ protocol, the transmitter sends packets to the receiver continuously and does not wait for acknowledgements from the receiver. The acknowledgement for each transmission block arrives after a round-trip delay of $N \times M$ packets (e.g., N blocks of M packets). During this interval, $N-1$ blocks of M packets have also been transmitted. For GBN ARQ, no buffering is available at the receiver. When a NAK is received for a particular packet, all the subsequent packets in the block, together will all packets in the subsequent $N-1$ blocks, are discarded by the receiver and must be resent by the transmitter.

Let S be the average number of accepted blocks prior to a NAK. S may be expressed as

$$\begin{aligned} S &= \sum_{k=1}^{\infty} k \cdot \left[\prod_{i=1}^{M}(1-P_i)\right]^k \left[1-\prod_{i=1}^{M}(1-P_i)\right] \\ &= [1-P_{\mathrm{S}}(M)] \sum_{k=1}^{\infty} k \cdot P_{\mathrm{S}}^k(M) \\ &= \frac{P_{\mathrm{S}}(M)}{1-P_{\mathrm{S}}(M)}. \end{aligned} \tag{2.17}$$

The channel utilization for GBN ARQ is

$$\eta = \frac{S}{N+S}\sum_{i=1}^{M} R_i + \frac{1}{N+S}\sum_{k=1}^{M-1}\left(\sum_{i=1}^{k} R_i\right)\frac{P_{\mathrm{S}}(k)}{1-P_{\mathrm{S}}(M)}. \tag{2.18}$$

Using the matrices

$$\bar{\mathbf{R}} = \begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_{M-1} \end{bmatrix} \qquad \bar{\mathbf{P}} = \begin{bmatrix} P_{\mathrm{S}}(1) \\ P_{\mathrm{S}}(2) \\ \vdots \\ P_{\mathrm{S}}(M-1) \end{bmatrix}, \tag{2.19}$$

the channel utilization may be expressed in matrix form as

$$\eta = \frac{S}{S+N} \sum_{i=1}^{M} R_i + \frac{1}{N+S} \cdot \frac{1}{1-P_{\mathrm{S}}(M)} \bar{\mathbf{R}}^{\mathcal{T}} \bar{\mathbf{V}} \bar{\mathbf{P}} \tag{2.20}$$

where $\bar{\mathbf{V}}$ is an upper triangular matrix consisting of all ones.

Note that the first term in (2.20) does not change with re-ordering of the channel index. The matrix form of the second term is also similar to the matrix form of (2.4) for SW-ARQ. Applying the same line of reasoning to this case produces the same necessary condition (2.11) for improved channel utilization.

2.3.3 Packet-to-Channel Assignment Rule for SR ARQ

In the SR ARQ multichannel protocol, the transmitter sends packets to the receiver continuously and re-sends only those packets that were negatively acknowledged. The packets at the receiver are out of order. Assuming a sufficiently large buffer at the receiver to reassemble the packets received out of order, the channel utilization can be expressed as

$$\eta = \sum_{i=1}^{M} R_i \cdot (1 - P_i). \tag{2.21}$$

Reordering the channel indexes in (2.21) does not effect the channel utilization. Thus, channel utilization is independent of the channel assignment for the SR ARQ protocol.

2.4 Simulation Results of Channel Utilization

2.4.1 Accuracy of the Channel Utilization Expressions

Computer simulations were used to assess the accuracy of the channel utilization expressions derived in Section 2.3. For our numerical example, we consider

a four channel system (i.e. $M = 4$) where the idle time of SW-ARQ is $D = 2$ block times and the round-trip delay of the GBN-ARQ is $N = 3$ block times. To illustrate the effect of different channel characteristics on the channel utilization, we adopt the same technique used in [37]: P_{i+1}/P_i is a constant (which we call r_P) for $i = 1, 2, \ldots, M-1$; R_{i+1}/R_i is a constant r_R for $i = 1, 2, \ldots, M-1$.

The results are summarized in Figure 2.3 where we see that the simulation results matched the analytical expressions exactly.

2.4.2 Comparison of Different Assignments in AWGN Channel

To illustrate the effect of different packet-to-channel assignment ordering rules, we compare our optimal rule (OR) with three other rules:

- *Dynamic Reverse Rule* (DRR): The packets are assigned across the parallel channels dynamiclly, while the channels are ordered in the reverse order relative to the ordering defined by OR.

- *Static Rule* (SR): Order the channels according to OR and retransmit a NAK'ed packet on the same channel as the original transmission.

- *Static Reverse Rule* (SRR): Order the channels in the reverse order as defined by OR and retransmit a NAK'ed packet on the same channel as the original transmission.

The rules are defined to illustrated the impact of "doing the wrong thing" (DRR) and "doing nothing" (SR and SRR). The gain of the optimal rule over the other rules is given as

$$G = \frac{\eta_{\mathrm{OR}} - \eta_{\mathrm{x}}}{\eta_{\mathrm{x}}} \tag{2.22}$$

where x is DRR, SR or SRR.

For the case where all the channels have the same transmission rate but different packet error probabilities, $r_R = 1$ and $\mathbf{P} = [\ P_1 \quad r_P P_1 \quad r_P^2 P_1 \quad r_P^3 P_1\ ]$ with $r_P > 1$. Figure 2.4 shows the channel utilization gain G as a function of P_1 and r_P, for SW-ARQ and GBN-ARQ.

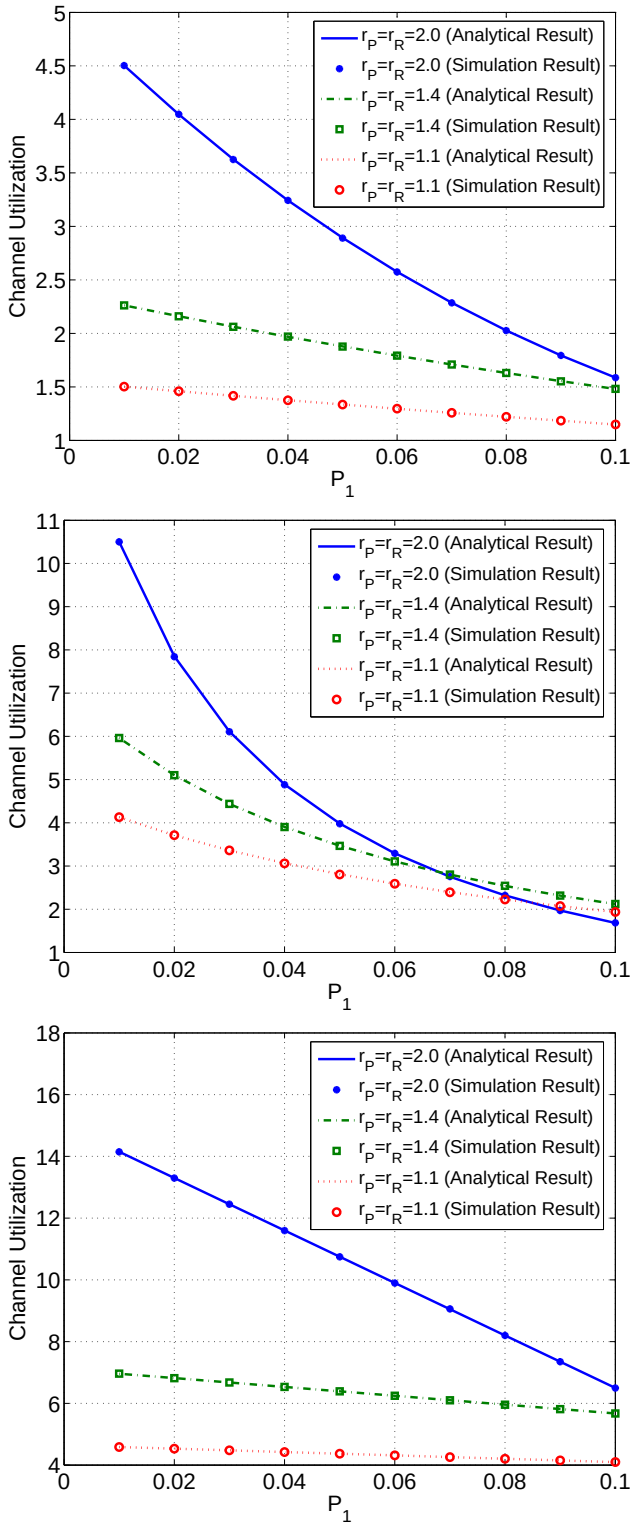


Figure 2.3: Comparison of computer simulations with analytical expressions for channel utilization for the three ARQ protocols: (top) stop-and-wait; (middle) go-back-N; (bottom) selective-repeat. In all cases, $R_1 = 1$ bit/symbol and r_R and r_P are kept to be equal to simplify the presentation.

For the case where all channels have the same packet error probability but different transmission rates, $r_P = 1$ and $\mathbf{R} = [\ R_1 \quad r_R R_1 \quad r_R^2 R_1 \quad r_R^3 R_1\ ]$ with $r_R < 1$. Figure 2.5 shows the channel utilization gain G as a function of P_1 and r_R for SW-ARQ and GBN-ARQ.

For the case where all channels have different transmission rates but the same bit error rates, the packet error rates are proportional to the transmission rates (this was the forth special case treated in Section 2.3.1). In this case $r_P = r_R = r > 1$. For the optimal rule,

$$\mathbf{R} = \begin{bmatrix} R_1 & rR_1 & r^2R_1 & r^3R_1 \end{bmatrix} \text{ and}$$
$$\mathbf{P} = \begin{bmatrix} P_1 & rP_1 & r^2P_1 & r^3P_1 \end{bmatrix}.$$

Figure 2.6 plots the channel utilization gain G as a function of r and P_1 for SW-ARQ and GBN-ARQ.

The simulation results for SW-ARQ and GBN-ARQ presented above suggest the following:

1. Assigning a packet to be retransmitted to the "worst" available channel produces the greatest reduction in over all channel utilization. This is demonstrated by the fact that the DRR consistently has the worst channel utilization.

2. The two "static" assignment rules do not achieve the channel utilization of the OR, which is a "dynamic" assignment rule. This suggests that a properly ordered dynamic assignment rule is needed to optimize channel utilization.

3. The gain of the OR over the other rules considered increases as the difference between the channels becomes more pronounced. This is illustrated by the fact that G increases as P_1 increases (see Figures 2.4 – 2.6), as r_P increases (see Figure 2.4), as r_R decreases (see Figure 2.5), or as r increases (see Figure 2.6).

The above simulations are done for static AWGN channel. In the next section we will show some analysis and simulations for quasi-static MIMO channel.

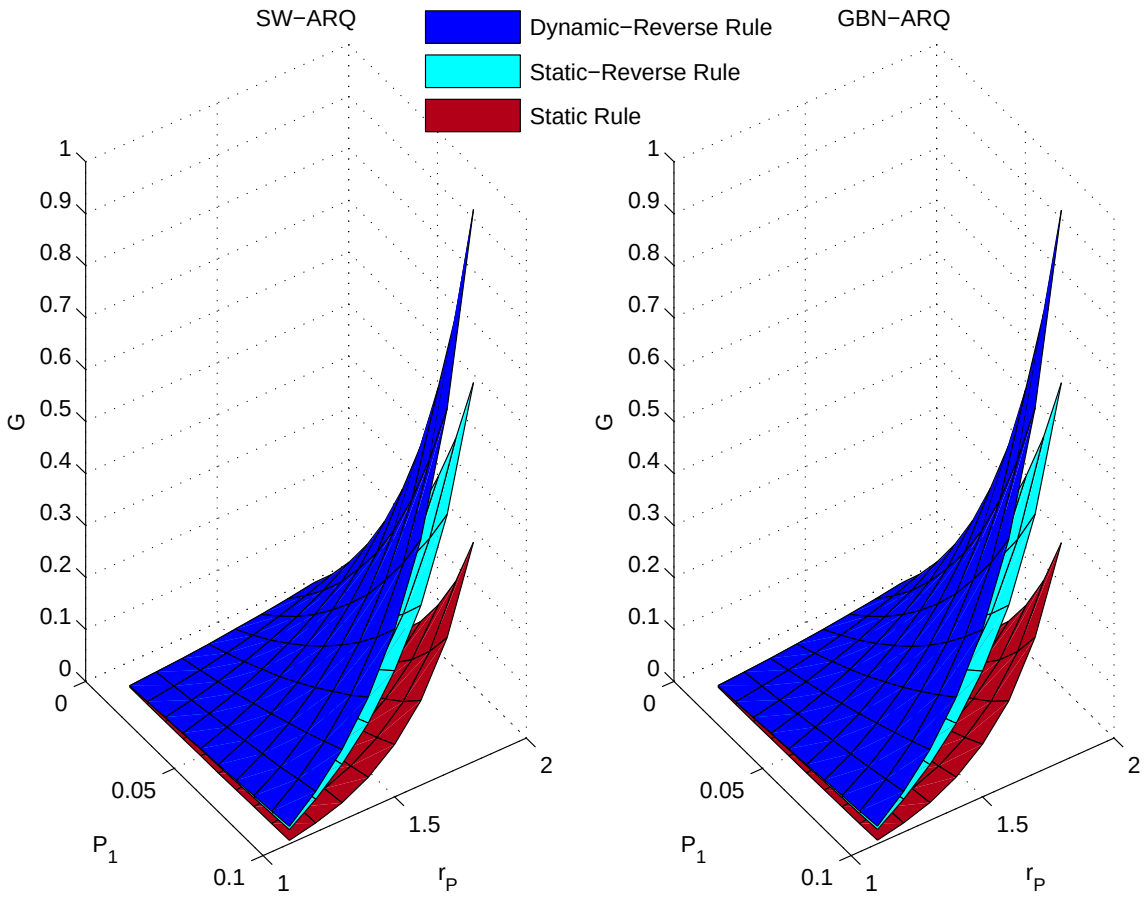


Figure 2.4: Channel utilization comparison for the case where the channels have the same transmission rate but different packet error probabilities.

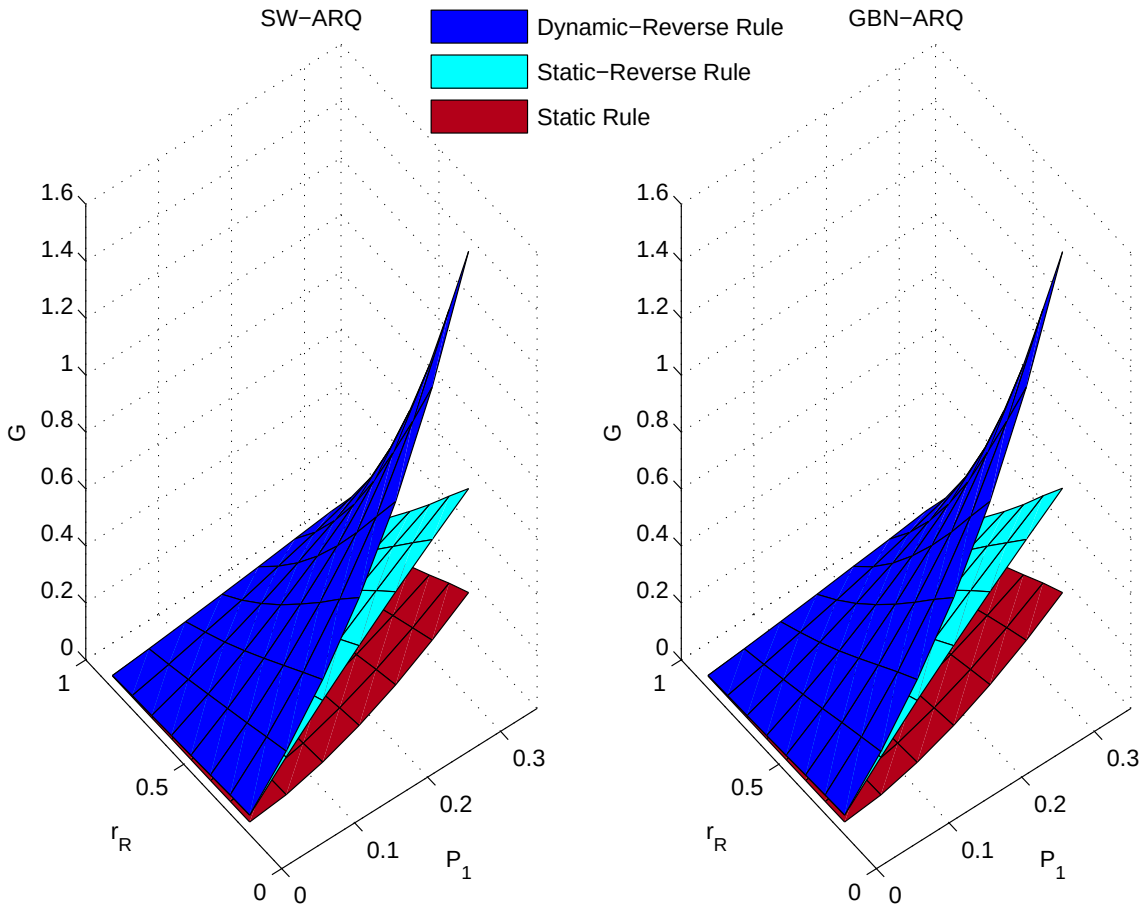


Figure 2.5: Channel utilization comparison for the case where the channels have the same packet error probability but different transmission rates.

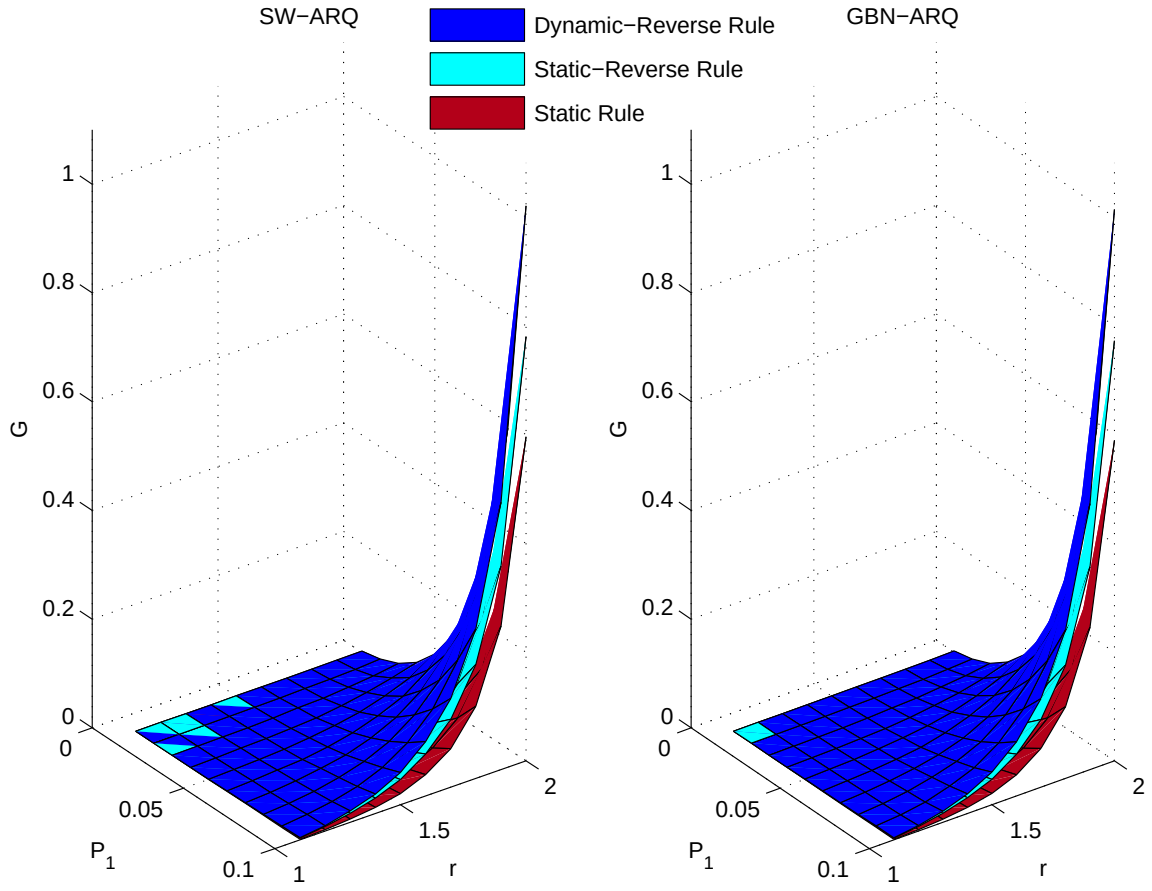


Figure 2.6: Channel utilization comparison for the case where all the channels have different packet error probabilities and different transmission rates.

2.4.3 Comparison of Different Assignments in Wireless Channels

Simulations were also performed for a MIMO system with 4 transmit antennas and 4 receive antennas using spatio-temporal vector coding [31]. Spatio-temporal vector coding over a frequency non-selective MIMO channel $\mathbf{H}$ is accomplished by computing the singular value decomposition of the channel matrix: $\mathbf{H} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^*$ and using the right singular vectors (columns of $\mathbf{V}$) as the bases for the transmitted sequences and the left singular vectors (columns of $\mathbf{U}$) as the matched filters. The vector of matched filter outputs may be expressed as

$$\mathbf{R} = \mathbf{U}^*\mathbf{H}\mathbf{V}\mathbf{Z} + \nu = \mathbf{\Lambda}\mathbf{Z} + \nu \tag{2.23}$$

where ν is the vector of noise samples and and $\mathbf{Z}$ is the vector of information symbols.

In this way, spatio-temporal vector coding creates rank($\mathbf{H}$) parallel communication channels. The gains of each of the channels is given by its singular values λ_n which is the element (n, n) in the matrix $\mathbf{\Lambda}$. Different information rates are assigned to each of the sub-channels using a spatio-temporal water-filling and bit-allocation solution to achieve capacity.

Simulations were performed using two 4×4 channel matrices: the first was the IID MIMO channel where $\mathbf{H}$ consists of 16 zero-mean unit-variance complex Gaussian random variables. The second channel was measured in an indoor environment as described in [48]. The channel matrix is assumed constant during one packet interval but varies from packet to packet.

A set of Gray-coded $\mathcal{M}$-PSK modulation schemes for $\mathcal{M} = 2, 4, 8, 16, 32, 64$ was used to provide transmission rates of 1, 2, 3, 4, 5, and 6 bits/symbol, respectively. The binary reflected Gray code described in [25] was used for the bit-to-symbol mapping. The symbols were indexed $0, 1, \ldots, M-1$ starting with the point $1 + j0$ and proceeding in the counter-clockwise direction.

Since this set of modulation schemes provide finite granularity in the transmission rates, the operations such as rounding the result of water-filling and bit allocation solution to a finite number may not be optimal. A bit loading algorithm [4] is

used in stead to allocate power and bit to the parallel channels subject to a pre-ARQ bit error rate constraint.

After the bit loading operation, the channels are ordered $\lambda_1 \geq \lambda_2 \cdots \geq \lambda_r$ so that the transmission rates of the r sub-channels satisfy $R_1 \geq R_2 \cdots \geq R_r$ while the packet error rate of the r channels satisfy $P_1 \geq P_2 \cdots \geq P_r$. Thus, this scenario matches the forth special case in Section 2.3. The packet-to-channel assignment rule developed in Section 2.3 requires the packets to be ordered from lowest transmission rate to highest transmission rate (that is, in the reverse order from above). The channel utilization using that assignment rule was simulated and designated η_{opt}.

To model the effect of "doing nothing," the channel ordering $\lambda_1 \geq \lambda_2 \cdots \geq \lambda_r$ was left in place and packets to be retransmitted were inserted in the next available channel without regarding for the channel number. The channel utilization for this case was also simulated and designated η_{static}.

The simulation results are summarized in Figures 2.7 and 2.8 for the SW ARQ protocol (with idle time $D = 2$) and the GBN ARQ protocol (with $N = 3$), respectively. In these plots, the channel utilization gain

$$G = \frac{\eta_{\text{opt}} - \eta_{\text{static}}}{\eta_{\text{static}}} \tag{2.24}$$

is plotted as a function of packet length.

Observe that substantial gains in channel utilization can be realized, especially as the packet length increases and the pre-ARQ bit error rate increases. The optimal assignment rule produces a channel utilization that is better than the channel utilization that results from doing nothing. The simulation results again suggest that for SW-ARQ and GBN-ARQ a properly ordered dynamic assignment rule is needed to optimize channel utilization.

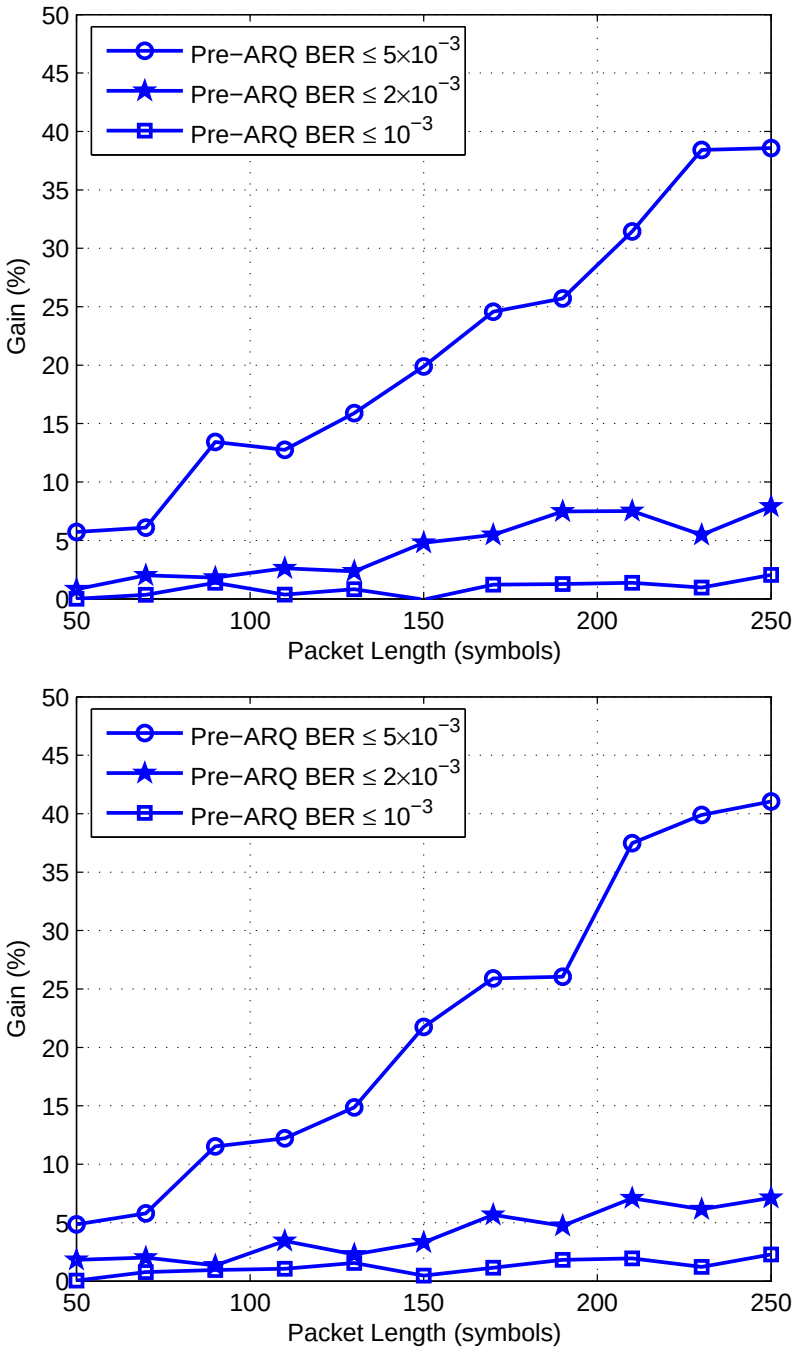


Figure 2.7: Performance comparison for the multichannel SW-ARQ retransmission protocol for a 4×4 MIMO channel using spatio-temporal coding: (top) IID channel; (bottom) BYU channel.

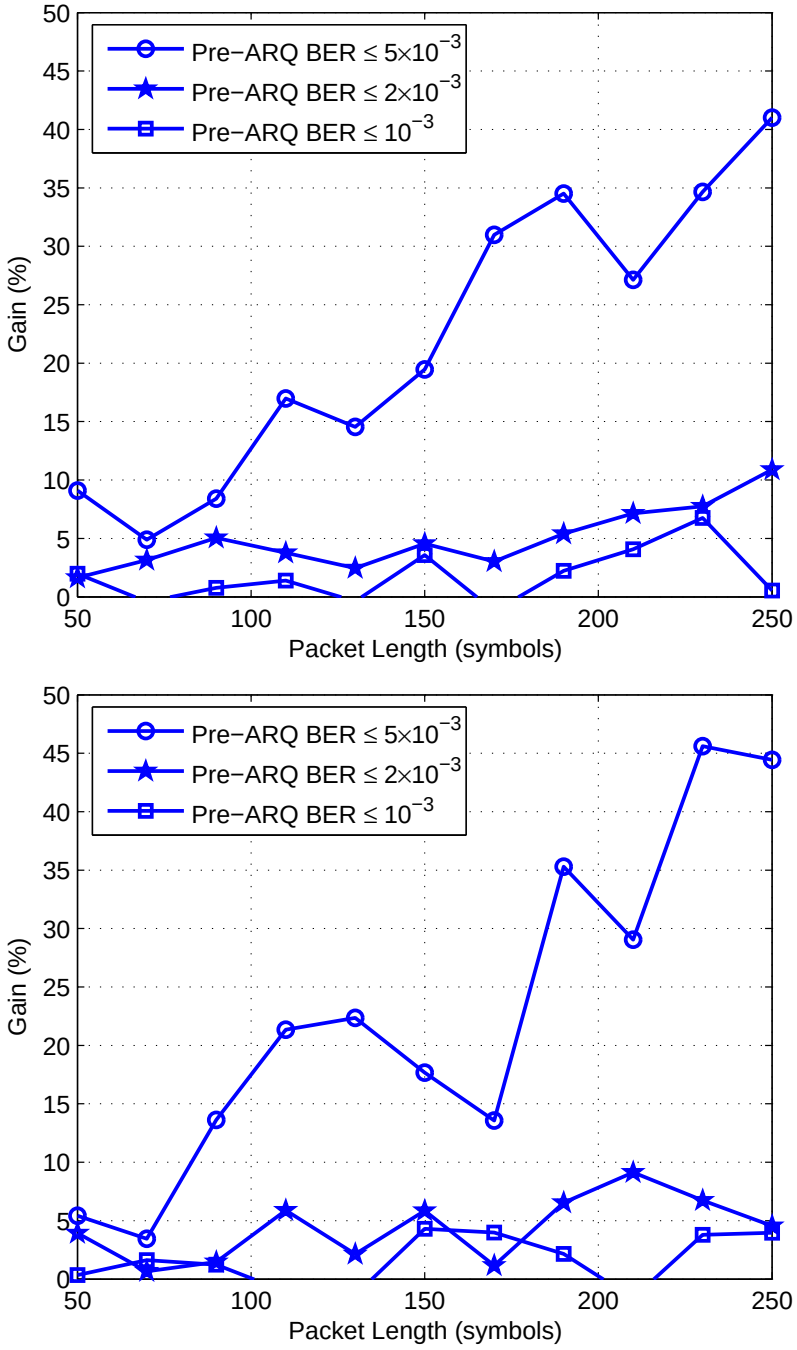


Figure 2.8: Performance comparison for the multichannel GBN-ARQ retransmission protocol for a 4×4 MIMO channel using spatio-temporal coding: (top) IID channel; (bottom) BYU channel.

2.5 Transmission Delay Analysis

The transmission delay performance of multichannel ARQ protocols have been studied in Chang and Yang's paper[5], where all the channels in the multichannel system are assumed to be identical. In this section, we will derive the transmission delay expressions for the ARQ protocols of multichannel system as shown in Figure 2.1. We will start with the case in which all the channels have the same transmission rate but different packet error rates. Then the result will be extended to the case in which all the channels have both different transmission rates and different packet error rates. Some simulation results will be presented in order to show that the packet-to-channel assignment rules we derived to maximize the channel utilization of the system will also minimize the transmission delay of the system.

2.5.1 Transmission Delay Expressions

Consider a system with M parallel channels between the transmitter and the receiver. The transmission rates of the M channels are the same. The round-trip signal propagation delay is assumed to be N block times. The transmission delay is measured from the instant at which a packet is transmitted for the first time till the time it is successfully received by the receiver.

Parallel Channels with Same Transmission Rate: SW-ARQ

In this section, we have extended the result given in [5], where all the channels are assumed to be identical, to the case that all the channels have the same transmission rate but different packet error rate. Similar to the process given in [5], we define $N_t(i)$ to be the expected number of transmissions that a packet has to go through, given that its first transmission is placed on channel i for $i = 1, \cdots, M$. Let $\mathbf{P}_i$ be the probability that the first transmission of an arbitrarily selected packet is placed on channel i. Then the average number of transmissions that an arbitrarily selected packet has to go through before it is successfully received by the receiver is $N_t = \sum_{i=1}^{M} \mathbf{P}_i N_t(i)$.

The leading packet is the packet whose retransmission depends only on whether its own transmission is erroneous, and its erroneous transmission will force all of its following packets to be retransmitted. According to our transmission scheme, the leading packet is the packet that is placed on channel number 1. Thus, $N_t(1)$ is just the average number of transmissions for the leading packet to go through, which is given as $N_t(1) = 1/(1 - P_1)$.

Notice that when a packet is transmitted on channel number 2, the errors of the leading packet (the packet transmitted on channel number 1) will cause the retransmission of that packet. When the leading packet is received successfully, which will take $N_t(1)$ transmissions on average, there are two cases of the packet on channel number 2: it is received successfully together with the leading packet; or it is rejected by the receiver. For the second case, in the next transmission, the packet initially transmitted on channel number 2 becomes the leading packet and will take $N_t(1)$ transmissions on average before it is received successfully. Based on the above analysis, we get

$$\begin{aligned} N_t(2) &= (1 - P_2)N_t(1) + P_2N_t(1) + P_2N_t(1) \\ &= \frac{1 - P_2}{1 - P_1} + \frac{P_2}{1 - P_1} + \frac{P_2}{1 - P_1} = \frac{1 + P_2}{1 - P_1}. \end{aligned} \tag{2.25}$$

Similarly,

$$\begin{aligned} N_t(3) &= (1 - P_2)(1 - P_3)Nt(1) + (1 - P_2)P_3N_t(1) + (1 - P_2)P_3N_t(1) \\ &\quad + P_2N_t(1) + P_2N_t(2) \\ &= \frac{(1 - P_2)(1 - P_3)}{1 - P_1} + \frac{(1 - P_2)P_3}{1 - P_1} + \frac{(1 - P_2)P_3}{1 - P_1} + \frac{P_2}{1 - P_1} + \frac{P_2(1 + P_2)}{1 - P_1} \\ &= \frac{1 + P_2 + P_3 + P_2^2 - P_2P_3}{1 - P_1}. \end{aligned} \tag{2.26}$$

If we keep on doing the same process, we get

$$N_t(i) \approx \frac{1 + \sum_{k=2}^{i} P_k + \sum_{k=2}^{i-1} P_k^2 + \sum_{\{k,l:2\le k<l\le i, k+l\le i+1\}} P_kP_l - \sum_{\{k,l:2\le <l\le i, k+l>i+1\}} P_kP_l}{1 - P_1} \tag{2.27}$$

where all the higher order terms with order ≥ 3 are ignored.

In order to obtain $\mathbf{P}_i$, we further define $\mathbf{Y}_i$ to be the probability that a new packet is transmitted on channel i in a block, given that there is at least one new packet attempted across channel $1, \cdots, M$ in that block. Under joint operation, we get

$$\begin{aligned}
\mathbf{Y}_1 &= \frac{\prod_{k=1}^{M}(1-P_k)}{1-P_1} = \prod_{k=2}^{M}(1-P_k), \\
&\vdots \\
\mathbf{Y}_i &= \frac{\prod_{k=1}^{M}(1-P_k) + \prod_{k=1}^{M-1}(1-P_k)P_M + \cdots + \prod_{k=1}^{M-i+1}(1-P_k)P_{M-i+2}}{1-P_1} \\
&= \prod_{k=2}^{M-i+1}(1-P_k), \text{ and} \\
&\vdots \\
\mathbf{Y}_M &= \frac{\prod_{k=1}^{M}(1-P_k) + \prod_{k=1}^{M-1}(1-P_k)P_M + \cdots + (1-P_1)P_2}{1-P_1} = 1. \qquad (2.28)
\end{aligned}$$

Let $\mathbf{Z} = \mathbf{Y}_1 + \cdots + \mathbf{Y}_M$, where $\mathbf{Z}$ represents the average number of new packets transmitted across channel $1, \cdots, M$ in a block. Then $\mathbf{P}_i = \mathbf{Y}_i/\mathbf{Z}$.

For SW-ARQ, the average transmission delay can be obtained as

$$\text{Delay}_{SW} = (N_t - 1) * NT + T \qquad (2.29)$$

where T is the block transmission time.

Parallel Channels with Same Transmission Rate: GBN-ARQ

For any packet whose first transmission is placed on channel i for $i = 1, \cdots, M$, the number of retransmissions depends not only on whether its own transmission is erroneous but also the transmission of the packets previous to this packet in the same block and the packets in the previous $N-1$ blocks. As shown in [5], these packets can be classified into two categories.

- Packets $1, \cdots, i, M+1, \cdots, M+i, \cdots, (N-1)M+1, \cdots, (N-1)M+i$, in which the occurrence of the leading failure results in adding extra N slots to the delay of the test packet. The total number of packets in this category is $i * N$,

each i packets in the same block contributing an average of $N_t(i) - 1$ leading failures. Thus, the total contribution to the average delay of the test packet is $N * [N_t(i) - 1] * N$ slots.

- Packets $i+1, \cdots, M, M+i+1, \cdots, 2M, \cdots, M(N-2)+i+1, \cdots, M(N-1)$, in which the occurrence of a leading error adds extra $N-1$ slots in the first round, and adds N slots starting from the second round to the delay of the test packet. The probability of having a leading failure in packets $i+1, \cdots, M$ is $f_{m-i} = 1 - \prod_{k=i+1}^{M}(1-P_k)$. Thus, the average contribution to the delay of the test packet $i+1, \cdots, M$ is $f_{m-i}(N-1) + [N_t(M) - N_t(i) - f_{m-i}]N \times (N-1)$. So the total contribution to the average delay of the test packet from this category is $(N-1)\{f_{m-i}(N-1) + [N_t(M) - N_t(i) - f_{m-i}]N\}$.

Let $D(i)$ be the average transmission delay of a packet whose first transmission takes place in a slot on channel i. From the description above, we get

$$D(i) = [N_t(i)-1]*N^2 + f_{m-i}(N-1)^2 + [N_t(M)-N_t(i)-f_{m-i}]*(N-1)*N+1. \quad (2.30)$$

For GBN-ARQ, the average transmission delay can be obtained as

$$\text{Delay}_{GBN} = \sum_{i=1}^{M} D(i)\mathbf{P}_i \times T. \quad (2.31)$$

Parallel Channels with Same Transmission Rate: SR-ARQ

Due to the fact that in SR-ARQ a retransmission request for a packet does not depend on the transmission status of other packets, the transmission delay of the multichannel system is just the average of the transmission delay of each channel.

$$\text{Delay}_{SR} = \left[\frac{(\sum_{i=1}^{M} P_i) \times N}{\sum_{i=1}^{M}(1-P_i)} + 1 \right] \times T. \quad (2.32)$$

Note that for selective repeat ARQ, packets arrive at the receiver out of order and have to be resequenced. The overall delay of SR-ARQ should be the summation of the transmission delay and the resequencing delay. Even though the packet-to-channel assignment rule does not affect the channel utilization and the transmission

delay of SR-ARQ as we have shown in equation 2.29, it is expected to affect the resequencing delay of the multichannel SR-ARQ protocol. We have not considered the resequencing delay in our analysis. Shacham and Shin [37] have described and analyzed a multichannel SR-ARQ protocol which provided lower resequencing delay over parallel channels with the same transmission rate but different packet error rates than assigning packets to channels statically.

Parallel Channels with Different Transmission Rates

Since the M parallel channels may have different transmission rates, if the transmission rate of the original transmission channel is higher than the transmission rate of the retransmission channels, then the packet to be retransmitted might be chopped into several segments and retransmitted on several channels. If the transmission rate of the original transmission channel is lower than the transmission rate of the retransmission channel, then the packet to be retransmitted might be combined with some other information to form a new packet. So it is nonsense to analyze the transmission delay of the whole packet transmitted on one particular channel. In order to simplify the analysis, we use the concept of *sub-packet*. First, we assume that the transmission rates of the M channels have a common factor R. Let the transmission rate of the mth channel be R_m for $m = 1, \cdots, M$.. Then the packet transmitted on the mth channel will be divided into $k_m = R_m/R$ sub-packets, where the information data within each sub-packet will always be transmitted or retransmitted together. The transmission delay is measured from the instant at which a *sub-packet* is transmitted for the first time till the time it is successfully received by the receiver.

Let $[k_1, k_2, \cdots, k_M]$ be the number of sub-packets transmitted on the M parallel channels. Let $K = \sum_{j=1}^{M} k_j$ be to total number of sub-packets transmitted on the M channels. The M parallel channels can be treated as K pseudo-channels, where the transmission rates of all pseudo-channels are the same, while the packet error rate of each pseudo-channel is equal to the packet error rate of the real channel it is located on. That is $R'_1 = R'_2 = \cdots = R'_K = R$, $P'_1 = \cdots = P'_{k_1} = P_1$,

$P'_{k_1+1} = \cdots = P'_{k_1+k_2} = P_2$, and so on. The result we got in Section 2.5.1 can be applied on the pseudo-channels since the transmission rates of the pseudo-channels are equal.

For SR-ARQ, the average transmission delay is

$$\text{Delay}_{SR} = \left[\frac{(\sum_{j=1}^{K} P'_j) \times N}{\sum_{j=1}^{K} (1 - P'_j)} + 1 \right] \times T. \tag{2.33}$$

According to our analysis in Section 2.5.1, the average delay of the sub-packet for SW-ARQ and GBN-ARQ will be given as (2.29) and (2.31) where the summation is taken from 1 up to K. We need to evaluate the expression of $N_t(j)$ and $\mathbf{Y}_j$ for each pseudo-channel $j = 1, \cdots, K$.

For each pseudo-channel, we have

$$\begin{aligned}
\mathbf{Y}_1 &= \prod_{k=2}^{M} (1 - P_k), \\
\mathbf{Y}_2 &= \prod_{k=2}^{N_2} (1 - P_k), \\
&\vdots \\
\mathbf{Y}_j &= \prod_{k=2}^{N_j} (1 - P_k), \text{ and} \\
&\vdots \\
\mathbf{Y}_M &= 1
\end{aligned} \tag{2.34}$$

where

$$N_j = \begin{cases} L - 1 & \exists L \in \{1, \cdots, M\} : k_L + \cdots + k_M < j, k_{L-1} + \cdots + k_M \geq j \\ M & \text{otherwise} \end{cases}$$

It is a little involved to get the expression of $N_t(j)$ for $j = 1, \cdots, K$, since it will depend on the relationship of the number of pseudo-channels on each real channel and the transmission states. Fortunately, for all the sub-packets transmitted on the pseudo-channels of one real channel, we have

$$N_t(j) \approx N_t(i), j\text{th pseudo-channel} \in i\text{th channel}, \tag{2.35}$$

with some difference on the second and higher order terms which can be ignored if the packet error rate is small.

2.5.2 Simulation Results of Transmission Delay

Accuracy of the Transmission Delay Expressions

Computer simulations were used to assess the accuracy of the transmission delay expressions derived in Section 2.5.1. For our numerical example, we consider a four channel system (i.e. $M = 4$) where the idle time of SW-ARQ is $D = 2$ block times and the round-trip delay of the GBN-ARQ is $N = 3$ block times. We will first examine the results for the parallel channel with the same transmission rate, where the transmission rates of all the channels are 1 bit/symbol, while the ratio of the packet error rates between adjacent channels P_{i+1}/P_i is a constant (which we call r_P) for $i = 1, 2, \ldots, M-1$.

The results are summarized in Figure 2.9. As we can see from Figure 2.9, for SW-ARQ and GBN-ARQ, the analytical expression and the simulation results diverge as the packet error rates getting larger. This is because we ignore the higher order terms with order ≥ 3 in (2.27). For small packet error rate, the simulation results matched the analytical expressions exactly. For SR-ARQ, the simulation results matched the analytical expressions exactly.

Next, we will examine the results for the parallel channel with different transmission rates, where we let: P_{i+1}/P_i be a constant (which we call r_P) for $i = 1, 2, \ldots, M-1$; R_{i+1}/R_i be a constant r_R for $i = 1, 2, \ldots, M-1$.

The results are summarized in Figure 2.10. For SW-ARQ and GBN-ARQ, the analytical expression and the simulation result diverge even faster as the packet error rates getting larger. This is because we ignore both the higher order terms with order ≥ 3 and some terms with order 2 in (2.35).

Comparison of Different Assignments in AWGN Channel

Instead of deriving the packet-to-channel assignment ordering rule that minimized the transmission delay for SW-ARQ and GBN-ARQ, we will compare the

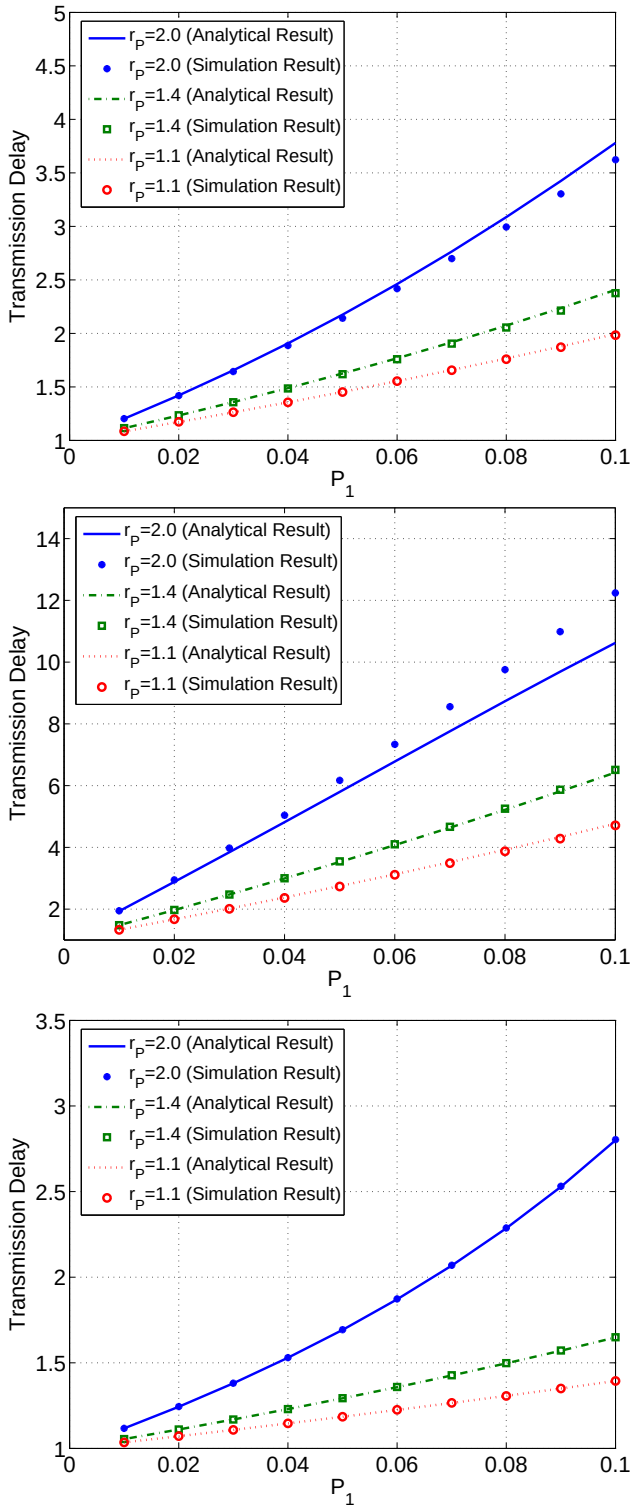


Figure 2.9: Comparison of computer simulations with analytical expressions for transmission delay for the three ARQ protocols: (top) stop-and-wait; (middle) go-back-N; (bottom) selective-repeat. In all cases, the parallel channels have the same transmission rate, $R = 1$ bits/symbol. The ratio of the packet error rate between adjcent channels, r_P, is kept to be a constant.

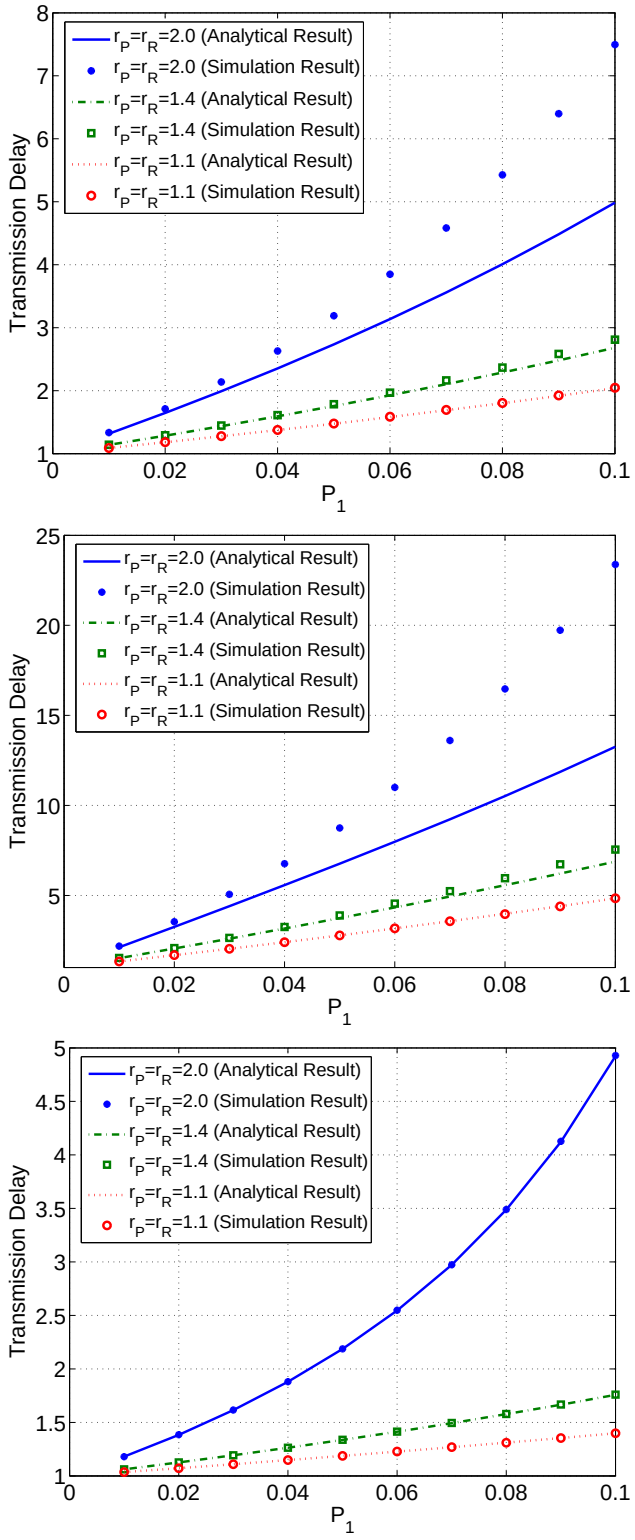


Figure 2.10: Comparison of computer simulations with analytical expressions for transmission delay for the three ARQ protocols: (top) stop-and-wait; (middle) go-back-N; (bottom) selective-repeat. In all cases, $R_1 = 1$ bit/symbol and $r_P = r_R$ to simplify the presentation.

optimal rule (OR) that maximize the channel utilization and the three other rules (DRR, SR, and SRR) we defined in Section 2.4.

The reduction of the optimal rule over the other rules on the transmission delay is given as

$$R = \frac{\text{Delay}_{\text{x}} - \text{Delay}_{\text{OR}}}{\text{Delay}_{\text{x}}} \tag{2.36}$$

where x is DRR, SR or SRR.

The three cases defined in Section 2.4.2 have been simulated. For the first case, Figure 2.11 shows the reduction of the transmission delay R as a function of P_1 and r_P for SW-ARQ and GBN-ARQ. For the second case, Figure 2.12 shows the reduction of the transmission delay R as a function of P_1 and r_R for SW-ARQ and GBN-ARQ. For the third case, Figure 2.13 shows the reduction of the transmission delay R as a function of r and P_1 for SW-ARQ and GBN-ARQ.

The simulation results presented above suggest the following:

1. The reductions of the transmission delay of OR over the other rules are always positive, which means the optimal rules we derived in 2.3 that maximize the channel utilization also minimize the transmission delay for SW-ARQ and GBN-ARQ.

2. Assigning a packet to be retransmitted to the "worst" available channel produces the greatest accretion in overall transmission delay. This is demonstrated by the fact that the DRR consistently has the worst transmission delay.

3. The two "static" assignment rules do not achieve the transmission delay of the OR, which is a "dynamic" assignment rule. This suggests that a properly ordered dynamic assignment rule is needed to optimize transmission delay.

4. The reduction of the OR over the other rules considered increases as the difference between the channels becomes more pronounced. This is illustrated by the fact that R increases as P_1 increases (see Figures 2.11 – 2.13), as r_P increases (see Figure 2.11), as r_R decreases (see Figure 2.12), or as r increases (see Figure 2.13).

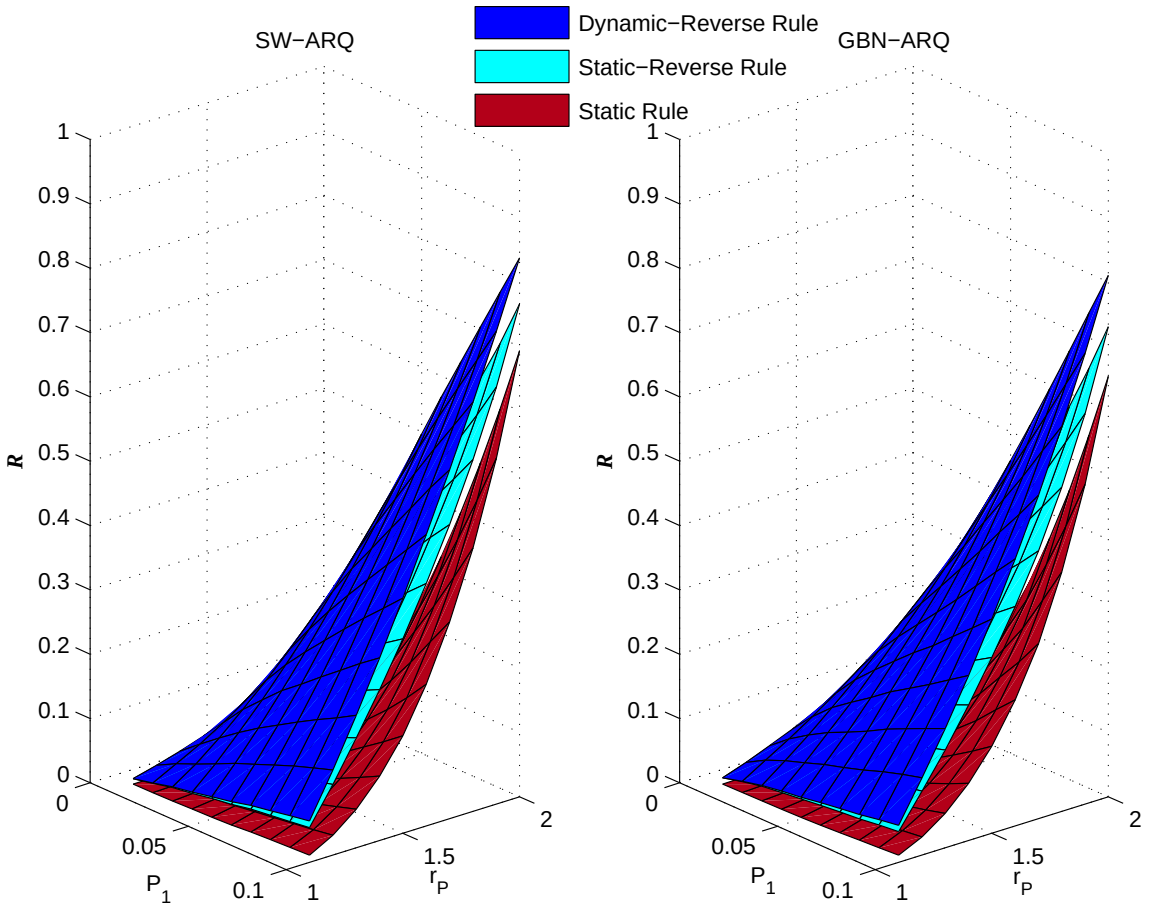


Figure 2.11: Transmission delay comparison for the case where the channels have the same transmission rate but different packet error probabilities.

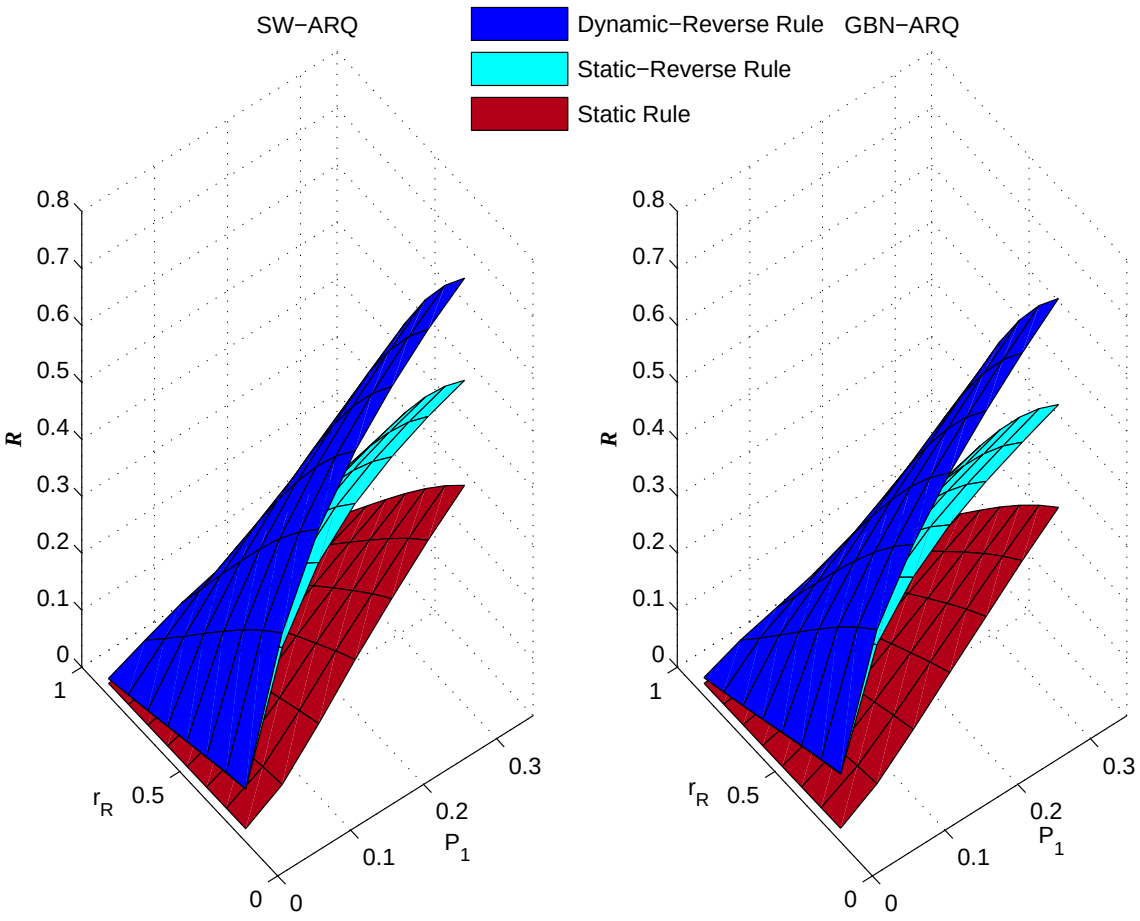


Figure 2.12: Transmission delay comparison for the case where the channels have the same packet error probability but different transmission rates.

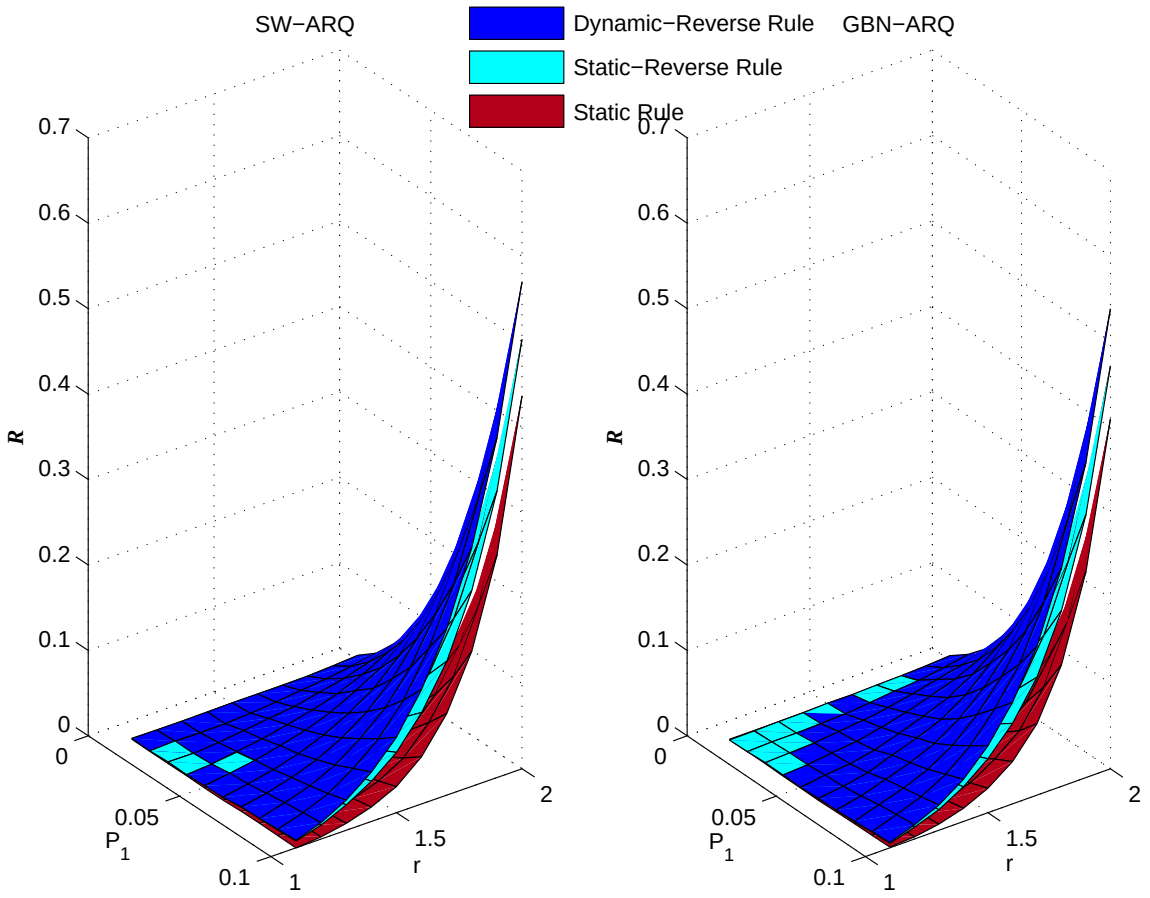


Figure 2.13: Transmission delay comparison for the case where all the channels have different packet error probabilities and different transmission rates.

2.6 Conclusions

In this chapter, we have shown that the SW and GBN retransmission protocols must be generalized when used in a multichannel communication system. The generalization takes the form of packet-to-channel assignment rules. A general condition governing the packet-to-channel assignment rules was derived and important special cases were pointed out. Simulation results were used to demonstrate the that 1) the packet-to-channel assignment impacts channel utilization and transmission delay when the channels are different and 2) the optimal assignment rule produces a channel utilization (transmission delay) that is better than the channel utilization (transmission delay) that results from doing nothing. The gain (reduction) increases as the difference between the channels becomes more pronounced.

Chapter 3

Type-I Hybrid-ARQ Using MTCM STVC for MIMO Systems

In Chapter 2, we have shown that in parallel multichannel communication systems, the conventional ARQ protocols for single channel communications have to be generalized in order to realize the full potential of parallel multi-channel communications. In this chapter, we will show the performance improvement of employing type-I hybrid-ARQ scheme in MIMO systems where we assume that the channel state information is available at both the transmitter and the receiver. An idealistic retransmission protocol that maximizes the channel utilization is described and analyzed.

3.1 Introduction

Due to the fact that a multiple-input multiple-output (MIMO) channel can offer a significant capacity gain over a traditional single-input single-output (SISO) channel, MIMO systems are today regarded as one of the most promising research areas of wireless communications. In order to approach the theoretical limit of such systems, techniques that exploit maximally the spatial diversity brought by the multiple antennas are used. However, the complexity of these techniques grows enormously as channel utilization approaches the theoretic channel capacity limit. It is well known that hybrid-ARQ techniques can be used to improve the bit error rate performance of coded signaling schemes without a significant increase in decoder complexity. The cost of this approach is throughput which will reduce the overall channel utilization. In this chapter, we combine type-I hybrid-ARQ techniques with Spatio-Temporal

Vector Coding (STVC) [31] over a quasi-static flat-fading MIMO channel. The complexity versus channel utilization trade-off demonstrates that hybrid-ARQ offers an effective solution to the complexity problem.

Tu [46] proposed a Multi-Channel Modulation (MCM) technique which is based on the concept of creating multiple mutually orthogonal subchannels over which independent streams of data can be sent without inter-channel interference. Raleigh [31] extended MCM to the MIMO channel and named it spatio-temporal vector-coding (STVC) and suggested the use of Multidimensional Trellis Coded Modulation (MTCM) [30, 10] for the coding scheme in the STVC structure.

In order to demonstrate the concept, we use the Yamamoto-Itoh algorithm given in [52]. Rasmussen and Wicker [32] proposed a method for modifying MTCM systems for use in type-I hybrid-ARQ protocols over AWGN and fading SISO channel based on the Yamamoto-Itoh algorithm [52].

In this chapter, we demonstrate how type-I hybrid-ARQ protocols can be used in the slowly varying MIMO channel using STVC. The MTCM-HARQ scheme provides improved channel utilization and reliability performance relative to that provided by a FEC scheme over slowly varying MIMO channel.

This chapter is organized as follows. The system model is described in Section 3.2. In Section 3.3 the performance of the protocol is considered for the very slowly varying MIMO channel. Upper and lower bounds for the decoded probability of error and retransmission will be derived. Section 3.4 contains a series of examples that include both analytical and simulation results. Conclusions are given in Section 3.5.

3.2 System Model

3.2.1 Channel Model

A discrete-time baseband channel model with M_T transmission antennas and M_R receiver antennas is considered. With the flat-fading assumption, the channel matrix $\mathbf{H}$ with dimension $(M_R \times M_T)$ describes the channel. Entries of $\mathbf{H}$, h_{ij} represent the equivalent baseband channel gains between jth transmit and ith receive antennas. The h_{ij} are usually modeled as normalized, circularly symmetric complex Gaussian

random variables. It is assumed that the channel varies very slowly, so that $\mathbf{H}$ can be assumed to be constant during one packet interval but vary from packet to packet.

In the discrete-time model, let $\mathbf{Z} = [Z_1, Z_2, \cdots, Z_{M_T}]^T$ be the transmitted symbols on M_T antennas, where

$$Z_n = [z_{n,1}, z_{n,2}, \cdots, z_{n,N+w}] \qquad n = 1 \cdots M_T, \tag{3.1}$$

N is the packet length, and w is the number of dummy symbols required to zero-out the encoder shift registers. The corresponding received symbols on the M_R receiver array, $\mathbf{R} = [R_1, R_2, \cdots, R_{M_R}]^T$, is

$$\mathbf{R} = \mathbf{HZ} + \mathbf{n} \tag{3.2}$$

where

$$R_i = [r_{i,1}, r_{i,2}, \cdots, r_{i,N+w}] \qquad i = 1 \cdots M_R \tag{3.3}$$

and $\mathbf{n}$ is the equivalent baseband noise whose elements are zero-mean complex Gaussian random variables noise with variance $\sigma^2 = N_0/2$.

3.2.2 Spatio-Temporal Vector Coding

STVC is best described using the singular value decomposition (SVD) of the channel matrix $\mathbf{H}$. Let $\mathbf{H} = \mathbf{U}_H \mathbf{\Lambda}_H \mathbf{V}_H^*$ be the SVD of $\mathbf{H}$, where $\lambda_{H,n}$ is the n-th singular value. If the channel state information is available at both the transmitter and the receiver, the STVC parallel channel [31] is written

$$\mathbf{R} = \mathbf{U}_H^* \mathbf{H} \mathbf{V}_H \mathbf{Z} + \nu = \mathbf{U}_H^* \mathbf{U}_H \mathbf{\Lambda}_H \mathbf{V}_H^* \mathbf{V}_H \mathbf{Z} + \nu = \mathbf{\Lambda}_H \mathbf{Z} + \nu. \tag{3.4}$$

Thus the MIMO channel can be considered as a set of parallel independent subchannels whose number equals the rank of $\mathbf{H}$ matrix, and the subchannel gains are given by the singular values $\lambda_{H,n}$.

In a parallel channel communication system, power and bit distributions algorithms [46], [31] are used to maximize data rate subject to a probability of error constraint. The rate maximizing water-filling solution and bit allocation for the STVC

channel becomes

$$\lambda_{Z,n} = \left(\xi - \frac{\sigma^2\alpha}{|\lambda_{H,n}|^2}\right)^+ \tag{3.5}$$

$$b_n = \log_2\left(1 + \frac{\lambda_{Z,n}|\lambda_{H,n}|^2}{\alpha\sigma^2}\right) \qquad \text{bit/symbol} \tag{3.6}$$

where α is the code gap [1] [31] which is determined by the probability of error of the system. The energy of the transmitted packet Z_n on each subchannel n is $\lambda_{Z,n}$. The energy in each subchannel at the receiver equals the energy at the transmitter $\lambda_{Z,n}$ multiplied by the channel gain $\lambda_{H,n}$.

Under the quasi-stationary channel condition the number of bits calculated by (3.6) is constant on each subchannel during one packet interval, so we use an MTCM encoder which can support b_n bits/symbol on subchannel n during one packet interval. The MTCM encoder, as illustrated in Figure 3.1, consists of three sections. These sections are the binary convolutional encoder, the multi-D signal set mapper, and the 2-D signal set mapper.

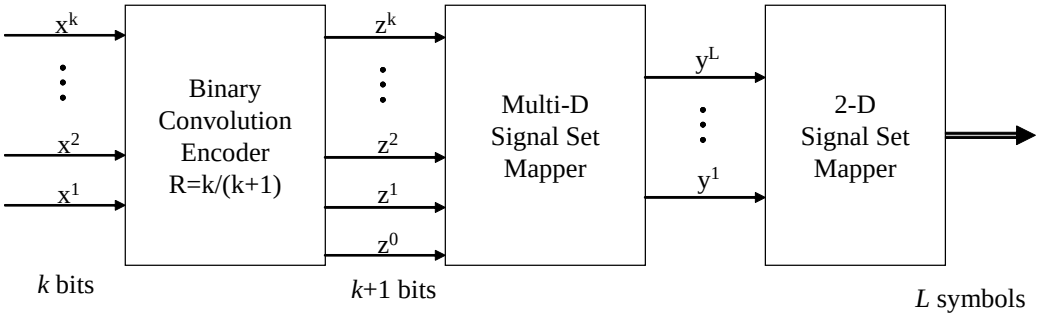


Figure 3.1: General multidimensional trellis code modulation encoder.

The *efficient rate*, R_{eff}, of the code is defined as the average number of information bits transmitted during each 2-D signal period T. With k bits input into the encoder and L 2-D symbols output, we have $R_{\text{eff}} = k/L$. The efficient

[1] The code gap is difference between the SNR required to achieve the target probability of error at the desired data rate and the SNR required to achieve a theoretical capacity equal to the desired data rate.

rate is the indicator of the bit resolution (granularity). It is not possible to achieve infinite granularity with MTCM codes. The granularity of possible bit allocations is determined by the dimensionality of the MTCM code [46].

3.2.3 MTCM and Type-I Hybrid-ARQ

The key to modifying an FEC decoder for use in a type-I hybrid-ARQ protocol is the identification of a source of reliability information within the decoding process [50]. The path metrics calculated during Viterbi decoding provide such a source of information. In [32] and [52], the surviving path and the best non-surviving path are compared at each node at each stage in the decoding process. If the difference in the path metrics of the two paths is smaller than a threshold u, then the survivor is declared unreliable. Once a surviving path is declared unreliable, it will remain so regardless of the results of later comparisons. If all survivors are labeled as unreliable at some level, then the decoder requests a retransmission of the packet.

A simple retransmission protocol is described as follows: At a certain transmission time, suppose a packet is transmitted over a subchannel which can support s bits/symbol. Now suppose the modified MTCM decoder produces a retransmission request for that packet. At the next transmission time, there are k =rank($\mathbf{H}$) subchannels available for the retransmission. For now, we only consider the case that the retransmission subchannel is chosen as the subchannel which can support exactly s bit/symbol. The other subchannels are used to transmit new packets. If there is no suitable subchannel, the packet will be stored in the buffer until a suitable subchannel is available. Clearly, this retransmission protocol uses the channel with the best efficiency even though it could have very poor delay performance. As such, the performance of this retransmission protocol is an upper bound on the performance of any real retransmission protocol.

3.3 Performance Analysis

Through an extension of the techniques used in Rasmussen *et al.* [32] and Biglieri *et al.* [3], upper and lower bounds for the reliability and channel utilization

performance of the scheme are derived for the case of the very slowly varying MIMO channel. The channel utilization, χ is defined as the average achievable data rate over the MIMO channel measured in bits/symbol.

3.3.1 Upper Bound on Bit Error Rate

Assume Z_n is the packet transmitted over subchannel n. According to equations (3.5) and (3.6), for fixed values of σ and α, the energy of packet Z_n and the number of bits that Z_n can support are functions of $\lambda_{H,n}$, the n-th singular value of the channel matrix $\mathbf{H}$. The number of bits calculated by equation (3.6) determines the properties of the MTCM encoder over the subchannel. That is, the free minimum squared Euclidean distance d^2_{free}, the number of nearest neighbors N_{free}, and the transfer function of the error state diagram $T(D, I)$ are also functions of $\lambda_{H,n}$.

Now consider a packet Z transmitted on certain subchannel, with transmit energy λ_Z and channel gain λ_H, which can support $b(\lambda_H)$ bits/symbol. The MTCM code can be characterized by $d^2_{\text{free}}(\lambda_H)$, $N_{\text{free}}(\lambda_H)$, and $T(D, \lambda_H)$. Let Z_L be a trellis path that diverges from Z and remerges after precisely L branches. The pairwise error probability $P\{Z \to Z_L|\lambda_H, u\}$ shall be defined as the probability that Z_L survives as a reliable path when compared with Z at the node at which that two paths remerge under the condition that the singular value of the channel is λ_H, i.e.,

$$P\{Z \to Z_L|\lambda_H, u\} = P\{\|\sqrt{\lambda_Z}\lambda_H Z - R\|^2 - \|\sqrt{\lambda_Z}\lambda_H Z_L - R\|^2 \geq u\} \quad (3.7)$$

where R is the received packet. Using well known techniques, we have

$$\begin{aligned} P\{Z \to Z_L|\lambda_H, u\} \leq & \mathrm{Q}\left(\sqrt{\frac{\lambda_Z\lambda_H^2 d^2_{\text{free}}(\lambda_H) + 2u}{2N_0}}\right) \\ & \times \exp\left(\frac{\lambda_Z\lambda_H^2 d^2_{\text{free}}(\lambda_H)}{4N_0}\right) \\ & \times \exp\left[-\frac{\lambda_Z\lambda_H^2\|Z - Z_L\|^2}{4N_0}\right]. \end{aligned} \quad (3.8)$$

A union bound argument can be used to derive an upper bound on P_e, the probability that an error event occurs during the decoding of a packet received on

a given transmission. Since trellis codes are in general nonlinear, the summation is taken over all possible transmitted code packets Z as well as over all possible error event lengths L:

$$P_e\left(\lambda_H, u\right) \leq \sum_{L=1}^{N_{\text{stages}}} \sum_{Z} \sum_{Z_L} P\{Z\} \cdot P\{Z \rightarrow Z_L | \lambda_H, u\} \tag{3.9}$$

where N_{stages} is the number of stages in the decoding trellis. Following [32], substituting (3.8) into (3.9) produces the upper bound

$$P_e\left(\lambda_H, u\right) \leq \mathrm{Q}\left(\sqrt{\frac{\lambda_Z \lambda_H^2 d_{\text{free}}^2(\lambda_H) + 2u}{2N_0}}\right) \times \exp\left(\frac{\lambda_Z \lambda_H^2 d_{\text{free}}^2(\lambda_H)}{4N_0}\right) T(D, \lambda_H)\Big|_{D=\gamma} \tag{3.10}$$

where

$$\gamma = \exp\left(-\frac{\lambda_Z \lambda_H^2}{4N_0}\right). \tag{3.11}$$

$T(D, \lambda_H)$ is used to enumerate the squared Euclidean distances $\|Z - Z_L\|^2$. We assume that the bit allocation algorithm is same for the first transmission and all subsequent retransmissions of the packet. (If not, the analysis will be more complex. The error probability is composed of two parts: the error probability on the first transmission and the error probability on any subsequent retransmissions. The error probability for a retransmitted packet is a function of the retransmission probability $P_{r,\text{ML}}$ discussed in the next Section.)

$P_e(u)$ can be calculated by averaging over λ_H:

$$P_e(u) \leq \int f(\lambda_H) P_e(\lambda_H, u) d\lambda_H \tag{3.12}$$

where $f(\lambda_H)$ is the distribution of the singular values of the channel matrix $\mathbf{H}$. The distribution of the singular values of the channel matrix can be obtained through channel measurement. In this dissertation, the evaluation of (3.12) are done using Monte Carlo simulations.

The upper bound on the probability of bit error in a packet received on a given transmission is thus

$$P_b(u) \leq \int \frac{f(\lambda_H)}{b(\lambda_H)} \mathrm{Q}\left(\frac{\sqrt{\lambda_Z \lambda_H^2 d_{\text{free}}^2(\lambda_H) + 2u}}{\sqrt{2N_0}}\right) \times \exp\left(\frac{\lambda_Z \lambda_H^2 d_{\text{free}}^2(\lambda_H)}{4N_0}\right) \frac{\partial}{\partial I} T(D, I, \lambda_H)\Big|_{I=1, D=\gamma} d\lambda_H \quad (3.13)$$

where γ is defined by (3.11) and $b(\lambda_H)$ is the number of bits transmitted on the channel with singular value λ_H. Note that since the channel gains $h_{i,j}$ are assumed known and are also assumed constant over the packet, the error probability is minimized by using a code designed to maximize d_{free}^2.

3.3.2 Lower Bound on Channel Utilization

According to [32], the following upper bound is obtained for the retransmission probability $P_r(\lambda_H, u)$ for the subchannel with singular value λ_H:

$$P_r(\lambda_H, u) \leq 1 - (1 - P_{r,\text{ML}}(\lambda_H, u))^{N_{\text{stages}}-1} \quad (3.14)$$

where

$$P_{r,\text{ML}}(\lambda_H, u) \leq \sum_{L=1}^{N_{\text{stages}}} \sum_{Z} P\{Z\} \sum_{Z_L} P\{Z \leftrightarrow Z_L | \lambda_H, u\} \quad (3.15)$$

is the probability that the comparison at the node through which the ML path passes results in a declaration of unreliability. The upper bound on $P_{r,\text{ML}}(\lambda_H, u)$ is [32]

$$P_{r,\text{ML}}(\lambda_H, u) \leq \left[\mathrm{Q}\left(\sqrt{\frac{\lambda_Z \lambda_H^2 d_{\text{free}}^2(\lambda_H) - 2u}{2N_0}}\right) + \mathrm{Q}\left(\sqrt{\frac{\lambda_Z \lambda_H^2 d_{\text{free}}^2(\lambda_H)}{2N_0}}\right)\right] \times \exp\left(\frac{\lambda_Z \lambda_H^2 d_{\text{free}}^2(\lambda_H)}{4N_0}\right) T(D, \lambda_H)\Big|_{D=\gamma} \quad (3.16)$$

where γ is defined by (3.11).

The channel utilization of the system is determined in part by the retransmission protocol as we have shown in Chapter 2. We assume that the bit allocation algorithm is same for the first transmission and all the retransmitted packets. Then the channel utilization of the MIMO system is the sum of the channel utilization of

each subchannel and averaged by the distribution of λ_H. If a selective-repeat scheme with infinite buffering is assumed [52], then the channel utilization of the MIMO system is bounded below by

$$\begin{aligned}\chi(u) &= \int \sum_{n=1}^{M_T} f(\lambda_{H,n}) \frac{N}{(N+w)} \cdot b(\lambda_{H,n}) \\ &\quad \times (1 - P_r(\lambda_{H,n}, u)) d\lambda_{H,n} \\ &\geq \int \sum_{n=1}^{M_T} f(\lambda_{H,n}) \frac{N}{(N+w)} \cdot b(\lambda_{H,n}) \\ &\quad \times (1 - P_{r,\mathrm{ML}}(\lambda_{H,n}, u))^{N_{\mathrm{stages}}-1} d\lambda_{H,n}. \end{aligned} \tag{3.17}$$

3.3.3 Lower Bound on Bit Error Rate

A lower bound on bit error rate can be obtained through the use of the probabilities of most significant error events [32]. Given a received packet R and the choice between Z and Z_d, the receiver will select the wrong packet if

$$\|\sqrt{\lambda_Z}\lambda_H Z - R\|^2 - \|\sqrt{\lambda_Z}\lambda_H Z_d - R\|^2 \geq u. \tag{3.18}$$

A lower bound on P_e is then derived as follows [32]

$$P_e(\lambda_H, u) \geq N_{\mathrm{free}}(\lambda_H) \mathrm{Q}\left(\frac{\lambda_Z \lambda_H^2 d_{\mathrm{free}}^2(\lambda_H) + u}{\sqrt{2N_0\lambda_Z}\lambda_H d_{\mathrm{free}}(\lambda_H)} \right) \tag{3.19}$$

and the probability of bit error in a received packet is bounded below by

$$P_b(u) \geq \int f(\lambda_H) \frac{N_{\mathrm{free}}(\lambda_H)}{b(\lambda_H)} \mathrm{Q}\left(\frac{\lambda_Z \lambda_H^2 d_{\mathrm{free}}^2(\lambda_H) + u}{\sqrt{2N_0\lambda_Z}\lambda_H d_{\mathrm{free}}(\lambda_H)} \right) d\lambda_H. \tag{3.20}$$

3.3.4 Upper Bound on Channel Utilization

Consider significant error events, the receiver will request a retransmission whenever the received packet R and the two suggested code packets Z and Z_d satisfy the condition

$$|\|\sqrt{\lambda_Z}\lambda Z - R\|^2 - \|\sqrt{\lambda_Z}\lambda_H Z_d - R\|^2| < u. \tag{3.21}$$

Using the same approach as before the following is obtained:

$$P_{r,\mathrm{ML}}(\lambda_H, u) \geq N_{\mathrm{free}}(\lambda_H)\mathrm{Q}\left(\frac{\lambda_Z\lambda_H^2 d_{\mathrm{free}}^2(\lambda_H) - u}{\sqrt{2N_0\lambda_Z}\lambda_H d_{\mathrm{free}}(\lambda_H)}\right) - N_{\mathrm{free}}(\lambda_H)\mathrm{Q}\left(\frac{\lambda_Z\lambda_H^2 d_{\mathrm{free}}^2(\lambda_H) + u}{\sqrt{2N_0\lambda_Z}\lambda_H d_{\mathrm{free}}(\lambda_H)}\right). \quad (3.22)$$

The channel utilization of the system is then upper bounded by

$$\chi(u) \leq \int \sum_{n=1}^{M_T} f(\lambda_{H,n})\frac{N}{(N+w)}b(\lambda_{H,n}) \times (1 - P_{r,\mathrm{ML}}(\lambda_{H,n}, u))^{N_{\mathrm{stages}}-1} d\lambda_{H,n}. \quad (3.23)$$

3.4 Analytical and Simulation Results

3.4.1 Tightness of Bounds

The following examples illustrate the tightness of the bounds and the performance improvements provided by modifying the MTCM decoder for type-I hybrid-ARQ error control over a slowly varying MIMO channel. In the simulation examples that follow, we use a set of six 4-dimensional 8-state trellis codes [30], each of which provides a different channel utilization measured in bits/symbol. The parity-check polynomials of these encoders written in octal are listed in Table 3.1, each of which provides a different R_{eff} measured in bits/T. These codes offer a bit granularity of 1/2 bit per 2-dimensional symbol. The number of transmitted symbols per block is $N = 64$. The transfer functions for the MTCM codes were derived using the method outlined in [27].

The channel is modeled as a simple 4×4 STVC channel. Since the bounds on channel utilization and error probability are a function of the random channel gains in the matrix $\mathbf{H}$, Monte Carlo simulations were used to produce the numerical values shown in the plots that follow.

Comparing actual performance to the bounds is not straightforward for an STVC system since an MTCM code is selected (from the set of available codes) to maximize the channel utilization while keeping the bit error rate constant. The tightness of the bounds is tested using the code gap α [31]. In general, the code gap decreases as the probability of error increases.

Table 3.1: Parity check polynomial of the four dimensional 8-state trellis codes.

R_{eff} (bit/T)	M-PSK	q	h^2	h^1	h^0	d^2_{free}	N_{free}	d^2_{next}	N_{next}
1.0	4	1	04	02	11	12	5	-	-
1.5	4	0	04	06	11	8	5	-	-
2.0	8	1	04	02	11	4	2	5.172	16
2.5	8	0	04	06	11	2.929	16	-	-
3.0	16	1	04	02	11	1.476	16	-	-
3.5	16	0	04	06	11	0.761	16	-	-

Simulations have been performed to test the tightness of the bounds. The simulations are based on the transmission of 10000 packets for $\alpha = 2$ dB, $\alpha = 3$ dB, and $\alpha = 4$ dB. For $\alpha = 5$ dB and $\alpha = 6$ dB, 40000 packets were sent. Figures 3.2 and 3.3 show bit error rate and channel utilization, respectively, as a function of the coding gap for SNR[2] = 10 dB and $u = 0.5$. As expected, the code gap decreases for increasing bit error rate as demonstrated in Figure 3.2. The relationship between channel utilization and code gap is more interesting. For large α, the channel utilization decreases with increasing α because the bit error rate decreases with increasing α as demonstrated before. For small α the channel utilization decreases with decreasing α since smaller α correspond to higher operating bit error rates which are accompanied by an increased number of retransmissions which reduce the channel utilization.

These figures demonstrate that the upper bound is much tighter than the lower bound. In the analytical results that follow, only the upper bound is used. It should be noted that the number of transmitted packets allows for reliable simulation data for the BER performance down to 10^{-5}. Bellow this threshold the reliability of the simulation data is degrading.

[2]SNR is defined as SNR= $P_T/(M_T M_R \sigma^2) \sum_{n=1}^{k=\text{rank}(\mathbf{H})} |\lambda_{H,n}|^2$

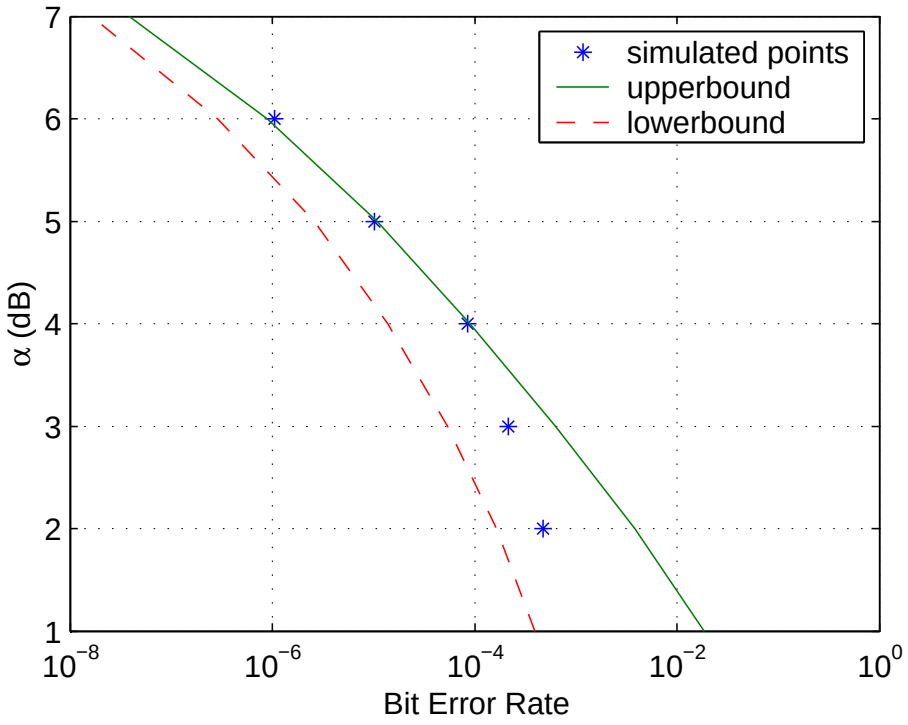


Figure 3.2: Code gap α versus decoded bit error rate for the set of MTCM codes used in the simulations. The upper bound and lower bounds for bit error rate are plotted for comparison. The system SNR is 10 dB and the MTCM decoder is modified for hybrid-ARQ error control using $u = 0.5$.

3.4.2 Analytical Results Using Water-Filling Algorithm

The upper bounds of channel utilization of the FEC and HARQ systems using practical coding schemes are compared with the information capacity for the discrete-time STVC channel which is given by [31]

$$C_{M_R,M_T} = \sum_{n=1}^{K} \log_2 \left(1 + \frac{\lambda_{Z,n}|\lambda_{H,n}|^2}{\sigma^2}\right) \qquad \text{bits/transmission} \tag{3.24}$$

where $\lambda_{Z,n}$ is found from the spatio-temporal water-filling solution

$$\lambda_{Z,n} = \left(\xi - \frac{\sigma^2}{|\lambda_{H,n}|^2}\right)^+. \tag{3.25}$$

Figures 3.4 and 3.5 show the upper bounds for χ as a function of SNR for $P_b = 10^{-6}$ and $P_b = 10^{-5}$, respectively, for different values of u ($u = 0$ corresponds to the FEC case). The information capacity for the 4×4 STVC is also plotted for comparison.

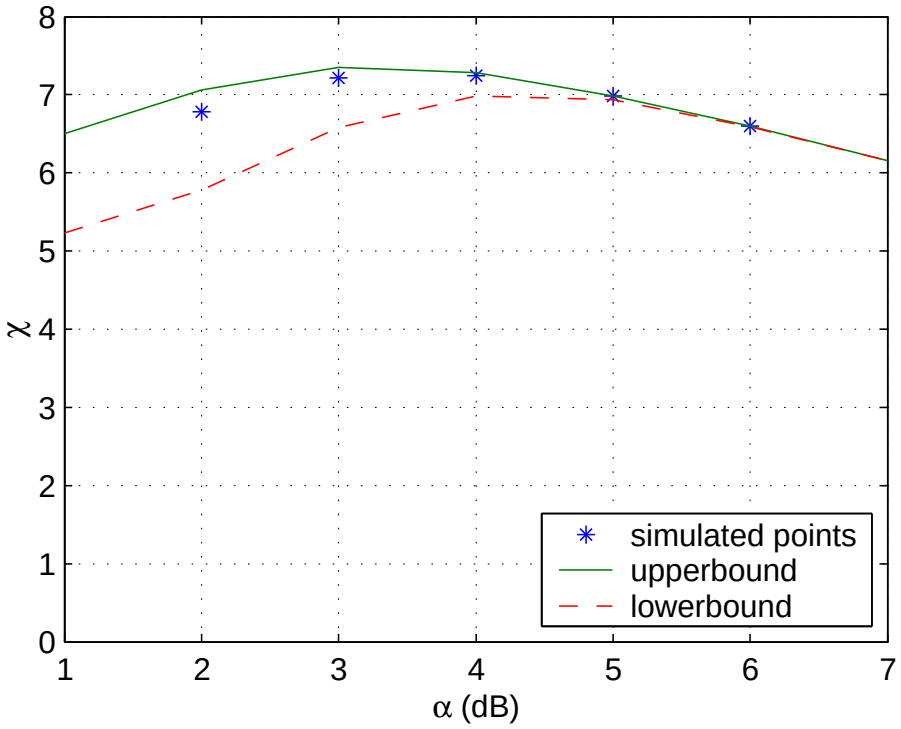


Figure 3.3: Channel utilization χ versus code gap α for the set of MTCM codes used in the simulations. The upper bound and lower bounds for channel utilization are plotted for comparison. The system SNR is 10 dB and the MTCM decoder is modified for hybrid-ARQ error control using $u = 0.5$.

The performance improvement relative to FEC is obvious. At SNR = 10 dB and bit error rate $P_b = 10^{-6}$, the code gap for the FEC system is 6 dB while for the hybrid-ARQ system using $u = 1.5$ it is 5.2 dB. In contrast, for an FEC system, the code gap can be reduced by either keeping the code fixed and allowing higher BERs or by keeping the BER fixed and using codes with more states. Figures 3.4 and 3.5 show that the code gap can be reduced by 0.8 dB with a ten-fold increase in BER from 10^{-6} to 10^{-5}, while [30] shows that the code gap can be reduced by about 1 dB by using a set of 64-state trellis codes. Thus HARQ systems allows us to decrease the code gap α without increasing complexity.

Note that it is not true, in general, that higher u offers smaller code gaps. Figure 3.5 shows that at a bit error rate of 10^{-5}, increasing u beyond 1.5 actually

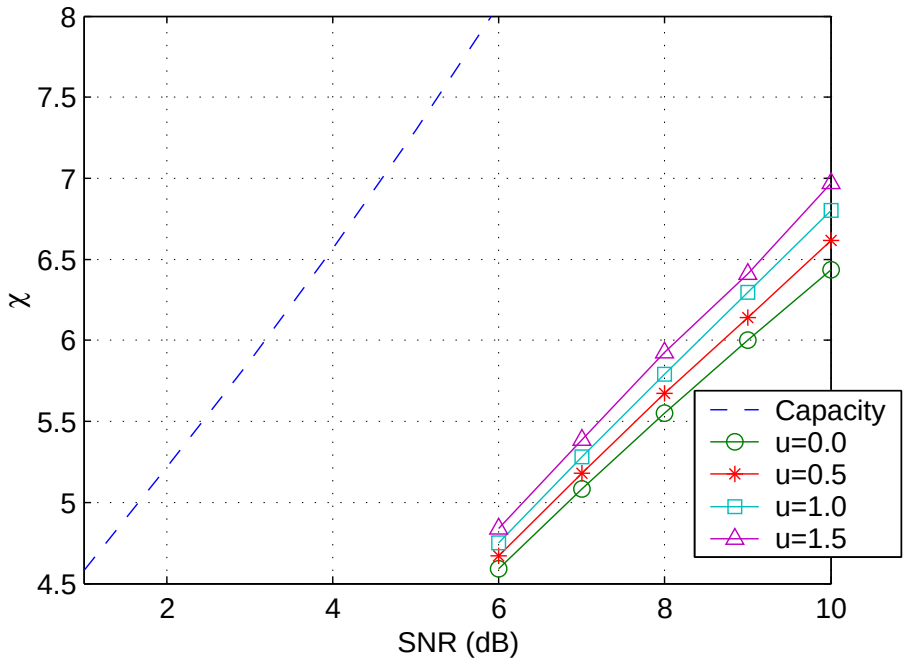


Figure 3.4: The upper bounds of channel utilization at $P_b = 10^{-6}$ using water-filling solution.

increases the code gap at SNR = 10 dB. This is a consequence of the increased number of retransmissions accompanying this value of u.

3.4.3 Analytical Results Using Bit-Loading Algorithm

For practical data transmission, the water-filling solution and bit allocation algorithm given in Section 3.2 is not optimal, due to the fact of finite granularity and non-unique code gap.

In reality, we can have only finite granularity. Suppose we have granularity Δ ($\Delta = 0.5$ in our case). *Ad hoc* operations such as rounding to an number $k \times \Delta$ of bits/symbol form the code for the method given in Section 3.2 may not be optimal. The code gap is not unique for different MTCM codes.

Because of finite granularity and non-unique code gap, bit-loading algorithm is used instead. One known optimal loading algorithm is the Hughes Hartog algorithm [4]. It initially assigns zero power to all subchannels and allocates power

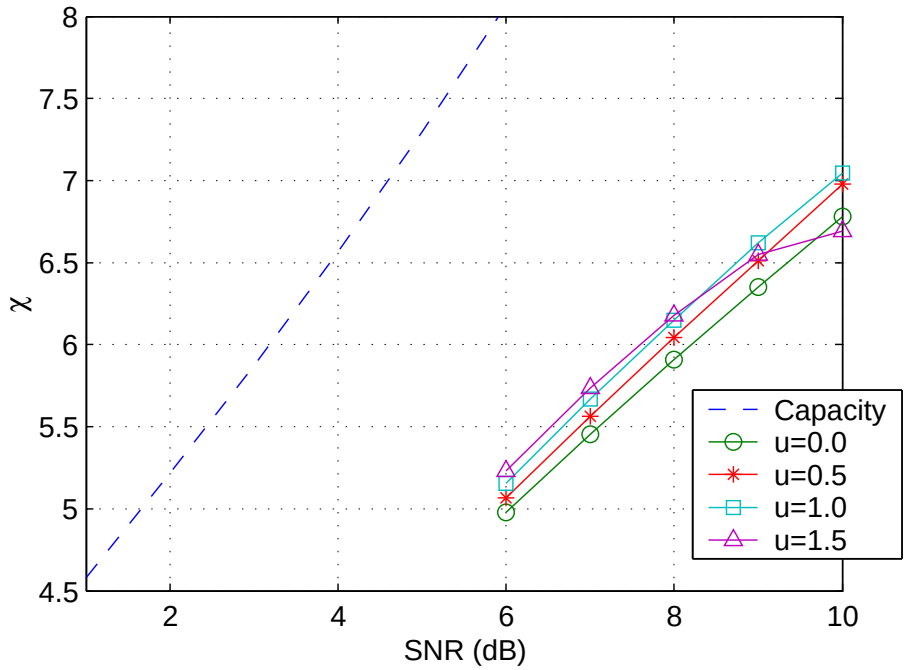


Figure 3.5: The upper bounds of channel utilization at $P_b = 10^{-5}$ using water-filling solution.

to the channel which needs the lowest amount to increase its rate by Δ bit/symbol. This process is iterated until the power budget is used or until the small remaining power cannot increase the rate of any channel.

In general, it is possible to use a two-dimensional look-up table, $S(P_e, s)$, that contains the required SNR for the desired BER (P_e) and the number of bits s. The code gap is included in the required SNR so that this bit loading algorithm allows non-unique code gap. The content of the look-up table are actually several sets of threshold for different desired BER (P_e). Instead of changing the α in the infinity granularity case to reach the desired BER, in the finite granularity case, we change the threshold to reach the desired BER. Once one value $S(P_e, s_k)$ in the look-up table is selected, the required transmit energy is simply

$$P_k(s_k) = \frac{S(P_e, s_k)\sigma^2}{|\lambda_{H,n}|^2}. \tag{3.26}$$

The look-up table $S(P_e, s)$ for our code set is shown in Table 3.2, where the SNR is

presented in dB. The upper bounds of channel utilization at $P_b = 10^{-5}$ are shown in Figure 3.6, which is compared with the result of the water-filling solution(dashed line). The channel utilization for the bit-loading algorithm is higher than the water-filling solution, which shows that the later is not optimal. Both of these two algorithms share the same pattern when μ is changing.

Table 3.2: The look-up table used in the bit-loading algorithm.

s (bit/symbol)	0.5	1	1.5	2	2.5	3	3.5
$P_e = 10^{-2}$	0.87	2.53	4.75	7.11	9.58	12.4	15.3
$P_e = 10^{-3}$	1.87	3.7	5.73	8.16	10.6	13.46	16.4
$P_e = 10^{-4}$	3.17	4.71	6.6	9.1	11.48	14.35	17.3
$P_e = 10^{-5}$	3.65	5.59	7.38	9.925	12.25	15.15	18.1

3.5 Conclusions

A system combining type-I hybrid-ARQ error control with adaptive-MTCM over a slowly varying MIMO channel has been analyzed. Upper and lower bounds for the decoded probability of error and retransmission were derived. The bound on the probability of retransmission was used to create a bound on the channel utilization, measured in bits/symbol. A simple, though ideal, retransmission scheme was defined and incorporated into the MTCM STVC system. Simulation results showed that the upper bound on decoded bit error rate and channel utilization is tight. The upper bound was used to demonstrate that using a set of relatively simple 8-state trellis codes, the type-I hybrid-ARQ modifications reduce the code gap by approximately 1 dB. This performance improvement was realized without a reduction in quality of service or an increase in code complexity. While for an FEC system, the code gap of the set of 8-state trellis codes can be reduced by about 1 dB by using a set of 64-state trellis codes. It appears that the use of more complex trellis codes, modified for type-I hybrid-ARQ, will provide even more substantial reductions in the code gap.

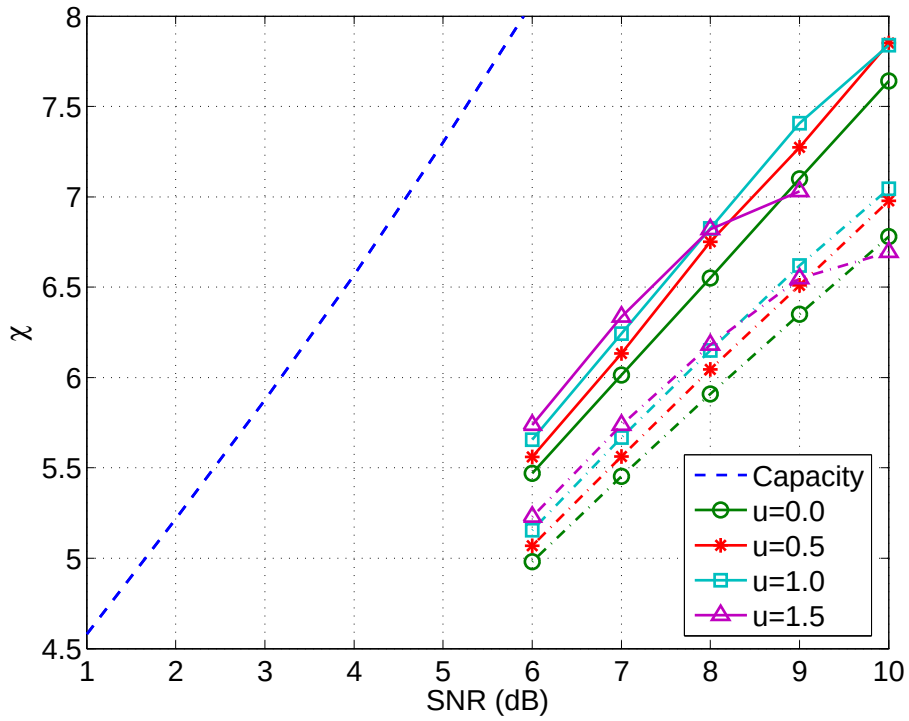


Figure 3.6: The upper bounds of channel utilization at $P_b = 10^{-5}$ using bit-loading algorithm.

Chapter 4

A Type-II Hybrid-ARQ Error Control for MSTTCs

In Chapter 2 and Chapter 3, we discussed how hybrid-ARQ error control can be employed in MIMO systems when the channel state information is known to both the transmitter and the receiver. When the channel state information is unknown to the transmitter, the MIMO channel can not be manipulated into parallel channels using singular value decomposition as we have shown in Chapter 2 and Chapter 3. The coding schemes, known as space-time coding, are used across both space and time to maximize the link performance. In this chapter, we address how hybrid-ARQ error control can be combined with multidimensional space-time trellis codes when the channel state information is known to the receiver but unknown to the transmitter.

4.1 Introduction

Hybrid automatic repeat request (ARQ) schemes combine forward error correction (FEC) code and retransmission scheme [49]. A proper combinatin of a FEC code and ARQ scheme provides lower frame error rate (FER), and higher throughput. Yu et al [54] have shown that hybrid-ARQ error control using different TCM outperform the hybrid-ARQ error control consisting of the same TCM for all transmission. By using the same state transition diagram for retransmission, an effficient decoding method was exploited.

Recently, space-time trellis codes (STTCs), proposed by Tarokh et al. [44], Guey et al. [13], have been widely studied as an FEC code used in MIMO communication systems for high data rate wireless transmission. In quasi-static flat fading

channel, the rank and the determinant criterions are used to optimize the code design [44]. Seok and Lee [22] presented a hybrid-ARQ scheme employing STTCs as the FEC codes. Each transmission in their HARQ scheme used different STTCs. These codes were found using a computer search. The hybrid-ARQ scheme, consisting of the optimal STTC for each transmission, outperforms the hybrid-ARQ scheme, consisting of the same STTC for all transmissions.

More recently, Ionescu et al. [18] [20] [19] introduced a Euclidean type distance and showed that the determinant criterion can be strengthened. The modified determinant criterion (equal eigenvalue criterion) is used to design a class of multidimensional space-time trellis codes (MSTTCs) for QPSK with two transmit antennas which outperforms the optimum STTCs with one receive antenna and outperforms most of the known STTCs with two receive antennas [20]. The multidimensional space-time codes were proved, by Yan and Ionescu [53], to be generalized coset codes, and thus geometrically uniform.

In this chapter, we consider the hybrid-ARQ scheme employing MSTTCs as the FEC codes. The multidimensional space-time codes used for retransmissions are designed using partition chains on the super-constellation that are different from the partition chain used for the original transmission. The partitioning used for retransmission is sub-optimal. The codes designed based on the sub-optimal partition chains are by themselves not optimal codes. But when combined with the code used for the original transmission, they provide better error control than using the same code for all transmissions. We will show that the MSTTCs based on the sub-optimal partitioning are also geometrically uniform. Geometrical uniformity of the MSTTCs makes the performance analysis simplified, since the distance profiles of geometrically uniform code are transparent to the choice of the reference codeword [11]. Theoretical performance analysis and simulation results show that hybrid-ARQ scheme, consisting different MSTTCs for each transmissions, outperforms the hybrid-ARQ scheme, consisting the same MSTTCs for all transmissions.

This chapter is organized as follows: Section 4.2 includes a brief review of multidimensional space-time trellis codes. Section 4.3 describes the hybrid-ARQ

scheme for MSTTCs and the system performance analysis. Code design criterion and the code design are given in Section 4.4. We also show in Section 4.4 the geometrical uniformity of the retransmission codes. Section 4.5 provides the simulation results and the conclusions are stated in Section 4.6.

4.2 Review of Multidimensional Space-Time Trellis Codes

4.2.1 Motivation and Design Criterions

Multidimensional space-time trellis codes were proposed by Ionescu et al. [18, 20, 19] on frequency non-selective fading MIMO channels. Consider a system with M_T transmit antennas and M_R receive antennas. Let l be the frame length. A codeword is the concatenation of all symbols sent over all of the M_T antennas during the corresponding l consecutive symbol epochs. The transmitted codeword starting at time t can be expressed as a $l \times M_T$ code matrix

$$\mathbf{S} = \begin{pmatrix} s_t^1 & s_t^2 & \dots & s_t^{M_T} \\ s_{t+1}^1 & s_{t+1}^2 & \dots & s_{t+1}^{M_T} \\ \vdots & \vdots & \ddots & \vdots \\ s_{t+l-1}^1 & s_{t+l-1}^2 & \dots & s_{t+l-1}^{M_T} \end{pmatrix}$$

where s_{t+i}^n, $i = 0, \cdots, l-1$, $n = 1, \cdots, M_T$, is the symbol transmitted on the nth antenna at time slot $t+i$. Let $\mathbf{H}$ be the $M_T \times M_R$ channel matrix, where the elements of $\mathbf{H}$, h_{ij}, represent the channel attenuation coefficients between transmit antenna i and receive antenna j. The $l \times M_R$ receive matrix is given by

$$\mathbf{Y} = \sqrt{E_s/M_T}\mathbf{S}\mathbf{H} + \mathbf{N} \tag{4.1}$$

where E_s is the total average energy available at the transmitter over a symbol period, and $\mathbf{N}$ is the additive noise.

Assume i.i.d. quasi-static Rayleigh fading with perfect channel state information at the receiver. Let $\mathbf{E}$ be the code matrix of another codeword. For a maximum-likelihood receiver, the probability of decoding $\mathbf{E}$ when $\mathbf{S}$ is transmitted

has been shown in [44] to be upper bounded by

$$P(\mathbf{S} \to \mathbf{E}) \leq \left(\prod_{i=1}^{r} \lambda_i \right)^{-1} (Es/2N_0)^{-rM_R} \tag{4.2}$$

where r is the rank of the difference matrix $\mathbf{D_{SE}} = \mathbf{S} - \mathbf{E}$, and λ_i are the eigenvalues of $\mathbf{D^*_{SE}D_{SE}}$. By the determinant criterion [44], one must maximize $\prod_{i=1}^{r} \lambda_i$. Suppose $l \geq M_T$, Ionescu [19, 20] has shown that by Hadamard's theorem [17], the eigenvalue product for some square, positive definite matrix $\mathbf{A} = [a_{ij}]$ assumes its maximum value if and only if $\mathbf{A}$ is diagonal. Once $\mathbf{D^*_{SE}D_{SE}}$ is diagonalized, the product of its diagonal elements is maximized if and only if they are rendered equal, that is $\lambda_i = \lambda$, for $i = 1, \cdots, M_T$, and their sum, $tr(\mathbf{D^*_{SE}D_{SE}}) = \lambda M_T$, is also maximized. It has also been shown in [19, 18] that $tr(\mathbf{D^*_{SE}D_{SE}}) = \sum_{i=1}^{r} \lambda_i = d^2(\mathbf{S}, \mathbf{E})$ is a squared Euclidean distance between code matrix $\mathbf{S}$ and $\mathbf{E}$.

Based on the above observation, Ionsecu proposed an equal eigenvalue criterion (EEC): In i.i.d., M_T transmit antenna, quasi-static Rayleigh fading with perfect CSI, an upper bound to the pairwise error probability is made as small as possible if and only if, for all pairs $\mathbf{S}$, $\mathbf{E}$, the Euclidean squared distances $tr(\mathbf{D^*_{SE}D_{SE}})$ are made as large as possible, and $\mathbf{D^*_{SE}D_{SE}} = (tr(\mathbf{D^*_{SE}D_{SE}})/M_T)\mathbf{I}_{M_T}$. For suboptimal codes the main diagonal elements of the matrices $\mathbf{D^*_{SE}D_{SE}}$ should be as close as possible to each other (or $tr(\mathbf{D^*_{SE}D_{SE}})/M_T$), and for which the row-wise sum of the absolute values of the elements off the main diagonal is as small as possible for each row. It is also shown that the pairwise error probability is lower bounded by

$$P(\mathbf{S} \to \mathbf{E}) \geq Q\left(\sqrt{\frac{E_s d^2(\mathbf{S}, \mathbf{E})}{2N_0}} \right). \tag{4.3}$$

The orthogonal space-time block codes (OSTBC) given in [43], which included the Alamouti scheme [1] as a special case, obey the the aforementioned (optimal) equal eigenvalue criterion structure. Consider a constellation $\mathcal{M}$ with 2^b elements. A OSTBC maps k symbols selected from the constellation to a $p \times M_T$ matrix. Elements of the $p \times M_T$ matrix will be transmitted over the M_T transmit antennas during p transmission epochs. Thus kb bits are sent during each p transmissions. Let $R = k/p$ be the rate of OSTBC, then the throughput of OSTBC is $b \times R$

bits/symbol/transmission. It has been asserted by [43] that, a generalized orthogonal design can be found with size (k, p)

$$(k,p) = \begin{cases} (2,2) & \text{if } M_T = 2 \\ (4,4) & \text{if } M_T = 3,4 \\ (8,8) & \text{if } M_T = 5,6,7,8 \end{cases} \tag{4.4}$$

for a real constellation such as PAM; a generalized complex orthogonal design can be found with size (k, p)

$$(k,p) = \begin{cases} (2,2) & \text{if } M_T = 2 \\ (3,4) & \text{if } M_T = 3,4 \\ (M_T, 2M_T) & \text{if } M_T = 5,6,7,8 \end{cases} \tag{4.5}$$

for a complex constellation such as PSK and QAM. Let $\mathcal{G}$ be the orthogonal design given in [43], such that $\mathcal{G}^*\mathcal{G} = a\mathbf{I}$ for some constant a. For a constellation $\mathcal{M}$ of size 2^b and (k, p) pair given in (4.4) and (4.5), there are $L = 2^{bk}$, $p \times M_T$ matrices generated by $\mathcal{G}$.

Orthogonal space-time block codes provide diversity gain only with no coding gain. Concatenating an encoder with OSTBCs would decrease the throughput, i.e., with 4PSK, two transmit antennas, the throughput would be less than 2 bits/symbol/transmission, even though the OSTBCs for two transmit antennas reach the full transmission rate ($R = 1$). Ionescu et al. [19, 20] presented a multidimensional space-time trellis codes (MSTTCs) designed with the equal eigenvalue criterion to achieve full diversity gain and provide coding gain without decreasing the overall throughput. The orthogonal space-time block codes are used as the building blocks in MSTTCs. In order to introduce coding gain and keep the throughput bk/p unchanged, one must have more than 2^{bk} matrices. This requires augmenting the optimal matrix set. The set of all $p \times M_T$ matrix building blocks used for MSTTCs can be regarded as a super-constellation $\mathbb{M}$. Some matrix pairs in the augmented set $\mathbb{M}$ will not obey the equal eigenvalue structure. Rather than enforcing equal eigenvalue criterion on all valid code matrix pairs $\mathbf{S}$ and $\mathbf{E}$, one can focus on those that dominate performance. The design goal of MSTTCs becomes insuring that difference code matrices

pertaining to an error event path (EEP) in the trellis of length $k \leq k'$ transitions (kp modulator symbols) be optimal for k' as large as possible, and as close to optimal as feasible for $k > k'$.

4.2.2 Construction of the MSTTCs

The design of a MSTTC is similar to the schemes proposed for multiple trellis coded modulation [10]. First, set partitioning of the super-constellation $\mathbb{M}$ with respect to maximize the distance between the elements in the subsets is performed. For each block of kb input bits, m bits are sent to the rate m/n convolutional encoder and generate n coded bits while $kb - m$ bits are uncoded as shown in Figure 4.1. Let $Q = 2^{kb+z}$, for z a positive integer, be the number of elements ($p \times M_T$ matrices) in the super-constellation $\mathbb{M}$. Since the $kb - m + n$ bits are used to select a $p \times M_T$ matrix from the super-constellation, thus the rate of the convolutional encoder has to satisfy the constraint $n - m = \log_2(Q) - kb$, i.e., $n - m = z$. The n coded bits are used to select one of the 2^n partitions of the super-constellation at the nth level of the super-constellation's partition tree. The $kb - m$ uncoded bits are used to select a $p \times M_T$ matrix within the designated partition. The number of states, q, in the convolutional encoder should satisfy $q > 2^m$. Generally, for fixed q, there are several options for MSTTC encoder. The option providing the best performance is usually chosen as the MSTTC.

As shown in (4.5), for a complex constellation $\mathcal{M}$ with more than two transmit antennas, the OSTBC does not have full rate. Some quasi-orthogonal space time block codes have been proposed with full rate for three and four transmit antennas with full diversity [40] and without full diversity [45]. Jafarkhani and Hassampour [21] propose a family of multidimensional space-time trellis codes using the quasi-orthogonal space time block codes as building blocks in the trellis code for four transmit antennas. In this Chapter, we will focus on the case of two transmit antennas with QPSK. This is the case considered by Ionescu et al. [19, 20] in his comparison of the MSTTCs and the STTCs with transmission rate 2 bits/symbol/transmission.

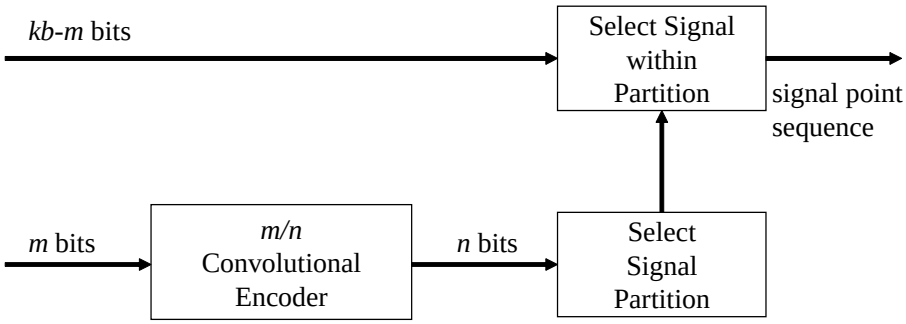


Figure 4.1: General Encoder of multidimensional space-time trellis.

For two transmission antennas and QPSK constellation, the (k, p) pair relationship is $(k, p) = (2, 2)$ as shown in equation (4.5). For two transmission antennas and complex constellation, the possible orthogonal designs are

$$\mathcal{G}_1 = \begin{bmatrix} a & b \\ b* & -a* \end{bmatrix} \text{ and}$$

$$\mathcal{G}_2 = \begin{bmatrix} a & b \\ -b* & a* \end{bmatrix}.$$

The super-constellation for MSTTCs generated by the orthogonal designs has $32 = 2^5$ 2×2 matrices whose entries are from 4PSK constellation. The matrices in the first and the second half of the super-constellation are generated by $\mathcal{G}_1$ and $\mathcal{G}_2$, respectively. Each 2×2 matrix M_i defines the 4PSK symbols to be sent over the $M_T = 2$ transmit

antennas, during two consecutive symbol epochs. The 32 matrices are:

$$\mathbf{M}_0 = \begin{bmatrix} 1 & 3 \\ 0 & 0 \end{bmatrix} \quad \mathbf{M}_8 = \begin{bmatrix} 3 & 3 \\ 0 & 2 \end{bmatrix} \quad \mathbf{M}_{16} = \begin{bmatrix} 3 & 1 \\ 0 & 0 \end{bmatrix} \quad \mathbf{M}_{24} = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$$

$$\mathbf{M}_1 = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} \quad \mathbf{M}_9 = \begin{bmatrix} 3 & 2 \\ 1 & 2 \end{bmatrix} \quad \mathbf{M}_{17} = \begin{bmatrix} 3 & 0 \\ 1 & 0 \end{bmatrix} \quad \mathbf{M}_{25} = \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix}$$

$$\mathbf{M}_2 = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix} \quad \mathbf{M}_{10} = \begin{bmatrix} 3 & 1 \\ 2 & 2 \end{bmatrix} \quad \mathbf{M}_{18} = \begin{bmatrix} 3 & 3 \\ 2 & 0 \end{bmatrix} \quad \mathbf{M}_{26} = \begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix}$$

$$\mathbf{M}_3 = \begin{bmatrix} 1 & 0 \\ 3 & 0 \end{bmatrix} \quad \mathbf{M}_{11} = \begin{bmatrix} 3 & 0 \\ 3 & 2 \end{bmatrix} \quad \mathbf{M}_{19} = \begin{bmatrix} 3 & 2 \\ 3 & 0 \end{bmatrix} \quad \mathbf{M}_{27} = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}$$

$$\mathbf{M}_4 = \begin{bmatrix} 0 & 3 \\ 0 & 1 \end{bmatrix} \quad \mathbf{M}_{12} = \begin{bmatrix} 2 & 3 \\ 0 & 3 \end{bmatrix} \quad \mathbf{M}_{20} = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix} \quad \mathbf{M}_{28} = \begin{bmatrix} 0 & 1 \\ 0 & 3 \end{bmatrix}$$

$$\mathbf{M}_5 = \begin{bmatrix} 0 & 2 \\ 1 & 1 \end{bmatrix} \quad \mathbf{M}_{13} = \begin{bmatrix} 2 & 2 \\ 1 & 3 \end{bmatrix} \quad \mathbf{M}_{21} = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix} \quad \mathbf{M}_{29} = \begin{bmatrix} 0 & 0 \\ 1 & 3 \end{bmatrix}$$

$$\mathbf{M}_6 = \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix} \quad \mathbf{M}_{14} = \begin{bmatrix} 2 & 1 \\ 2 & 3 \end{bmatrix} \quad \mathbf{M}_{22} = \begin{bmatrix} 2 & 3 \\ 2 & 1 \end{bmatrix} \quad \mathbf{M}_{30} = \begin{bmatrix} 0 & 3 \\ 2 & 3 \end{bmatrix}$$

$$\mathbf{M}_7 = \begin{bmatrix} 0 & 0 \\ 3 & 1 \end{bmatrix} \quad \mathbf{M}_{15} = \begin{bmatrix} 2 & 0 \\ 3 & 3 \end{bmatrix} \quad \mathbf{M}_{23} = \begin{bmatrix} 2 & 2 \\ 3 & 1 \end{bmatrix} \quad \mathbf{M}_{31} = \begin{bmatrix} 0 & 2 \\ 3 & 3 \end{bmatrix}.$$

The mapping between the matrix elements and the 4PSK constellation is given in Figure 4.2.

For all $i, j \in \{0, \cdots, 15\}$, $i \neq j$, $(\mathbf{M}_i - \mathbf{M}_j)^*(\mathbf{M}_i - \mathbf{M}_j)$ has equal eigenvalues, from the set $\{2, 4, 6, 8\}$; the same holds for all $i, j \in \{16, \cdots, 31\}$, $i \neq j$. For all $i \in \{0, \cdots, 15\}$, $j \in \{16, \cdots, 31\}$, $(\mathbf{M}_i - \mathbf{M}_j)^*(\mathbf{M}_i - \mathbf{M}_j)$ does not have equal eigenvalues. The super-constellation is first partitioned into two subsets (index set $\{0, \cdots, 15\}$ and $\{16, \cdots, 31\}$). Then each half is partitioned into subsets based on the eigenvalues of various $(\mathbf{M}_i - \mathbf{M}_j)^*(\mathbf{M}_i - \mathbf{M}_j)$. For simple expression, we generate the partitioning chain, as shown in Figure 4.3, from Table I in [53]. In Figure 4.3, the indexes of the matrices in the super-constellation have been used and the elements in the subsets are ordered by the indexes of the uncoded bits.

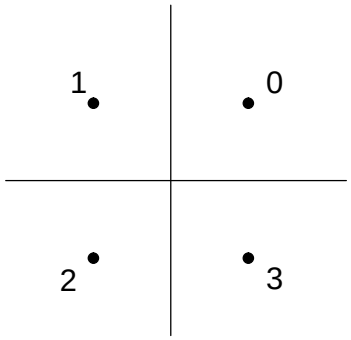


Figure 4.2: Indexing for the 4PSK constellation points.

For QPSK, we have $b = 2$. For two transmit antenna $M_T = 2$. Since the super-constellation has $Q = 32 = 2^5 = 2^{bk+1}$ elements, the rate of the convolutional encoder has to be $m/(m+1)$. All possible MSTTC encoders for QPSK with two transmit antennas are shown in Figure 4.4 with the possible memory v and number of state $q = 2^v$.

For the 8-state MSTTC, there are three candidates, while for the 16-state MSTTC, there are four candidates. The 8- and 16-state MSTTCs for QPSK with two transmit antennas with the best performance have been given in [20], where the 8-state MSTTC is of the form (b) in Figure 4.4, while the 16-state MSTTC is of the form (c) in Figure 4.4. The MSTTCs outperform the optimum STTCs (of equal complexity[1]) with one receive antenna and outperform most of the known STTCs with two receive antennas. It is shown in [53], that MSTTCs are geometrically uniform. The encoder and the trellis diagrams of the 8-state MSTTC for QPSK with two transmit antennas are given in Figure 4.5 and Figure 4.6, respectively. The encoder and the trellis diagrams of the 16-state MSTTC are given in Figure 4.7 and Figure 4.8, respectively.

[1]Complexity is indicated by the product between the number of states and the number of transitions emerging from each state, normalized to one modulator symbol epoch.

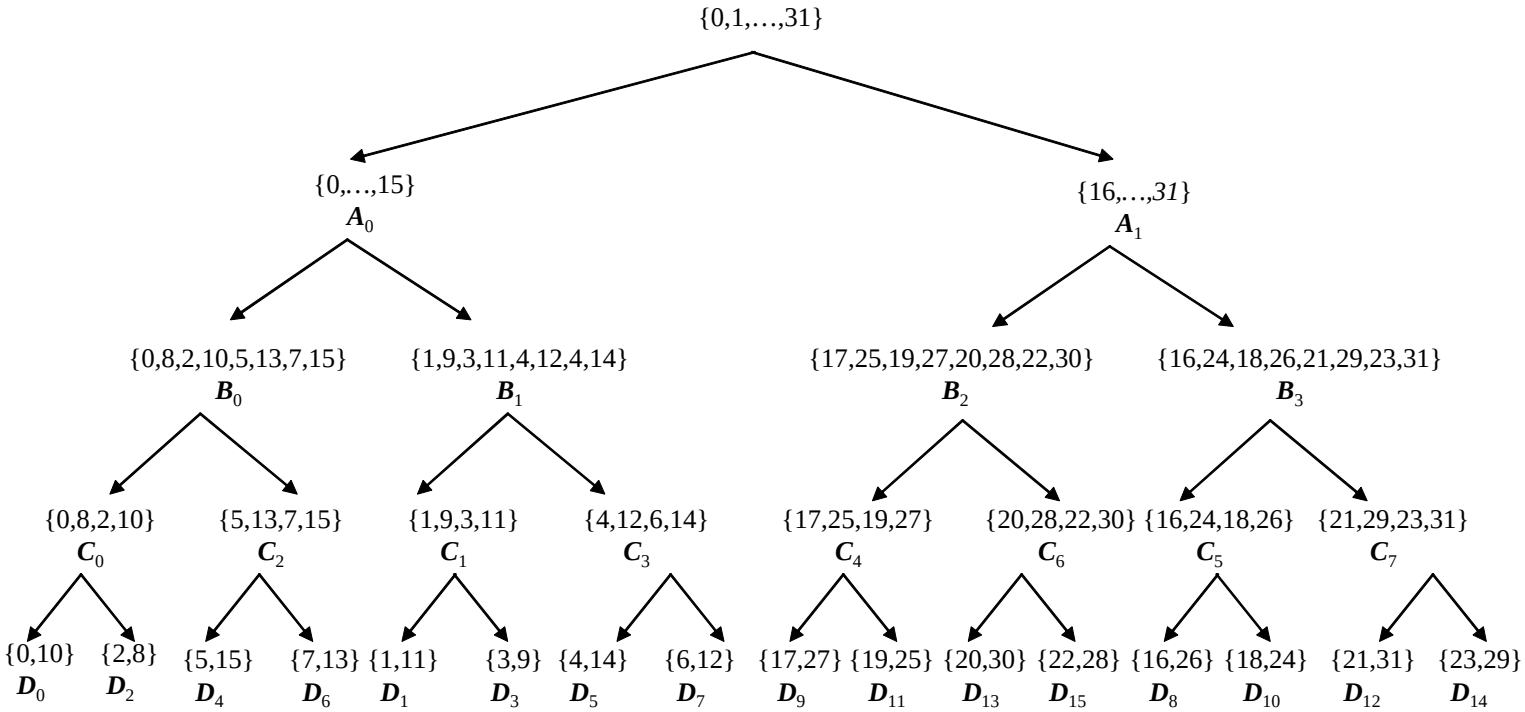


Figure 4.3: The super-constellation partition chain of the MSTTCs. The elements in each subset are ordered by increasing decimal value of corresponding uncoded bit pair.

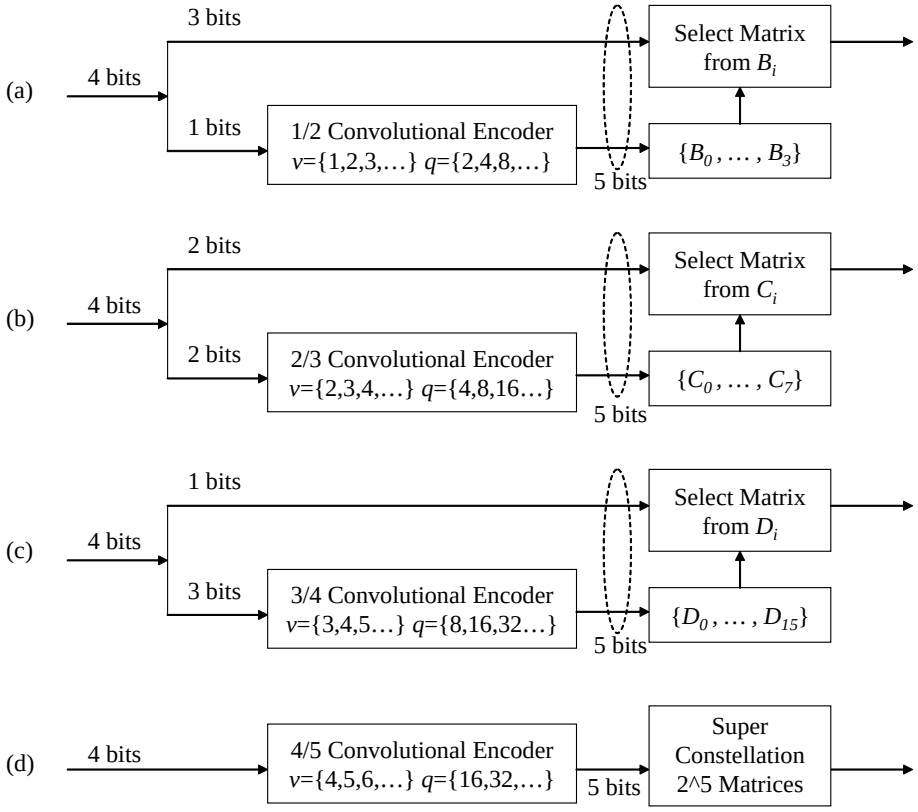


Figure 4.4: Possible Encoders of multidimensional space-time trellis code for QPSK with two transmit antennas.

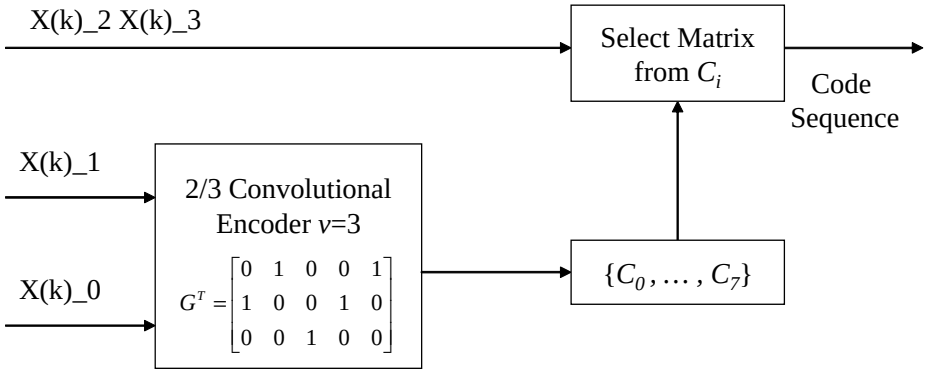


Figure 4.5: Encoder for the best 8-state multidimensional space-time trellis code.

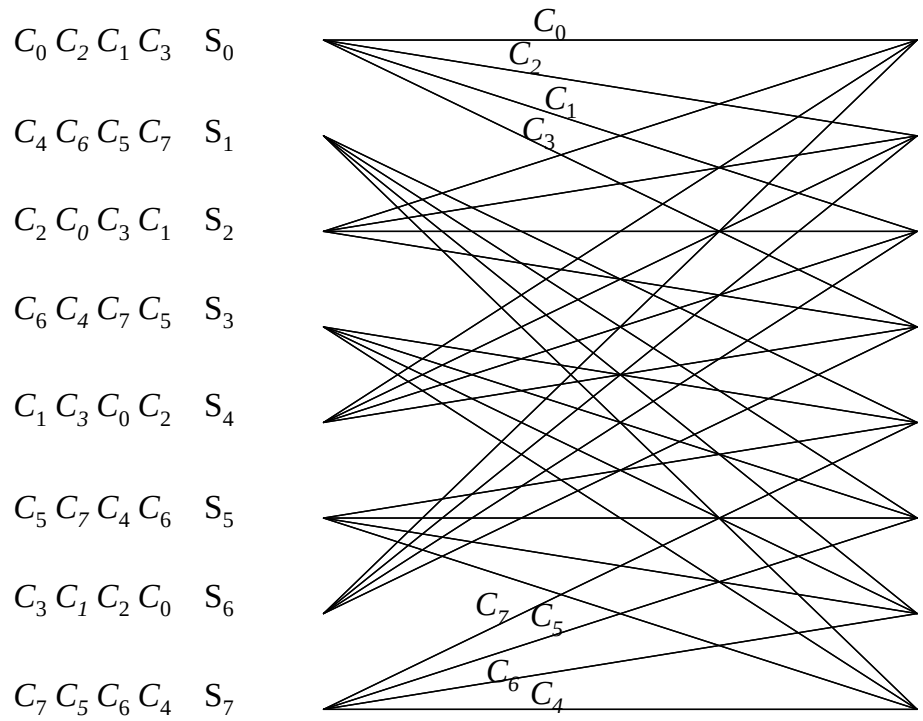


Figure 4.6: Trellis diagram of the 8-state MSTTC.

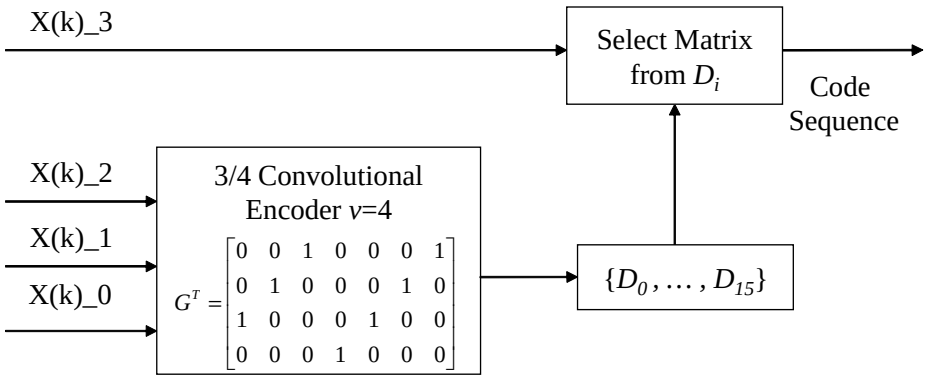


Figure 4.7: Encoder for the best 16-state multidimensional space-time trellis code.

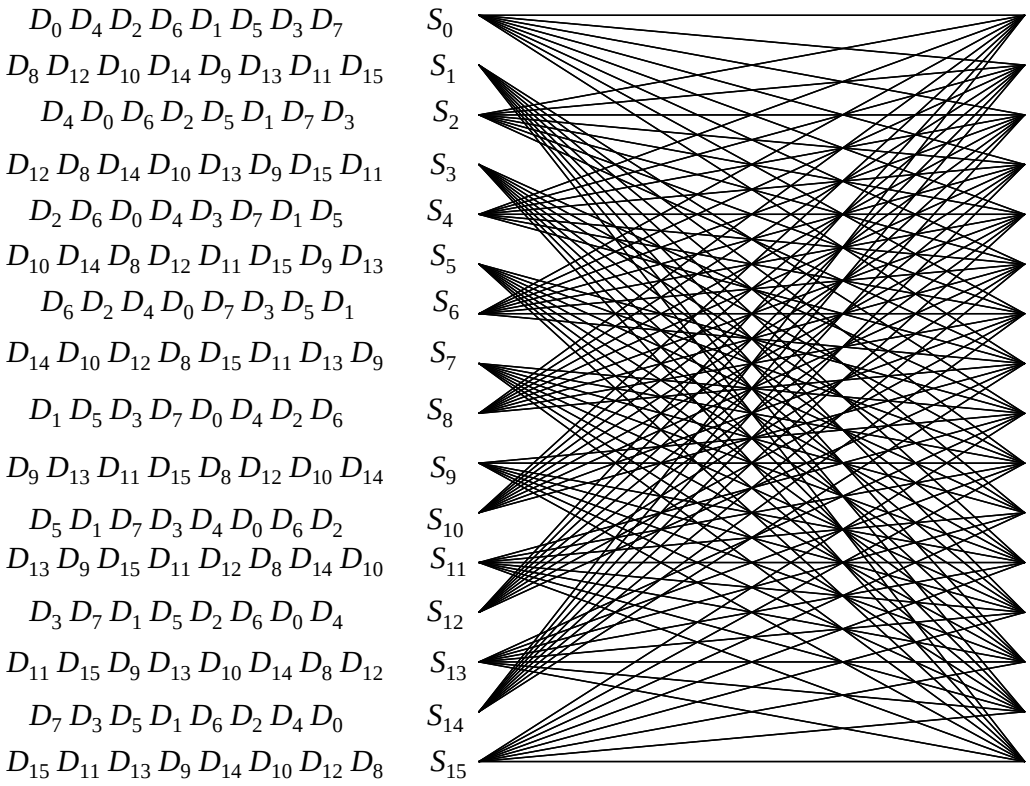


Figure 4.8: Trellis diagram of the 16-state MSTTC.

4.2.3 An Example

In this section, we will take the 8-state MSTTC as an example and illustrate how the encoder and the decoder work. With the transmission rate of 2 bits/symbol/transmission, there are 4 bits corresponding to two consecutive symbol epochs. The first two bits are sent to the convolutional encoder and the encoder outputs a label in $\{0, \cdots, 7\}$ to select a signal partition from $\{\mathcal{C}_0, \cdots, \mathcal{C}_7\}$ (Figure 4.3). The coded bits will also determine the state transition on the trellis. The second two bits are uncoded and are used to select a 2×2 matrix from the partition chosen by the coded data bits. A sequence of data bits are mapped to a sequence of 2×2 square sub-matrices, i.e., a codeword. Note that in the trellises shown in Figure 4.6 and Figure 4.8, transitions from even states use only matrices from first half of the super-constellation, while transitions from odd states use only matrices from second half of the super-constellation. As a consequence, $\mathbf{D}_{\mathbf{SE}}^{*}\mathbf{D}_{\mathbf{SE}}$ has equal eigenvalues for any difference matrix $\mathbf{D}_{\mathbf{SE}}$ corresponding to error event path of length $k \leq 2$ (i.e., up to four 4PSK symbols).

Consider the all zero input sequence

$$d_0 = [0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]. \tag{4.6}$$

The all zero coded bits select the solid black line as shown in Figure 4.9. The four matrices on the solid black line are the elements of sub-constellation $\mathcal{C}_0$. The all zero uncoded bits select the first matrix in the sub-constellation $\mathcal{C}_0$, which is circled in Figure 4.9. Let us consider an other sequence

$$d_1 = [\underline{0, 1}, 1, 1, \underline{1, 0}, 0, 0, \underline{1, 1}, 1, 0, \underline{0, 0}, 1, 1]. \tag{4.7}$$

The convolutional encoder input is $[01, 10, 11, 00]$; the output is $[010, 101, 000, 110]$. The decimal equivalent of the output sequence is $[2, 5, 0, 6]$, which represent the sub-constellation index of each state transition. The trellis path determined by the coded bits is shown as the dashed black line in Figure 4.9. The four matrices on the first segment of the dashed black line belong to subset

$$\mathcal{C}_2 = \left\{ \begin{bmatrix} 0 & 2 \\ 1 & 1 \end{bmatrix} \quad \begin{bmatrix} 2 & 2 \\ 1 & 3 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 \\ 3 & 1 \end{bmatrix} \quad \begin{bmatrix} 2 & 0 \\ 3 & 3 \end{bmatrix} \right\}. \tag{4.8}$$

The uncoded bits [11] selected the forth matrix

$$\begin{bmatrix} 2 & 0 \\ 3 & 3 \end{bmatrix}. \tag{4.9}$$

The four matrices on the second segment of the dashed black line belong to subset

$$\mathcal{C}_5 = \left\{ \begin{bmatrix} 3 & 1 \\ 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} \quad \begin{bmatrix} 3 & 3 \\ 2 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix} \right\}. \tag{4.10}$$

The uncoded bits [00] selected the first matrix

$$\begin{bmatrix} 3 & 1 \\ 0 & 0 \end{bmatrix}. \tag{4.11}$$

The four matrices on the third segment of the dashed black line belong to subset

$$\mathcal{C}_0 = \left\{ \begin{bmatrix} 1 & 3 \\ 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 3 & 3 \\ 0 & 2 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix} \quad \begin{bmatrix} 3 & 1 \\ 2 & 2 \end{bmatrix} \right\}. \tag{4.12}$$

The uncoded bits [10] selected the second matrix

$$\begin{bmatrix} 3 & 3 \\ 0 & 2 \end{bmatrix}. \tag{4.13}$$

The four matrices on the forth segment of the dashed black line belong to subset

$$\mathcal{C}_6 = \left\{ \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 0 & 1 \\ 0 & 3 \end{bmatrix} \quad \begin{bmatrix} 2 & 3 \\ 2 & 1 \end{bmatrix} \quad \begin{bmatrix} 0 & 3 \\ 2 & 3 \end{bmatrix} \right\}. \tag{4.14}$$

The uncoded bits [11] selected the forth matrix

$$\begin{bmatrix} 0 & 3 \\ 2 & 3 \end{bmatrix}. \tag{4.15}$$

So the index of the QPSK sequence transmitted on the first and the second antenna are $[2, 3, 3, 0, 3, 0, 0, 2]^T$ and $[0, 3, 1, 0, 3, 2, 3, 3]^T$, respectively. Given the channel coefficients between the transmit and receive antennas and the additive noise, the received

signal corresponding to the sequence d_1 is

$$\begin{bmatrix} -0.3095 + 0.0018i \\ -1.3051 + 1.8240i \\ -0.0746 - 0.2637i \\ -1.0528 - 0.7063i \\ -1.3153 + 1.5568i \\ -0.1388 + 0.2471i \\ -1.0331 - 0.3079i \\ -0.1665 + 1.0929i \end{bmatrix} = \begin{bmatrix} -1-i & +1+i \\ +1-i & +1-i \\ +1-i & -1+i \\ +1+i & +1+i \\ +1-i & +1-i \\ +1+i & -1-i \\ +1+i & +1-i \\ -1-i & +1-i \end{bmatrix} \begin{bmatrix} -0.6210 + 0.1817i \\ -1.0755 + 0.0310i \end{bmatrix} + \begin{bmatrix} -0.0947 + 0.4297i \\ -0.2559 + 0.4740i \\ -0.5025 - 0.0489i \\ 0.2972 + 0.3429i \\ -0.2661 + 0.2068i \\ -0.3536 - 0.1808i \\ 0.2731 - 0.7797i \\ 0.0045 - 0.0001i \end{bmatrix}. \tag{4.16}$$

Decoding using the Viterbi decoder is performed in two steps

1. At each branch in the trellis, the decoder compares the received signal to each of signals allowed for that branch. The metric is taken as the Frobenius norm between the received signal and the product of H and the possible transmit signal. The branch is labeled with a metric of the allowed signal closest to the received signal. In other word, the branch metric is the minimum Forbenius norm between the allowed signal and the received signal.

2. The Viterbi algorithm is then applied to the trellis, with the surviving partial paths corresponding to the path with the lowest accumulated metric.

The metrics at each branch in the trellis are shown in Table 4.1 with the branch metric being underlined. The Viterbi Decoding is shown in Figure 4.10 The highest partial path metrics are labeled at each node, and the corresponding surviving paths are denoted by the black lines. The non-surviving paths are denoted by grey lines.

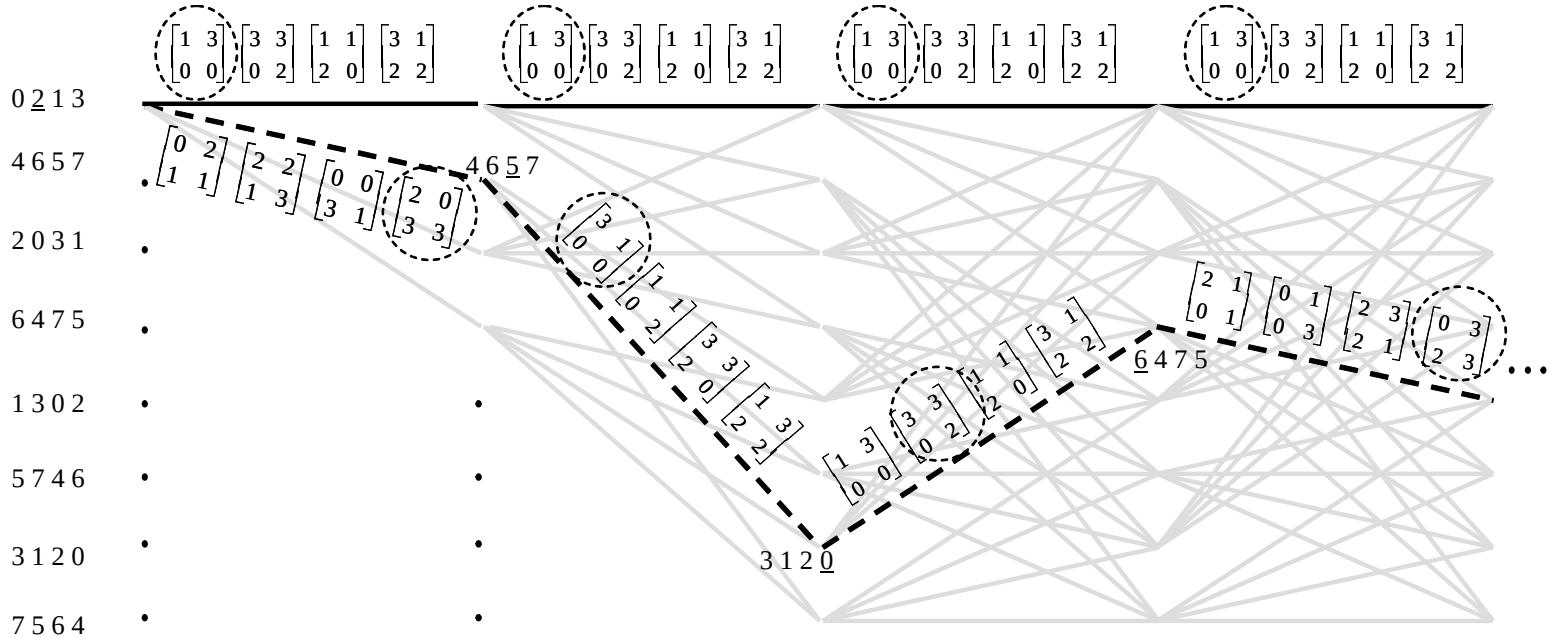


Figure 4.9: Trellis diagram of the 8-state MSTTC.

Table 4.1: The metrics of the Viterbi decoder.

Stage 1	Metric	Stage 2	Metric	Stage 3	Metric	Stage 4	Metric
$S_0 \to S_0$	[8.32 6.62 9.93 8.24]	$S_0 \to S_0$	[0.56 6.45 3.22 9.11]	$S_0 \to S_0$	[5.73 0.27 14.50 9.03]	$S_0 \to S_0$	[6.62 3.33 7.73 4.44]
$S_0 \to S_1$	[16.07 7.21 9.35 0.48]	$S_0 \to S_1$	[5.39 4.99 4.67 4.28]	$S_0 \to S_1$	[7.57 7.44 7.32 7.20]	$S_0 \to S_1$	[9.54 8.36 2.71 1.53]
$S_0 \to S_2$	12.49 10.80 5.76 4.07]	$S_0 \to S_2$	[2.25 8.14 1.53 7.42]	$S_0 \to S_2$	[10.24 4.77 9.99 4.53]	$S_0 \to S_2$	[10.59 7.30 3.76 0.47]
$S_0 \to S_3$	[11.90 3.04 13.52 4.66]	$S_0 \to S_3$	[3.71 3.30 6.37 5.97]	$S_0 \to S_3$	[3.06 2.94 11.83 11.70]	$S_0 \to S_3$	[5.57 4.39 6.68 5.49]
		$S_1 \to S_4$	[1.74 6.62 3.05 7.93]	$S_1 \to S_4$	[5.55 10.43 4.34 9.21]	$S_1 \to S_4$	[5.87 3.74 7.32 5.20]
		$S_1 \to S_5$	[3.60 2.20 7.47 6.07]	$S_1 \to S_5$	[12.87 10.64 4.12 1.89]	$S_1 \to S_5$	[10.39 3.75 7.32 0.68]
		$S_1 \to S_6$	[0.46 5.33 4.33 9.21]	$S_1 \to S_6$	[9.32 14.19 0.57 5.44]	$S_1 \to S_6$	[8.13 6.00 5.06 2.94]
		$S_1 \to S_7$	[4.88 3.48 6.19 4.79]	$S_1 \to S_7$	[9.11 6.87 7.89 5.66]	$S_1 \to S_7$	[8.12 1.48 9.58 2.94]
		$S_2 \to S_0$	[5.39 4.99 4.68 4.28]	$S_2 \to S_0$	[7.57 7.44 7.32 7.20]	$S_2 \to S_0$	[9.54 8.36 2.71 1.53]
		$S_2 \to S_1$	[0.56 6.45 3.22 9.11]	$S_2 \to S_1$	[5.73 0.27 14.50 9.03]	$S_2 \to S_1$	[6.62 3.33 7.73 4.44]
		$S_2 \to S_2$	[3.71 3.30 6.37 5.97]	$S_2 \to S_2$	[3.06 2.93 11.83 11.70]	$S_2 \to S_2$	[5.57 4.39 6.68 5.49]
		$S_2 \to S_3$	[2.25 8.14 1.53 7.42]	$S_2 \to S_3$	[10.24 4.77 9.99 4.53]	$S_2 \to S_3$	[10.59 7.30 3.76 0.47]
		$S_3 \to S_4$	[3.59 2.20 7.47 6.07]	$S_3 \to S_4$	[12.87 10.64 4.12 1.89]	$S_3 \to S_4$	[10.39 3.75 7.32 0.68]
		$S_3 \to S_5$	[1.74 6.62 3.05 7.93]	$S_3 \to S_5$	[5.55 10.43 4.34 9.21]	$S_3 \to S_5$	[5.87 3.74 7.32 5.20]
		$S_3 \to S_6$	[4.87 3.48 6.18 4.79]	$S_3 \to S_6$	[9.11 6.87 7.89 5.66]	$S_3 \to S_6$	[8.12 1.48 9.58 2.94]
		$S_3 \to S_7$	[0.46 5.33 4.33 9.21]	$S_3 \to S_7$	[9.32 14.19 0.58 5.44]	$S_3 \to S_7$	[8.13 6.00 5.06 2.94]
				$S_4 \to S_0$	[10.24 4.77 9.99 4.53]	$S_4 \to S_0$	[10.59 7.30 3.76 0.47]
				$S_4 \to S_1$	[3.06 2.94 11.83 11.70]	$S_4 \to S_1$	[5.57 4.39 6.68 5.49]
				$S_4 \to S_2$	[5.73 0.27 14.50 9.03]	$S_4 \to S_2$	[6.62 3.33 7.73 4.44]
				$S_4 \to S_3$	[7.57 7.44 7.32 7.20]	$S_5 \to S_3$	[9.54 8.36 2.71 1.53]
				$S_5 \to S_4$	[9.32 14.19 0.58 5.44]	$S_5 \to S_4$	[8.13 6.00 5.06 2.94]
				$S_5 \to S_5$	[9.11 6.87 7.89 5.66]	$S_5 \to S_5$	[8.12 1.48 9.58 2.94]
				$S_5 \to S_6$	[5.55 10.43 4.34 9.21]	$S_5 \to S_6$	[5.87 3.74 7.32 5.20]
				$S_5 \to S_7$	[12.87 10.64 4.13 1.89]	$S_5 \to S_7$	[10.39 3.75 7.32 0.68]
				$S_6 \to S_0$	[3.06 2.94 11.83 11.70]	$S_6 \to S_0$	[5.57 4.39 6.68 5.49]
				$S_6 \to S_1$	[10.24 4.77 9.99 4.53]	$S_6 \to S_1$	[10.59 7.30 3.76 0.47]
				$S_6 \to S_2$	[7.57 7.44 7.32 7.20]	$S_6 \to S_2$	[9.54 8.36 2.71 1.53]
				$S_6 \to S_3$	[5.73 0.27 14.50 9.03]	$S_6 \to S_3$	[6.62 3.33 7.73 4.44]
				$S_7 \to S_4$	[9.11 6.87 7.89 5.66]	$S_7 \to S_4$	[8.12 1.48 9.58 2.94]
				$S_7 \to S_5$	[9.32 14.19 0.57 5.44]	$S_7 \to S_5$	[8.13 6.00 5.06 2.94]
				$S_7 \to S_6$	[12.87 10.64 4.12 1.89]	$S_7 \to S_6$	[10.39 3.75 7.32 0.68]
				$S_7 \to S_7$	[5.55 10.43 4.34 9.21]	$S_7 \to S_7$	[5.87 3.74 7.32 5.20]

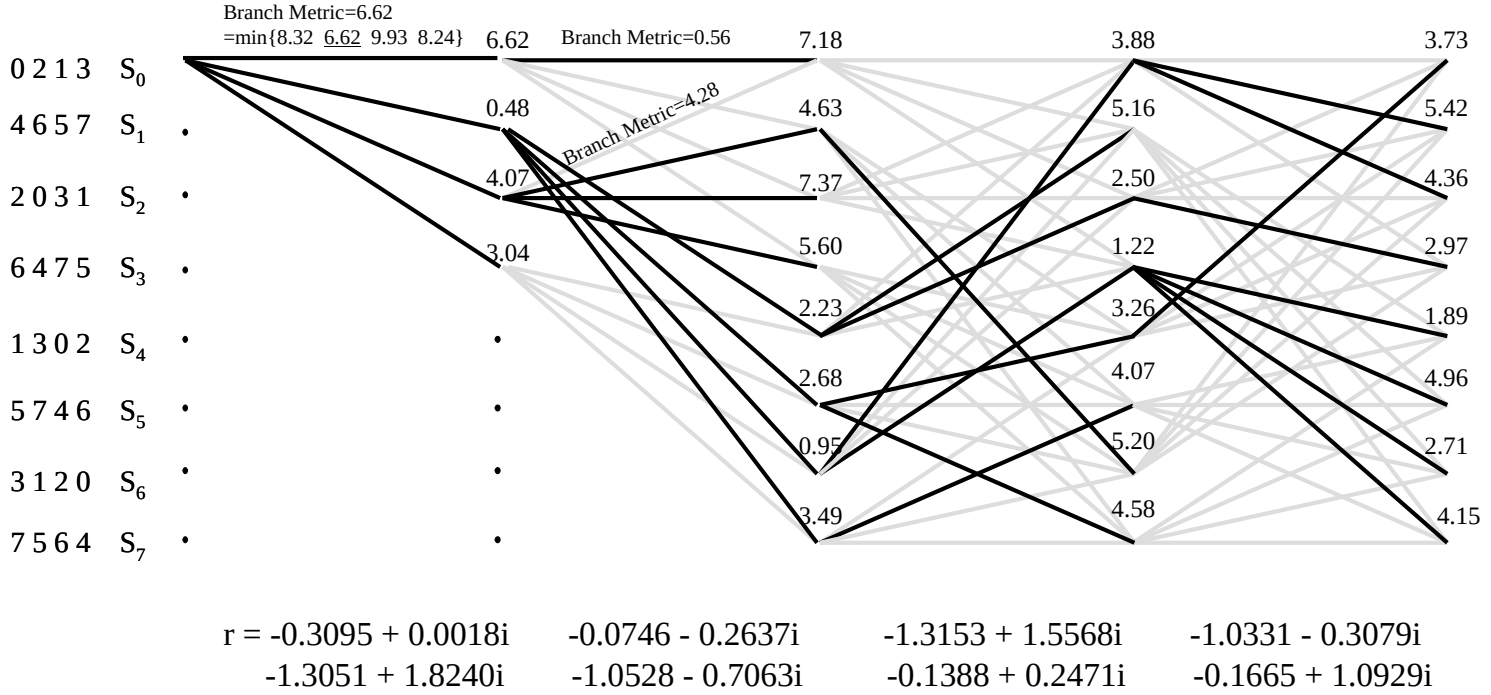


Figure 4.10: Viterbi Decoding.

4.3 System Performance of Hybrid-ARQ using MSTTCs

We denote the MSTTC codeword of the τ^{th} transmission as

$$\mathbf{S}^{\tau} = \begin{pmatrix} \mathbf{s}^{\tau}(1) \\ \mathbf{s}^{\tau}(2) \\ \vdots \\ \mathbf{s}^{\tau}(l/M_T) \end{pmatrix}$$

where $\mathbf{s}^{\tau}(k)$, $k = 1, \cdots, l/M_T$, is the $M_T \times M_T$ transmitted matrix symbol over the $[(k-1) * M_T + 1]^{\text{th}}$ to the kM_T^{th} symbol period of the τ^{th} transmission. The channel is assumed to be quasi-static flat fading, i.e., the channel remains constant during a frame and vary from one frame to another. The received signal matrix of the τ^{th} transmission

$$\mathbf{Y}^{\tau} = \begin{pmatrix} \mathbf{y}^{\tau}(1) \\ \mathbf{y}^{\tau}(2) \\ \vdots \\ \mathbf{y}^{\tau}(l/M_T) \end{pmatrix}$$

is given by

$$\mathbf{Y}^{\tau} = \sqrt{E_s/M_T}\mathbf{S}^{\tau}\mathbf{H}^{\tau} + \mathbf{N}^{\tau} \tag{4.17}$$

where $\mathbf{y}^{\tau}(k)$ is the $M_T \times M_R$ receive matrix corresponding to $\mathbf{s}^{\tau}(k)$, E_s is the total average energy available at the transmitter over a symbol period, $\mathbf{H}^{\tau}$ is the $M_T \times M_R$ channel matrix during the τ^{th} transmission, and $\mathbf{N}^{\tau}$ is the additive noise during the τ^{th} transmission. The elements of the channel matrix are modeled as independent samples of a zero-mean complex Gaussian random variable with variance 0.5 per dimension. The additive noise on each receive antenna at each symbol time is modeled as independent samples of a zero-mean complex Gaussian random variable with variance $N_0/2$ per dimension.

In order to use the same Viterbi decoder at the receiver, the MSTTCs for each transmission must have the same state transition diagram [54, 22]. After the J^{th} transmission, consecutively received sequences are combined and decoded by a single

Viterbi decoder. The branch metrics at time k is given by

$$\sum_{\tau=1}^{J} \|\mathbf{y}^{\tau}(k) - \sqrt{E_s/M_T}\hat{\mathbf{s}}^{\tau}(k)\mathbf{H}^{\tau}\|^2 \tag{4.18}$$

where $\|\cdot\|$ is the Frobenius norm and $\hat{\mathbf{s}}^{\tau}(k)$ denotes the possible transmitted signal of the corresponding trellis state over the τ^{th} transmission. The path with the smallest sum of branch metric gives the estimated codeword. For conventional code combining HARQ scheme, the same codes are used for each transmission, i.e., $\mathbf{s}^1(k) = \mathbf{s}^2(k) = \cdots = \mathbf{s}^J(k)$, for $k = 1, \cdots, l/M_T$. We would like to design a set of different codes for different transmission, such that the over all performance is better than the conventional HARQ scheme.

Let $\mathbf{S}^1, \cdots, \mathbf{S}^J$ be the codewords corresponding to the information data $I_{\mathbf{S}}$ of the J transmissions. We consider the matrix

$$\mathcal{S} = \begin{pmatrix} \mathbf{S}^1 & 0 & \dots & 0 \\ 0 & \mathbf{S}^2 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & \mathbf{S}^J \end{pmatrix}$$

where 0 denotes the all-zero $l \times M_T$ matrix. The input output relation of the codewords associated with $I_{\mathbf{S}}$ can be expressed as

$$\mathcal{Y} = \sqrt{E_s/M_T}\mathcal{S}\mathcal{H} + \mathcal{N} \tag{4.19}$$

where

$$\mathcal{Y} = \begin{pmatrix} \mathbf{Y}^1 \\ \vdots \\ \mathbf{Y}^J \end{pmatrix} \qquad \mathcal{H} = \begin{pmatrix} \mathbf{H}^1 \\ \vdots \\ \mathbf{H}^J \end{pmatrix} \qquad \mathcal{N} = \begin{pmatrix} \mathbf{N}^1 \\ \vdots \\ \mathbf{N}^J \end{pmatrix}.$$

The ML detection for $I_{\mathbf{S}}$ is given by

$$\hat{I_{\mathbf{S}}} = \arg \min_{I_{\mathbf{S}} \to \mathcal{S}} \|\mathcal{Y} - \sqrt{E_s/M_T}\mathcal{S}\mathcal{H}\|^2. \tag{4.20}$$

Now consider another information data $I_{\mathbf{E}}$ corresponding to the codewords $\mathbf{E}^1, \cdots, \mathbf{E}^J$. Let

$$\mathcal{E} = \begin{pmatrix} \mathbf{E}^1 & 0 & \dots & 0 \\ 0 & \mathbf{E}^2 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & \mathbf{E}^J \end{pmatrix}.$$

The conditional pairwise error probability (PEP) between $I_{\mathbf{S}}$ and $I_{\mathbf{E}}$ is given by

$$P(I_{\mathbf{S}} \to I_{\mathbf{E}}|\mathcal{H}) = P\left(\|\mathcal{Y} - \sqrt{E_s/M_T}\mathcal{S}\mathcal{H}\|^2 > \|\mathcal{Y} - \sqrt{E_s/M_T}\mathcal{E}\mathcal{H}\|^2\right). \tag{4.21}$$

By using the same manipulation of [19], average pairwise error probability is lower bounded by

$$P(I_{\mathbf{S}} \to I_{\mathbf{E}}) \geq Q\left(\sqrt{\frac{E_s d^2(\mathcal{S}, \mathcal{E})}{2N_0}}\right) \tag{4.22}$$

$$= Q\left(\sqrt{\frac{E_s \sum_{\tau=1}^{J} d^2(\mathbf{S}^\tau, \mathbf{E}^\tau)}{2N_0}}\right) \tag{4.23}$$

where $d(A, B)$ is an Euclidean type distance on space-time code defined in [19] as $d(A, B) = (\sum_{i=1}^{q} \sigma_i^2(A - B))^{1/2}$, where $q = \text{rank}(A - B)$ and σ_i, $i = 1 \cdots, q$ are the singular values of $A - B$. The derivation from equation (4.22) to equation (4.23) comes from the fact that $d^2(\mathcal{S}, \mathcal{E}) = \sum_{i=1}^{Q} \sigma_i^2(\mathcal{S} - \mathcal{E}) = \sum_{\tau=1}^{J} \sum_{m=1}^{M(\tau)} \sigma_m^2(\mathbf{S}^\tau - \mathbf{E}^\tau) = \sum_{\tau=1}^{J} d^2(\mathbf{S}, \mathbf{E})$, where $Q = \text{rank}(\mathcal{S} - \mathcal{E})$, $M(\tau) = \text{rank}(\mathbf{S}^\tau - \mathbf{E}^\tau)$ for $\tau = 1, \cdots, J$, and $Q = \sum_{\tau=1}^{J} M(\tau)$.

The frame error rate (FER) is obtained by averaging the union of all the error events. FER performance of MSTTC is dominated by the squared Euclidean distance over the shortest error event path (EEP), i.e., d_{free}^2. Let $N(d_{\text{free}})$ be the average number of sequences that are distance d_{free} from the transmitted sequence. The FER is approximated by

$$P_e \approx N(d_{\text{free}}) Q\left(\sqrt{\frac{E_s d_{\text{free}}^2}{2N_0}}\right). \tag{4.24}$$

In Section 4.4, we will design a set of codes such that the over all minimal distance between the codewords is maximized. We assume that errors are perfectly detected and the feed back channel is error free.

4.4 Code Design of the Hybrid ARQ scheme

According to equation (4.24), increasing the overall d^2_{free} after J transmissions can reduce the frame error rate. For the conventional code combining HARQ scheme, which uses the same code for all transmissions, the d^2_{free} increases at a constant rate upon successive retransmission, i.e, the d^2_{free} of the combined code is a multiple of the d^2_{free} of the code used for each transmission. By designing different (supplementary) MSTTCs for each transmission, it is possible to improve the rate by which d^2_{free} increases with the number of transmissions,

The concept of supplementary code design for hybrid-ARQ scheme was proposed by Yu et al. [54] for SISO channel communication, where a supplementary trellis-coded modulation (TCM) codes were obtained through the computer search. In Section 4.4.1, we will show that the optimal MSTTCs for each transmission can be obtained using different partition chain of the super-constellation for each transmission.

4.4.1 Super-Constellation Partition Chains and Code Design

As shown in equation (4.24), the frame error rate (FER) performance is dominated by the squared Euclidean distance over the shortest error event path (EEP), i.e., d^2_{free}. For the 8-state MSTTC, d^2_{free} occurs over the parallel transitions (the transitions connecting a pair of states directly); while for the 16-state MSTTC, d^2_{free} occurs over the non-parallel transitions (two paths emerging from the same state and then merging into another state in at least two trellis stages).

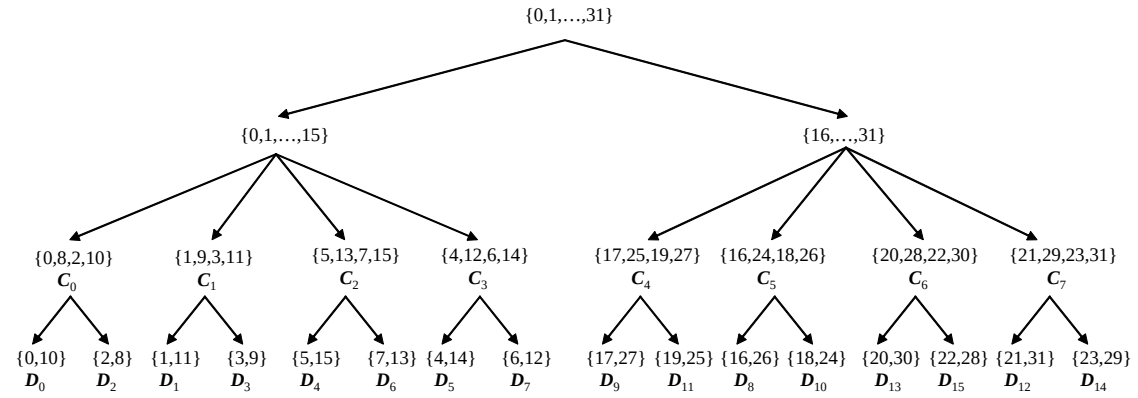
The TCM codes of hybrid-ARQ scheme over SISO channel given in [54] are designed using computer search to maximize the over all free distance after retransmission. Their algorithm only works for the case that the d^2_{free} occurs over the

non-parallel transitions For hybrid-ARQ schemes using MSTTCs, instead of following the idea of Yu et al. given in [54], we come up with a code design procedure for retransmissions using a sub-optimal partition chain. As will shown in Section 4.4.2, the new code designed using the sub-optimal partition chain is also geometrically uniform.

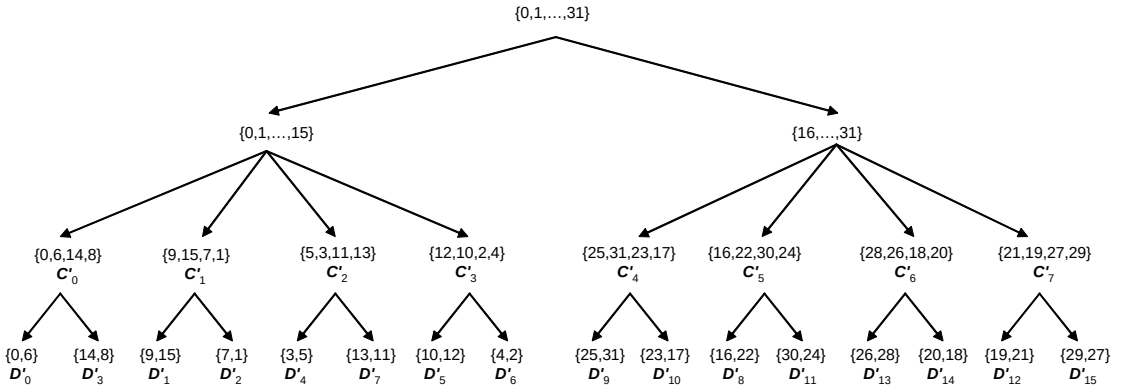
The sub-optimal partitioning of the first, the second and the third transmissions are given in Figure 4.11. In part (a) of Figure 4.11, C_i, $i = 0, \cdots, 7$, are used in the 8-state MSTTCs for the initial transmission; D_i, $i = 0, \cdots, 15$, are used in the 16-state MSTTCs for the initial transmission. In part (b) of Figure 4.11, C_i', $i = 0, \cdots, 7$, are used in the 8-state MSTTCs for the second transmission; D_i', $i = 0, \cdots, 15$, are used in the 16-state MSTTCs for the second transmission. In part (c) of Figure 4.11, C_i'', $i = 0, \cdots, 7$, are used in the 8-state MSTTCs for the third transmission; D_i'', $i = 0, \cdots, 15$, are used in the 16-state MSTTCs for the third transmission. We could keep on design new partition for later on retransmissions or use the partitioning repeatly. In this dissertation, we will just stop after three transmissions. For the subsets C_i, $i = 0, \cdots, 7$, the intra set distance is 8 or 16, while for the subsets C_i' and C_i'', $i = 0, \cdots, 7$, the intra set distance is 12 or 8 correspondingly. The minimal squared Euclidean distance (SED) within each subset is kept unchanged (minimal SED=8). And the overall minimal SED after retransmissions is increased. The design procedure based on the sub-optimal partition chain works fine for the case where d^2_{free} occurs over parallel transitions.

For the subsets D_i, $i = 0, \cdots, 15$, the intra set distance is 16, while for the subsets D_i' and D_i'', $i = 0, \cdots, 15$, the intra set distance is 12. The intra set minimal SED is decreased while the inter set minimal SED is increased which is acceptable since d^2_{free} of the 16-state MSTTC does not occurs over parallel transitions, i.e, the intra set distance does not dominate the FER performance. The design procedure based on the sub-optimal partition chain works fine for the case where d^2_{free} occurs over non-parallel transitions.

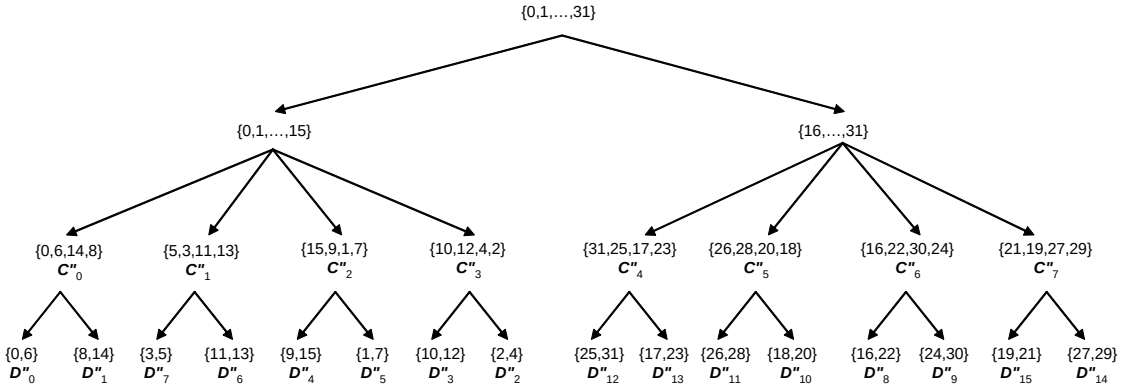
Before we compare the performance of the conventional hybrid-ARQ scheme and the new scheme, we would like to show that the codes designed base on the



(a)



(b)



(c)

Figure 4.11: The super-constellation partition chain of the initial, the second, and the third transmission for QPSK with two transmit antennas.

sub-optimal partition chain are geometrically uniform, which will simplify the performance analysis since the distance profiles are transparent to the choice of a reference codeword if a code is geometrically uniform [11].

4.4.2 Geometrical Uniformity of the Codes

An outline proof will be presented for the 8-state MSTTC used for the second transmission. The proof of the other codes is similar. Following the notation of [53], let R_1, R_2, and R_3 denote the reflection operation about the x, y axes and the origin respectively. And let R_0 denotes no reflection. let $U(\mathbb{M})$ be the generating group for $\mathbb{M}$ as defined in [53]. Consider the group

$$U(\mathbb{M}') = \left\{ \begin{bmatrix} R_0 & R_0 \\ R_0 & R_0 \end{bmatrix} \quad \begin{bmatrix} R_2 & R_3 \\ R_3 & R_2 \end{bmatrix} \quad \begin{bmatrix} R_1 & R_3 \\ R_3 & R_1 \end{bmatrix} \quad \begin{bmatrix} R_3 & R_0 \\ R_0 & R_3 \end{bmatrix} \right\} \tag{4.25}$$

as the generating group of C_0'. The sub-optimal partition of the thirty-two code matrices into the eight cosets given in part (b) of Figure 4.11 is induced by the factor group $U(\mathbb{M})/U(\mathbb{M}')$, where $\mathbb{M}'$ is chosen to be C_0'.

The cosets of $U(\mathbb{M})/U(\mathbb{M}')$ that correspond to $C_0' \cdots C_7'$ are

$$U', \begin{bmatrix} R_3 & R_2 \\ R_2 & R_3 \end{bmatrix} U', \begin{bmatrix} R_2 & R_2 \\ R_2 & R_2 \end{bmatrix} U', \begin{bmatrix} R_1 & R_0 \\ R_0 & R_1 \end{bmatrix} U', \tag{4.26}$$

$$\begin{bmatrix} R_0 & R_1 \\ R_2 & R_3 \end{bmatrix} U', \begin{bmatrix} R_3 & R_3 \\ R_0 & R_0 \end{bmatrix} U', \begin{bmatrix} R_2 & R_3 \\ R_0 & R_1 \end{bmatrix} U', \text{ and } \begin{bmatrix} R_1 & R_1 \\ R_2 & R_2 \end{bmatrix} U', \tag{4.27}$$

respectively, where U' is short for $U(\mathbb{M}')$. Let

$$f(a_2, a_1, a_0) = \begin{bmatrix} R_0 & R_1 \\ R_2 & R_3 \end{bmatrix}^{a_2} \begin{bmatrix} R_2 & R_2 \\ R_2 & R_2 \end{bmatrix}^{a_1} \begin{bmatrix} R_3 & R_2 \\ R_2 & R_3 \end{bmatrix}^{a_0}, \tag{4.28}$$

then $f : (\mathbf{Z}_2)^3 \to U(\mathbb{M})/U(\mathbb{M}')$ is a isomorphism between $\mathcal{A} = (\mathbf{Z}_2)^3$ and $U(\mathbb{M})/U(\mathbb{M}')$. Under this isomorphism, the of $U(\mathbb{M}')$ corresponding to $C_0' \cdots C_7'$ are the images of (0,0,0), (0,0,1), (0,1,0), (0,1,1), (1,0,0), (1,0,1), (1,1,0), (1,1,1) respectively. The label map $m : \mathcal{A} \to \mathbb{M}/\mathbb{M}'$ is an isometric labeling with the induced one-to-one map of $U(\mathbb{M})/U(\mathbb{M}') \to \mathbb{M}/\mathbb{M}'$. So the 8-state MSTTC based on the sub-optimal coset

partitioning given in part (b) of Figure 4.11 is a generalized coset codes in the sense of [11], and there by geometrically uniform. In a similar way, one can show that the other MSTTCs are also geometrically uniform.

4.4.3 Performance Comparison

Geometrical uniformity of the MSTTCs based on the optimal and the sub-optimal coset partitioning make the performance analysis simplified, since the distance profiles of geometrically uniform code are transparent to the choice of the reference codeword.

The combined d^2_{free}'s of the conventional and the new hybrid-ARQ scheme are summarized in Table 4.2 for 8- and 16-state MSTTCs. The '>' sign in the table represents the distance contribution by the matrix pairs which do not satisfy the equal eigenvalue condition.

Table 4.2: Comparison of the d^2_{free}.

q	initial	2nd transmission		3rd transmission	
		conventional	new	conventional	new
8	8	16	20	24	32
16	>8	>16	24	>24	40

The overall d^2_{free} after retransmission(s) using different codes that is optimal over each transmission is larger than the d^2_{free} using unique code for all transmissions.

4.5 Simulation Results

A system with 2 transmit-antenna and 1 receive-antenna is used. The frame length is 130 symbols. The transmission rate is 2 bits/symbol/transmission. Figure 4.12 and 4.13 show the FER performance of the 8-state and 16-state MSTTCs employing two types of HARQ schemes. HARQ I represents the conventional code combining hybrid-ARQ scheme using unique code for all transmissions. HARQ II

represents the code combining hybrid-ARQ scheme using the different code designed for each transmission.

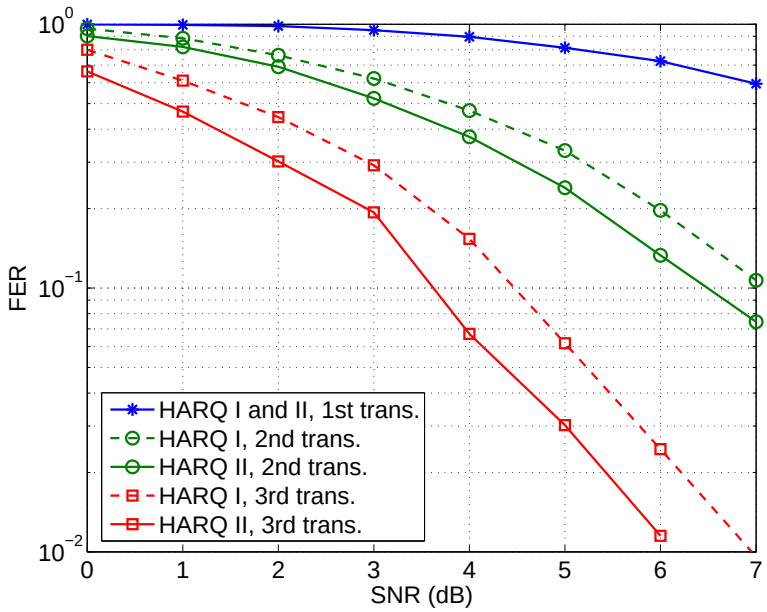


Figure 4.12: Frame error rate for the 8-state MSTTCs.

4.6 Conclusions

As the simulation result indicated, hybrid-ARQ scheme using different code further improves the FER performance of the conventional code combining by about 0.5 dB after 2^{nd} transmission, and by about 0.8 dB at after 3^{rd} transmission.

The sub-optimal partition chain of the super-constellation is proposed for the hybrid-ARQ scheme using MSTTCs. The hybrid-ARQ scheme, consisting of the optimal MSTTC for each transmission, outperforms the hybrid-ARQ scheme, consisting of the same MSTTC for all transmissions.

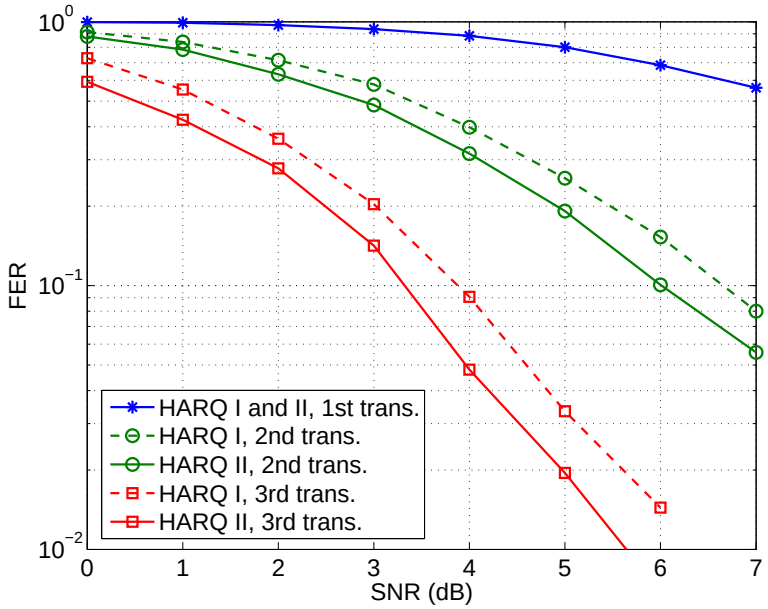


Figure 4.13: Frame error rate for the 16-state MSTTCs.

Chapter 5

Conclusions

In this dissertation, issues about using ARQ in MIMO communication systems have been addressed. Historically, ARQ techniques have been designed for SISO communication systems. This work analyzed and developed the ARQ retransmission protocols and some error control codes for ARQ combined MIMO communication systems. The major contribution can be divided into several parts.

5.1 Contributions

In Chapter 2, the channel utilization and transmission delay expressions are derived for a communication link consisting of multiple parallel channels with different transmission rates and different packet error rates. Generalized retransmission ARQ protocols for SW and GBN ARQ are proposed to improve the channel utilization and transmission delay when applied to multiple parallel channels.

Chapter 3 examines the performance improvement of employing a type-I hybrid-ARQ scheme in MIMO systems where we assume that the channel state information is available at both the transmitter and the receiver. Spatio-temporal Vector coding [31] has been used to convert the MIMO channel into parallel channels. Upper and lower bounds for the decoded probability of error and retransmission are derived. These bounds are used to demonstrate that the type-I hybrid-ARQ modifications reduce the code gap without a reduction in quality of service or an increase in code complexity.

Chapter 4 presents a hybrid-ARQ scheme employing the multidimensional space-time code (MSTTC) as the FEC code. The concept of sub-optimal partition

chain of the super-constellation is proposed. The sub-optimal partition chains are used to design the retransmission error control codes which serve as supplementary codes of the previous transmissions.

5.2 Areas of Future Work

The delay of SR-ARQ consists of the transmission delay and the resequencing delay. We have not consider the resequencing delay in Chapter 2. Even though the packet-to-channel assignment rule does not affect the channel utilization and the transmission delay of SR-ARQ as we have shown in Chapter 2, but it is expected to effect the resequencing delay of the multichannel SR-ARQ protocol. Shacham and Shin [37] described and analyzed a modified SR ARQ protocol for used over parallel channels with the same transmission rate but different packet error rates. It will be our future work to extend the result to the case that all the channels have different transmission rates and different packet error rates.

We assume slow fading channel over all this dissertation. The performance and code design of ARQ technique in fast fading channels can be studied.

The super-constellation given in Chapter 4 is for MIMO communication system with 2 transmit antennas and QPSK. We might want to extend the super-constellation and the partition method of the super-constellation to the case of any number of transmit antennas and any constellation.

Bibliography

[1] S. Alamouti. A simple transmit diversity technique for wireless communications. *IEEE Journal on Selected Areas in Communications*, 16(8):1451–1458, October 1998.

[2] M. Anagnostou and E. Protonotarios. A selective repeat ARQ protocol with n channels. *Proc. GLOBECOM*, pages 1445–1449, 1984.

[3] E. Biglieri et al. *Introduction to trellis-coded modulation with application*. Macmillan Publishing Company, 1991.

[4] J. Bingham. Multicarrier modulation for data transmission: an idea whose time has come. *IEEE Communications Magazine*, 28(5):5 – 14, May 1990.

[5] J.-F. Chang and T.-H. Yang. Multichannel ARQ protocols. *IEEE Transactions on Communications*, 41(4):592–598, April 1993.

[6] D. Dardari. A uniform power and constellation size bit-loading scheme for OFDM based WLAN systems. *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, 2(5-8):1220–1224, September 2004.

[7] Z. Ding and M. Rice. Throughput analysis of ARQ protocols for parallel multi-channel communications. In *IEEE Transactions on Wireless Communications*. Accepted for Publication.

[8] Z. Ding and M. Rice. Type-I hybrid-ARQ using mtcm spatio-temporal vector coding for MIMO systems. *Proceedings of the IEEE International Conference on Communications*, 4:2758 – 2762, May 2003.

[9] Z. Ding and M. Rice. Throughput analysis of ARQ protocols for parallel multi-channel communications. *Proceedings of the IEEE Global Communications Conference*, 3:1279 – 1283, November 28 – December 2 2005.

[10] D. Divsalar and M. Simon. Multiple trellis coded modulation (MTCM). *IEEE Transactions on Communications*, 36(4):410–419, April 1988.

[11] J. Forney, G.D. Geometrically uniform codes. *IEEE Transactions on Information Theory*, 39(5):1491–1512, September 1991.

[12] G. J. Foschini. Layered space-time architecture for wireless communication in a fading environment when using multiple antennas. *Bell Labs Technical Journal*, 1(2):41–59, Autumn 1996.

[13] J.-C. G. Guey, M. Fitz, M. Bell, and W.-Y. Kuo. Signal design for transmitter diversity wireless communicationsystems over rayleigh fading channels. *IEEE Transactions on Communications*, 47(4):527–537, April 1999.

[14] B. Hochwald et al. Systematic design of unitary space-time constellations. *IEEE Transactions on Information Theory*, 46(6):1962–1973, September 2000.

[15] B. Hochwald and T. Marzetta. Unitary space-time modulation for multiple-antenna communications in rayleigh flat fading. *IEEE Transactions on Information Theory*, 46(2):543–564, March 2000.

[16] B. Hochwald and W. Sweldens. Differential unitary space-time modulation. *IEEE Transactions on Communication*, 48(12):2041–2052, December 2000.

[17] R. Horn and C. Johnson. *Matrix Analysis*. Cambridge Univ. Press, New York, 1994.

[18] D. Ionescu. New results on space-time code design criteria. *IEEE Wireless Communication and Networking Conference, 1999. WCNC*, 2:684–687, September 1999.

[19] D. Ionescu. On space-time code design. *IEEE Transactions on Wireless Communication*, 2(1):20–28, January 2003.

[20] D. Ionescu, K. Mukkavilli, Z. Yan, and J. Lilleberg. Improved 8- and 16-state space-time codes for 4psk with two transmit antennas. *IEEE Communication letters*, 5(7):301–303, July 2001.

[21] H. Jafarkhani and N. Hassanpour. Super-quasi-orthogonal space-time trellis codes for four transmit antennas. *IEEE Proceedings on Vehicular Technology Conference, 2002*, 1:247–251, September 2002.

[22] Y. S. Jung and J. H. Lee. Hybrid-ARQ scheme employing different space-time trellis codes in slow fading channels. *IEEE Transactions on Wireless Communications*, 4(1):215–227, January 2005.

[23] S. Kasturia, J. Aslanis, and J. Cioffi. Vector coding for partial-response channels. *IEEE Transactions on Information Theory*, 36(4):741 – 762, July 1990.

[24] T. Koike, H. Murata, and S. Yoshida. Hybrid ARQ scheme suitable for coded MIMO transmission. *IEEE International Conference on Communications*, 5:2919–2923, June 2004.

[25] J. Lassing, E. Ström, E. Agrell, and T. Ottosson. Computation of the exact bit-error rate of coherent M-ary PSK with Gray code bit mapping. *IEEE Transactions on Communications*, 51(11):1758–1760, November 2003.

[26] J. Lechleider. The feasibility of using adaptive transmitters to suppress crosstalk. In *Proceedings of the IEEE International Conference on Communications*, volume 1, pages 548 – 551, Boston, MA, June 11-14 1989.

[27] J. P. Odenwalder. Optimal decoding of convolutional codes. *Ph.D. dissertation, University of California, Los Angeles*, 1970.

[28] E. Onggosanusi, A. Dabak, Y. Hui, and G. Jeong. Hybrid ARQ transmission and combining for MIMO systems. *IEEE International Conference on Communications*, 5:3205–3209, May 2003.

[29] L. Pierrugues, O. Moreno, P. Duvaut, and F. Ouyang. DMT performance improvement based on clustering modulation, applications to ADSL. *IEEE International Conference on Communications*, 5(20-24):2807–2811, June 2004.

[30] S. S. Pietrobon, R. H. Deng, and G. Ungerboeck. Trellis-Coded multidimensional phase modulation. *IEEE Transactions on Information Theory*, 36(1):63–89, Jan 1990.

[31] G. G. Raleigh and J. M. Cioffi. Spatio-temporal coding for wireless communication. *IEEE Transactions on Communications*, 46(3):357–366, March 1998.

[32] L. K. Rasmussen and S. B. Wicker. The performance of type-I trellis coded hybrid-ARQ protocols over AWGN and slowly fading channels. *IEEE Transactions on Information Theory*, 40(2):418–428, March 1994.

[33] B. Saltzberg. Performance of an efficient parallel data transmission system. *IEEE Transactions on Communication Technology*, 15(6):805–811, December 1967.

[34] H. Samra and Z. Ding. Sphere decoding for retransmission diversity in MIMO flat-fading channels. *IEEE International Conference on Acoustics, Speech, and Signal Processing, 2004. (ICASSP'04)*, 4:585–588, May 2004.

[35] H. Samra and Z. Ding. New MIMO ARQ protocols and joint detection via sphere decoding. *IEEE Transactions on Signal Processing*, 54(2):473–482, Feburary 2006.

[36] N. Shacham. Packet resequencing in reliable transmission over parallel channels. *Proc. Int. Conf. Commun.*, pages 557–562, 1987.

[37] N. Shacham and B. C. Shin. A selective-repeat-ARQ protocol for parallel channels and its resequencing analysis. *IEEE Transactions on Communications*, 40(4):773–782, April 1992.

[38] D. J. C. J. Shu Lin. *Error control coding, fundamentals and applications.* Prentice-Hall, Inc., 1983.

[39] R. V. Sonalkar. Bit- and power-allocation algorithm for symmetric operation of DMT-based DSL modems. *IEEE Transactions on Communications*, 50(6):902–906, June 2002.

[40] W. Su and X. Xia. Signal constellations for quasi-orthogonal space-time block codes with full diversity. *IEEE Transactions on Information Theory*, 50(10):2331–2347, October 2004.

[41] H. Sun, J. Manton, and Z. Ding. Progressive linear precoder optimization for MIMO packet retransmissions. *IEEE Journal on Selected Areas in Communications*, 24(3):448–456, March 2006.

[42] H. Sun, H. Samra, Z. Ding, and J. Manton. Constrained capacity of linear precoded ARQ in MIMO wireless systems. *IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005. (ICASSP'05)*, 3:1023–1026, September 2005.

[43] V. Tarokh, H. Jasarkhani, and A. R. Calederbank. Space-time block codes from orthogonal design. *IEEE Transactions on Information Theory*, 45(5):1456–1467, July 1999.

[44] V. Tarokh, N. Seshadri, and A. R. Calderbank. Space-time codes for high data rate wireless communication: performance criterion and code construction. *IEEE Transactions on Information Theory*, 44(2):744–765, March 1998.

[45] O. Tirkkonen, A. Boariu, and A. Hottinen. Minimal non-orthogonality rate 1 space-time block code for 3+ Tx antennas. *IEEE Sixth International symposium on Spread Spectrum Techniques and Applications*, 2(6-8):429–432, September 2000.

[46] J. C. Tu. Theory, design and application of multi-channel modulation for digital communications. *Ph.D. dissertation, Stanford Univ., Stanford, CA*, 1991.

[47] A. Van Nguyen and M. Ingram. Hybrid ARQ protocols using space-time codes. *IEEE Proceedings on Vehicular Technology Confernece, 2001*, 4:2364–2368, October 2001.

[48] J. Wallace, M. Jensen, L. Swindlehurst, and B. Jeffs. Experimental characterization of the MIMO wireless channel: data acquisition, analysis, and modeling,. *IEEE Transactions on Wireless Communications*, 2:335–343, March 2003. Data sets are available on-line at http://www.ee.byu.edu/wireless/data/probe_narrow.cgi.

[49] S. Wicker. *Error Control Systems for Digital Communication and Storage*. Prentice-Hall, Upper Saddle River, NJ, 1995.

[50] S. B. Wicker. An adaptive type-I hybrid-ARQ technique using the Viterbi algorithm. *IEEE Military Communication Conference*, pages 15.5.1–15.5.5, October 1988.

[51] W.-C. Wu, S. Vassiliadis, and T.-Y. Chung. Performance analysis of multi-channel ARQ protocols. *Proceedings of the 36th Midwest Symposium on Circuits and Systems*, 2:1328–1331, August 1993.

[52] H. Yamamoto and K. Itoh. Viterbi decoding algorithm for convolutional codes with repeat request. *IEEE Transactions on Information Theory*, 26(5):540–547, September 1980.

[53] Z. Yan and D. Ionescu. Geometrical uniformity of a class of space-time trellis codes. *IEEE Transactions on Information Theory*, 50(12):3343–3347, December 2004.

[54] J. Yu, Y. Li, H. Murata, and S. Yoshida. Hybrid-ARQ scheme using different TCM for retransmission. *IEEE Transactions on Communications*, 48(10):1609–1613, October 2000.

[55] J. Yun, W. Jeong, and M. Kavehrad. Throughput analysis of selective repeat ARQ combined with adaptive modulation for fading channels. *MILCOM Proceedings*, 1:710–714, October 2002.

[56] J. Yun, W. Jeong, and M. Kavehrad. Throughput performance analysis for automatic-repeat-request techniques as combined with adaptive rate transmission. *Radio and Wireless Conference, 2002*, pages 11–14, August 2002.

[57] H. Zheng, A. Lozano, and M. Haleem. Multiple ARQ processes for mimo systems. *The 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, 3:1023–1026, September 2002.