Euclidean Space Codes as Space-Time Block Codes

by

Anne On-Yi Pak

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Master of Science in Electrical Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2001

Author .....
Department of Electrical Engineering and Computer Science
May 11, 2001

Certified by .....
Vahid Tarokh
Associate Professor
Thesis Supervisor

Accepted by .....
Arthur C. Smith
Chairman, Department Committee on Graduate Students

Euclidean Space Codes as Space-Time Block Codes

by

Anne On-Yi Pak

Submitted to the Department of Electrical Engineering and Computer Science
on May 11, 2001, in partial fulfillment of the
requirements for the degree of
Master of Science in Electrical Engineering

Abstract

The emergence of wireless communications has transformed the concept of "any time, any where" telephony and Internet access into a reality. However, techniques for improving or maintaining acceptable levels of channel integrity and information reliability that have conventionally been applied to wirelined systems are no longer appropriate for the wireless medium. Specifically, wireless channels are plagued with the existence of multipath fading, which often degrade the signal-to-noise ratio (SNR) at the receiver and make correct signal detection and decoding more difficult.

Diversity techniques are commonly applied to such systems to ameliorate the effects of the fading. This thesis deals with the simultaneous exploitation of spatial and temporal diversity through space-time codes. Space-time codes are effective means of achieving transmit diversity as well as exploiting the fact that the capacity of a multiple transmit and receive antenna system grows *at least* linearly with the number of transmit antennas.

The theory and fundamentals of Euclidean space forward error correction (FEC) codes have been well explored and established. The construction technique presented in this thesis aims to extend the performance improvements of these codes by providing additional coding gain and introducing diversity gain, both of which are inherent to space-time codes. A detailed description of the encoding and decoding mechanisms for space-time codes based on Euclidean space channel codes are presented as well as the diversity order achieved by them. Simulation results will verify that Reed-Muller RM(1,3) and extended Hamming codes achieve diversity orders of 2 and 3, respectively.

Thesis Supervisor: Vahid Tarokh
Title: Associate Professor

Acknowledgments

This section is supposed to acknowledge all the people who contributed to the production and completion of this thesis; however, in the finite space to which I have been confined, I find this to be a nearly impossible task. For me, producing this thesis was more than just an academic exercise. It was an invaluable exploration of my intellectual limits, and it also played a crucial role in the development of my emotional maturity and strength of character. And so I thank all who made this journey so incredibly exciting.

To my wonderful advisor, Vahid. Thank you for leaving AT&T and coming to MIT. It goes without saying that without you, I would not have a thesis, but what is more important is that I do not think I could have finished this project without your relentless encouragement, endless patience and infinite wisdom.

To Professors George Verghese and Vincent Chan. Thank you both immensely for all the support and assistance you provided me last summer when I was looking for a new advisor. You provided with the hope and courage I needed and for that I am most grateful. MIT is very lucky to be in the company of two such dedicated and caring faculty members.

To my family. Thank you for your love and support, and for encouraging me to leave my sheltered existence that was Maryland to pursue my dreams and aspirations academically as well as emotionally.

To Irina. Thank you, thank you, thank you. You were a source of inspiration for me from the very beginning. Your zeal for learning, enthusiasm in exploration and thirst for knowledge are contagious. Your smile, laughter and everyday cheer made the days go by so effortlessly and joyously. I am honored that you chose me to be your friend, will cherish the memories we have made and look forward to making many more of them in the years to come. Good luck to you in all your future endeavors!

To my better half, Joe. Although we have known each other for less than a year, I feel like I have known you forever. You have no idea how much I value our friendship, appreciate your relentless support, treasure your companionship and cherish your

love. You kept me optimistic, hopeful and most importantly, sane when times were tough and you kept me level-headed, focussed, and motivated when times were good. All my love...

Special thanks to Chan-Vee Chong and Andy Wang both of whom provided me with invaluable advice, guidance, and words of wisdom. Thank you for your time, patience and willingness to teach me. You both are amazing teachers, incredible people and wonderful cooking partners. I will miss our 6PM dinner dates next year!

Unending gratitude goes to my dear friend Jin Fan. Thank you for putting up with my mood swings, my fits of anxiety and frustration, and my frequent crying spells. But more importantly, thank you for being such a wonderful roommate, friend and confident.

Many thanks go to the rest of my friends both at MIT and back in Maryland. You know who you are. To my friends at MIT, thank you for making my experiences here so incredible and memorable. You turned homesickness into a 24-hour bug and you helped alleviate many of the stresses and frequent pangs of anxiety I experienced while taking classes and doing research. I will cherish all the wonderful memories we made the past two years and hope we will not drift too far apart in the future. And to my friends back in Maryland, thank you for keeping in touch with me despite the long distance, and for making Maryland a place I always felt welcomed and looked forward to returning to.

Two final groups of people I would like thank are the faculty and students in the Laboratory for Information and Decision Systems (LIDS) and the residents of Ashdown House. The efficiency and kindness of the staff members in LIDS Headquarters and the willingness of LIDS students to openly and freely share ideas contributed in countless ways to the integrity, accuracy and completion of this work. The diversity, friendliness and family-like living environment created by the residents of Ashdown transformed it from being merely a dormitory into a "home away from home".

Lastly, I would like to thank the American Society for Engineering Education for their generous sponsorship and financial support of this project.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Wireless communications provide the necessary mobility for providing communication and, more recently, Internet access to more places, in effect, extending the accessibility of such services beyond the realm of the public switched telephone network (PSTN). The current megatrend in mobile communications is the merging of mobile communication technology with the Internet to give birth to a whole new suite of mobile Internet applications [1]. Furthermore, with already more than one billion subscribers worldwide, subscribership to wireless services is expected to exceed those of wireline services by the year 2005 [2]. With the integration of communication into information technology, the UMTS Forum expects that by 2010, the 90 million mobile subscribers in Europe using mobile multimedia services will make up 60% of total data traffic [1].

Parallel to increasing demand for wireless accessibility is a growing list of wireless applications that are constantly evolving and developing. The mobile cellular phone handset no longer simply places and accepts phone calls but now acts a portable office, providing data through Internet access and multimedia capabilities. In fact, it is expected that by 2004, the increased functionality of mobile handsets will cause tetherless access to the Internet to surpass PC access and that mobile terminals will eventually replace PC's as the principle man-machine interface of the future [1]. Figures 1-1 and 1-2 [1] illustrate projected trends in telecommunications and information technology.

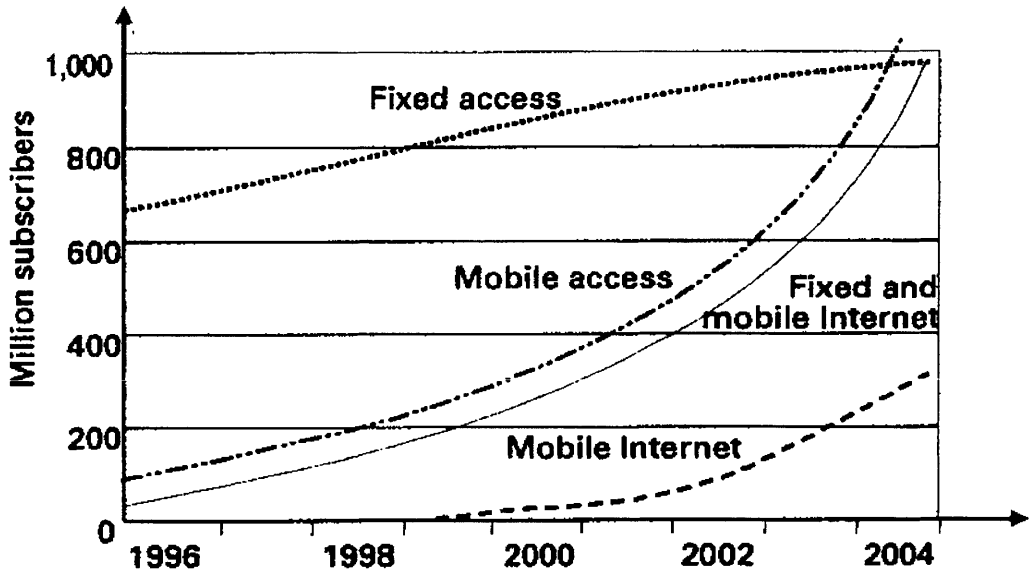

Figure 1-1: Traffic growth for different access systems

Providing means to support these additional services is the main objective of many third generation (3G) wireless standards such as GSM EDGE, WCDMA and CDMA2000. Other wireless applications include wireless local area networks (LANs), digital broadcasting, and wireless local loops.

1.1 Motivation

Incorporation and support for these new wireless applications introduces a whole host of new challenges that must be addressed and resolved. Several such challenges are as follows:

High bandwidth efficiency: Bandwidth is a scarce and consequently, expensive commodity and should be utilized efficiently and cost-effectively.

High data rate: Multimedia traffic often include applications with different data rate requirements. Supporting these applications over existing frequency band designations requires the use of higher-level modulation in order to achieve higher data rates.

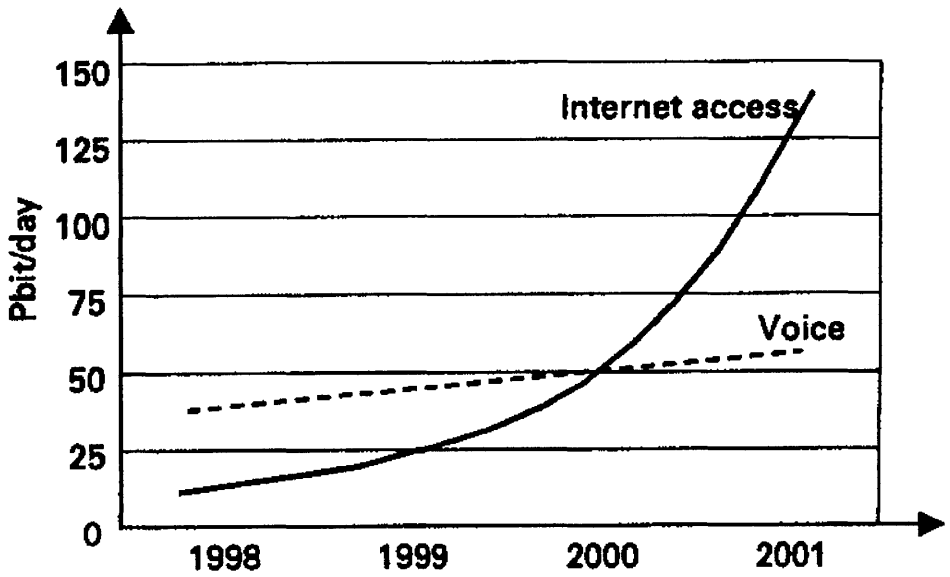

Figure 1-2: Trends in voice and data services

Resource allocation: Inherent to multimedia applications are differences in quality of service (QoS) tolerances; for example, a voice application may be more tolerant of errors but less tolerant of delay in contrast to data applications, which impose antithesis requirements.

Mobility: In addition to multipath fading, which already plagues ***stationary*** wireless environments, the element of relative motion between the transmitter and receiver makes the physical channel time-varying and hence factors such as Doppler shift must be taken into account.

Portability: Especially for wireless systems that support mobile handheld devices such as cellular phones or portable radios, the size, weight, shape and battery life of the devices themselves are quite important. Low power consumption by all physical components as well as highly power-efficient coding schemes should be implemented to prolong the life of devices.

Privacy and security: In addition to minimizing interception of voice calls by other wireless users, the introduction of data applications warrants the provision of

data security and encryption capabilities.

Evidently, providing support for multimedia applications over wireless channels has introduced many new challenges and problems of interest to both academia and industry. This thesis aims at providing means for increasing the effective data rate over wireless channels by making transmissions over such a medium more robust to noise and other distortions. This is achieved by using space-time codes, which are codes that employ both spatial and temporal diversity techniques. Diversity techniques exploit certain channel characteristics, which would otherwise be detrimental to conventional receivers for wireless systems.

1.2 Thesis Outline

The remaining chapters of this thesis are organized as follows. Each chapter assumes basic knowledge of the fundamental concepts in the field, but can be read independently from each other.

Chapter 2 reviews the new challenges and characterizing parameters associated with modelling wireless channels. It also reviews the concept of diversity and provides a brief review of transmit and receive diversity techniques.

Chapter 3 reviews concepts associated with the new code paradigm that are space-time block codes.

Chapter 4 introduces the new code construction technique of space-time block codes from Euclidean space block codes and presents an analysis of the diversity order of the space-time codes resulting from the new construction.

Chapter 5 details the simulation written to evaluate the performance of this new construction scheme and discusses findings derived from simulation results.

Chapter 6 summarizes the main ideas of the thesis and suggests directions for future research endeavors.

Chapter 2

Wireless Channel Conditions

In contrast to wired systems, the channel for wireless communications is time-varying and, if mobility is a feature, highly dynamic. Due to complexities created by the presence of multipath components, modelling the channel as additive white Gaussian noise (AWGN) is no longer appropriate. Consequently, new parameters must be introduced to more accurately describe the channel of a wireless communications system.

However, before proceeding to these introductions, it should be noted that the AWGN channel model is still useful for establishing an upper bound on system performance. In a wireless environment, there exists, in addition to the thermal noise of AWGN, multipath components, which often have the effect of attenuating the received signal. Furthermore, since radio channels are a shared-medium, there is multiple access interference from other users, which can obscure the original signal. Naturally, the bit error rate (BER) for a given signal-to-noise (SNR), resulting from the incorporation of these latter two factors in the channel model, is expected to be greater than the BER for the simple AWGN channel model. Hence, if multipath mitigation or diversity techniques are employed, then the resulting BER is bound to improve and approach the limit set by the BER of the AWGN channel, thus explaining why the AWGN channel typically upper bounds system performance [3].

Cell Classification	Size	Comments
Large cells	< 30 km	no LOS component
Small cells	< 1 km	LOS is probable
Microcells	< 100 m	LOS is a feature

Table 2.1: Line-of-sight dependency on cell size

2.1 Propagation models

Unlike for wired communications, the propagation path between a transmitter-receiver (T-R) pair can be time-varying and unpredictable. There exists three principle mechanisms that affect the propagation of electromagnetic waves through such time-varying channels [4]:

Reflection: An electromagnetic wave experiences a change in direction when impacting obstacles whose dimensions are significantly larger than the wavelength of the wave.

Diffraction: Electromagnetic waves bend and spread out around objects with sharp irregularities such as edges.

Scattering: Electromagnetic waves disperse when they traverse through dense mediums composing of obstacles with dimensions smaller than the wavelength of the wave; for example, foliage, street signs and lamp posts can induce scattering.

The received signal manifests these three effects in two ways: large-scale path loss and small-scale fading. Furthermore, the presence of an LOS component, which depends on cell size and shape, can profoundly affect the propagation characteristics of the system [3]. Table 2.1 [4] lists some typical numerical guidelines.

2.1.1 Mobile Radio Propagation: Large-Scale Path loss

The large-scale propagation model captures the macroscopic variations in signal strength over T-R separation distances, typically on the order of several hundreds

Environment	Path Loss Exponent, n
Free space	2
Urban area cellular radio	2.7 to 3.5
Shadowed urban cellular radio	3 to 5
In building line-of sight	1.6 to 1.8
Obstructed in building	4 to 6
Obstructed in factories	2 to 3

Table 2.2: Path-loss exponents for different environments

to thousands of meters [4]. The signal strength is affected by two factors: distance-dependent path loss and log-normal shadowing.

Log-distance Path Loss Model

For both indoor and outdoor radio channels, signals experience attenuation in signal strength over long distances. This attenuation can be described as a function of the T-R separation distance, d and a path-loss exponent, n. Table 2.2 lists the typical path loss exponents n for various mobile radio environments [4]. The average path loss in decibels is [4]:

$$\overline{PL}[dB] = \overline{PL}(d_0) + 10n \log \frac{d}{d_o}, \tag{2.1}$$

where d_0 is the close-in reference distance, which depends on the coverage of the cellular system. For large-coverage cellular systems, d_0 is 1 km; for microcellular systems, d_0 is typically smaller, such as 1 m or 100 m [4].

Log-normal Shadowing

Equation (2.1) does not capture the effects of obstacles in the propagation path. The random shadowing effects that occur over a large number of measurement locations, which are characterized by the same T-R separation distance but have dissimilar shadow clutter distributions in the propagation path, can be described by a log-normal distribution. The resulting effect is called log-normal shadowing [4]. Consequently, a

more complete description of the path loss is:

$$\overline{PL}(d)[dB] = \overline{PL}(d_0) + 10n \log \frac{d}{d_o} + X_\sigma. \tag{2.2}$$

$\overline{PL}(d)$ is log-normally distributed with a mean equal to the mean distance-dependent expression, given by equation (2.1). The randomness of the fading is captured by X_σ, which is a log-normal zero-mean random variable (RV) with a variance in decibels on the order of 6 dB.

2.1.2 Mobile Radio Propagation: Small-Scale Fading

Due to obstacles in the path of propagation between the transmitter and receiver, the original signal gets deflected in various directions, and consequently, multiple time-delayed versions of the signal arrive superimposed at the receiver. These replicas are often attenuated in amplitude and shifted in phase with respect to the original signal. The result can be either constructive or destructive interference, which becomes manifested as signal enhancement or fading, respectively.

Since the replicas arrive at the receiver at different times than the original signal, the spread in time over which "echoes" of the original signal can still be discerned is quantified by the *multipath delay spread* in the time-domain and by the *coherence* ***bandwidth***, in the frequency-domain.

Furthermore, since the environment of the propagation path may be changing, the channel is characterized as being time-varying. Additional quantities used to parameterize the channel under these conditions include *coherence* ***time*** and *Doppler spread.* The coherence time is the maximum delay between two replicas of a signal, such that the correlation between them is small. Hence, two replicas separated by a time difference less than the coherence time will be highly correlated. The relationships between these quantities are outlined in Table 2.3.

The terminology is admittedly confusing. Channel characteristics depend on the relationship of the bit or symbol period T to either the multipath spread or coherence time. An attempt to clarify these relationship is as follows [4]:

Terminology and Relationships		
Term	*Symbol*	*Relationship*
Multipath spread	T_m	reciprocals
Coherence Bandwidth	Δf_c	
Coherence time	Δt_c	reciprocals
Doppler Spread	W_{DS}	

Table 2.3: Wireless channel parameters

Small-Scale fading	
Based on multipath delay spread	
Flat Fading	**Frequency Selective Fading**
1. BW of signal $< \Delta f_c$	1. BW of signal $> \Delta f_c$
2. $T_m <$ Symbol period	2. $T_m >$ Symbol period
3. No channel-induced ISI	3. Channel-induced ISI

Table 2.4: Small-scale fade dependency on multipath delay spread

Flat fading: The fading characteristics of the channel are constant over the bandwidth of the signal. Furthermore, the multipath delay spread is smaller than the symbol period; therefore, the existence of multipath does not cause adjacent symbols or their "echoes" to interfere with each other.

Frequency Selective fading: Different frequencies within the bandwidth of the signal are faded to different degrees than other frequencies. Moreover, due to a multipath spread greater than one symbol period, adjacent symbols sent in time will overlap at the receiver resulting in channel-induced intersymbol interference (ISI).

A channel can change due to relative motion between the transmitter and receiver. Consequently, it can be categorized as exhibiting either fast or slow fading, depending on the speed of the moving terminal.

Fast fading: If the mobile is moving quickly relative to one symbol period, then the motion of the mobile will affect the fading characteristics acting on the transmitted signal.

Small-Scale fading	
Based on Doppler spread	
Fast Fading	**Slow Fading**
1. BW of signal $< W_{DS}$	1. BW of signal $> W_{DS}$
2. $\Delta t_c <$ Symbol period	2. $\Delta t_c >$ Symbol period
3. Channel variations faster than baseband signal variations	3. Channel variations slower than baseband signal variations

Table 2.5: Small-scale fade dependency on Doppler spread

Slow fading: If the mobile is moving slowly relative to one symbol period, then the effects of the motion can be neglected and the mobile can be treated as being virtually stationary.

Tables 2.4 and 2.5 summarize the main points behind these channel qualifiers.

Hence, in a flat fading channel lacking LOS with local scatterers, the fading due to multipath is modelled by a complex Gaussian RV. Written in polar form, the amplitude of the RV is Rayleigh-distributed and the phase is uniformly-distributed. This is the channel model that is assumed for the analyses and simulations presented in this thesis.

To effectively combat distortions due to multipath, the conventional technique of increasing transmit power, as used in AWGN channels, is no longer sufficient. In fact, it was found that for AWGN channels, the SNR need only be increased by 1-2 dB; however, for multipath channels, increases on the order of 10 dB are necessary to achieve comparable performance improvements [5]. Often, system requirements prohibit such large expenditures of power and/or bandwidth, hence motivating the need to find alternative means for achieving reliable transmissions in a multipath fading channel [5].

A logical solution to this problem would be to provide the receiver with replicas of the original signal that could be manipulated or combined in such a way that the original signal can be decoded more reliably. These replicas, coined *diversity*, come in several forms, the most common being in time, frequency and space.

2.2 Diversity

Diversity is best exploited under channel conditions that best promote independence between the fade coefficients affecting each signal replica. Often, sending these replicas is compared to the act of interleaving bits within a repetition code to break up burst errors, hence obtaining independent errors [6]. As aforementioned, there are three principle forms of diversity:

- *Temporal Diversity* - typically used in conjunction with time interleaving and channel coding, the transmitted signals are sent multiple *times* to the receiver [5]. This technique is most effective when the difference in signal replica transmission times exceed the coherence ***time*** $(\Delta t)_c$ of the channel; signals transmitted within one coherence time of each other will be have a nonzero correlation and consequently experience dependent fading [6]. Consequently, it is desirable to have a fast-fading channel such that the wait time between transmissions can be minimized. This is an especially important requirement for delay sensitive applications such as voice applications.

- *Frequency Diversity* - the replicas of the same signal are modulated up to multiple carrier frequencies and are transmitted simultaneously or in parallel. The key to this technique is having the carrier separation exceed the coherence ***bandwidth*** $(\Delta f)_c$ of the channel. This ensures that the replicated transmissions are independent and allows the multipaths to be resolved and effectively processed by the optimum wideband signal receiver, the RAKE receiver [6].

- *Spatial or Antenna Diversity* - multiple transmission and receive antennas are used to provide redundancy in the spatial domain. The receive antennas should be spaced sufficiently far apart to ensure that multipath components of the received signals experience independent propagation delays [6]. Based on the Jakes model, which is appropriate for receivers surrounded by many local scatterers, such as a remote unit, this minimum separation is on the order of half

a wavelength[1]. However, for basestations, which are typically erected at higher elevations, there are far less scatterers, so separations on the order of ten wavelengths may be needed [5]. Furthermore, in contrast to both temporal and frequency diversity, spatial diversity does not require additional expenditure in time or bandwidth for its implementation because all signals are transmitted simultaneously and over the same frequency bands [7].

In many cases where there is Rayleigh fading and the channel is either non-frequency selective or is slowly time-varying, frequency and temporal diversity cannot each be individually exploited. Consequently, spatial diversity is needed. Smart coding over multiple transmit antennas exploit transmit diversity and using multiple receive antennas automatically result in full receive diversity. The next section will elaborate on these two notions.

2.2.1 Transmit versus Receive Diversity

As previously mentioned, there are two types of diversity that can contribute to the overall diversity gain of a multiple input, multiple output (MIMO) system with multiple transmit and receive antennas. They are transmit and receive diversity, respectively. In a practical system where physical limitations are a factor in system design, using antenna arrays are typically only employed at the basestation because incorporating multiple antennas on a small handheld device such as a cellular phone is often unrealistic. In this case, downlink transmissions will benefit from transmit diversity and uplink transmissions will benefit from receive diversity [8].

Receive Diversity Techniques

Full receive diversity has been historically exploited using multiple receive antennas employing maximal receive ratio combining (MRRC) techniques. This optimum linear combining technique uses diversity to combat multipath fading and to suppress

[1]Physical limitations in current system designs often prevent the incorporation of multiple receive antennas on platforms the size of cellular phones. Consequently, the use of antenna arrays is typically implemented at the basestation.

interference, thereby increasing the capacity of the system while also improving its performance [10].

Essentially, a signal is sent from the single transmit antenna and is received by multiple receive antennas. The receive antennas are spaced at least half a wavelength apart to promote independence between replicas [8] or, in other words, to ensure that each replica of the original signal undergoes independent fading. MRRC weights and combines the received signals to maximize the signal-to-interference plus noise ratio (SINR) at the receiver. Assuming perfect channel estimation, these weights are typically equal to the complex conjugate of the complex Gaussian fading coefficient associated with the channel between each transmit-receive antenna pair.

In other words, if the signal s_0 was broadcast to all M receive antennas and the fade coefficients corresponding to each transmit-receive antenna pair is characterized by $\alpha_m|_{m=1,...,M}$, then

$$r_m = \alpha_m s_0 + \eta_m \qquad \forall m = 1, ..., M$$

where η_m represents the circularly symmetric additive white Gaussian noise at the m^{th} receive antenna. Optimum linear combining or MRRC would yield the sufficient statistic

$$\hat{s}_0 = \sum_{m=1}^{M} \alpha_m^* r_m$$

MRRC is an example of one technique that captures the full receive diversity benefit of multiple receive antennas [12]. There exists numerous other techniques for extracting receive diversity in single input/multiple output channels. The principle ingredients for these techniques include using coding and/or modulation in conjunction with array processing and beamforming.

Receive diversity is a topic that has been well-understood and established. In contrast, techniques for deriving transmit diversity have been relatively underdeveloped.

Transmit Diversity Techniques

Transmit diversity systems use processing at the transmitter to spread the information across multiple transmit antennas [10]. Historically, these systems have been viewed as being more difficult to design than those that exploit receive diversity because in the absence of feedback, the transmitter has nearly zero knowledge about the channel.

Transmit diversity schemes typically allow each transmit antenna to send a signal different from that sent from other antennas. Hence, a linear combination of these different signals is received by the single receive antenna. These combination coefficients are unknown at the transmitter but it is assumed they can be perfectly estimated at the receiver [8].

In other words, if $s_n|_{n=1,\ldots,N}$ are the different signals simultaneously sent during one channel access by each of the N transmit antennas, then the received signal is

$$r = \sum_{n=1}^{N} \alpha_n s_n + \eta.$$

The α_n's are assumed to be known at the receiver but not at the transmitter.

The first scheme that used linear processing at the transmitter to spread the information across multiple transmit antennas was introduced by Wittneben [13]. This technique was a variation on SIMULCAST networks, which broadcasted the same signal from antennas mounted on *different* basestations [2]. It was shown that sending the same signal over each transmit antenna yielded no diversity gain; however, if conventional or non-bandwidth expanding basestation modulation techniques such as those presented by Wittneben are employed over SIMULCAST, then diversity gains equal to that of MRRC can be achieved.

Seshadri and Winters [14] presented two signaling schemes that also achieved transmit diversity. The first one, a special case of the scheme presented by Wittenben, delays transmissions from successive antennas by multiples of the symbol interval T. For example if $s_1(t)$ is sent from antenna 1, then the signal sent from antenna 2 is

[2]Note, these techniques can be applied to transmit diversity systems that assume the transmit antennas are on the *same* basestation.

$s_1(t-T)$ and the signal sent from antenna i is $s_1\big(t-(i-1)T\big)$. The purpose of these delays is to artificially introduce multi-path distortion to the receiver, thereby making the channel frequency selective. It was shown in [14] that a maximum likelihood sequence estimator (MLSE) can resolve these deliberate multipaths at the receiver and realize a diversity benefit of N if N transmit antennas are used [15].

The second scheme derives diversity advantage from channel coding. A rate $\frac{k}{N}$ channel code encodes k information bits into a N-symbol codeword and transmits these code symbols over non-overlapping time slots from different antennas using time multiplexing. The maximum likelihood decoder reduces to correlating received signal vector with every possible codeword and selecting the codeword that yields the highest correlation. If the channel code has minimum Hamming distance $d_{min} \leq N$, then maximum likelihood decoder provides diversity d_{min} [14]. A side-effect of imposing orthogonality on the transmission of code symbol is an N-fold reduction in bandwidth efficiency.

However, if the coding at the transmitter and receiver are designed carefully, the orthogonality requirement can be somewhat relaxed. The resulting scheme will extract diversity as well as coding advantage without sacrificing bandwidth efficiency. Space-time codes, which are designed for the generalized MIMO system, are examples of such codes.

Space-time codes: Joint exploitation of transmit and receive diversity

Space-time codes use transmit diversity techniques over multiple transmit antennas and linear combining techniques similar to MRRC at the receive antennas to extract both transmit and receive diversity. However, unlike some of the transmit diversity techniques presented previously, these schemes do not require bandwidth expansion.

The block coding technique for two transmit and M receive antennas introduced by Alamouti [5] provided a fundamental basis for the more generalized space-time block codes based on orthogonal designs described by Tarokh, Jafarkhani and Calderbank [16], which essentially extended Alamouti's designs to N transmit antennas. These space-time codes exploited full transmit diversity and were characterized by

decoders based on simple linear processing techniques.

Naguib, Seshadri and Calderbank [10] proposed a scheme for using space-time codes and exploiting the spatial and temporal structure of these codes to develop minimum mean-squared error (MMSE) interference suppression and maximum likelihood (ML) techniques. They demonstrated how the joint space-time block coding and MMSE interference suppression could be beneficial to high capacity and high data rate applications. One example of such a technique is to form a concatenated code where the outer code is a conventional channel error correcting code and the inner code is the space-time code. The inner space-time code provided means for interference suppression and also provided the outer code with soft decisions which the channel decoder could use to provide further protection against channel errors. In fact, this concept of concatenating codes forms the basis of the code construction being presented in this thesis.

2.3 Summary

Several important parameters commonly used to describe a wireless channel were presented in this chapter. Through their definition and discussion, it became apparent that unless the effects of multipath were mitigated, they would hinder the possibility for reliable communication over wireless mediums. Hence, the method of applying diversity techniques was introduced. Several forms of diversity were discussed, but the majority of the discussion focused on the types and merits of antenna or spatial diversity techniques. Ultimately, the concept of joint spatial and temporal diversity exploitation was presented through a discussion of space-time coding.

Chapter 3

Space-time codes

Space-time codes are typically employed over multiple input, multiple output (MIMO) systems. Suppose there are N transmit and M receive antennas. During P consecutive transmission slots or channel accesses , N signals are transmitted simultaneously from each of the N transmit antennas. It can be assumed that inter-transmission times will exceed the coherence time of the channel and that the antennas will be spaced sufficiently far apart so that the signals transmitted from the different antennas during the different time slots will experience independent fading.

Each receive antenna detects the superposition of the N transmitted signals, attenuated by large-scale and small-scale fading and distorted by noise. The signals received at each receive antenna are all statistically similar; however, since the fading characteristics between each transmit and receive antenna pair are different and independent from each other, realizations of each actual received signal will be different.

Figure 3-1 is a block diagram illustrating the basic steps involved in space-time coding [17]. A realistic channel model is the flat fading channel, which holds the path gains constant over a block of multiple channel accesses, but changes from block to block. Let $s_n(t)$ denote the signal transmitted from the n^{th} transmit antenna $(n = 1, ..., N)$ and let $r_m(t)$ denote the signal received at the m^{th} receive antenna

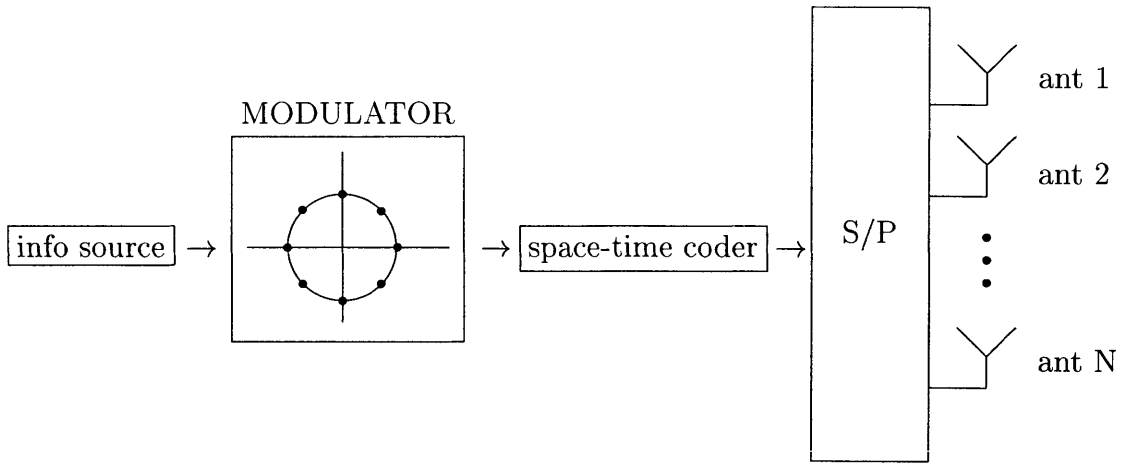

Figure 3-1: Space-time encoder block diagram

$(m = 1, ..., M)$.

$$r_m(t) = \sum_{n=1}^{N} \alpha_{n,m} s_n(t) + \eta_m(t) \tag{3.1}$$

$\alpha_{n,m}$ is the path gain associated with the channel between the n^{th} transmit antenna and m^{th} receive antenna. They are independent complex Gaussian RVs with variance $\frac{1}{2N}$ per real dimension.

$\eta_m(t)$ is the zero-mean complex Gaussian noise associated with the m^{th} receive antenna. The variance of this noise is $\frac{1}{2SNR}$ per real dimension.

This ensures that $r_m(t)$ will have a signal-to-noise ratio of SNR. To see this, note that the complex variance of $\alpha_{n,m}$ and $\eta_m(t)$ are $\frac{1}{N}$ and $\frac{1}{SNR}$, respectively and $s_n(t)$ is assumed to have unit energy. Hence, since the $\alpha_{n,m}$'s are independent, the power in the signal part of $r_m(t)$ is $\sum_{n=1}^{N} \frac{1}{N} = 1$ and the power in the noise part of $r_m(t)$ is $\frac{1}{SNR}$.

3.1 An capacity argument for using space-time codes

The capacity and performance of systems with one transmit antenna and multiple receive antennas or multiple transmit antennas and one receive antenna have been well established. However, as shown through the independent efforts of Telatar [11],

# tx antennas	# rx antennas	Capacity
1	M	$C_{1,M} \sim \log M$
N	1	$C_{N,1} \sim \log(1 + SNR)$
N	N	$C_{N,N} \sim N$
N	M	$C_{N,M} \sim \min(N, M)$

Table 3.1: Capacity limits

and Foschini and Gans [12], the capacity of a MIMO system far exceeds that of single antenna systems. More importantly, if the number of transmit antennas N does not exceed the number of receive antennas M, then the capacity of such a system is *at least linearly proportional* to N [7].

Generally, when comparing the performance of systems using multiple transmit and/or receive antennas, it is more appropriate to use outage capacity instead of Shannon capacity, which is a limit derived for additive white Gaussian noise channels.

Telatar derived the outage capacity under the assumption that fading was constant over each channel use, but changed independently between successive channel uses. Foschini and Gans derived the outage capacity under the assumption of block fading. Both outage capacities and error exponents were derived for multiple transmit antenna systems with Gaussian channel noise [7]. Note that though the outage capacities in the limit that $N \to \infty$ for both first and third system configurations in Table 3.1 tend to infinity, the capacity of the MIMO system grows significantly faster. In contrast, the capacity of the second configuration saturates at $\log_2(1 + SNR)$ and becomes independent of N.

Intuitively, there should exist some threshold value of N above which adding transmit antennas will not increase the capacity of the system. This question was addressed in [7], where it was shown that there is little to gain terms of outage capacity when $N > 4$. The reader is encouraged to read [7] for a more comprehensive proof of this claim. Furthermore, it is evident from the standard formula for Shannon capacity that a 3 dB increase in SNR only increments the transmission rate by one bit/cycle; however, for MIMO systems, a similar increase in SNR increments the

transmission rate by N bits/cycle. From these three arguments, the merits of using multiple antennas at both transmitter and receiver become apparent.

Outage capacity outline an outage region, which exhibits the following three properties [8]:

1. The edge of the region is a surface of constant mutual information [1] corresponding to a target transmission rate.

2. Within the region, reliable communication is not possible at the desired transmission rate.

3. Outside the region, reliable communication is possible at or above the desired transmission rate.

Figure 3-2 illustrates the outage regions for various vector-coding and scalar-coding schemes [8]. The lines outline surfaces of constant mutual information for various (α_1, α_2) pairs, where α_1 (abscissa) and α_2 (ordinate) are independent fade coefficients associated with two parallel channels.

Vector-coded antenna systems use vector-valued codebooks, where the i^{th} element in the N-symbol vector codeword specifies what is to be transmitted over the i^{th} transmit antennas. Space-time codes are examples of vector-codes. Scalar-coded codebooks use conventional codes in conjunction with linear preprocessing at the input to the antenna array on the transmitter side. The preprocessing converts a multiple-input/single-output system into a single-input/single-output system. Examples of preprocessing techniques include time and frequency multiplexing, time and frequency shifting, and randomized time-weighting as described in [8].

It was shown by Narula *et al.* that using vector-coding is optimal in the sense that they approach the performance limits of a multiple input/single output channel. In fact, if I_x represents the average mutual information achieved by each transmit

[1] Mutual information $I(X;Y) = H(X) - H(X \mid Y)$ [9]. It is a measure of achievable rates for reliable communications [8].

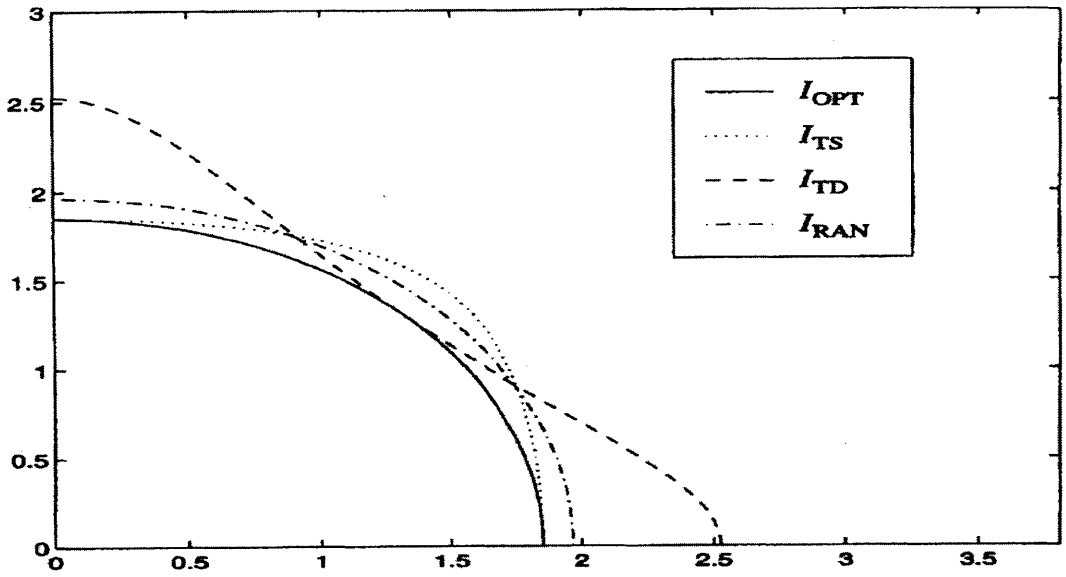

Figure 3-2: The outage region for a target transmission rate of 1 nat per symbol

antenna for coding scheme x where

$$x = \begin{cases} OPT & : \text{ vector-coding} \\ TD/FD & : \text{ time/frequency division multiplexing} \\ TS/FS & : \text{ time/frequency shifting} \\ RAN & : \text{ randomized time weighting} \end{cases}$$

it was proven in [8] and further illustrated in Figure 3-2 that in general,

$$\begin{aligned} I_{TD/FD} &\leq I_{OPT} \\ I_{TS/FS} &\leq I_{OPT} \\ I_{RAN} &\leq I_{OPT} \end{aligned} \tag{3.2}$$

Furthermore, let C_{RAY} represent the capacity of a single antenna system with Rayleigh fading (i.e no spatial diversity) and C_{AWGN} represent the capacity of a AWGN channel at a specific average SNR level. It has been shown that the capacity of vector-coded systems under conditions of infinite spatial diversity and fixed tempo-

ral diversity approaches C_{AWGN}, and that the capacity of scalar-coded systems under the same conditions approaches C_{RAY} [8]. These findings further support the claim made earlier that AWGN channels upperbound the performance of MIMO systems.

Typically,vector-coding require computationally intensive and inefficient decoder algorithms. Hence, there has been considerable interest in scalar-coding, which is sub-optimal to vector-coding but provide low-complexity decoding algorithms [8]. However, the recent works of Foschini and Gans as well as Tarokh, Seshadri and Calderbank have given birth to vector-codes which do not suffer from complex encoding and decoding strategies.

Space-time codes are the vector-codes developed by Tarokh *et al.* that are appropriate to MIMO systems and provide an essential stepping stone for approaching its capacity limits. In fact, it was shown that using space-time diversity techniques could potentially achieve a significant fraction of the calculated outage capacity limits for the MIMO system [7]. There are two flavors of space-time codes: block codes and trellis codes. This thesis focuses on introducing a construction for the former.

3.2 An improved performance argument for using space-time codes

Similar to coding and shaping gains associated with Euclidean space codes, performance improvements derived from space-time codes are quantified by the coding and diversity gains they provide.

Diversity gain or diversity order DG is used to quantify how many replicas of the original signal were successfully extracted from the received signal. Note, however, that it is not a benefit derived merely from using multiple transmit and receive antennas.

As shown by Alamouti [5], it is very easy to exploit full receive diversity using maximal ratio receive combining (MRRC) techniques; however, extracting diversity advantage from using multiple transmit antennas is not so trivial a task.

For an encoding scheme that achieves transmit diversity order $R \leq N$ in a N transmit, M receive antenna system, the overall diversity order is $R * M$. The maximum diversity order DG_{max} for such a system is $N * M$.

Space-time codes are examples of codes that can extract transmit diversity order when used in conjunction with multiple transmit antennas. The importance of achieving high diversity order is best shown through the generalized expression for probability of error, P_e. Specifically, it has been shown that for Rayleigh fading channels,

$$P_e \sim \frac{1}{SNR^{DG}} \tag{3.3}$$

Hence, increases in DG exponentially improve system performance.

3.3 Space-time block codes

Channel codes are typically designed to encode the original information symbols in such a way as to improve their robustness against channel noise and distortions. The codewords produced by (n, k) *binary* channel codes for Gaussian channels are 2^k n-tuples that can be geometrically represented as vectors or discrete code points in the subspace of the vector space S, which is the space consisting of all 2^n of the possible binary n-tuples. The AWGN is represented as a spherical noise cloud centered about each code point, where the radius of the noise sphere is proportional to the variance of the noise [6].

The optimal maximum likelihood decoder for a Gaussian channel simplifies to a minimum distance rule, which essentially takes the received n-tuple and decodes it into the codeword to which it is the closest. Consequently, to minimize the probability that a received n-tuple will fall into the noise sphere of codeword $\mathbf{e}$ and hence be decoded as $\mathbf{e}$ if codeword $\mathbf{c}$ is the actual transmitted codeword, it is desirable to have these code points spaced as far apart as possible. To this end, minimum Hamming distance is the parameter to be maximized for binary codes [6].

For space-time codes, the design criteria are slightly more complex; however, as will soon become evident, the increased complexity in design criteria does not affect the complexity of the encoding and decoding algorithms. In fact, employing simple encoding and decoding schemes are two of the design criteria for space-time codes. The other design criteria are related to maximizing the diversity and coding gains of the code. However, before stating them, it is important to introduce the notation for space-time codes that will be used throughout the rest of this thesis.

3.3.1 Notation and Terminology

Space time codes are symbolically represented by a carefully designed transmission matrix, H, and can be optimally decoded by a maximum likelihood decoder based on linear processing at the receiver. As seen from Figure 3-1, the output of the coder is not a matrix, but is instead merely a serial stream of symbols that are fed into a serial-to-parallel converter. The matrix notation introduced here is simply a mathematical construction used to facilitate the representation and manipulation of the code. Figure 3-3 illustrates this point.

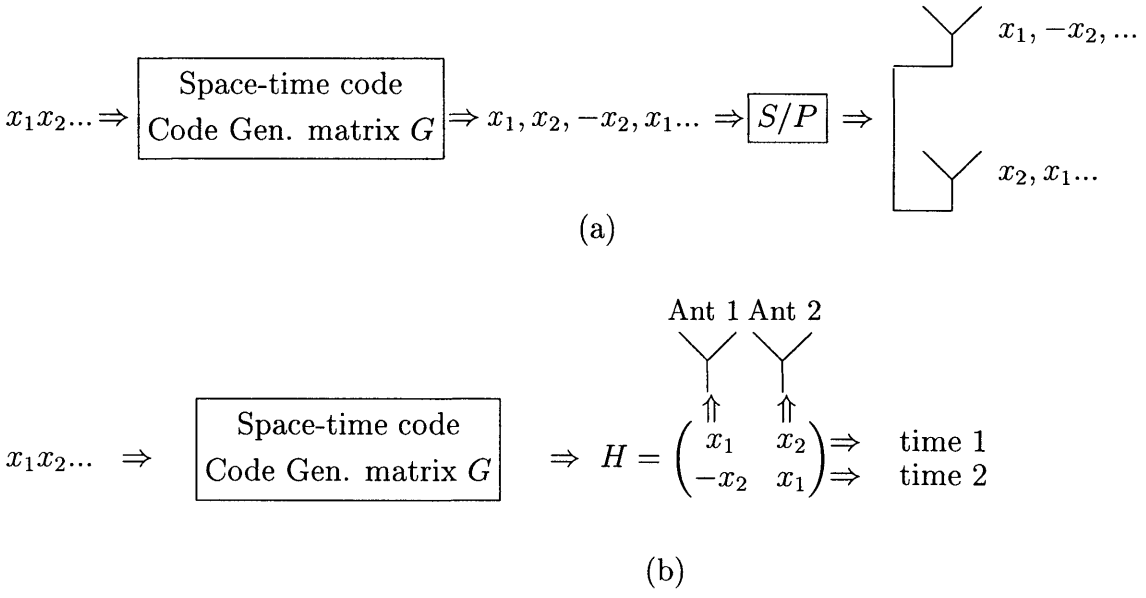

Figure 3-3: (a) Actual encoder procedures (b) Matrix representation of encoder procedures

In general, any space-time codeword can be written as a PxN codeword or trans-

mission matrix H such as the following:

$$\mathbf{H} = \begin{pmatrix} c_{1,1} & c_{1,2} & \dots & \dots & c_{1,N} \\ c_{2,1} & c_{2,2} & \dots & \dots & c_{2,N} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ c_{P,1} & c_{P,2} & \dots & \dots & c_{P,N} \end{pmatrix}$$

If $c_{i,j}$[2] is the code symbol to be sent *during* the i^{th} time slot *over* the j^{th} transmit antenna ($j = 1, ..., N$), then over each time slot $t = 1, 2, ..., P$, a single frame of symbols $c_{t,1}, c_{t,2}, ..., c_{t,N}$ will be sent simultaneously from each of the N transmit antennas.

To avoid confusion, the following naming conventions have been established:

Codeword : This term can be used interchangeably with 'space-time[3] codeword', 'transmission matrix' or 'space-time codeword matrix', such as H above.

Code Generator Matrix, G: The PxN transformation matrix applied to the input vector to the space-time coder to generate the codeword.

Symbol frame: The set of space-time code symbols sent simultaneously over the N transmit antennas during *one* time slot: $c_{t,1}, c_{t,2}, ..., c_{t,N}$. Note that this is merely one row within H.

Space-time code error matrix : (STC error matrix, for short) The difference between the transmission matrices $H(\mathbf{c})$ and $H(\mathbf{e})$ needed to send the codewords $\mathbf{c}$ and $\mathbf{e}$, respectively.

$$\mathbf{B}(\mathbf{c}, \mathbf{e}) = H(\mathbf{c}) - H(\mathbf{e}) = \begin{pmatrix} c_{1,1} - e_{1,1} & c_{1,2} - e_{1,2} & \dots & \dots & c_{1,N} - e_{1,N} \\ c_{2,1} - e_{2,1} & c_{2,2} - e_{2,2} & \dots & \dots & c_{2,N} - e_{2,N} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ c_{P,1} - e_{P,1} & c_{P,2} - e_{P,2} & \dots & \dots & c_{P,N} - e_{P,N} \end{pmatrix}$$

[2] $c_{i,j}$ are not the original information symbols. They are symbols selected from a baseband signal constellations with 2^b elements.

[3] Henceforth, all codewords are assumed to be space-time codewords, unless otherwise specified.

Note that if the input to the space-time coder is

$$\mathbf{x} = x_1, x_2, ..., x_s, x_{s+1}, ..., x_{s+s}, ..., x_{Ks+1}, ..., x_{Ks+s},$$

then define **transmission block** as the $(N * P)$-symbol long codeword associated with any s-symbol[4] input sequence. For instance, for input sequence, $\mathbf{x}$, there will be K such transmission blocks: one for encoding $x_1, ..., x_s$, one for encoding $x_{s+1}, ..., x_{s+s}$ and so forth.

3.3.2 Design Criteria for Space-time Codes

Having established the mathematical notation for space-time codes, the design criteria related to diversity and code gains for space-time codes over Rayleigh fading channels, as first introduced by Tarokh, *et al.* [7], can now be stated.

Given two codewords

$$\mathbf{c} = c_{1,1}, c_{1,2}, ..., c_{1,N}, c_{2,1}, c_{2,2}, ..., c_{2,N}, ..., c_{P,1}, c_{P,2}, ..., c_{P,N}, \tag{3.4}$$

and

$$\mathbf{e} = e_{1,1}, e_{1,2}, ..., e_{1,N}, e_{2,1}, e_{2,2}, ..., e_{2,N}, ..., e_{P,1}, e_{P,2}, ..., e_{P,N}, \tag{3.5}$$

the transmit diversity order of this code is the *minimum* rank R ($\leq N$) of its STC error matrix $B(\mathbf{c}, \mathbf{e})$ when taken over all possible $(\mathbf{c}, \mathbf{e})$ codeword pairs. The overall diversity order DG is hence $R*M$ when M receive antennas is assumed. Consequently, $B(\mathbf{c}, \mathbf{e})$ must be full rank ($R = N$) to achieve full or maximum diversity order $DG_{max} = N * M$.

The coding gain CG is measured with respect to an uncoded system operating with the same diversity gain and is related to maximizing the determinant of the matrix $A(\mathbf{c}, \mathbf{e}) = B(\mathbf{c}, \mathbf{e})B^*(\mathbf{c}, \mathbf{e})$, where $B^*(\mathbf{c}, \mathbf{e})$ is the Hermitian (conjugate transpose) of

[4]In many cases $s = N$: number of transmit antennas

$B(\mathbf{c}, \mathbf{e})$.

As found in [7], if you define $P(\mathbf{c} \rightarrow \mathbf{e})$ as the probability of transmitting codeword $\mathbf{c}$ and deciding in favor of codeword $\mathbf{e}$, then

$$P(\mathbf{c} \rightarrow \mathbf{e}) \leq \frac{1}{(CG * SNR)^{DG}} \tag{3.6}$$

Evidently, error performance is significantly improved through high diversity and coding gains. Interestingly, space-time codes based on orthogonal designs are examples of codes that achieve maximum diversity order for a given number of transmit and receive antennas. Readers interested in learning more about these codes are encouraged to read [16].

3.4 The Encoding Scheme for Space-Time Codes

The following lists the ingredients needed for the construction of a space-time codeword:

1. Baseband transmission signal constellation A with 2^b elements.
2. N transmit antennas, M receive antennas.
3. PxN transmission matrix H.
4. P time slots per transmission block; i.e. H has P rows.
5. k information symbols input into the space-time coder.

The last two criteria imply a code rate $R = k/P$.

During any single time slot, $k * b$ bits are sent through k information symbols $s_1, ..., s_k$ chosen from signal constellation A. Given generator matrix G, whose elements are typically indeterminates $x_1, ..., x_k$, set $x_i = s_i$ to produce the transmission matrix H. For example, suppose $G = \begin{pmatrix} x_1 & x_2 \\ -x_2^* & x_1^* \end{pmatrix}$, which implies $k = P = N = 2$. The transmission matrix would be $H = \begin{pmatrix} s_1 & s_2 \\ -s_2^* & s_1^* \end{pmatrix}$, where s_1 and s_2 are the constellation symbols sent into the space-time coder.

As shown in Figure 3-3(b), each row in H is sent simultaneously from each of the N transmit antennas, during the appropriate time slot.

3.5 The Decoding algorithm for Space-Time Codes

For purposes of illustration, the decoding algorithm will be derived for the simple example given in the previous section. From detection theory, we know that the optimal receiver for a channel with Gaussian noise is a maximum likelihood detector, which reduces to minimizing the following minimum distance-type decision metric over all possible values of s_1 and s_2:

$$\sum_{m=1}^{M}(|r_{1,m}-\alpha_{1,m}s_1-\alpha_{2,m}s_2|^2+|r_{2,m}-(-\alpha_{1,m}s_2^*)-\alpha_{2,m}s_1^*|^2) \tag{3.7}$$

Expanding and removing terms that are common to all metrics regardless of choice of s_1 and s_2, the previous expression reduces to

$$\left.\begin{aligned}-\sum_{m=1}^{M}[r_{1,m}\alpha_{1,m}^*s_1^*+r_{1,m}^*\alpha_{1,m}s_1+r_{2,m}\alpha_{2,m}^*s_1+r_{2,m}^*\alpha_{2,m}s_1^*]+|s_1|^2\sum_{m=1}^{M}\sum_{n=1}^{2}|\alpha_{n,m}|^2+\\ -\sum_{m=1}^{M}[r_{1,m}\alpha_{2,m}^*s_2^*+r_{1,m}^*\alpha_{2,m}s_2-r_{2,m}\alpha_{1,m}^*s_2-r_{2,m}^*\alpha_{1,m}s_2^*]+|s_2|^2\sum_{m=1}^{M}\sum_{n=1}^{2}|\alpha_{n,m}|^2\end{aligned}\right\}\begin{aligned}(a)\\(b)\end{aligned}$$

Notice that all terms in (a) depend solely on s_1 and all terms in (b) depend solely on s_2; therefore, you can minimize the entire metric by minimizing each term independently over all possible s_1 and s_2.

This metric is equivalent to using

$$\left|[\sum_{m=1}^{M}(r_{1,m}\alpha_{1,m}^*+r_{2,m}^*\alpha_{2,m})]-s_1\right|^2+|s_1|^2\sum_{m=1}^{M}\sum_{n=1}^{2}|\alpha_{n,m}|^2 \tag{3.8}$$

to decode s_1 and using

$$\left|[\sum_{m=1}^{M}(r_{1,m}\alpha_{2,m}^{*} - r_{2,m}^{*}\alpha_{1,m})] - s_2\right|^2 + |s_2|^2 \sum_{m=1}^{M}\sum_{n=1}^{2}|\alpha_{n,m}|^2 \tag{3.9}$$

to decode s_2.

If the baseband signal constellation has equal energy, such as M-PSK, then the second terms in equations 3.8 and 3.9 drop out and the metrics for decoding s_1 and s_2 reduce further to

$$\left|[\sum_{m=1}^{M}(r_{1,m}\alpha_{1,m}^{*} + r_{2,m}^{*}\alpha_{2,m})] - s_1\right|^2 \tag{3.10}$$

and

$$\left|[\sum_{m=1}^{M}(r_{1,m}\alpha_{2,m}^{*} - r_{2,m}^{*}\alpha_{1,m})] - s_2\right|^2, \tag{3.11}$$

respectively.

As shown above, the decoder reduces to simple linear processing of the received signals $r_{1,m}$ and $r_{2,m}$ and their complex conjugates.

3.6 Summary

Through both information theoretic and performance improvement arguments, space-time codes show great promise in aiding efforts of providing reliable communication capabilities over wireless mediums. Many of the fundamental concepts and terminology associated with space-time codes were introduced in this chapter. Furthermore, a brief glance into their encoding and decoding algorithms was provided.

Chapter 4

Concatenating Euclidean Space Codes with Space-Time Codes

4.1 The Transmission Model

Similar to before, the channel is modelled as exhibiting block fading. The transmitter and receivers are each equipped with N and M antennas, respectively. Without loss of generality, choose the expected signal power at the receiver to be 1. To meet this constraint, choose the variance of the complex Gaussian RV used to capture the effects of Rayleigh fading and uniform phase distortions to have a variance of 1 per complex dimension or equivalently, 0.5 per real dimension. Furthermore, normalize the average energy of the signal constellation to be $\frac{1}{N}$[1].

The signal at receive antenna $m|_{m=1,\ldots,M}$ at time t is

$$r_m(t) = \sum_{n=1}^{N} \alpha_{n,m} d_n(t) + \eta_m(t), \tag{4.1}$$

where $\{d_n(t)|_{n=1,\ldots,N}\}$ is a frame of constellation symbols to be sent simultaneously from the N transmit antennas.

[1]NOTE: Normalizing all signals to have equal energy implies an assumption that all signals from the transmit side experience the same large-scale path loss; therefore, for simplicity, large-scale fading is ignored.

To obtain a specific SNR level, choose the variance of the independent, zero-mean complex Gaussian noise samples, $\eta_m(t)$ to have a variance of $\frac{1}{SNR}$ per complex dimension, or equivalently $\frac{1}{2SNR}$ per real dimension.

4.2 The Construction

Before introducing the construction of the space-time code, let us establish the notational conventions for matrices and vectors that are to be used, as follows:

A^* : Hermitian (conjugate, transpose) of A.

A^T : transpose of A.

$A_{i,.}$: i^{th} row of A.

$A_{.,j}$: j^{th} column of A.

$A_{i,j}$: (i^{th}, j^{th}) element of A.

Given N transmit antennas, choose the code generation matrix G to be a unitary PxN complex matrix, such that $G^*G = I$. Let C_{ES} denote the constituent, rate R Euclidean-space (linear block[2]) code, where the codeword length for this code is L. Code C_{ES} will consist of 2^{R*L} codewords of length L, where each segment of b code bits will select a symbol from a signal constellation with 2^b elements. Figure 4-1 illustrates the encoding process. Note that the construction is essentially a concatenated code, where the information symbols are first encoded by a Euclidean-space channel code and then encoded further by the space-time code.

4.2.1 Interface between the Euclidean-space coder and the space-time coder

In Figure 4-1, $\mathbf{c}$ is the stream of codewords chosen from codebook C_{ES} that are first modulated and then fed into the space-time coder. Hence, a L-bit codeword

[2]A review of relevant linear block code theory is presented in section 5.2.1

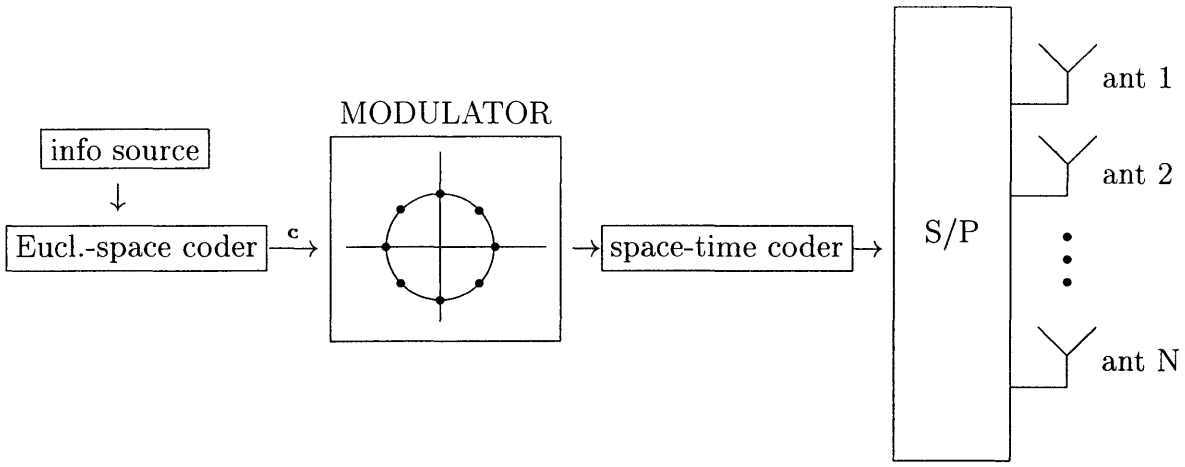

Figure 4-1: Block diagram for a space-time coder based on a Euclidean space code

will result in $J = \frac{L}{b}$ modulation symbols if a modulation with 2^b elements is used. For example, if the modulation is binary phase-shift keying (BPSK, $b = 1$), then $J = L$ and if the modulation is quadrature phase-shift keying (QPSK, $b = 2$), then $J = \frac{L}{2}$. If the number of modulation symbol per codeword, J is greater than the number of transmit antennas N (i.e. $J = x * N$), then each block of J modulation symbols must be further fragmented into segments containing N modulation symbols. Consequently, x transmission blocks will be needed to send a single codeword from code C_{ES} in its entirety. For instance, suppose the codeword $\mathbf{c}$ is of the following form:

$$\mathbf{c} = c_1, c_2, ..., c_L$$

Since $L = b*J = b*x*N$, $\mathbf{c}$ can be rewritten as the following stream of modulation symbols $\mathbf{d}$:

$$\mathbf{d} = d_{1,1}, d_{1,2}, ..., d_{1,N}, d_{2,1}, ..., d_{2,N}, ..., d_{x,1}, ..., d_{x,N}$$

where $d_{i,m}$ is the modulation symbol representing b codeword bits that is transmitted over the $i^{th}|_{i=1,...,x}$ transmission block and $m^{th}|_{m=1,...,M}$ transmit antenna. This being the case, let block fading imply that the path gains are constant over all x transmission blocks.

4.2.2 Choice of Code Generator Matrix G

Thus far, the only requirement on G for our construction is that it be a PxN complex unitary matrix. One implication of this requirement is that $G^*G = I$, which implies that the columns of G are mutually orthonormal complex P-dimensional complex vectors. This property of G will later prove to be key in the simplification the decoding process.

Unlike the code generator matrices introduced earlier, the elements of G now take on numerical values and are no longer indeterminants. Using the matrix notation previously introduced , a codeword is generated by creating a NxN diagonal matrix from the modulation symbols of $\mathbf{d}$, and then right-multiplying G by this diagonal matrix. The effect of this matrix multiplication is to modulate each column of G by one modulation symbol. Specifically, during the i^{th} transmission block, the n^{th} column of G will be multiplied by $d_{i,n}|_{n=1,...,N}$. An example is illustrated in Figure 4-2.

$$
\begin{array}{c}
d_{i,1}, d_{i,2}, d_{i,3}, d_{i,4} \\
\downarrow \\
\begin{pmatrix} d_{i,1} & 0 & 0 & 0 \\ 0 & d_{i,2} & 0 & 0 \\ 0 & 0 & d_{i,3} & 0 \\ 0 & 0 & 0 & d_{i,4} \end{pmatrix} \\
\downarrow \\
\frac{1}{\sqrt{4}} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} d_{i,1} & 0 & 0 & 0 \\ 0 & d_{i,2} & 0 & 0 \\ 0 & 0 & d_{i,3} & 0 \\ 0 & 0 & 0 & d_{i,4} \end{pmatrix} \rightarrow \frac{1}{\sqrt{4}} \begin{pmatrix} d_{i,1} & d_{i,2} & d_{i,3} & d_{i,4} \\ d_{i,1} & -d_{i,2} & d_{i,3} & -d_{i,4} \\ d_{i,1} & d_{i,2} & -d_{i,3} & -d_{i,4} \\ d_{i,1} & -d_{i,2} & -d_{i,3} & d_{i,4} \end{pmatrix}
\end{array}
$$

Figure 4-2: Encoding using matrix multiplication

Recall that $d_{i,j}$ is an element from a signal constellation with 2^b points; therefore, if $J \geq N$, then $i = 1, ..., \frac{L}{b*N}$.

This construction does not place any further restrictions on the choice of G; however, if N is a power of 2, then appropriately scaled Hadamard or Fourier matrices can be used for G and consequently, the encoding and decoding algorithms can be made

very computationally efficient through use of fast Hadamard or Fourier transform techniques.

4.2.3 Encoding

Although previous sections provided previews into the encoding process for the our construction, it will be formally stated here.

Using $R * L$ bits of data, one of the 2^{R*L} codewords in code C_{ES} is chosen. As before, let this codeword be rewritten as

$$\mathbf{d} = d_{1,1}, d_{1,2}, ..., d_{1,N}, d_{2,1}, ..., d_{2,N}, ..., d_{x,1}, ..., d_{x,N},$$

where $d_{i,j}$ are modulation symbols representing b channel code bits.

If x transmission blocks are needed to send one codeword from C_{ES}, then let

$$\{H(i)|_{i=1,...,x} = G(d_{i,1}, d_{i,2}, ..., d_{i,N})\} \tag{4.2}$$

denote a set of space-time codeword matrices where the $n^{th}|_{n=1,...,N}$ column in each individual transmission matrix is $d_{i,n}G_{.,n}$. Each matrix is computed every P time slots; i.e. $H(i)$ is computed at time $t = (i-1)P + 1$.

At time $t = (i-1)P + p$, matrix element $H_{p,n}(i)|_{i=1,...,x}$ is sent over the $p^{th}|_{p=1,...,P}$ time slot and $n^{th}|_{n=1,...,N}$ transmit antenna.

The rate of transmission is $\frac{R*L}{P*x}\frac{\text{number of input bits}}{\text{number of channel uses needed}} = \frac{R*N*b}{P}$ bits per channel use or equivalently $\frac{R*N}{P}$ symbols per channel use.

Figure 4-3 illustrates the encoding process with the simple 2-antenna case, using BPSK as the modulation scheme.

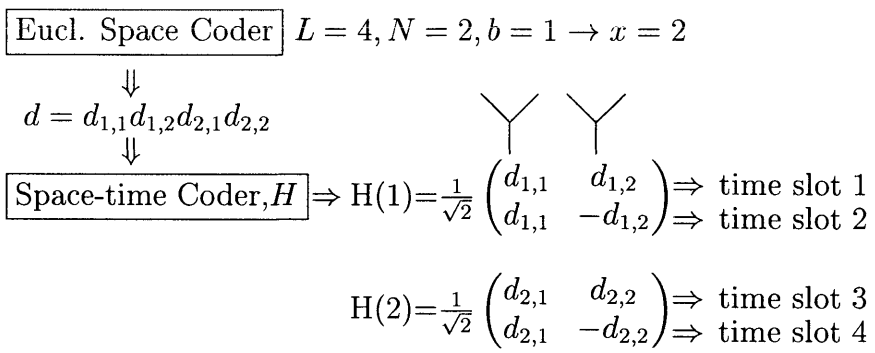

Figure 4-3: Encoding example using a unitary 2x2 Hadamard $G = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$

4.2.4 Decoding

The signal received at the $m^{th}|_{m=1,\ldots,M}$ receive antenna is of the form

$$r_{t,m} = \sum_{n=1}^{N} \alpha_{n,m} d_{t,n} + \eta_{t,m}, \tag{4.3}$$

Assuming perfect channel estimation; i.e. the fading coefficients $\alpha_{n,m}|_{n=1,\ldots,N}^{m=1,\ldots,M}$ can be recovered without error, then phase distortions can be eliminated[3], and coherent detection can be used at the receivers. Hence, the maximum likelihood detector under the Gaussian noise condition reduces to finding the sequence of modulation symbols $\mathbf{d}$ that minimizes the following decision metric:

$$\sum_{i=1}^{x} \sum_{p=1}^{P} \sum_{m=1}^{M} |r_m\big((i-1)P + p\big) - \sum_{n=1}^{N} \alpha_{n,m} G_{p,n} d_{i,n}|^2 \tag{4.4}$$

when taken over all possible codewords

$$\mathbf{d} = d_{1,1}, d_{1,2}, \ldots, d_{1,N}, d_{2,1}, \ldots, d_{2,N}, \ldots\ldots, d_{x,1}, \ldots, d_{x,N}.$$

Define $\mathbf{R}_{i,m}$ to be a vector of the P signals received during the $i^{th}|_{i=1,\ldots,x}$ trans-

[3]Even if $\alpha_{n,m}$ cannot be perfectly recovered, they generally vary slow enough for a phase-locked loop to track them.

mission block at the $m^{th}|_{m=1,...,M}$ receive antenna. Specifically,

$$\mathbf{R}_{i,m} = \big[r_m\big((i-1)P+1\big), r_m\big((i-1)P+2\big), ..., r_m\big((i-1)P+P\big)\big]$$

Equation 4.4 simplifies to

$$\sum_{i=1}^{x}\sum_{m=1}^{M}|\mathbf{R}_{i,m} - \sum_{n=1}^{N}\alpha_{n,m}G_{.,n}d_{i,n}|^2 \tag{4.5}$$

$$= \sum_{i=1}^{x}\sum_{m=1}^{M}(\mathbf{R}_{i,m} - \sum_{n=1}^{N}\alpha_{n,m}G_{.,n}d_{i,n})(\mathbf{R}^*_{i,m} - \sum_{n=1}^{N}\alpha^*_{n,m}G^*_{.,n}d^*_{i,n}) \tag{4.6}$$

$$= \sum_{i=1}^{x}\sum_{m=1}^{M}(\underbrace{|\mathbf{R}_{i,m}|^2}_{(a)} - \sum_{n=1}^{N}\alpha_{n,m}G_{.,n}\mathbf{R}^*_{i,m}d_{i,n} - \sum_{n=1}^{N}\alpha^*_{n,m}G^*_{.,n}\mathbf{R}_{i,m}d^*_{i,n} + \sum_{n=1}^{N}|\alpha_{n,m}|^2|d_{i,n}|^2\underbrace{|G_{.,n}|^2}_{(b)}) \tag{4.7}$$

(a) is independent of the codewords and can be eliminated. (b) equals 1 because of the property that $G^*G = I$.

Introducing the notation $A^i_{n,m} = G^*_{.,n}\mathbf{R}_{i,m}$, this simplifies to

$$\sum_{i=1}^{x}\sum_{n=1}^{N}(-\sum_{m=1}^{M}A^{i*}_{n,m}\alpha_{n,m}d_{i,n} - \sum_{m=1}^{M}A^i_{n,m}\alpha^*_{n,m}d^*_{i,n} + \sum_{m=1}^{M}|\alpha_{n,m}|^2|d_{i,n}|^2), \tag{4.8}$$

which is equivalent to the following metric:

$$\sum_{i=1}^{x}\sum_{n=1}^{N}\Big(|(\sum_{m=1}^{M}A^i_{n,m}\alpha^*_{n,m}) - d_{i,n}|^2 + |d_{i,n}|^2(\sum_{m=1}^{M}|\alpha_{n,m}|^2 - 1)\Big) \tag{4.9}$$

To see this, expand equation (4.9) as follows:

$$\sum_{i=1}^{x}\sum_{n=1}^{N}\Big(\big((\sum_{m=1}^{M}A^{i}_{n,m}\alpha^{*}_{n,m})-d_{i,n}\big)\big((\sum_{m=1}^{M}A^{i*}_{n,m}\alpha_{n,m})-d^{*}_{i,n}\big)+(\sum_{m=1}^{M}|d_{i,n}|^2|\alpha_{n,m}|^2-|d_{i,n}|^2)\Big)$$

$$=\sum_{i=1}^{x}\sum_{n=1}^{N}\Big(\underbrace{(\sum_{m=1}^{M}A^{i}_{n,m}\alpha^{*}_{n,m})(\sum_{m=1}^{M}A^{i*}_{n,m}\alpha_{n,m})}_{\text{independent of the codeword}}-(\sum_{m=1}^{M}A^{i*}_{n,m}\alpha_{n,m}d_{i,n})$$
$$-(\sum_{m=1}^{M}A^{i}_{n,m}\alpha^{*}_{n,m}d^{*}_{i,n})+|d_{i,n}|^2+\sum_{m=1}^{M}|d_{i,n}|^2|\alpha_{n,m}|^2-|d_{i,n}|^2\Big)$$

Equation (4.9) can be further simplified if all elements of the signal constellation have equal energy; i.e. $|d_{i,n}|^2$ equal $\forall\ {}^{i=1,\ldots,x}_{n=1,\ldots,N}$. Upon invoking this additional condition on the choice of constellation, the metric becomes

$$\sum_{i=1}^{x}\sum_{n=1}^{N}|(\sum_{m=1}^{M}A^{i}_{n,m}\alpha^{*}_{n,m})-d_{i,n}|^2 \tag{4.10}$$

In summary, the decoding process involves first computing $A^{i}_{n,m}|_{i=1,\ldots,x}\forall\ {}^{n=1,\ldots,N}_{m=1,\ldots,M}$, and subsequently, $\sum_{m=1}^{M}A^{i}_{n,m}\alpha^{*}_{n,m}$ and sending these sufficient statistics to the *Euclidean-space* decoder for C_{ES}. This Euclidean-space decoder turns out to be a minimum-distance decoder such as the Viterbi algorithm with branch metrics of the form

$$\Big(|(\sum_{m=1}^{M}A^{i}_{n,m}\alpha^{*}_{n,m})-d_{i,n}|^2+|d_{i,n}|^2(\sum_{m=1}^{M}|\alpha_{n,m}|^2-1)\Big).$$

Consequently, this construction introduces a simple method for turning any Euclidean-space channel code into a space-time code without creating additional complexities in the decoder. In fact, the space-time code inherits the same decoding technique as the constituent Euclidean-space channel code. This is quite a remarkable property, which makes implementation of this construction very appealing.

Figure 4-4 illustrates the steps in the decoding process for the simplified case of a

single receive antenna ($M = 1$).

$$\mathbf{R}_{i,1} = \begin{pmatrix} r_{(i-1)P+1} \\ r_{(i-1)P+2} \\ \vdots \\ r_{(i-1)P+P} \end{pmatrix} \Rightarrow \boxed{G^*} \Rightarrow \begin{pmatrix} \alpha^*_{1,1} \\ \alpha^*_{2,1} \\ \vdots \\ \alpha^*_{N,1} \end{pmatrix} \Rightarrow \mathbf{R}^{modified}_{i,1}$$

linear processing at receiver

co-phases outputs for coherent detection

Find $\mathbf{R}^{modified}_{i,1}$ for $i = 1, ..., x$ and send the collective set of metric values to the Euclidean-space decoder of the constituent Euclidean-space block code.

$\mathbf{R}_{i,1}$ is the vector of received signals over the P time slots of the i^{th} transmission block

Figure 4-4: Block Diagram for the Decoder of the Space-time Code

Figure 4-5 illustrates why the property that G is a unitary matrix simplifies the decoding process.

$$\begin{matrix} \begin{pmatrix} \alpha & \beta & \gamma & \theta \end{pmatrix} \\ \frac{1}{\sqrt{4}} \begin{pmatrix} d_{1,1} & d_{1,2} & d_{1,3} & d_{1,4} \\ d_{1,1} & -d_{1,2} & d_{1,3} & -d_{1,4} \\ d_{1,1} & d_{1,2} & -d_{1,3} & -d_{1,4} \\ d_{1,1} & -d_{1,2} & -d_{1,3} & d_{1,4} \end{pmatrix} \\ \Downarrow \end{matrix}$$

$$\underbrace{\begin{pmatrix} r_1 = \frac{1}{\sqrt{4}}(\alpha d_{1,1} + \beta d_{1,2} + \gamma d_{1,3} + \theta d_{1,4}) + n_1 \\ r_2 = \frac{1}{\sqrt{4}}(\alpha d_{1,1} - \beta d_{1,2} + \gamma d_{1,3} - \theta d_{1,4}) + n_2 \\ r_3 = \frac{1}{\sqrt{4}}(\alpha d_{1,1} + \beta d_{1,2} - \gamma d_{1,3} - \theta d_{1,4}) + n_3 \\ r_4 = \frac{1}{\sqrt{4}}(\alpha d_{1,1} - \beta d_{1,2} - \gamma d_{1,3} + \theta d_{1,4}) + n_4 \end{pmatrix}}_{\mathbf{R}_{1,1}} \Rightarrow \underbrace{\begin{pmatrix} |\alpha|^2 d_{1,1} + \frac{1}{\sqrt{4}}\alpha^*(n_1 + n_2 + n_3 + n_4) \\ |\beta|^2 d_{1,2} + \frac{1}{\sqrt{4}}\beta^*(n_1 - n_2 + n_3 - n_4) \\ |\gamma|^2 d_{1,3} + \frac{1}{\sqrt{4}}\gamma^*(n_1 + n_2 - n_3 - n_4) \\ |\theta|^2 d_{1,4} + \frac{1}{\sqrt{4}}\theta^*(n_1 - n_2 - n_3 + n_4) \end{pmatrix}}_{\mathbf{R}^{modified}_{1,1}}$$

Figure 4-5: Decoder example for 4 transmit antennas and 1 receive antenna using the 4x4 Hadamard matrix for G and codeword length $L = 4$

Note that multiplying the received signals on the left by G^* has the effect of decoupling from each other the signals transmitted from the other transmit antennas. The end result is that the i^{th} element in $\mathbf{R}^{modified}_{1,1}$ is a function of only $d_{1,i}$, as opposed to being a function of the superposition of **all** the signals $(d_{1,1}, ..., d_{1,4})$. This is a direct

benefit from the columns of G being orthonormal, a property inherent to unitary matrices. Also, note that the noise terms in $\mathbf{R}_{1,1}^{modified}$ are still uncorrelated due to the orthogonality of the columns of G.

4.3 Analysis of Diversity Order

As mentioned in Section 3.3.2, the diversity of space-time codes is related to the minimum rank of the STC error matrix $B(\mathbf{d}, \mathbf{e})$ when taken over all possible $(\mathbf{d}, \mathbf{e})$ modulation symbol sequence pairs. However, unlike space-time codes based on orthogonal designs, which always yielded full diversity order, the transmit diversity of space-time codes based on unitary matrices is always $R \leq N$. Hence, being that this construction is based on a concatenated code principle, the diversity order is derived from both the space-time code as well as the underlying Euclidean-space channel block code.

Let us examine the STC error matrix for the construction introduced here. The STC error matrix must be found for the *entire* symbol sequence, $\mathbf{d}$, hence if the number of modulation symbols needed to send a codeword from C_{ES} is longer than the number of transmit antennas; e.g. $J = xN$, then x transmission blocks will be needed to transmit $\mathbf{d}$ in its entirety and hence the error matrix will contain x copies of the code generating matrix G, each copy being scaled by a different segment of symbols from $\mathbf{d}$.

Consider the case of two distinct modulation symbol sequences $\mathbf{d}$ and $\mathbf{e}$ derived from two different codewords from C_{ES}, which are of the following form:

$$\mathbf{d} = d_{1,1}, d_{1,2}, ..., d_{1,N}, d_{2,1}, ..., d_{2,N}, ..., d_{x,1}, ..., d_{x,N}$$

and

$$\mathbf{e} = e_{1,1}, e_{1,2}, ..., e_{1,N}, e_{2,1}, ..., e_{2,N}, ..., e_{x,1}, ..., e_{x,N}, .$$

The space-time code used to encode each are of the form:

$$\big(G(d_{1,1}, d_{1,2}, ..., d_{1,N}), G(d_{2,1}, d_{2,2}, ..., d_{2,N}), ..., G(d_{x,1}, d_{x,2}, ..., d_{x,N})\big)$$

and

$$\big(G(e_{1,1}, e_{1,2}, ..., e_{1,N}), G(e_{2,1}, e_{2,2}, ..., e_{2,N}), ..., G(e_{x,1}, e_{x,2}, ..., e_{x,N})\big),$$

respectively, where notation for $G(d_{i,1}, d_{i,2}, ..., d_{i,N})$ was introduced in equation (4.2). The STC error matrix is then given by

$$B(\mathbf{d}, \mathbf{e}) = \big(G(d_{1,1}, d_{1,2}, ..., d_{1,N}) - G(e_{1,1}, e_{1,2}, ..., e_{1,N}), ..., G(d_{x,1}, d_{x,2}, ..., d_{x,N}) - G(e_{x,1}, e_{x,2}, ..., e_{x,N})\big)$$

or equivalently,

$$B(\mathbf{d}, \mathbf{e}) = \big(G(d_{1,1} - e_{1,1}, d_{1,2} - e_{1,2}, ..., d_{1,N} - e_{1,N}), ..., G(d_{x,1} - e_{x,1}, d_{x,2} - e_{x,2}, ..., d_{x,N} - e_{x,N})\big)$$

Hence, as previously noted, G appears x times in $B(\mathbf{d}, \mathbf{e})$. The scaling factor for $n^{th}|_{n=1,...,N}$ column of G in its $i^{th}|_{i=1,...,x}$ appearance is the Euclidean distance $(d_{i,n} - e_{i,n})$ between the modulation symbols ($d_{i,n}$ and $e_{i,n}$).

The rank of a matrix is a count of how many of its *non-zero* columns are linearly independent. Since the code generation matrices for our construction are chosen to be unitary, then by definition, the NxN code generation matrix will have rank N. However, for $B(\mathbf{d}, \mathbf{e})$, the n^{th} column of G could potentially degenerate to zero if $(d_{i,n} - e_{i,n}) = 0$. However, since this column of G appears x times in $B(\mathbf{d}, \mathbf{e})$, then the diversity order of the space-time code is reduced only if *all* appearances of this column is driven to zero; i.e. iff $(d_{i,n} - e_{i,n}) = 0 \; \forall \; i = 1, ..., x$. Hence, the rank of $B(\mathbf{d}, \mathbf{e})$ is the number of columns of G **not** scaled by zero, when viewed over all x possible appearances of them.

Since scalings of the n^{th} column of G depend solely on the Euclidean distances

between the sequences $d_{1,n}, d_{2,n}, ..., d_{x,n}$ and $e_{1,n}, e_{2,n}, ..., e_{x,n}$, let

$$\delta_i(\mathbf{d}, \mathbf{e}) = \begin{cases} 0 & : \quad (d_{1,n}, d_{2,n}, ..., d_{x,n}) = (e_{1,n}, e_{2,n}, ..., e_{x,n}) \\ 1 & : \quad else. \end{cases}$$

The motivation behind this notational definition is that if

$$(d_{1,n}, d_{2,n}, ..., d_{x,n}) = (e_{1,n}, e_{2,n}, ..., e_{x,n}),$$

then $(d_{i,n} - e_{i,n}) = 0 \ \forall \ \ i = 1, ..., x$ and hence the n^{th} column of G cannot contribute to rank of $B(\mathbf{d}, \mathbf{e})$. Hence, the rank of $B(\mathbf{d}, \mathbf{e})$ is $\sum_{i=1}^{N} \delta_i(\mathbf{d}, \mathbf{e})$.

Before presenting an example, it is interesting to note that if $J = N$ and BPSK modulation is used, then the rank of the STC error matrix and hence the transmit diversity order of the space-time code is the same as the minimum distance parameter, d_{min}, of the constituent Euclidean-space channel block code. To see this, note that if $L = N$, then $x = 1$, hence each column of G appears only *once* in the error matrix.

Example 1: $L = 2N$, BPSK Modulation ($b = 1$)

For simplicity, choose $N = 2 \Rightarrow L = 4$. $G = \frac{1}{\sqrt{2}}\left(\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right)$

$$\begin{array}{c} \mathbf{d} = d_{1,1}, d_{1,2}, d_{2,1}, d_{2,2} \\ \underbrace{\mathbf{e} = e_{1,1}, e_{1,2}, e_{2,1}, e_{2,2}} \\ \Downarrow \end{array}$$

$$B(\mathbf{d}, \mathbf{e}) = \frac{1}{\sqrt{2}} \begin{pmatrix} d_{1,1} - e_{1,1} & d_{1,2} - e_{1,2} & d_{2,1} - e_{2,1} & d_{2,2} - e_{2,2} \\ d_{1,1} - e_{1,1} & -(d_{1,2} - e_{1,2}) & d_{2,1} - e_{2,1} & -(d_{2,2} - e_{2,2}) \end{pmatrix}$$

$$\begin{array}{cc} \Downarrow & \Downarrow \\ G(d_{1,1} - e_{1,1}, d_{1,2} - e_{1,2}) & G(d_{2,1} - e_{2,1}, d_{2,2} - e_{2,2}) \end{array}$$

Table 4.1 enumerates all possible combinations of equality and inequality between the symbols of $\mathbf{d}$ and $\mathbf{e}$ and the resulting diversity order. In essence, the diversity order will be

0: if the scalings for *both* appearances of each column of G are zero,

$d_{1,1}, e_{1,1}$	$d_{1,2}, e_{1,2}$	$d_{2,1}, e_{2,1}$	$d_{2,2}, e_{2,2}$	DG	$d_{1,1}, e_{1,1}$	$d_{1,2}, e_{1,2}$	$d_{2,1}, e_{2,1}$	$d_{2,2}, e_{2,2}$	DG
$=$	$=$	$=$	$=$	0	$\neq$	$=$	$=$	$=$	1
$=$	$=$	$=$	$\neq$	1	$\neq$	$=$	$=$	$\neq$	2
$=$	$=$	$\neq$	$=$	1	$\neq$	$=$	$\neq$	$=$	1
$=$	$=$	$\neq$	$\neq$	2	$\neq$	$=$	$\neq$	$\neq$	2
$=$	$\neq$	$=$	$=$	1	$\neq$	$\neq$	$=$	$=$	2
$=$	$\neq$	$=$	$\neq$	1	$\neq$	$\neq$	$=$	$\neq$	2
$=$	$\neq$	$\neq$	$=$	2	$\neq$	$\neq$	$\neq$	$=$	2
$=$	$\neq$	$\neq$	$\neq$	2	$\neq$	$\neq$	$\neq$	$\neq$	2

Table 4.1: Diversity Gain/Order (DG) when $L = 2N$

$d_{1,1}, e_{1,1}$	$d_{1,2}, e_{1,2}$	$d_{1,3}, e_{1,3}$	$d_{1,4}, e_{1,4}$	DG	$d_{1,1}, e_{1,1}$	$d_{1,2}, e_{1,2}$	$d_{1,3}, e_{1,3}$	$d_{1,4}, e_{1,4}$	DG
$=$	$=$	$=$	$=$	0	$\neq$	$=$	$=$	$=$	1
$=$	$=$	$=$	$\neq$	1	$\neq$	$=$	$=$	$\neq$	2
$=$	$=$	$\neq$	$=$	1	$\neq$	$=$	$\neq$	$=$	2
$=$	$=$	$\neq$	$\neq$	2	$\neq$	$=$	$\neq$	$\neq$	3
$=$	$\neq$	$=$	$=$	1	$\neq$	$\neq$	$=$	$=$	2
$=$	$\neq$	$=$	$\neq$	2	$\neq$	$\neq$	$=$	$\neq$	3
$=$	$\neq$	$\neq$	$=$	2	$\neq$	$\neq$	$\neq$	$=$	3
$=$	$\neq$	$\neq$	$\neq$	3	$\neq$	$\neq$	$\neq$	$\neq$	4

Table 4.2: Diversity Gain/Order (DG) when $L = N$

1: if the scaling for at least one appearance of only *one* column of G is nonzero,

2: if the scalings for at least one appearance of *both* columns of G are nonzero.

Example 2: $L = N$, BPSK Modulation ($b = 1$)

For simplicity, choose $N = 4 \Rightarrow L = 4$. $G = \frac{1}{\sqrt{4}} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}$

$$\underbrace{\begin{matrix} \mathbf{d} = d_{1,1}, d_{1,2}, d_{1,3}, d_{1,4} \\ \mathbf{e} = e_{1,1}, e_{1,2}, e_{1,3}, e_{1,4} \end{matrix}}$$
$$\Downarrow$$
$$B(\mathbf{d}, \mathbf{e}) = \frac{1}{\sqrt{4}} \begin{pmatrix} d_{1,1} - e_{1,1} & d_{1,2} - e_{1,2} & d_{1,3} - e_{1,3} & d_{1,4} - e_{1,4} \\ d_{1,1} - e_{1,1} & -(d_{1,2} - e_{1,2}) & d_{1,3} - e_{1,3} & -(d_{1,4} - e_{1,4}) \\ d_{1,1} - e_{1,1} & d_{1,2} - e_{1,2} & -(d_{1,3} - e_{1,3}) & -(d_{1,4} - e_{1,4}) \\ d_{1,1} - e_{1,1} & -(d_{1,2} - e_{1,2}) & -(d_{1,3} - e_{1,3}) & d_{1,4} - e_{1,4} \end{pmatrix}$$
$$\Downarrow$$
$$G(d_{1,1} - e_{1,1}, d_{1,2} - e_{1,2}, d_{1,3} - e_{1,3}, d_{1,4} - e_{1,4})$$

Table 4.2 enumerates all possible relationship combinations between the symbols of $\mathbf{d}$ and $\mathbf{e}$ and the resulting diversity order. However, in contrast to the previous

example, each column of G now only appears **once** in the STC error matrix. Consequently, the diversity order will be automatically reduced by one if the scaling for one column of G is zero. Recall that in the previous example, the scaling for one appearance of a column can be zero, but as long as the scaling for the other appearance of the column is *nonzero*, the diversity is not reduced. In essence, the diversity order for this example will be

0: if $\mathbf{d} = \mathbf{e}$,

1: if the scaling for any one column of G is nonzero,

2: if the scalings for any two columns of G are nonzero,

3: if the scalings for any three columns of G are nonzero,

4: if $\mathbf{d} \neq \mathbf{e}$.

Consequently, since in this case, the transmit diversity order is the same as the minimum distance of the constituent Euclidean-space channel code, using a constituent code with large d_{min} would maximize the diversity advantage of the concatenated code.

4.4 Summary

Building on the fundamental concepts of space-time codes introduced in the previous chapter, a new code construction of concatenating Euclidean space channel codes with space-time codes was presented. The simplicity of the encoding and decoding processes made this technique very appealing in that the potential gains in capacity and communication reliability can be obtained without adding significant processing complexity. Moreover, the diversity order of codes constructed in this way was shown to depend indirectly on the Hamming distance properties of the constituent Euclidean space channel code. Two examples of this will be provided in the next chapter.

Chapter 5

Simulation Description and Results

The C-based simulations written in support of this code construction technique aim to verify performance improvements expected from both the coding and diversity gains achieved by these codes. If the construction is viewed as a concatenated code, then the outer code is a Euclidean-space code and the inner code is the space-time code.

The outer Euclidean-space channel code already provides some coding gain over uncoded signals. The concatenated nature of this construction scheme should supplement the coding gain of this channel code and provide further performance improvements by way of diversity gains inherent to space-time codes. Although concatenated codes typically provide performance advantage at the expense of increased decoding complexity [6], it has already been demonstrated that the decoding complexity for our code construction is not increased by the concatenation. In fact, the decoder remains the same as that of the constituent Euclidean-space channel code, requiring only simple linear processing on the received signals prior to decoding.

This chapter is organized as follows:

Section 5.1 provides a brief overview of the simulation structure.

Section 5.2 briefly describes the two Euclidean-space channel codes that were simulated and their respective Euclidean-space decoders.

Section 5.3 presents simulation results and draw conclusions about the performance

gains of the construction scheme.

5.1 Simulation Description

Two scenarios were modeled via simulation. Both scenarios assumed $N > 1$ transmit antennas, $M = 1$ receive antennas [1]. Each simulation modelled systems of rate 1 bit/sec/Hz. The two scenarios are as follows:

1. Information bits are BPSK modulated but not channel nor space-time (ST) coded. (baseline case)

2. Information bits are rate $\frac{1}{2}$ channel-coded, QPSK modulated, and then space-time coded as prescribed by the proposed construction technique.

The two rate $\frac{1}{2}$ channel codes simulated are First-order Reed Muller RM(1,3) codes, and extended Hamming codes, both of which are (8,4,4) linear block codes.

The channel used for all scenarios simulated is modeled to be a flat Rayleigh fading channel, distorted by complex additive white Gaussian noise. Specifically, if $CN(0, \sigma^2)$ denotes a complex Gaussian distribution with zero mean and variance σ^2 per complex dimension (or $\frac{\sigma^2}{2}$ per real dimension) and $s_n|_{n=1,\ldots,\mathrm{N}}$[2] is the symbol transmitted from n^{th} transmit antenna, then the received signal after passing through this channel is of the form

$$r = \sum_{n=1}^{\mathrm{N}} h_n s_n + \eta$$

where $h_i \sim CN(0,1)$ and $\eta \sim CN(0, \frac{1}{SNR})$. To ensure r has an effective signal-to-noise ratio of SNR, s_i is normalized to have energy $\frac{1}{\mathrm{N}}$.

The receivers for the uncoded BPSK system consists merely of a threshold detector. In contrast, the receivers for the coded cases first linearly process the space-time

[1] Note that the theory behind this construction technique is applicable to systems with more than one receive antenna; however, no simulation support was provided for these configurations in this thesis.

[2] For the baseline case, s_n are BPSK symbols. For the coded cases, the s_n are complex QPSK symbols.

coded QPSK symbols and then perform some additional processing on the resulting symbols to convert them into a form appropriate for the Euclidean-space channel decoders, which have been designed for real-valued and not complex-valued inputs.

The next section will provide further detail on the channel decoders used. For both channel codes, an additional step of inverse mapping the decoded codewords into information sequences was needed prior to calculating the bit-error rate.

5.2 Euclidean-space Channel Codes and their Decoders

Prior to delving into the details of the encoder and decoders of the two specific channel codes simulated for this thesis, it would be useful to review the fundamentals behind linear block codes and other related topics.

5.2.1 Review of Linear Block Codes

Definition: Linear Block Codes

(n, k) Linear blocks codes operate on the principle of mapping k input symbols onto n code word symbols. If the alphabet for these symbols are elements of the Galois field (GF(2)), then these codes are called *binary* linear block codes. Henceforth, all linear block codes will be assumed to be binary and, from this point on, will be referred to simply as code C [18].

In contrast to regular block codes, *linear* block codes are special in that the code C forms a vector subspace over GF(2), which means that it exhibits all the properties of a subspace, including linearity. This implies:

- C is required to contain the all-zero vector or codeword.
- C can be represented by a k x n generator matrix G, whose rows span the subspace of C.

- C is associated with a $(n-k)$ x n parity check matrix H, whose rows span a subspace that is orthogonal to that of C [18].

To encode using a (n, k) linear block code, right multiply a row vector with k information bits by the generator matrix G and the result will be a n-element row vector containing the codeword.

Definition: Perfect codes

C is considered to be a perfect code if its redundancy parameter $r = n - k$ meets the Hamming bound with equality. The Hamming bound is defined as follows: A Hamming sphere of radius t contains all possible received vectors that are at a Hamming distance t from a specific code word within the vector space of dimension n. If the volume of the resulting Hamming sphere is $V_q(n, t)$, where

$$V_q(n,t) = \sum_{j=0}^{t} \binom{n}{j} (q-1)^j,$$

then the Hamming bound is $r \geq \log_q V_q(n, t)$ [18].

Meeting this bound with equality implies that all the Hamming spheres do not overlap and completely fill out the n-dimensional vector space spanned by C. Hence a perfect q-ary code of length n can correct t errors using only minimal bits of redundancy.

Definition: Hadamard Matrices [18]

Hadamard matrices, H_n, of order n are n x n matrices of +1s and -1s, with the property that $H_n H_n^T = H_n^T H_n = nI$. This implies the rows and columns of H_n are mutually orthogonal to each other and are of length n. As shown in the previous chapter, this latter property was very useful and was exploited in our code construction.

Furthermore, H_n only exists for values of n that are multiples of 1, 2, 4, or multiples of 4. The Hadamard matrices implemented in this thesis were constructed using the Sylvester Construction method, which starts with $H_1 = [1]$ and recursively generates

higher order Hadamard matrices as follows:

$$H_{2n} = \begin{pmatrix} H_n & H_n \\ H_n & -H_n \end{pmatrix}$$

5.2.2 First-order Reed-Muller Codes

Reed-Muller codes are not strong channel codes, but they have the advantage of having simple decoders. They were used extensively in the 1960s and 1970s on space missions, and have recently returned to the limelight due to their extremely high-speed maximum-likelihood decoding algorithms, which is an attractive quality to many optical communication applications [18].

The (n, k, d) parameters for r^{th}-order Reed-Muller codes $RM(r, m)$ are as follows [18]:

$$\begin{aligned} n &= 2^m \\ k &= \sum_{j=0}^{r} \binom{m}{j} \\ d &= 2^{m-r} \end{aligned}$$

For first-order Reed-Muller codes $RM(1, m)$, these parameters simplify to

$$\begin{aligned} n &= 2^m \\ k &= \binom{m}{0} + \binom{m}{1} = 1 + m \\ d &= 2^{m-1} \end{aligned}$$

Encoder

The generator matrix for first-order Reed-Muller codes is a $(m + 1)$ x (2^m) matrix. The easy way to generate this matrix is to make the first column of the generator matrix equal to the column vector consisting of a one followed by m zeros. Subsequent columns are simply the previous column incremented by 1 using modulo 2 arithmetic.

For example, the generator matrix for $RM(1,3)$ is

$$G = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}$$

Decoder

The maximum likelihood decoding algorithm for a Gaussian channel can be simplified to performing a correlation operation on the received signal against all possible codewords in the codebook. Conveniently, the rows of a N x N Hadamard matrix H_N and its complement matrix $-H_N$ enumerate all the codewords of a $RM(1, m = \log_2 N)$ code. Hence, the correlation operation can be accomplished by multiplying the received N-symbol column vector on the left by H_N. Fast Hadamard transforms are computationally efficient techniques for performing precisely this task.

The elements of the vector resulting from this multiplication are the correlation values between each row of H_N with the received vector. If the i^{th} element of this vector stores the correlation value with the largest absolute magnitude, the codeword is decoded into either the i^{th} row of H_N or $-H_N$, depending on the sign of this correlation value.

5.2.3 (7,4,3) Hamming Codes

Hamming codes are pioneers in the field of error correcting codes. They were introduced in April 1950 and first used for forward error correction in long-distance telephony. They are perfect codes capable of single-error correction and are parame-

terized as follows [18]:

$$\begin{aligned} \text{code word length, n} &= 2^m - 1 \\ \text{number of information bits, k} &= 2^m - m - 1 \\ \text{minimum distance, d} &= 3 \\ \text{number of parity bits, r} &= n - k = m \\ \text{error correcting capability, t} &= 1 \text{ bit} \end{aligned}$$

Encoder

The generator matrix for this codes is derived by first constructing the parity-check matrix using the following procedure:

1. Construct a matrix with $(n-k)$-rows, whose columns enumerate all 2^{n-k} possible binary $(n-k)$-tuples excluding the all-zero vector. For instance, for a (7,4,3) Hamming code, this initial matrix would look like

$$\begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}$$

2. Searching from the first column to the last, pick out the columns containing only one nonzero element and move them to the end of the matrix. Arrange those columns so that they resemble a $(n-k)$ x $(n-k)$ identity matrix. The resulting matrix is the parity check matrix H.

$$\begin{pmatrix} \mathbf{0} & \mathbf{0} & 0 & \mathbf{1} & 1 & 1 & 1 \\ \mathbf{0} & \mathbf{1} & 1 & \mathbf{0} & 0 & 1 & 1 \\ \mathbf{1} & \mathbf{0} & 1 & \mathbf{0} & 1 & 0 & 1 \end{pmatrix} \qquad H = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}$$

3. If P denotes the set of columns that do not constitute the $(n-k)$ x $(n-k)$ identity matrix, then construct the generator matrix G by appending the transpose of

P onto the end of a k x k identity matrix.

$$P = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \end{pmatrix} \qquad P' = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \qquad G = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}$$

Decoder

The Hamming distance decoder for this code involves computing a syndrome for the received word using the parity check matrix; however, for our code construction, a Euclidean distance decoder was required. Consequently, a Viterbi trellis decoder based on a technique developed by Wolf [20] was implemented. For details of this algorithm, the reader is encouraged to read [20]; however, a summary of the key steps in the algorithm will be presented here.

Any (n, k) linear block code over $GF(q)$ can be maximum likelihood soft-decision decoded using a Viterbi decoder with complexity q^{n-k}, which is a significant savings over any correlation decoder which correlates the received word against each of q^k possible codewords [20].

For a (n, k) binary $(q = 2)$ linear block code C , there are 2^k possible codewords. A trellis is a useful way for keeping track of all 2^k codewords because each codeword corresponds to a distinct path traversal through the trellis.

Prior to introducing the algorithm for constructing the trellis for this code, let us establish the following notational conventions:

- If H is the $(n - k)$ x n parity matrix for C, let $\vec{h}_i|_{i=1,\ldots,n}$ denote the i^{th} column of H.

- If d specifies the depth within a trellis, then for a (n, k) code, d ranges from 0 to n.

- The states of the trellis are $(n-k)$-tuples, and there are q^{n-k} such $(n-k)$-tuples

that should be listed sequentially.

- I_d is the subset of all possible states "active" at depth d.

- $\vec{s_i}(d)$ is the "active" state node at depth d corresponding to state $i \in I_d$.

ALGORITHM

1. At $d = 0$, the only active state is the zeroth state; i.e., $I_0 = \vec{0}$.

2. To define the trellis for $d = 1, ..., n$, the states that become "activated" at depth d is found by taking each state in I_{d-1} and applying the following formula:

$$\vec{s_l}(d) = \vec{s_i}(d-1) + \alpha_j \vec{h}_d \tag{5.1}$$

   for all $i \in I_{d-1}, j = 0, ..., q-1$ and $\alpha_j = j \in GF(q)$. The set of nodes derived from applying this formula to all the "active" states in I_{d-1} contribute to I_d.

   In drawing the trellis, a line is drawn from each state in I_{d-1} to each new state it"activates" by the formula above, and it is labeled by the appropriate value of α_j.

3. Any paths that do not terminate in the zeroth state at depth n are removed. Only q^k distinct paths should remain, and these paths correspond to the q^k valid codewords.

After constructing the appropriate trellis for code C, standard Viterbi decoding algorithms can be applied over them to decode a real-valued vector that represents the received signals. For an example of a trellis generated by the algorithm outlined above, and for detail on applying the Viterbi decoding algorithm over these trellises, the reader is encouraged to read [20].

Extended Hamming codes (8,4,4)

To promote consistency between all simulations, a rate $\frac{1}{2}$ channel code used in conjunction with a QPSK modulation scheme must be implemented to maintain a transmission rate of 1 bit/sec/Hz. Unfortunately, Hamming codes are rate $\frac{4}{7}$. However, when the "natural" length of a linear block code is unsuitable for a given application, its length can be altered by various means. One way of increasing the length of the codeword is a process called **extending**. To create a rate $\frac{1}{2}$ code from the $\frac{4}{7}$ Hamming code, the code is extended by appending an additional redundant coordinate. Hence the (7,4,3) Hamming code becomes an (8,4,4) extended Hamming code. Notice that not only does the codeword length increase, but so does the minimum Hamming weight.

The conventional method for extending Hamming codes is to add a row of ones and the column vector $[0, 0, 0, ..., 0, 1]^T$ to the parity-check matrix for the Hamming code [18]. However, to retain the property of producing a systematic code, the extended Hamming code will have the following generator (G) and parity check (H) matrices:

$$G = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \end{pmatrix} \quad H = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

The encoding and decoding algorithms for this extended Hamming code remain the same as those used for the Hamming (7,4,3) code.

5.3 Simulation Results

To provide a basis for comparison, the performance curves for systems achieving diversity orders 1, 2, and 3 are shown in Figure 5-1. The systems used to generate these curves consisted of uncoded BPSK symbols being sent over one transmit antenna and being received over one, two, and three receive antennas, respectively. As previously

mentioned, full receive diversity can always be extracted from using multiple receive antennas. MRRC is the linear processing performed on the received signals to produce estimates of the transmitted signals.

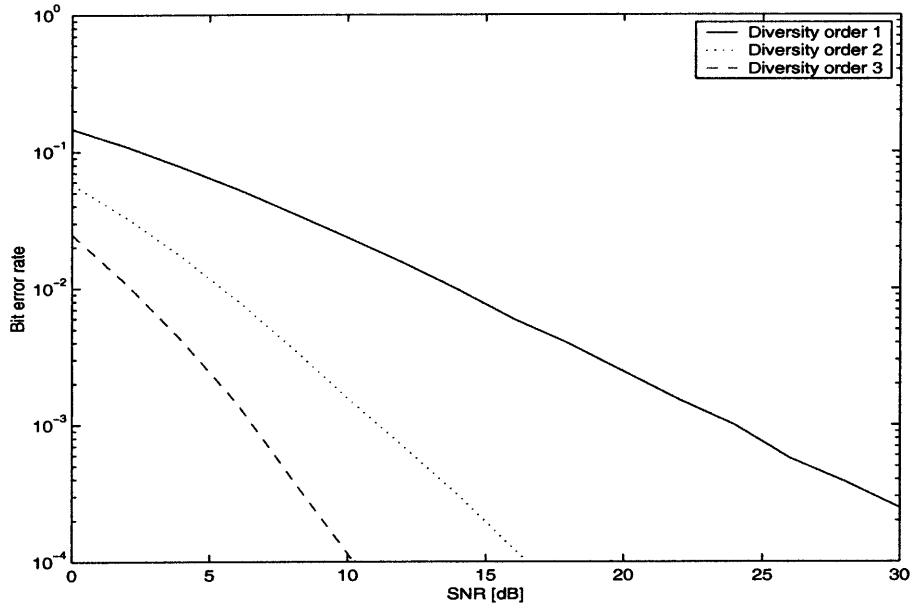

Figure 5-1: Diversity gains achieved using 1, 2 and 3 receive antennas for uncoded BPSK systems.

Notice that the slopes of each curve is approximately equal to the negative of the diversity order. To see this, note that from equation 3.3,

$$\begin{aligned} P_e &\sim \frac{1}{SNR^{DG}} \\ \log P_e &\sim -DG * \log SNR \end{aligned}$$

The two communication schemes being compared against uncoded BPSK are shown Figures 5-2 and 5-3. Each of the three simulations were tested using 800,000 information bits, which translated to either 800,000 BPSK symbols or 800,000 QPSK symbols. Four transmit and one receive antenna were used in simulations employing space-time codes.

The same block fading was applied in each scenario. Specifically, the fading coefficients were held constant over four channel uses, which translated to either four

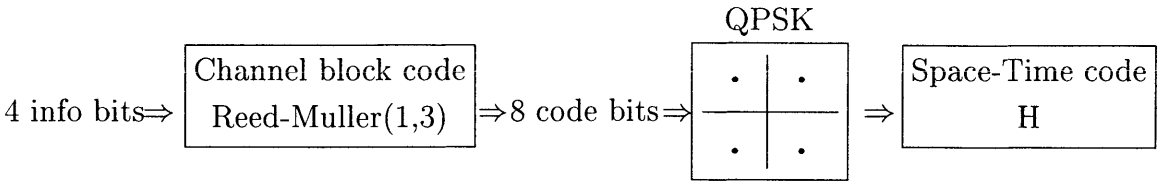

Figure 5-2: Simulation scenario #1

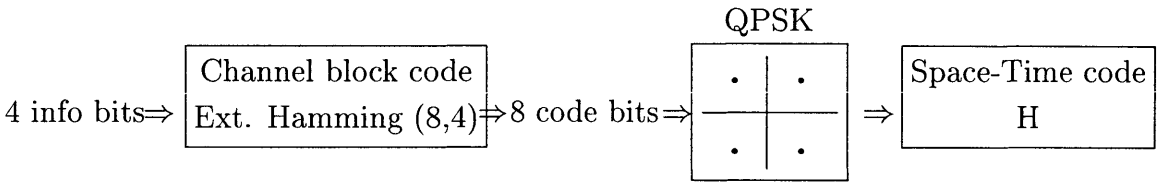

Figure 5-3: Simulation scenario #2

BPSK symbols or four QPSK symbols. However, the implications of this channel condition are different for each simulation. These differences are as follows:

1. In the baseline case, the fading coefficients are held constant over four uncoded and **uncorrelated** BPSK symbols.

2. In the coded cases, the fading coefficients over each of the four transmit antennas are held constant for the four time slots over which four QPSK symbols are transmitted. Since each transmit antenna merely sends the same codeword symbol (or some scaled version of it), each QPSK symbol is faded by a coefficient that is held constant over the four time slots. Recall that each set of four QPSK symbols encapsulate the eight bits constituting to one codeword. Hence, effectively, each *pair* of consecutive channel code bits (constituting to one QPSK symbol) experiences the same fade coefficients over a block of four channel accesses, but is attenuated by a coefficient that is different from any other pair of channel code output bits being transmitted from other transmit antennas.

Figures 5-4 shows the improvement in relative bit error rate performance when comparing uncoded BPSK with channel and space-time coded QPSK.

Comparing the slopes of the three curves in Figure 5-4 to those in Figure 5-1, the following three observations can be made:

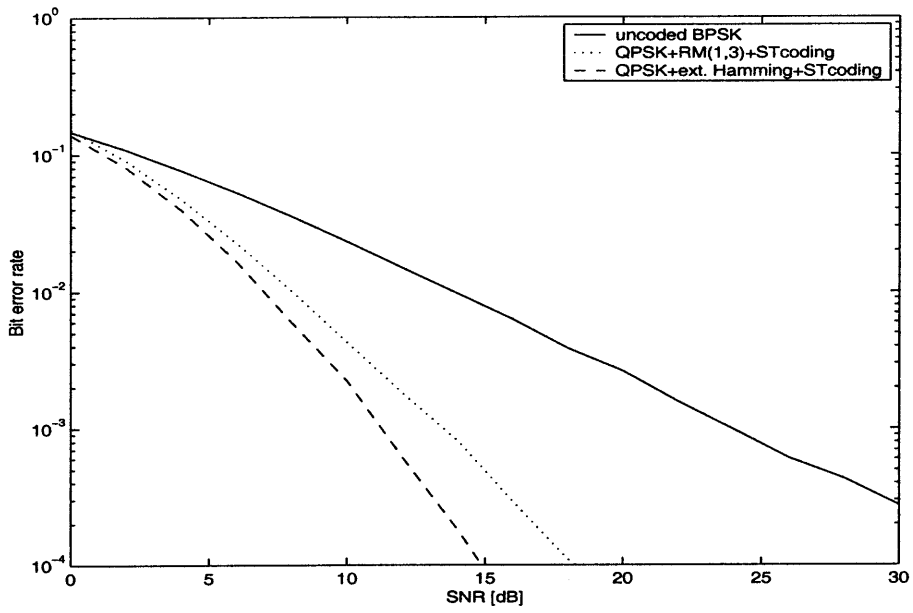

Figure 5-4: Diversity gains achieved by transmitting channel block coded QPSK over 4 transmit antennas.

1. The uncoded BPSK system achieves a diversity order of 1.

2. The RM(1,3) channel coded QPSK system achieves a diversity order of 2 when space-time coding is employed.

3. The extended Hamming channel coded QPSK system achieves a diversity order of 3 when space-time coding is employed.

These simulation results agree perfectly with theoretical predictions. The uncoded BPSK system employs a single antenna at both the transmit and receive stations. Since each BPSK symbol within a block, defined previously as four BPSK symbols, is attenuated by the same fading coefficient, it is not surprising to observe no diversity gain or a diversity order of 1 for this system.

In contrast, when the information bits are first channel coded and then space-time coded, diversity advantage can be achieved. The channel code amortizes each information bit over the codeword bits; consequently, using a rate $\frac{1}{2}$ channel code spreads the four input information bits over eight codeword bits. Therefore, these

eight codeword bits are correlated with each other by virtue of them constituting to a codeword selected from a carefully designed codebook consisting of only 2^4 out of 2^8 possible eight-tuples.

The channel between each of the four transmit antennas to the receive antenna are characterized by four independent fade coefficients, which are held constant over four transmissions. Each pair of adjacent codeword bits is mapped to a QPSK symbol prior to the space-time coding. Using the encoding scheme described in Section 4.2.3, each antenna is dedicated to sending the same QPSK symbol over the four transmission slots. As a result, each pair of codeword bits and equivalently different parts of a single codeword experiences different degrees of fade.

Since the codeword bits are correlated, they can provide information about each other at the receiver in the following way: Suppose of the four transmit antennas, three of them experience deep fades. Had the bits constituting to the QPSK symbols been uncorrelated, it would be difficult, if not impossible, for the other six bits to be recovered. However, since the eight bits being sent over this channel form a codeword and are consequently correlated, the two recoverable bits at the fourth antenna can be used to derive some information about the other six codeword bits.

A repetition code can best illustrate the benefits of this scheme. Consider a repetition code where one input bit produces a codeword which is merely the input repeated four times. This is a rate $\frac{1}{4}$ repetition code. Furthermore, suppose the space-time coder uses four transmit antennas, each transmitting one code bit, which in this case is a copy of the input bit, but this time, only *one* transmission slot is used. This is a rate $\frac{1}{4}$ space-time code. Compare the performance of this scheme to the simple case of sending the input bit uncoded; i.e. directly over the channel, from a single transmit antenna, bypassing the repetition channel coding.

Suppose in the latter scenario, the fade coefficient associated with the single transmit antenna, α is small, indicating a deep fade. It would be very difficult for the decoder at the receiver to recover this information bit.

In contrast, suppose in the former scenario, the fade coefficients over the four transmit antennas are α,β, γ, and θ. Assume α, β, γ, and θ are independent. Con-

sequently, if α, β, and γ are small but θ is large, then the bit sent over the fourth antenna could reliably be recovered and be used to recover the other three bits because it is correlated with those bits. Specifically, in the case of the repetition code, since all four code bits are equal, the correlation between them is maximal. Therefore, the bit recovered from the fourth channel can be used to determine exactly what the other three bits are because they are all supposed to be the same. This scheme achieves full transmit diversity order of 4.

It is now apparent why the diversity advantage from using the construction scheme presented here is derived from the fact that each QPSK symbol is faded by independent fade variables and from the fact that the bits being mapped into QPSK symbols are correlated. This also explains why the diversity order of the space-time code resulting from this construction relies so heavily on the underlying Euclidean-space channel code.

In fact, it is easy to show that without the channel code, there is no diversity gain. Intuitively, this is a reasonable claim because without channel coding, the bits being sent would be uncorrelated and hence the benefits of having independent fades over the different transmit antennas cannot be realized at the receiver. A more rigorous proof follows.

From [7], the diversity order of a space-time code is related to the minimum rank of the space-time code error matrix $B(\mathbf{c}, \mathbf{e})$ when taken over all possible (**c**,**e**) pairs. Consider the 4 x 4 Hadamard transmission matrix, where the four input symbols modulating this transmission matrix are denoted as $\mathbf{c} = c_1, c_2, c_3, c_4$. Compare this against the transmission matrix modulated by quadruple $\mathbf{e} = e_1, e_2, e_3, e_4$. Assuming BPSK modulation, let $c_i, e_i \in (\pm b)$. The error matrix is as follows:

$$B(\mathbf{c}, \mathbf{e}) = \begin{pmatrix} c_1 - e_1 & c_2 - e_2 & c_3 - e_3 & c_4 - e_4 \\ c_1 - e_1 & -(c_2 - e_2) & c_3 - e_3 & -(c_4 - e_4) \\ c_1 - e_1 & c_2 - e_2 & -(c_3 - e_3) & -(c_4 - e_4) \\ c_1 - e_1 & -(c_2 - e_2) & -(c_3 - e_3) & (c_4 - e_4) \end{pmatrix}$$

Suppose the symbols constituting **c** and **e** were uncorrelated; i.e. randomly gener-

ated. The minimum distance between the two quadruples would be 1. As an example, suppose $\mathbf{c} = (b, b, b, b)$ and $\mathbf{e} = (b, b, b, -b)$. The resulting error matrix would be

$$B(\mathbf{c}, \mathbf{e}) = \begin{pmatrix} 0 & 0 & 0 & 2b \\ 0 & 0 & 0 & -2b \\ 0 & 0 & 0 & -2b \\ 0 & 0 & 0 & 2b \end{pmatrix}$$

The rank of this matrix is 1 and since it is the minimum possible rank when taken over all possible (**c**,**e**) pairs, the diversity order is 1.

In contrast, consider the scenario when the symbols constituting **c** and **e** *are* correlated; for instance, when they are part of a codeword. Analysis of this scenario leads directly to the reason why Simulations #1 and #2 yielded diversity orders greater than 1.

Simple proofs supporting the diversity orders shown by the simulations for the two channel codes tested are presented next. However, it is first useful to establish some notational conventions that can facilitate the analysis.

The output of the channel coders are bits that are to be paired and then sent to a QPSK modulator. The mapping used is:

$$\begin{aligned} (1,1) &\Rightarrow (+\frac{1}{\sqrt{2}}, +\frac{1}{\sqrt{2}}) \Rightarrow a \\ (0,1) &\Rightarrow (-\frac{1}{\sqrt{2}}, +\frac{1}{\sqrt{2}}) \Rightarrow b \\ (0,0) &\Rightarrow (-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}) \Rightarrow c \\ (1,0) &\Rightarrow (+\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}) \Rightarrow d \end{aligned}$$

5.3.1 Diversity order of Simulation #1

Simulation #1 employed a Reed-Muller RM(1,3) channel code. This is an (8,4,4) linear block code. The eight-bit codewords and the resulting four QPSK symbols generated are shown in Table 5.1.

Input bits				Codeword bits								QPSK symb.				distance to (c,c,c,c)
0	0	0	0	0	0	0	0	0	0	0	0	c	c	c	c	0
0	0	0	1	1	1	1	1	1	1	1	1	a	a	a	a	4
0	0	1	0	1	0	1	0	1	0	1	0	d	d	d	d	4
0	0	1	1	0	1	0	1	0	1	0	1	b	b	b	b	4
0	1	0	0	1	1	0	0	1	1	0	0	a	c	a	c	2
0	1	0	1	0	0	1	1	0	0	1	1	c	a	c	a	2
0	1	1	0	0	1	1	0	0	1	1	0	b	d	b	d	4
0	1	1	1	1	0	0	1	1	0	0	1	d	b	d	b	4
1	0	0	0	1	1	1	1	0	0	0	0	a	a	c	c	4
1	0	0	1	0	0	0	0	1	1	1	1	c	c	a	a	4
1	0	1	0	0	1	0	1	1	0	1	0	b	b	d	d	4
1	0	1	1	1	0	1	0	0	1	0	1	d	d	b	b	4
1	1	0	0	0	0	1	1	1	1	0	0	c	a	a	c	2
1	1	0	1	1	1	0	0	0	0	1	1	a	c	c	a	2
1	1	1	0	1	0	0	1	0	1	1	0	d	b	b	d	4
1	1	1	1	0	1	1	0	1	0	0	1	b	d	d	b	4

Table 5.1: Codewords and associated QPSK symbols for RM(1,3)

Since the minimum distance after the channel code bits are mapped into QPSK symbols is 2, the minimum rank of the error matrix resulting from this code construction will also be 2 and hence the diversity order achieved by this scenario is 2, as supported by the performance curve generated by Simulation #1.

5.3.2 Diversity order of Simulation #2

Simulation #2 employed an extended Hamming channel code. This is also an (8,4,4) linear block code. The eight-bit codewords and the resulting four QPSK symbols generated are shown in Table 5.2.

Since the minimum distance after the channel code bits are mapped into QPSK symbols is 3, the minimum rank of the error matrix resulting from this code construction will also be 3 and hence the diversity order achieved by this scenario is 3, as supported by the performance curve generated by Simulation #2.

Input bits				Codeword bits								QPSK symb.				distance to (c,c,c,c)
0	0	0	0	0	0	0	0	0	0	0	0	c	c	c	c	0
0	0	0	1	0	0	0	1	1	1	1	0	c	b	a	d	3
0	0	1	0	0	0	1	0	0	1	1	1	c	d	b	a	3
0	0	1	1	0	0	1	1	1	0	0	1	c	a	d	b	3
0	1	0	0	0	1	0	0	1	0	1	1	b	c	d	a	3
0	1	0	1	0	1	0	1	0	1	0	1	b	b	b	b	4
0	1	1	0	0	1	1	0	1	1	0	0	b	d	a	c	3
0	1	1	1	0	1	1	1	0	0	1	0	b	a	c	d	3
1	0	0	0	1	0	0	0	1	1	0	1	d	c	a	b	3
1	0	0	1	1	0	0	1	0	0	1	1	d	b	c	a	3
1	0	1	0	1	0	1	0	1	0	1	0	d	d	d	d	4
1	0	1	1	1	0	1	1	0	1	0	0	d	a	b	c	3
1	1	0	0	1	1	0	0	0	1	1	0	a	c	b	d	3
1	1	0	1	1	1	0	1	1	0	0	0	a	b	d	c	3
1	1	1	0	1	1	1	0	0	0	0	1	a	d	c	b	3
1	1	1	1	1	1	1	1	1	1	1	1	a	a	a	a	4

Table 5.2: Codewords and associated QPSK symbols for extended Hamming codes (8,4,4)

Chapter 6

Conclusions

This thesis presented a code construction which jointly exploited spatial and temporal diversity through space-time block coding and also provided coding gain via channel coding. This chapter summarizes the research findings and provides direction for future research.

6.1 Summary

Multipath fading is an ailment that plagues the wireless medium. However, using diversity techniques, chances for obtaining reliable communication can be increased in spite of the presence of multipath. In particular, it has already been shown that space-time codes employed over MIMO systems can benefit from the multipath environment as well as provide significant performance improvements over single antenna systems. Consequently, provisions for reasonably reliable communication and throughput at high data rates makes support for multimedia Internet applications for wireless devices a closer reality.

Although space-time block codes provide both coding and diversity gains over uncoded systems, further coding gains can be achieved by concatenating space-time block codes with Euclidean space channel block codes. As shown by both theoretical as well as simulation findings, the diversity order of the resulting concatenated code is related to combinatorial properties of the outer Euclidean space channel code.

The construction of the code is quite simple. First the information bits are channel coded and then the channel code bits are modulated and sent to the space-time coder. The decoder merely involves simple linear processing of the received signals prior to regular Euclidean-space decoding for the constituent channel code.

Hence, through simple encoding and low-complexity decoding algorithms, *any* Euclidean space block code can be turned into a space-time code, while providing significant performance improvements over uncoded systems as well as systems employing only channel coding.

Although simulations for multiple receive antenna configurations were not including in this thesis, many of the theoretical and mathematical derivations particularly for the decoder side are generalizable to systems with multiple antennas at *both* the transmitter and receiver. Furthermore, similar performance improvements are expected for comparisons of systems with higher spectral efficiency such as $\rho = 2$ or 3 bits/sec/Hz.

6.2 Future Work

Currently, the code generator matrix, G has been constrained to be complex unitary. Results of this choice are simplicity and low complexity at the decoder as well as maintenance of uncorrelated noise terms after the linear processing. However, it would be more interesting to investigate the ramifications of applying this technique to transmission matrices of a more generalized form. Naturally, the decoder will be more complex; i.e., it will no longer reduce to mere linear processing of the received signals because the noises would no longer be uncorrelated. However, perhaps any increase in decoder complexity can be offset or justified by provisions of benefits in the form of increased capacity and system performance.

Bibliography

[1] W. Mohr and R. Becher, "Mobile Communications Beyond Third Generation," *Proc. IEEE VTC 2000*, vol. 2, pp. 654-661, Sept. 2000.

[2] The World of Wireless Communications http://www.wow-com.com

[3] R. Steele, *Mobile Radio Communication*, New York: IEEE Press, 1992.

[4] T. Rappaport, *Wireless Communications: Principles & Practice*, New Jersey: Prentice Hall PTR, 1996.

[5] S. Alamouti, "A Simple Transmit Diversity Technique for Wireless Communications," *IEEE J. Select. Area Commun.*, vol. 16, pp. 1451-1458, Oct. 1998.

[6] J. Proakis, *Digital Communications*, New York: McGraw-Hill International Editions, 1995.

[7] V. Tarokh, N. Seshadri, and A. Calderbank, "Space-Time Codes for High Data Rate Wireless Communication: Performance Criterion and Code Construction," *IEEE Trans. Inform.* Theory, vol. 44, pp. 744-765, March 1998.

[8] A. Narula, M. Trott, and G. Wornell , "Performance Limits of Coded Diversity Methods for Transmitter Antenna Arrays," *IEEE Trans. Inform. Theory*, vol. 45, pp. 2418-2433, Nov. 1999.

[9] T. Cover and J. Thomas, *Elements of Information Theory*, New York: John Wiley & Sons, 1999.

[10] A. Naguib, N. Seshadri, and A. Calderbank, "Applications of Space-Time Block Codes and Interference Suppression for High Capacity and High Data Rate Wireless Systems," *Confer. on Signals, Systems & Computers, 1998,* vol. 2, pp. 1803 -1810, Nov. 1998.

[11] E. Telatar, "Capacity of multi-antenna Gaussian channels," *AT&T-Bell Labs Internal Tech. Memo.*, June 1995.

[12] G. Foschini and M. Gans, "On Limits of Wireless Communications in a Fading Environment when Using Multiple Antennas," *Wireless Personal Communications*, vol. 6, pp. 311-335, March 1998.

[13] A. Wittneben, "Basestation modulation diversity for digital SIMULCAST," *Proc. IEEE VTC'91*, vol. 1, pp. 848-853, May 1991.

[14] N. Seshadri and J. Winters, "Two Signaling Schemes for Improving the Error Performance of Frequency-Division-Duplex (FDD) Transmission System Using Transmitter Antenna Diversity," *Int. J. Wireless Inform. Networks*, vol. 1 pp. 49-60, Jan. 1994.

[15] J. Winters, "The Diversity Gain of Transmit Diversity in Wireless Systems with Rayleigh Fading," *IEEE Trans. Veh. Technol.*, vol. 47, pp. 119-123, Feb. 1998.

[16] V. Tarokh, H. Jafarkhani, and A. Calderbank, "Space-Time Block Codes from Orthogonal Designs," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1456-1467, July 1999.

[17] V. Tarokh, H. Jafarkhani, and A. Calderbank, "Space-Time Block Coding for Wireless Communications: Performance Results," *IEEE J. Select. Area Commun.*, vol. 17, pp.451-460, March 1999.

[18] S. Wicker, *Error Control Systems for Digital Communication and Storage*, New Jersey: Prentice Hall, 1995.

[19] F. MacWilliams and J. Sloane, *The Theory of Error-Correcting Codes*, New York: North-Holland, 1996.

[20] J. Wolf, "Efficient Maximum Likelihood Decoding of Linear Block Codes Using a Trellis," *IEEE Trans. Inform. Theory*, vol. 24, pp. 76-80, Jan. 1978.

[21] S. Haykin, *Communication Systems*, New York: John Wiley & Sons, Inc., 1994.