Fall 1-31-2012

High rate space time code with linear decoding complexity for multiple transmitting antennas

Amir Laufer

New Jersey Institute of Technology

Copyright Warning & Restrictions

The copyright law of the United States (Title 17, United States Code) governs the making of photocopies or other reproductions of copyrighted material.

Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be "used for any purpose other than private study, scholarship, or research." If a, user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of "fair use" that user may be liable for copyright infringement,

This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve violation of copyright law.

Please Note: The author retains the copyright while the New Jersey Institute of Technology reserves the right to distribute this thesis or dissertation

Printing note: If you do not wish to print this page, then select "Pages from: first page # to: last page #" on the print dialog screen

ABSTRACT

HIGH RATE SPACE TIME CODE WITH LINEAR DECODING COMPLEXITY FOR MULTIPLE TRANSMITTING ANTENNAS

by
Amir Laufer

The multipath nature of the wireless channel, results in a superposition of the signals of each path at the receiver. This can lead to either constructive or destructive interference. Strong destructive interference is frequently referred to as deep fade and may result in temporary failure of communication due to the severe drop in the channel's signal-to-noise ratio (SNR). To avoid this situation, signal diversity might be introduced. When having more than one antenna at the transmitter and / or receiver, forming a Multiple-Input Multiple-Output (MIMO) channel, spatial diversity can be employed to overcome the fading problem. Space time block codes (STBC) have been shown to be used well with the MIMO channel. Each type of STBC is designed to optimize a different criteria such as rate and diversity, while other characteristics of the code are its error performance and decoding computational complexity. The Orthogonal STBC (OSTBC) family of codes is known to achieve full diversity as well as very simple implementation of the Maximum Likelihood (ML) decoder. However, it was proven that, with complex symbol constellation one cannot achieve a full rate code when the number of transmitting antennas is larger than two. Quasi OSTBC are codes with full rate but with the penalty of more complex decoding, and in general does not achieve full diversity.

In this work, new techniques for OSTBC transmission / decoding are explored, such that a full rate code can be transmitted and decoded with linear complexity. The Row Elimination Method (REM) for OSTBC transmission is introduced, which basically involves the transmission of only part of the original OSTBC codeword, resulting in a full rate code termed Semi-Orthogonal STBC (SSTBC). Novel decoding

scheme is presented, such that the SSTBC decoding computational complexity remains linear although the transmitted codeword is not orthogonal anymore. A new OSTBC, that complies with the new scheme's requirements, is presented for any number of transmit antennas. The performance of the new scheme is studied under various settings, such as system with limited feedback and multiple antennas at the receiver.

The general decoding techniques presented for STBC, assume perfect channel knowledge at the receiver. It was shown, that the performance of any STBC system is severely degraded due to partial channel state information, results from imperfect channel estimation. To minimize the performance loss, one may lengthen the training sequences used for the channel estimation which, inevitably, results in some rate loss. In addition, complex decoding schemes can be used at the receiver to jointly decode the data while enhancing the channel estimation. It is suggested in this work to apply adaptive techniques to mitigate the performance loss without the penalty of additional rate loss or complex decoding. Namely, the bootstrap algorithm is used to further refine the received signals, resulting in better effective rate and performance in the presence of channel estimation errors. Modified implementations for the bootstrap's weights calculation method are also presented, to improve the convergence rate of the algorithm, as well as to maintain a very low computational burden.

HIGH RATE SPACE TIME CODE WITH LINEAR DECODING COMPLEXITY FOR MULTIPLE TRANSMITTING ANTENNAS

by
Amir Laufer

A Dissertation
Submitted to the Faculty of
New Jersey Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Electrical Engineering

Department of Electrical and Computer Engineering, NJIT

January 2012

APPROVAL PAGE

HIGH RATE SPACE TIME CODE WITH LINEAR DECODING COMPLEXITY FOR MULTIPLE TRANSMITTING ANTENNAS

Amir Laufer

Dr. Yeheskel Bar-Ness, Dissertation Advisor — Date
Distinguished Professor, Department of Electrical and Computer Engineering , NJIT

Dr. Alexander M. Haimovich, Committee Member — Date
Professor, Department of Electrical and Computer Engineering , NJIT

Dr. Ali Abdi, Committee Member — Date
Associate Professor, Department of Electrical and Computer Engineering , NJIT

Dr. Osvaldo Simeone, Committee Member — Date
Assistant Professor, Department of Electrical and Computer Engineering , NJIT

Dr. Laurence B. Milstein, Committee Member — Date
Professor, Department of Electrical and Computer Engineering , University of California at San Diego (UCSD)

BIOGRAPHICAL SKETCH

Author: Amir Laufer

Degree: Doctor of Philosophy

Date: January 2012

Undergraduate and Graduate Education:

- Doctor of Philosophy in Electrical Engineering,
  New Jersey Institute of Technology, Newark, NJ, 2012
- Master of Science in Electrical and Electronic Engineering,
  Tel Aviv University, Tel Aviv, Israel, 2006
- Bachelor of Science in Electrical and Electronic Engineering,
  Tel Aviv University, Tel Aviv, Israel, 2004

Major: Electrical Engineering

Presentations and Publications:

A. Laufer and Y. Bar-Ness, "Full Rate Space Time Codes for Large Number of Transmitting Antennas with Linear Complexity Decoding," *Wireless Personal Communications*, Springer, Vol. 57, pp. 465-480, Apr. 2011.

A. Laufer and Y. Bar-Ness, "Bootstrap Decoding for the Alamouti Space-Time Scheme with Imperfect Channel Estimation," *Wireless and Optical Communication Conference* WOCC 2011.

A. Laufer and Y. Bar-Ness, "Improved Bootstrap Decoding Scheme for Space Time Codes with Imperfect Channel Estimation," *Conference on Information Science and Systems* CISS 2011.

A. Laufer and Y. Bar-Ness, "Linear Computational Complexity Decoding for Semi Orthogonal Full Rate Space Time Codes," *IEEE Wireless Communications and Networking Conference* WCNC 2011.

A. Laufer and Y. Bar-Ness, "Adaptive Decoding for Space Time Codes with Imperfect Channel Estimation, Using the Bootstrap Algorithm," *IEEE Broadband Wireless Access (BWA) Workshop*, GLOBECOM 2010.

A. Laufer and Y. Bar-Ness, "Full Rate Space Time Codes for Large Number of Transmitting Antennas with Linear Complexity Decoding and High Performance," *IEEE Information Theory Workshop* ITW 2009.

A. Jain, A. Laufer and Y. Bar-Ness, "On Converting OSTC scheme from Non-null rate to Full-rate with better error performance," *Wireless Communication and Sensor Networks* WCSN 2008.

M. Shi, Y. Bar-Ness, A. Laufer and W. Su, "Fourth order cumulants in distinguishing single carrier from OFDM signals," *MILCOM* 2008.

A. Laufer and Y. Bar-Ness, "Improved transmission scheme for orthogonal space time codes," *Conference on Information Science and Systems* CISS 2008.

A. Laufer, A. Leshem, and H. Messer, "Game Theoretic Aspects of Distributed Spectral Coordination with Application to DSL Networks," arXiv:cs/0602014, 2005.

A. Laufer, A. Leshem, "Distributed Spectrum Coordination and the Prisoner's Dilemma," *IEEE symposium on new frontiers in dynamic spectrum access networks* DYSPAN 2005.

To my parents, Itzhak and Raya,

my children, Eli, Dassi, Yonatan, Nati, Shira and Ayala,

and above all, to my wife, Michali.

ACKNOWLEDGMENT

New Jersey Institute of Technology provided me with the opportunities to expand my horizons, and grow, both professionally and personally. I have to thank my advisor, Prof. Yeheskel Bar-Ness for his uncompromising guidance and thrive for accomplishments. I've learned a lot from him, both from his deep and wide research experience and knowledge, as well as from being a role model for an unconditionally dedicated professional.

I would like to acknowledge the members of my dissertation committee, Prof. Alexander Haimovich, Prof. Ali Abdi and Prof. Osvaldo Simeone of NJIT, and Prof. Laurence Milstein of University of California at San Diego for their time and effort. Funding of my dissertation was provided mainly by the Ross Fellowship Memorial Fund, and the Elisha Yegal Bar-Ness Endowed Fellowship. I would like to thank Dr. Marino Xanthos and Ms. Clarisa Gonzalez-Lenahan from the office of graduate studies and Mr. Jeffrey Grundy of international students office for their essential help. A special, deep hearted, gratitude is reserved for Ms. Marlene Toeroek, for her all-around, dedicated help during my stay at NJIT.

Finally, I would like to thank my parents, Itzhak and Raya, for educating me all these years, and for their unconditional support and encouragement. I thank my wife's mother, Chaya, for her share of assistance and moral support. Special thanks, from the depths of my heart, goes to Yossi and Ruchi Stern, our local adopting family, for enabling all of this. To my children, Eli, Dassi, Yonatan, Nati, Shira and Ayala, for the joy you bring to my life and for reminding me, on a daily basis, what is truly important in life.

Last, but not least, to my loving and beloved wife, Michali, because it would not have happened without her support, encouragement, patience, and love. I am forever indebted to her and I thank her for constantly supporting my endeavors, never doubting my abilities, and for being, always, my pillar, my joy and my guiding light.

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

Chapter **Page**

TABLE OF CONTENTS
(Continued)

Chapter **Page**

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

BPSK	Binary Phase-Shift Keying
EVC	Equivalent Virtual Cahnnel
ICE	Imperfect Channel Estimation
LSTC	Layered Space Time Code
MIMO	Multiple Input Multiple Output
MISO	Multiple Input Single Output
ML	Maximum Liklihood
OSTBC	Orthogonal Space Time Block Code
PEP	Pairwise Error Probability
QAM	Quadrature Amplitude Modulation
QSTBC	Quasi Orthogonal Space Time Block Code
REM	Row Elimination Method
SER	Symbol Error Rate
SIMO	Single Input Multiple Output
SNR	Signal to Noise Ratio
SSTBC	Semi Orthogonal Space Time Block Code
STBC	Space Time Block Code
STC	Space Time Code
STTC	Space Time Trellis Code
ZF	Zero Forcing

CHAPTER 1

INTRODUCTION

Optimal design and successful deployment of high-performance wireless networks, present a number of technical challenges. These include, regulatory limits on usable radio frequency spectrum, and a complex time-varying propagation environment affected by fading and multipath. In order to meet the growing demand for higher data rates at better communication reliability, boldly innovative techniques that improve both spectral efficiency and link reliability are called for. Use of multiple antennas at the receiver and transmitter in a wireless network is a rapidly emerging technology, that promises higher data rates at longer ranges without consuming extra bandwidth or transmit power. This technology, popularly known as smart antenna technology, offers a variety of leverages which, if exploited correctly, can enable multiplicative gains in network performance.

1.1 Fading Channel Model

1.1.1 Multipath Propagation

In a wireless environment, the surrounding objects, such as houses, building or trees, act as reflectors of radio waves (Figure 1.1). These obstacles produce reflected waves with attenuated amplitudes and phases. When a signal is transmitted, multiple reflected waves of the transmitted signal will arrive at the receiving antenna from different directions with different propagation delays. These reflected waves are called multipath waves [1]. Due to the different arrival angles and times, the multipath waves at the receiver site have different phases. When they are collected by the receiver antenna at any point in space, they may combine either in a constructive or a destructive way, depending on the random phases. The sum of these multipath

components forms a spatially varying standing wave field. The mobile unit moving through the multipath field will receive a signal which can vary widely in amplitude and phase. When the mobile unit is stationary, the amplitude variations in the received signal are due to the movement of surrounding objects in the radio channel. The amplitude fluctuation of the received signal is called *signal fading*. It is caused by the time-variant multipath characteristics of the channel.

Because of the multiplicity of factors involved in propagation in a wireless environment, it is convenient to apply statistical techniques to describe signal variations. In a narrowband system, the transmitted signals usually occupy a bandwidth smaller than the channels coherence bandwidth, which is defined as the frequency range over which the channel fading process is correlated. That is, all spectral components of the transmitted signal are subject to the same fading attenuation. This type of fading is referred to as frequency nonselective or *frequency flat*. On the other hand, if the transmitted signal bandwidth is greater than the channel coherence bandwidth, the spectral components of the transmitted signal with a frequency separation larger than the coherence bandwidth are faded independently. The received signal spectrum becomes distorted, since the relationships between various spectral components are not the same as in the transmitted signal. This phenomenon is known as *frequency selective* fading.

1.1.2 Rayleigh Fading

Consider the transmission of a single tone with a constant amplitude. In a typical land mobile radio channel, one may assume that the direct wave is obstructed and the mobile unit receives only reflected waves. When the number of reflected waves is large, according to the central limit theorem, two quadrature components of the received signal are uncorrelated Gaussian random processes with a zero mean and variance $\sigma^2 = 1$. As a result, the envelope of the received signal at any time instant undergoes a

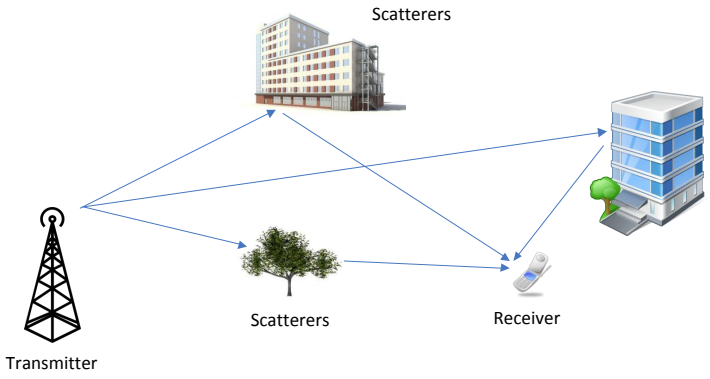

Figure 1.1 Multipath environment.

Rayleigh probability distribution and its phase obeys a uniform distribution between $-\pi$ and π. The probability density function (pdf) of the Rayleigh distribution is given by [2]

$$f_x(x) = \begin{cases} \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}} & x \geq 0 \\ 0 & x < 0 \end{cases} \tag{1.1}$$

The probability density function for a normalized Rayleigh distribution is shown in Figure 1.2.

In terms of the coefficients (amplitude and phase) variation speed, fast and slow fading channels are considered. For slow fading, it is assumed that the fading coefficients are constant during a frame and vary from one transmission frame to another, which means that the symbol period is small compared to the channel coherence time. The slow fading is also referred to as *quasi-static fading*. In a fast fading channel, the fading coefficients are constant within each symbol period and vary from one symbol to another.

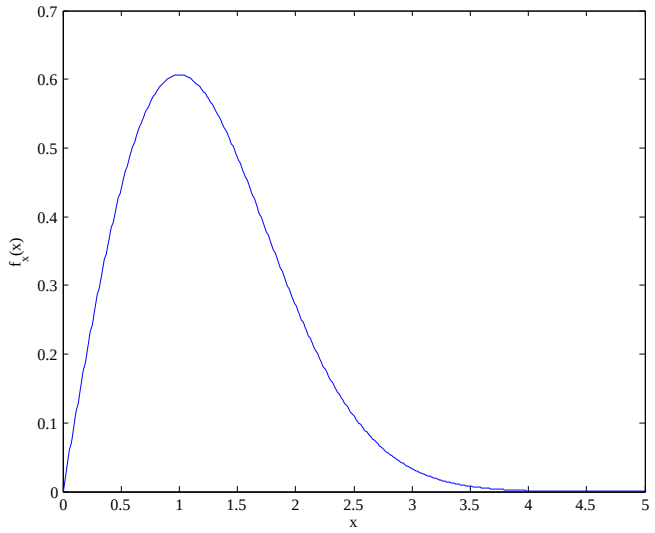

Figure 1.2 Normalized Rayleigh r.v. ($\sigma = 1$) probability density function.

1.2 Multiple Antenna Channel Model

Smart antenna technology provides a wide variety of options, ranging from single-input, multiple-output (SIMO) architectures that collect more energy to improve the signal to noise ratio (SNR) at the receiver, to multiple-input, multiple-output (MIMO) architectures that open up multiple data pipes over a link. The number of inputs and outputs here refers to the number of antennas used at the transmitter and receiver, respectively. Figure 1.3 shows a typical MIMO system with M_t transmit antennas and M_r receive antennas. The transmitter (Tx) encodes and modulates the information bits to be conveyed to the receiver and maps the signals to be transmitted across space (M_t transmit antennas) and time. The receiver (Rx) processes the signals received on each of the M_r receive antennas according to the transmitters signaling strategy and demodulates and decodes the received signal.

Different smart antenna architectures provide different benefits which can be broadly classified as array gain, diversity gain and multiplexing gain. The signaling strategy at the transmitter and the corresponding processing at the receiver are

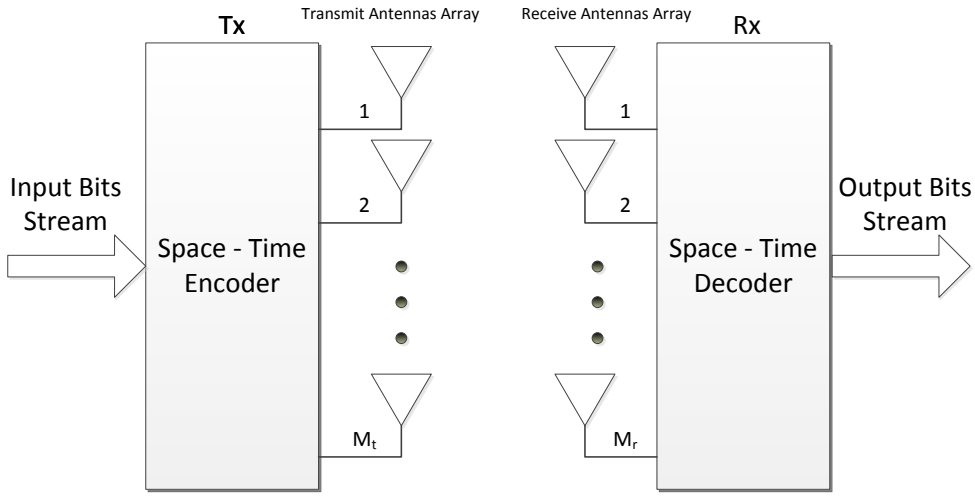

Figure 1.3 Typical Multiple-Input, Multiple-Output (MIMO) system.

designed based on link requirements (data rate, range, reliability etc.). For example, in order to increase the point to point spectral efficiency (in bits/sec/Hz) between a transmitter and receiver, multiplexing gain is required which is provided by the MIMO architecture. The signaling strategy also depends on the availability of channel information at the transmitter. For example, MIMO does not require channel knowledge at the transmitter, although it enjoys improved performance if channel information is available. Starting with a simple signal model, the basic smart antenna benefits namely array gain, diversity gain and multiplexing gain will then be discussed in greater detail.

Consider a MIMO system with M_t transmit antennas and M_r receive antennas as shown in Figure 1.3. For simplicity only flat fading is considered, i.e., the fading is not frequency selective. When a signal, s is launched from the i^{th} transmit antenna, each of the M_r receive antennas sees a complex-weighted version of the transmitted signal. The signal received at the j^{th} receive antenna is denoted by $h_{ji}s$, where h_{ji} is the channel response between the i^{th} transmit antenna and the j^{th} receive antenna. The vector $[h_{1j}\ h_{2j}\ \cdots\ h_{M_rj}]^T$ is the signature induced by the i^{th} transmit antenna

across the receive antenna array. It is convenient to denote the MIMO channel ($\mathbf{H}$) in matrix notation as shown below.

$$\mathbf{H} = \begin{pmatrix} h_{11} & h_{12} & \cdots & h_{1M_t} \\ h_{21} & h_{22} & \cdots & h_{2M_t} \\ \vdots & \vdots & \ddots & \vdots \\ h_{M_r1} & h_{M_r2} & \cdots & h_{M_rM_t} \end{pmatrix} \tag{1.2}$$

The channel matrix $\mathbf{H}$ defines the input-output relation of the MIMO system and is also known as the channel transfer function. If a signal vector $\mathbf{x} = [x_1\ x_2\ \cdots\ x_{M_t}]^T$ is launched from the transmit antenna array (x_i is launched from the i^{th} transmit antenna) then the signal received at the receive antenna array, $\mathbf{y} = [y_1\ y_2\ \cdots\ y_{M_r}]^T$ can be written as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{1.3}$$

where $\mathbf{n}$ is the $M_r \times 1$ noise vector consisting of independent complex-gaussian distributed elements with zero mean and variance σ_n^2 (white noise). Note that the above channel matrix can be interpreted as a snapshot of the wireless channel at a particular frequency and at a specific instant of time. When there is rich multipath with a large delay spread, $\mathbf{H}$ varies as a function of frequency. Likewise, when the scatterers are mobile and there is a large doppler spread, $\mathbf{H}$ varies as a function of time. With sufficient antenna separation at the transmit and receive arrays, the elements of the channel matrix $\mathbf{H}$ can be assumed to be independent, zero-mean, complex gaussian random variables (Rayleigh fading) with unit variance in sufficiently rich multipath. This model is popularly referred to as the i.i.d Gaussian MIMO channel. In general, if antennas are separated by more than half the carrier wavelength ($\frac{\lambda}{2}$) [3], the channel fades can be modeled as independent Gaussian random variables.

1.2.1 Array Gain

Consider a SIMO system with one transmit antenna and two receive antennas as shown in Figure 1.4. The two receive antennas see different versions, y_1 and y_2, of the same transmitted signal, x. The signals y_1 and y_2 have different amplitudes

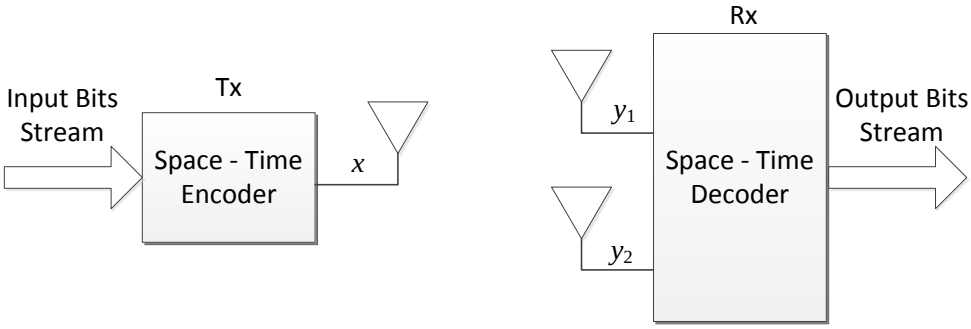

Figure 1.4 1×2 Single-Input, Multiple-Output (SIMO) system.

and phases as determined by the propagation conditions. If the channel is known to the receiver, appropriate signal processing techniques can be applied to combine the signals y_1 and y_2 coherently so that the resultant power of the signal at the receiver is enhanced, leading to an improvement in signal quality. More specifically, the SNR at the output is equal to the sum of the SNR on the individual links. This result can be extended to systems with one transmit antenna and more than two receive antennas as follows

$$\mathbf{w}^*\mathbf{y} = \mathbf{w}^*\mathbf{h}x + \mathbf{w}^*\mathbf{n} \tag{1.4}$$

where the optimal $M_r \times 1$ linear receive filter is $\mathbf{w} = \mathbf{h}$, and the maximum SNR is proportional to the channel norm $\|\mathbf{h}\|^2 = \sum_{m=1}^{M_r} |h_m|^2$, where $\|\mathbf{h}\|^2$ is the Frobenius norm. The average increase in receive signal power at the receiver is equal to $E\{\|\mathbf{h}\|^2\}$ and is defined as *array gain* and is proportional to the number of receive antennas.

Array gain can also be exploited in systems with multiple antennas at the

transmitter by using beamforming. Extracting the maximum possible array gain in such systems requires channel knowledge at the transmitter, so that the signals may be optimally processed before transmission. An example of transmit beamforming for $1 \times M_t$ MISO systems (Figure 1.5) is shown below

$$y = \mathbf{h}^*(\mathbf{w}x) + n \tag{1.5}$$

The optimal normalized $M_t \times 1$ transmit filter is $\mathbf{w} = \mathbf{h}/\|\mathbf{h}\|$. Analogous to the SIMO case, the array gain in MISO systems with channel knowledge at the transmitter is equal to $E\{\|\mathbf{h}\|^2\}$ and is proportional to the number of transmit antennas. The array gain in MIMO systems depends on the number of transmit and receive antennas and is a function of the dominant singular value of the channel.

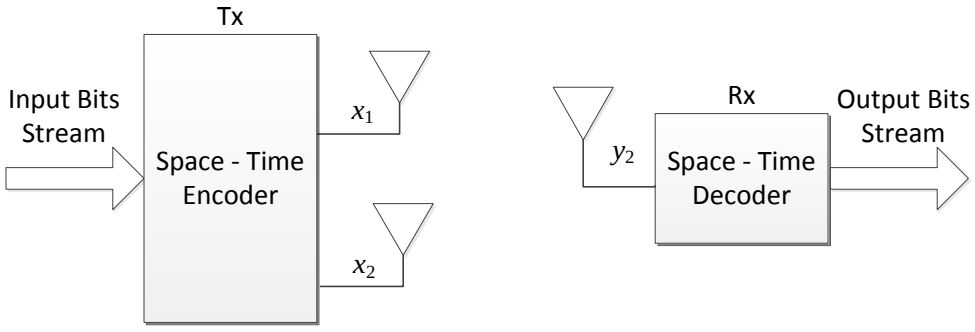

Figure 1.5 2×1 Multiple-Input, Single-Output (MISO) system.

1.2.2 Diversity Gain

Signal power in a wireless channel fluctuates (or fades) with time/frequency/space. When the signal power drops dramatically, the channel is said to be in a fade. Diversity is used in wireless systems to combat fading. The basic principle behind diversity is to provide the receiver with several looks at the transmitted signal over independently fading links (or diversity branches). As the number of diversity

branches increases, the probability that at any instant of time one or more branch is not in a fade increases. Thus diversity helps stabilize a wireless link. Diversity is available in SISO links in the form of time or frequency diversity. The use of time or frequency diversity in SISO systems often incurs a penalty in data rate due to the utilization of time or bandwidth to introduce redundancy. The introduction of multiple antennas at the transmitter and/or receiver provides spatial diversity, the use of which does not incur a penalty in data rate while providing the array gain advantage discussed earlier. There are two forms of spatial diversity receive and transmit diversity.

Receive diversity applies to systems with multiple antennas only at the receiver (SIMO systems) [4]. Figure 1.4 illustrates a system with receive diversity. Signal x is transmitted from a single antenna at the transmitter. The two receive antennas see independently faded versions, y_1 and y_2, of the transmitted signal, x. The receiver combines these signals using appropriate signal processing techniques so that the resultant signal exhibits much reduced amplitude variability (fading) as compared to either y_1 or y_2. The amplitude variability can be further reduced by adding more antennas to the receiver. The diversity in a system is characterized by the number of independently fading diversity branches, also known as the diversity order. The diversity order of the system in Figure 1.4 is two and in general is equal to the number of receive antennas, M_r , in a SIMO system.

Transmit diversity is applicable when multiple antennas are used at the transmitter and has become an active area for research in the past years [5],[6]. Extracting diversity in such systems does not necessarily require channel knowledge at the transmitter. However, suitable design of the transmitted signal is required to extract diversity. Space-time coding [7],[8] is a powerful transmit diversity technique that relies on coding across space (transmit antennas) and time to extract diversity. Figure 1.5 shows a generic transmit diversity scheme for a system with two transmit

antennas and one receive antenna. At the transmitter, signals x_1 and x_2 are derived from the original signal to be transmitted, x, such that the signal x can be recovered from either of the received signals y_{11} or y_{21}. The receiver combines the received signals in such a manner that the resultant output exhibits reduced fading when compared to y_{11} or y_{21}. The diversity order of this system is two and in general is equal to the number of transmit antennas, M_t , in a MISO system.

Utilization of diversity in MIMO systems requires a combination of receive and transmit diversity described above. A MIMO system consists of $M_t \times M_r$ SISO links. If the signals transmitted over each of these links experience independent fading, then the diversity order of the system is given by $M_t \times M_r$. Thus the diversity order in a MIMO system scales linearly with the product of the number of receive and transmit antennas. Mathematically, diversity is defined to be equal to the slope of the symbol error rate (SER) versus SNR graph. This will be shown in greater detail in the following derivation.

The vector equation in (1.3) can be written as the following matrix equation

$$\mathbf{Y} = \mathbf{HX} + \mathbf{N} \tag{1.6}$$

where the channel input $\mathbf{X}$ is an $M_t \times T$ codeword spanning T sample times, the channel output is the $M_r \times T$ matrix $\mathbf{Y}$ observed on M_r receive antennas over T sample times and the receiver noise is the $M_r \times T$ matrix $\mathbf{N}$.

Consider two $M_t \times T$ codewords $\mathbf{X}^{(i)}$ and $\mathbf{X}^{(j)}$ that are transmitted over M_t transmit antennas across T sample times. If $\mathbf{X}^{(i)}$ was transmitted, the probability that $\mathbf{X}^{(j)} \neq \mathbf{X}^{(i)}$ is detected for a given realization of the channel $\mathbf{H}$ is equal to the following

$$\begin{aligned} \mathrm{PEP}_{|\mathbf{H}} &= \mathrm{Prob}\left(\mathbf{X}^{(i)} \to \mathbf{X}^{(j)}\right) \\ &= \mathrm{Prob}\left(\|\mathbf{Y} - \mathbf{H}\mathbf{X}^{(j)}\|_F^2 \leq \|\mathbf{Y} - \mathbf{H}\mathbf{X}^{(i)}\|_F^2\right) = Q\left(\sqrt{D_{ij}\frac{\mathrm{SNR}}{2}}\right) \end{aligned} \tag{1.7}$$

where $Q(x) = \int_x^\infty \exp(-\frac{t^2}{2})dt$ is the complementary error function, $D_{ij} = \|\mathbf{H}\left(\mathbf{X}^{(i)} - \mathbf{X}^{(j)}\right)\|_F^2$ is the pairwise Euclidean distance at the receiver and $\mathrm{SNR} = \frac{E_s}{N_0}$ is the ratio of the total transmitted signal power to the noise power per receive antenna.

This conditional pairwise error probability (PEP) is a function of the channel realization. Since the transmitter does not know the channel, the best it can do is optimize a criterion that takes channel statistics into account. One popular criterion is the average PEP, i.e., the average of the conditional PEP over channel statistics. It is difficult to compute the expectation of the expression in (1.7). A simpler alternative is to compute the average of a tight upper bound, in particular the Chernoff upper bound

$$\mathrm{PEP}_{|\mathbf{H}} = Q\left(\sqrt{D_{ij}\frac{\mathrm{SNR}}{2}}\right) \leq e^{-D_{ij}\frac{\mathrm{SNR}}{4}} \tag{1.8}$$

For the i.i.d. Gaussian channel, the average Chernoff bound simplifies to the following as derived in [5],[7]

$$\mathrm{PEP} \leq \left(\frac{1}{\det\left(I_{M_t} + \frac{\mathrm{SNR}}{4}\left(\mathbf{X}^{(i)} - \mathbf{X}^{(j)}\right)\left(\mathbf{X}^{(i)} - \mathbf{X}^{(j)}\right)^H\right)}\right)^{M_r} = \left(\frac{1}{\prod_{l=1}^{L}\left(1 + \frac{\mathrm{SNR}}{4}\sigma_l^2\right)}\right)^{M_r} \tag{1.9}$$

where *det* is the determinant of a square matrix, $\{\sigma_l\}_{l=1}^{L}$ are the nonzero eigenvalues of the distance matrix $\Delta_{ij} = (\mathbf{X}^{(i)} - \mathbf{X}^{(j)})(\mathbf{X}^{(i)} - \mathbf{X}^{(j)})^H$ and L is its rank. Taking the limit at high SNR,

$$\text{PEP} \leq \left(\frac{\text{SNR}}{4}\right)^{-M_r L} \left(\left(\prod_{l=1}^{L} \sigma_l^2\right)^{\frac{1}{L}}\right)^{-M_r L} \tag{1.10}$$

and taking the logarithm of both sides, one obtain

$$\log \text{PEP} \leq -M_r L \left(\log\left(\frac{\text{SNR}}{4}\right) + \log\left(\left(\prod_{l=1}^{L} \sigma_l^2\right)^{\frac{1}{L}}\right)\right) \tag{1.11}$$

Consider the logarithm of the PEP in (1.11). The right hand side is clearly linear in the logarithms of SNR and the product of squared singular values of the difference matrix. In addition, the slope of the r.h.s. is a product of the number of receive antennas and the rank of the difference matrix. The *diversity gain* of the space-time codebook is defined to be the minimum value of L over all pairs of codewords. For a given diversity gain, the *coding gain* is defined to be the minimum of the product $\left(\prod_{l=1}^{L} \sigma_l^2\right)^{\frac{1}{L}}$ over all pairs of codewords.

Performance of space-time codes is usually illustrated by plotting the SER versus SNR on a logarithmic scale. Since the PEP is closely related to SER, (1.11) is a good approximation to SER especially at high SNRs. Figure 1.6 illustrates the effect of each code metric on the SER curve. Diversity gain affects the asymptotic slope of the SER versus SNR graph - greater the diversity, the faster the SER drops with SNR. Coding gain affects the horizontal shift of the graph - greater the coding gain, the greater the shift to the left.

1.2.3 Multiplexing Gain

The key differentiating advantage of MIMO systems is practical throughput enhancement which is not provided by SIMO or MISO systems. This leverage is referred as multiplexing gain and it can be realized through a technique known as spatial multiplexing [9]. Figure 1.7 shows the basic principle of spatial multiplexing

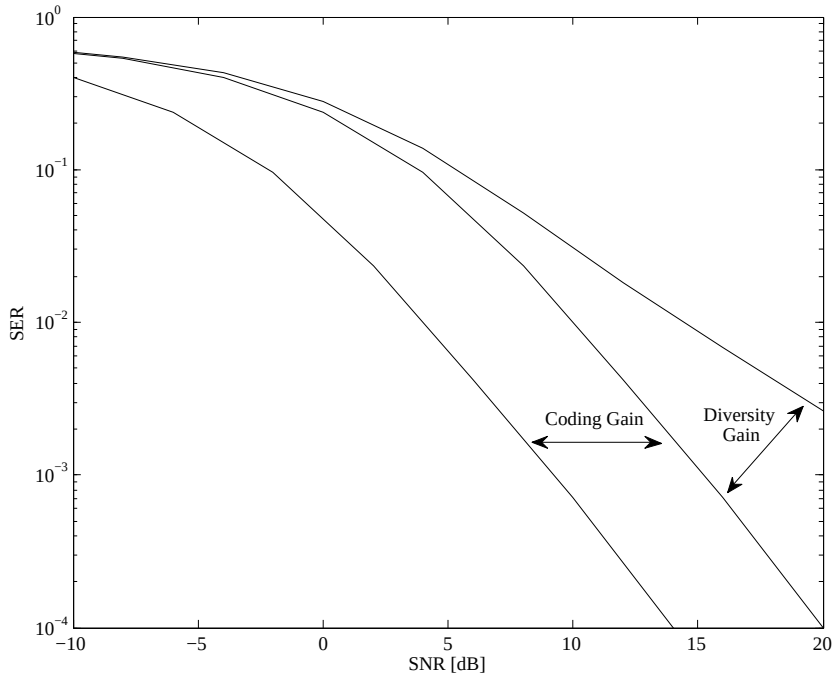

Figure 1.6 Effect of diversity gain and coding gain on the symbol error rate curve.

for a system with two transmit and two receive antennas. The symbol stream to be transmitted is split into two half-rate sub-streams and modulated to form the signals x_1 and x_2 that are transmitted simultaneously from separate antennas. Under favorable channel conditions, the spatial signatures of these signals (denoted by $[y_{11}\ y_{12}]^T$ and $[y_{21}\ y_{22}]^T$) induced at the receive antennas are well separated (ideally orthogonal). The receiver can then extract the two sub-streams, x_1 and x_2, which it combines to give the original symbol stream, x.

This can be mathematically expressed as the theoretical channel capacity as derived in [10],[11]. Channel capacity of the memoryless MIMO channel in (1.3) is defined to be the instantaneous mutual information which is a function of the channel realization as follows

$$C_{|\mathbf{H}} = \log\det\left(\mathbf{I}_{M_r} + \mathrm{SNR}\mathbf{H}\mathbf{K}_{\mathbf{X}}\mathbf{H}^*\right) \tag{1.12}$$

When the channel is square and orthogonal ($\mathbf{HH}^* = \mathbf{I}$), then with an i.i.d. input distribution ($\mathbf{K_X} = \frac{1}{M_t}\mathbf{I}_{M_t}$), (1.12) reduces to

$$C_\perp = M_t \log\left(1 + \frac{1}{M_t}\text{SNR}\right) \tag{1.13}$$

Hence, $M = M_t = M_r$ parallel channels are created within the same frequency bandwidth for no additional transmit power. Capacity scales linearly with number of antennas for increasing SNR, i.e., capacity increases by M b/s/Hz for every 3 dB increase in SNR. In general, it can be shown that an orthogonal channel of the form described above maximizes the Shannon capacity of a MIMO system. For the i.i.d fading MIMO channel model described earlier, the channel realizations become approximately orthogonal when the number of antennas used is very large. When the number of transmit and receive antennas is not equal, $M_t \neq M_r$, the increase in capacity is limited by the minimum of M_t and M_r. This increase in channel capacity is called *multiplexing gain*.

Having discussed the key advantages of smart antenna technology, it should be noted that it may not be possible to exploit all the leverages simultaneously in a smart antenna system. This is because the spatial degrees of freedom are limited and engineering tradeoffs must be made between each of the desired benefits. The optimal signaling strategy is a function of the wireless channel properties and network requirements.

1.3 Fundamentals of Space Time Codes

In the previous section, it was shown that the information capacity of wireless communication systems can be increased considerably by employing multiple transmit and receive antennas. For a system with a large number of transmit and receive antennas and an independent flat fading channel known at the receivers, the capacity

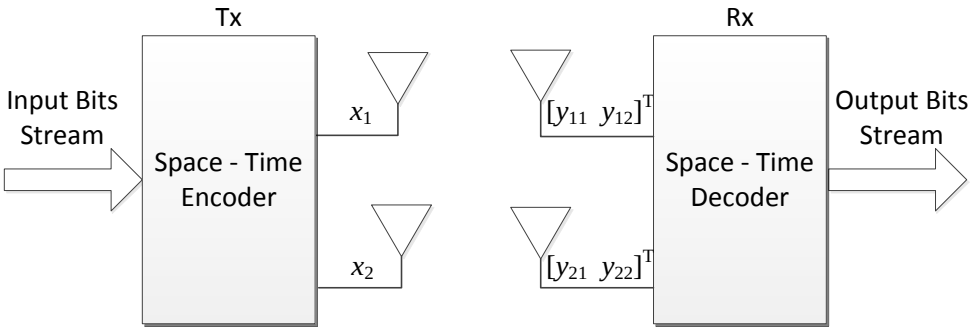

Figure 1.7 2×2 MIMO with spatial multiplexing.

grows linearly with the minimum number of antennas.

An effective and practical way to approaching the capacity of multiple-input multiple-output (MIMO) wireless channels is to employ space-time coding (STC) [7]. Space-time coding is a coding technique designed for use with multiple transmit antennas. Coding is performed in both spatial and temporal domains to introduce correlation between signals transmitted from various antennas at various time periods. The spatial-temporal correlation is used to exploit the MIMO channel fading and minimize transmission errors at the receiver. Space-time coding can achieve transmit diversity and power gain over spatially uncoded systems without sacrificing the bandwidth. There are various approaches in coding structures, including space-time block codes (STBC), space-time trellis codes (STTC), space-time turbo trellis codes and layered space-time (LST) codes. A central issue in all these schemes is the exploitation of multipath effects in order to achieve high spectral efficiencies and performance gains.

1.3.1 General Structure

Consider a baseband space-time coded communication system with M_t transmit antennas and M_r receive antennas, as shown in Figure 1.3. The transmitted data

are encoded by a space-time encoder. At each time instant t, a block of b binary information symbols, denoted by

$$\mathbf{c}_t = \left(c_t^1\ ,\ c_t^2\ ,\ \ldots\ ,\ c_t^b\right) \tag{1.14}$$

is fed into the space-time encoder. The space-time encoder initially maps the block of b binary input data into k symbols, $(s_1, s_2, \ldots, s_k)$, drawn from some signal constellation. The k data symbols are then encoded into M_t symbols $x_t^i\ ,\ 1 \leq i \leq M_t$, where

$$x_t^i = \sum_{l=1}^{k} \alpha_{li} s_l + \beta_{li} s_l^* \qquad ; \qquad \alpha_{li}, \beta_{li} \in \mathbb{Z} \tag{1.15}$$

Any STC is totaly defined by the α_{li}, β_{li}'s, which maps the k data symbols to the transmitted symbols. The encoded data is applied to a serial-to-parallel (S/P) converter producing a sequence of M_t parallel symbols, arranged into an $M_t \times 1$ column vector

$$\mathbf{x}_t = \left[x_t^1\ ,\ x_t^2\ ,\ \ldots\ ,\ x_t^{M_t}\right]^T \tag{1.16}$$

The M_t parallel outputs are simultaneously transmitted by M_t different antennas, whereby symbol $x_t^i\ ,\ 1 \leq i \leq M_t$, is transmitted by antenna i and all transmitted symbols have the same duration of T_s seconds. The vector of coded modulation symbols from different antennas, as shown in (1.16), is called a *space-time symbol*.

At the receiver, the signal at each of the M_r receive antennas is a noisy superposition of the M_t transmitted signals degraded by channel fading (see (1.3)). At time t , the received signal at antenna $j, j = 1, 2, ..., M_r$, denoted by y_t^j,is given by

$$y_t^j = \sum_{i=1}^{M_t} h_{ji} x_t^i + n_t^j \tag{1.17}$$

where n_t^j is the noise component of receive antenna j at time t , which is an independent sample of the zero-mean complex Gaussian random variable with the one sided power spectral density of N_0.

1.3.2 Design Criteria

To define the space time code design criteria, the performance analysis derived in Section 1.2.2 should be recalled. To that end, it is assumed that the transmitted data frame length is T symbols for each antenna. Defining a $T \times M_t$ *space-time codeword matrix*, obtained by arranging the transmitted sequence in an array, as

$$X = \begin{pmatrix} x_{11} & \cdots & x_{1M_t} \\ \vdots & \ddots & \vdots \\ x_{T1} & \cdots & x_{TM_t} \end{pmatrix} \tag{1.18}$$

where

$$x_{mi} = \sum_{l=1}^{k} \alpha_{lmi} s_l + \beta_{lmi} s_l^* \qquad ; \qquad \alpha_{lmi}, \beta_{lmi} \in \mathbb{Z} \tag{1.19}$$

is the symbol transmitted from the ith transmit antenna at the mth time slot. Recalling the error probability expression from (1.10)

$$\text{PEP} \leq \left(\frac{\text{SNR}}{4}\right)^{-M_r L} \left(\left(\prod_{l=1}^{L} \sigma_l^2 \right)^{\frac{1}{L}} \right)^{-M_r L} \tag{1.20}$$

it is clear that the error probability is a strong function of the codeword matrix structure. Namely, the performance is a function of L and $sigma_l^2$ which are the minimum rank and the nonzero eigenvalues respectively of the distance matrix $\Delta_{ij} = (\mathbf{X}^{(i)} - \mathbf{X}^{(j)})(\mathbf{X}^{(i)} - \mathbf{X}^{(j)})^H$ of the closest two codeword. This leads us to the following design criteria;

- Maximize the minimum rank L of matrix Δ_{ij} over all pairs of distinct codewords.

- Maximize the minimum product, $\prod_{l=1}^{L} \sigma_l^2$, of the matrix Δ_{ij} along the pairs of distinct codewords with the minimum rank.

Note that $\prod_{l=1}^{L} \sigma_l^2$ is the absolute value of the sum of determinants of all the principal $L \times L$ cofactors of matrix Δ_{ij} [7]. This criteria set is referred to as *rank and determinant criteria*. The minimum rank L of matrix Δ_{ij} over all pairs of distinct codewords is called the minimum rank of the space-time code. To maximize the minimum rank L, means to find a space-time code with the full rank of matrix Δ_{ij}, e.g., $L = M_t$. However, the full rank is not always achievable due to the restriction of the code structure.

1.4 Space Time Block Codes

After discussing the basics of space-time codes in general, the space-time block codes (STBC) are presented, which are in the main focus of this work. The general structure of STBC will be presented followed by the introduction of the the Alamouti code, which is a simple two-branch transmit diversity scheme. The key feature of the scheme is that it achieves a full diversity gain with a simple maximum-likelihood decoding algorithm.

1.4.1 General Structure

Figure 1.8 shows an encoder structure for space-time block codes. In general, a space-time block code is defined by an $T \times M_t$ transmission matrix $\mathbf{X}$. Here M_t represents the number of transmit antennas and T represents the number of time periods for transmission of one block of coded symbols.

Assuming that the signal constellation consists of 2^b points. At each encoding operation, a block of kb information bits are mapped into the signal constellation to select k modulated signals $s_1, s_2, \ldots, s_k$, where each group of b bits selects a constellation signal. The k modulated signals are encoded by a space-time block

encoder to generate M_t parallel signal sequences of length T according to the transmission matrix X. These sequences are transmitted through M_t transmit antennas simultaneously in T time periods. In the space-time block code, the number of symbols the encoder takes as its input in each encoding operation is k. The number of transmission periods required to transmit the space-time coded symbols through the multiple transmit antennas is T. In other words, there are T space-time symbols transmitted from each antennas for each block of k input symbols. The rate of a space-time block code is defined as the ratio between the number of symbols the encoder takes as its input and the number of space-time coded symbols transmitted from each antenna. It is given by

$$R = k/T \tag{1.21}$$

The entries of the transmission matrix $\mathbf{X}$ are linear combinations of the k modulated symbols s_1 , s_2, ... , s_k and their conjugates s_1^* , s_2^*, ... , s_k^*. The ith row of $\mathbf{X}$ represents the symbols transmitted simultaneously through M_t transmit antennas at time i, while the jth column of $\mathbf{X}$ represents the symbols transmitted from the jth transmit antenna consecutively in T transmission periods. The jth column of $\mathbf{X}$ is regarded as a space-time symbol transmitted at time j. The element of $\mathbf{X}$ in the ith row and jth column, x_{ij} , $i = 1, 2, \ldots, M_t$, $j = 1, 2, \ldots, T$, represents the signal transmitted from the ith antenna at time j. The structure and properties of $\mathbf{X}$ define the STBC as will be in explained in the following sections.

1.4.2 The Alamouti 2×2 Code

The Alamouti scheme is historically the first space-time block code to provide full transmit diversity for systems with two transmit antennas [12]. It is worthwhile to mention that delay diversity schemes [13] can also achieve a full diversity, but they introduce interference between symbols and complex detectors are required at the

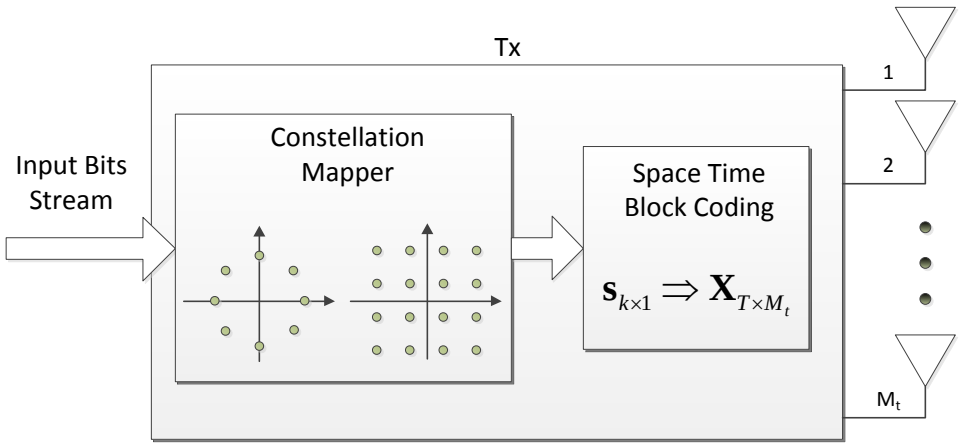

Figure 1.8 General structure of space time block coding encoder.

receiver. In this section, the Alamoutis transmit diversity technique is introduced, including encoding and decoding algorithms and its performance.

Figure 1.9 shows the block diagram of the Alamouti space-time encoder. Assuming that an M-ary modulation scheme is used. In the Alamouti space-time encoder, each group of b information bits is first modulated, where $b = \log_2 M$. Then, the encoder takes a block of two modulated symbols s_1 and s_2 in each encoding operation and maps them to the transmit antennas according to a code matrix given by

$$\mathbf{X} = \begin{pmatrix} s_1 & s_2 \\ -s_2^* & s_1^* \end{pmatrix} \tag{1.22}$$

The encoder outputs are transmitted in two consecutive transmission periods from two transmit antennas. During the first transmission period, two signals s_1 and s_2 are transmitted simultaneously from antenna one and antenna two, respectively. In the second transmission period, signal $-s_2^*$ is transmitted from transmit antenna one and signal s_1^* from transmit antenna two, where s_1^* is the complex conjugate of s_1. It is clear that the encoding is done in both the space and time domains. Let us denote

the transmit sequence from antennas one and two by $\mathbf{x}^1$ and $\mathbf{x}^2$, respectively,

$$\begin{aligned} \mathbf{x}^1 &= [s_1 \ -s_2^*] \\ \mathbf{x}^2 &= [s_2 \quad s_1^*] \end{aligned} \tag{1.23}$$

The key feature of the Alamouti scheme is that the transmit sequences from the two

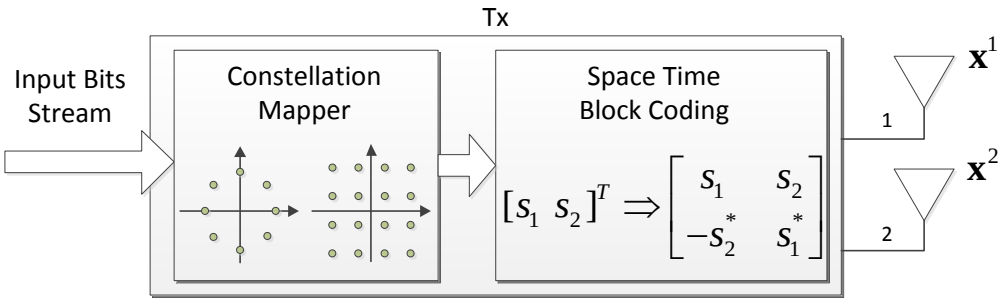

Figure 1.9 Alamouti's 2×2 space time block code encoder.

transmit antennas are orthogonal, since the inner product of the sequences $\mathbf{x}^1$ and $\mathbf{x}^2$ is zero, i.e.,

$$\langle \mathbf{x}^1, \mathbf{x}^2 \rangle = \mathbf{x}^1 (\mathbf{x}^2)^H = s_1 s_2^* - s_2^* s_1 = 0 \tag{1.24}$$

The code matrix has the following property

$$\mathbf{X^H} \cdot \mathbf{X} = \begin{pmatrix} |s_1|^2 + |s_2|^2 & 0 \\ 0 & |s_1|^2 + |s_2|^2 \end{pmatrix} = |s_1|^2 + |s_2|^2 I_2 \tag{1.25}$$

where I_2 is the 2×2 identity matrix.

Assuming that one receive antenna is used at the receiver. The fading channel coefficients from the first and second transmit antennas to the receive antenna at time t are denoted by $h_1(t)$ and $h_2(t)$, respectively. Assuming that the fading coefficients are constant across two consecutive symbol transmission periods, the received signals

over two consecutive symbol periods, can be expressed as

$$\begin{aligned} y_1 &= h_1 s_1 + h_2 s_2 + n_1 \\ y_2 &= h_1 s_2^* + h_2 s_1^* + n_2 \end{aligned} \tag{1.26}$$

where n_1 and n_2 are independent complex variables with zero mean and power spectral density $N_0/2$ per dimension, representing the additive white Gaussian noise samples.

If the channel fading coefficients, h_1 and h_2, can be perfectly recovered at the receiver, the decoder will use them as the channel state information (CSI). Assuming that all the signals in the modulation constellation are equiprobable, a maximum likelihood decoder chooses a pair of signals $(\hat{s}_1, \hat{s}_2)$ from the signal modulation constellation to minimize the distance metric

$$\begin{aligned} &d^2(y_1, h_1\hat{s}_1 + h_2\hat{s}_2) + d^2(y_2, -h_1\hat{s}_2^* + h_2\hat{s}_1^*) \\ &\quad = |y_1 - h_1\hat{s}_1 - h_2\hat{s}_2|^2 + |y_2 + h_1\hat{s}_2^* - h_2\hat{s}_1^*|^2 \end{aligned} \tag{1.27}$$

over all possible values of $\hat{s}_1$ and $\hat{s}_2$. Substituting (1.26) into (1.27), the maximum likelihood decoding can be represented as

$$(\hat{s}_1, \hat{s}_2) = \arg\min_{(\hat{s}_1, \hat{s}_2) \in \mathcal{C}} (|h_1|^2 + |h_2|^2 - 1)(|\hat{s}_1|^2 + |\hat{s}_2|^2) + d^2(r_1, \hat{s}_1) + d^2(r_2, \hat{s}_2) \tag{1.28}$$

where $\mathcal{C}$ is the set of all possible modulated symbol pairs $(\hat{s}_1, \hat{s}_2)$, r_1 and r_2 are two decision statistics constructed by combining the received signals with channel state information. The decision statistics are given by

$$\begin{aligned} r_1 &= h_1^* y_1 + h_2 y_2^* \\ r_2 &= h_2^* y_1 - h_1 y_2^* \end{aligned} \tag{1.29}$$

Substituting y_1 and y_2 from (1.26), into (1.29), the decision statistics can be written as,

$$\begin{aligned} r_1 &= (|h_1|^2 + |h_2|^2)s_1 + h_1^* n_1 + h_2 n_2^* \\ r_2 &= (|h_1|^2 + |h_2|^2)s_2 + h_2^* n_1 - h_1 n_2^* \end{aligned} \tag{1.30}$$

For a given channel realization h_1 and h_2, the decision statistics r_1 , $i = 1, 2$, is only a function of s_i , $i = 1, 2$. Thus, the maximum likelihood decoding rule (1.28) can be separated into two independent decoding rules for s_1 and s_2, given by

$$\hat{s}_1 = \arg\min_{\hat{s}_1 \in \mathcal{S}} (|h_1|^2 + |h_2|^2 - 1)|s_1|^2 + d^2(r_1, \hat{s}_1) \tag{1.31}$$

and

$$\hat{s}_2 = \arg\min_{\hat{s}_2 \in \mathcal{S}} (|h_1|^2 + |h_2|^2 - 1)|s_2|^2 + d^2(r_2, \hat{s}_2) \tag{1.32}$$

respectively. This important attribute will be referred later as *symbol by symbol* decoding. Decoding in a symbol by symbol fashion greatly reduces the complexity of the de-mapping at the decoder since it allows a linear computational complexity in the size of the signal constellation.

It will now be shown that due to the orthogonality between the sequences coming from the two transmit antennas, the Alamouti scheme can achieve the full transmit diversity of $M_t = 2$. Consider any two distinct codewords $\mathbf{X}$ and $\hat{\mathbf{X}}$ generated by the inputs (s_1, s_2) and $(\hat{s}_1, \hat{s}_2)$, respectively, where $(s_1, s_2) \neq (\hat{s}_1, \hat{s}_2)$. The codeword difference matrix is given by

$$\mathbf{X} - \hat{\mathbf{X}} = \begin{pmatrix} s_1 - \hat{s}_1 & s_2 - \hat{s}_2 \\ -s_2^* - (-\hat{s}_2^*) & s_1^* - \hat{s}_1^* \end{pmatrix} \tag{1.33}$$

Since the rows of the code matrix are orthogonal, the rows of the codeword difference matrix are orthogonal as well. The codeword distance matrix is given by

$$\Delta = (\mathbf{X}-\hat{\mathbf{X}})(\mathbf{X}-\hat{\mathbf{X}})^H = \begin{pmatrix} |s_1-\hat{s}_1|^2+|s_2-\hat{s}_2|^2 & 0 \\ 0 & |s_1-\hat{s}_1|^2+|s_2-\hat{s}_2|^2 \end{pmatrix} \quad (1.34)$$

Since $(s_1, s_2) \neq (\hat{s}_1, \hat{s}_2)$, it is clear that the distance matrices of any two distinct codewords have a full rank of two. In other words, the Alamouti scheme can achieve a full transmit diversity of $M_t = 2$. The determinant of the distance matrix Δ is given by

$$\det(\Delta) = \left(|s_1-\hat{s}_1|^2+|s_2-\hat{s}_2|^2\right)^2 \quad (1.35)$$

It is obvious from (1.34) that for the Alamouti scheme, the codeword distance matrix has two identical eigenvalues. The minimum eigenvalue is equal to the minimum squared Euclidean distance in the signal constellation. This means that for the Alamouti scheme, the minimum distance between any two transmitted code sequences remains the same as in the uncoded system. Therefore, the Alamouti scheme does not provide any coding gain relative to the uncoded modulation scheme.

1.4.3 Orthogonal STBC

The Alamouti scheme achieves the full diversity with a very simple maximum-likelihood decoding algorithm. The key feature of the scheme is orthogonality between the sequences generated by the two transmit antennas. This scheme was generalized to an arbitrary number of transmit antennas by applying the theory of orthogonal designs. The generalized schemes are referred to as *orthogonal space-time block codes* (OSTBC) [8]. The OSTBC can achieve the full transmit diversity specified by the number of the transmit antennas M_t, while allowing a very simple maximum-likelihood decoding algorithm, based only on linear processing of the received signals.

In order to achieve the full transmit diversity of M_t , the transmission matrix $\mathbf{X}$ is constructed based on orthogonal designs such that

$$\mathbf{X}^H\mathbf{X} = c(|s_1|^2 + |s_2|^2 + \cdots + |s_k|^2)\mathbf{I}_{M_t} \tag{1.36}$$

where c is a constant, $\mathbf{X}^H$ is the Hermitian of $\mathbf{X}$ and $\mathbf{I}_{M_t}$ is an $M_t \times M_t$ identity matrix. A more general definition of OSTBC is that the product $\mathbf{X}^H\mathbf{X}$ satisfies

$$\mathbf{X}^H\mathbf{X} = \mathbf{D} \tag{1.37}$$

where $\mathbf{D}$ is a diagonal matrix. The diagonal entries may not be identical, resulting in different error rates for different symbols. The codeword $\mathbf{X}$ may be square or none-square, based on the number of transmitting antennas (i.e., the codeword number of columns).

Note that when orthogonal designs are applied to construct space-time block codes, the rows of the transmission matrix X are orthogonal to each other. This means that in each block, the signal sequences from any two transmit antennas are orthogonal. For example, assuming that $\mathbf{x}_i = (x_{i1}, x_{i2}, \ldots, x_{iT})$ is the transmitted sequence from the ith antenna, $i = 1, 2, \ldots, M_T$, one have

$$\langle \mathbf{x}_i, \mathbf{x}_j \rangle = \mathbf{x}_i\mathbf{x}_j^H = \sum_{t=1}^{T} x_{it}x_{jt}^* = 0 \ , \ i \neq j \ , \ i, j \in \{1, 2, \ldots, M_t\} \tag{1.38}$$

where $\langle \cdot \rangle$ denotes the inner product. The orthogonality enables to achieve the full transmit diversity for a given number of transmit antennas. In addition, it allows the receiver to decouple the signals transmitted from different antennas and consequently, a simple maximum likelihood decoding, based only on linear processing of the received signals.

The decoding algorithm for OSTBC is similar to the decoding scheme of the Alamouti's code. For simplicity, consider a system with only one receiving antenna

and M_t transmitting antennas such that the output vector is given by

$$\mathbf{y}_{T\times 1} = \mathbf{X}_{T\times M_t}\mathbf{h}_{M_t\times 1} + \mathbf{n}_{T\times 1} \tag{1.39}$$

Initially, a linear combination of the received signals is being performed to produce a set of k samples. Each Sample r_i , $1 \leq i \leq k$ is a linear combination of the received signals y_j , $1 \leq j \leq T$ and the channel coefficients h_l , $1 \leq l \leq M_t$. This step is to reverse the space-time encoding. While in the encoding the k data symbols are mapped into the transmitted codeword $\mathbf{X}$, here this mapping rule is used in reverse to extract k samples out of the received vector $\mathbf{y}$. Each of the k samples r_i is a noisy version of its corresponded data symbol s_i. It is important to note that due to the orthogonality of the codeword, the noise terms of the combined samples r_i are uncorrelated. Thus, enables an efficient implementation of the ML decoder to recover the data.

The symbol transmission rate is given by the ratio of the number of transmitted data symbols k and the number of time slots needed for the transmission T or the codeword rows dimension. The maximal rate that can be achieved by OSTBC with complex data is given by the following theorem [14]

Theorem 1.4.1 *For any given of transmit antenna $M_t = 2m-1$ and $M_t = 2m$ with $m \in \mathcal{N}$ and $m \neq 0$, the rate of complex OSTBC is given by*

$$R \leq \frac{m+1}{2m} \tag{1.40}$$

In [15], the authors provided a closed form design method for complex OSTBC with rate $R = \frac{m+1}{2m}$ for $2m-1$ or $2m$ antennas. Table 1.1 gives the achievable rates for the square orthogonal STC, for any number of transmit antennas. These designs are referred to as rectangular designs as they impose large delay on the system. For example, for 8 transmit antennas, rate $\frac{5}{8}$ code was proposed with delay of 56 time slots. The rate converges to $\frac{1}{2}$ as the number of transmitting antennas M_t increases.

Table 1.1 Symbols per Channel Use of Complex Orthogonal Designs

N (**number of transmit antennas**)	**code rate** $\frac{symbols}{channeluse}$
2	1
3,4	3/4
5,6	2/3
7,8	5/8
9,10	3/5
11,12	7/12
13,14	4/7
15,16	9/16
$(2^{k-1}+1), 2^k$	$(k+1)/2k$

This is the main drawback of OSTBC which led to the design of other STC schemes who enjoy full rate transmission even for large number of transmit antennas.

In Chapter 2, where this work's main contribution of *semi-orthogonal* STBC is introduces, several examples of OSTC for various number of transmitting antennas will be presented. To summarize the properties of OSTBC -

- Full diversity.
- Simple, symbol by symbol, linear ML decoder implementation.
- Can't achieve full rate for $M_t > 2$.

1.4.4 Quasi Orthogonal STC

In [16],[17],[18],[19] so called Quasi Orthogonal Space-Time Block Codes (QSTBC) have been introduced as a new family of STBCs. These codes achieve full data rate at the expense of a slightly reduced diversity. In the proposed quasi-orthogonal

code designs, the columns of the transmission matrix are divided into groups. While the columns within each group are not orthogonal to each other, different groups are orthogonal to each other. Using quasi-orthogonal design, pairs of transmitted symbols can be decoded independently and the loss of diversity in QSTBC is due to some coupling term between the estimated symbols.

In general, there is no formal definition for the QSTBC family, but its main property can be written as

$$\mathbf{X}^H\mathbf{X} = c(|s_1|^2 + |s_2|^2 + \cdots + |s_k|^2)\mathbf{I}_{M_t} + Q \tag{1.41}$$

where the first term on the left hand side is similar to the one in OSTBC and the second term, $\mathbf{Q}$, is a sparse matrix with only off-diagonal elements. In addition, these elements are assume to be much smaller than the elements on the diagonal.

Since the transmitted symbols are coupled in pairs, the ML decoding scheme is more computationally demanding because it involves minimization on a vector space. Symbol by symbol decoding scheme can be implemented by using the zero forcing (ZF) decoder. This decoder decoupled the symbols by forcing zero interference between symbols. Inevitably, this comes with the penalty of noise enhancement hence, it is suboptimal decoding scheme. In addition, the calculation of the ZF filter usually involves a matrix inversion which adds substantial amount of calculations to the decoding process. Sphere decoding [20] was also proposed as a way to reduce the decoding complexity of the ML decoder. Although its attractive theoretical benefits, there still exists the issues of the choice of the appropriate covering radius of the lattice and the determination of the incremental/decremental radius values, and the complexity can be still quite high if the number of transmit antennas is large. It is widely known that the sphere decoding has intrinsically variable complexity and throughput [21] and this makes it not very suitable for hardware implementations.

In chapter 3, where the effect of imperfect channel estimation on the

performance of STBC is discussed, a detailed example of QSTBC for 4 transmission antennas will be presented. To summarize the properties of QSTBC -

- Full Rate.
- Complex pairwise ML decoding or suboptimal, symbol by symbol, ZF decoding.
- Doesn't achieve full diversity.

1.4.5 Equivalent Virtual Channel

The channel output of an STBC system is given by

$$\mathbf{y} = \mathbf{X}\mathbf{h} + \mathbf{n} \tag{1.42}$$

Since this is not a linear system model (in the transmitted data signals s_is), it is of interest to transform the representation of the channel output to comply with the regular linear system. To that end, the *equivalent virtual channel* (EVC) is introduced. Any STBC can be written as follows

$$\tilde{\mathbf{y}} = \mathbf{H}\mathbf{s} + \tilde{\mathbf{n}} \tag{1.43}$$

where $\mathbf{s}$ is the data symbols vector and $\mathbf{H}$ is the equivalent virtual channel matrix. In this form, the MISO channel output of Equation (1.42 is represented as the output of a virtual MIMO channel, given by $\mathbf{H}$. The vectors $\tilde{\mathbf{y}}$ and $\tilde{\mathbf{n}}$ are equal to the vectors $\mathbf{y}$ and $\mathbf{n}$ respectively, up to the conjugation of some of the vector's entries.

To illustrate the EVC model, consider the $2times2$ Alamouti code described in Section 1.4.2. The channel output for this code is given in Equation (1.26),

$$\begin{aligned} y_1 &= h_1 s_1 + h_2 s_2 + n_1 \\ y_2 &= h_1 s_2^* + h_2 s_1^* + n_2 \end{aligned} \tag{1.44}$$

Taking the conjugate for the second vector entry, y_2, one obtain

$$\begin{aligned} y_1 &= h_1 s_1 + h_2 s_2 + n_1 \\ y_2^* &= h_1^* s_2 + h_2^* s_1 + n_2^* \end{aligned} \tag{1.45}$$

or after rearrangement in a vector form

$$\begin{pmatrix} y_1 \\ y_2^* \end{pmatrix} = \begin{pmatrix} h_1 & h_2 \\ h_2^* & -h_1^* \end{pmatrix} \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} + \begin{pmatrix} n_1 \\ n_2^* \end{pmatrix} \tag{1.46}$$

which comply with the structure given in Equation (1.43). The matrix $\mathbf{H}$, given by

$$\mathbf{H} = \begin{pmatrix} h_1 & h_2 \\ h_2^* & -h_1^* \end{pmatrix} \tag{1.47}$$

is the EVC matrix for the Alamouti code.

It is obvious that, the matrix $\mathbf{H}$ depends on the structure of the code and the channel coefficients. The concept of the EVC simplifies the analysis of the STBC transmission scheme, and will be used extensively throughout this work.

1.5 Other STCs

Space-time block codes can achieve a maximum possible diversity advantage with a simple decoding algorithm. It is very attractive because of its simplicity. However, no coding gain can be provided by space-time block codes. In addition, the STBC is not aiming on providing multiplexing gain as there is no rate increase beyond 1. There are many more types of STC designed to maximize different attribute (rate, diversity, decoding complexity, etc.) In this section, two additional STC families are presented, namely, the space time trellis codes and the layered STC.

1.5.1 Space Time Trellis Codes

Space time trellis codes (STTC) developed as a joint design of error control coding, modulation, transmit and receive diversity, which is able to combat the effects of fading. STTC was first introduced in [7]. It was widely discussed and explored in the literature as STTC can simultaneously offer a substantial coding gain, spectral efficiency, and diversity improvement on flat fading channels. For space-time trellis codes, the encoder maps binary data to modulation symbols, where the mapping function is described by a trellis (or convolutional) diagram. The redundancy over space and time induced by the trellis code is used at the receiver end to reconstruct the transmitted data.

STTC are able to provide both coding gain and diversity gain and have a better bit-error rate performance. However, being based on trellis codes, they are more complex to encode and decode, where the decoding scheme relies on a Viterbi decoder as the implementation of the ML decoder. Space-time trellis codes have a potential drawback that the maximum likelihood decoder complexity grows exponentially with the number of bits per symbol, thus limiting achievable data rates.

1.5.2 Layered Space Time Codes

The layered space-time (LST) architecture, proposed in [9], can attain a tight lower bound on the MIMO channel capacity. The distinguishing feature of this architecture is that it allows processing of multidimensional signals in the space domain by 1-D processing steps, where 1-D refers to one dimension in space. The method relies on powerful signal processing techniques at the receiver and conventional 1-D channel codes. In the originally proposed architecture, M_t information streams are transmitted simultaneously, in the same frequency band, using M_t transmit antennas. The receiver uses $M_r = M_t$ antennas to separate and detect the M_t transmitted signals. The separation process involves a combination of interference suppression

and interference cancelation. The separated signals are then decoded by using conventional decoding algorithms developed for (1-D)-component codes, leading to much lower complexity compared to maximum likelihood decoding. Though in the original proposal the number of receive antennas, denoted by M_r, is required to be equal or greater than the number of transmit antennas, the use of more advanced detection/decoding techniques enables this requirement to be relaxed to $M_r \geq 1$.

The original LSTC receiver [22] is based on a combination of interference suppression and cancelation. Conceptually, each transmitted sub-stream is considered in turn to be the desired symbol and the remainder are treated as interferers. These interferers are suppressed by a zero forcing (ZF) approach. This detection algorithm produces a ZF based decision statistics for a desired sub-stream from the received signal vector $\mathbf{r}$, which contains a residual interference from other transmitted sub-streams. Subsequently, a decision on the desired sub-stream is made from the decision statistics and its interference contribution is regenerated and subtracted out from the received vector $\mathbf{r}$. Thus $\mathbf{r}$ contains a lower level of interference and this will increase the probability of correct detection of other sub-streams. The ZF strategy is only possible if the number of receive antennas is at least as large as the number of transmit antennas. Other decoding techniques for the LST codes are use the concept of QR matrix decomposition. This process transforms the channel matrix to an equivalent upper triangular matrix, which in turn used for the interference suppression and cancelation process. The complexity of the LSTC receivers grows linearly with the data rate.

1.6 Outline of the Dissertation

In this dissertation two aspect of STBC are considered. The first is an enhance method for transmitting OSTBC based STBC that achieves higher rate. This topic is covered in Chapter 2 and consist of a method for transforming an OSTBC codeword

to a Semi-Orthogonal STBC (SSTBC) by removing a portion of the codeword. A decoding scheme that preserve the simple, linear computational complexity is also introduced and analyzed. Further improvements for the proposed decoding scheme are also presented which make the SSTBC a powerful coding technique that achieves full rate, maintains the highly desired linear computational complexity decoding while enjoying a very good performance (error rate).

The results of this chapter are contained in

- A. Laufer and Y. Bar-Ness, "Improved transmission scheme for orthogonal space time codes," *Conference on Information Science and Systems* CISS 2008.
- A. Jain, A. Laufer and Y. Bar-Ness, "On Converting OSTC scheme from Non-null rate to Full-rate with better error performance," *Wireless Communication and Sensor Networks* WCSN 2008.
- A. Laufer and Y. Bar-Ness, "Full Rate Space Time Codes for Large Number of Transmitting Antennas with Linear Complexity Decoding and High Performance," *IEEE Information Theory Workshop* ITW 2009.
- A. Laufer and Y. Bar-Ness, "Linear Computational Complexity Decoding for Semi Orthogonal Full Rate Space Time Codes," *IEEE Wireless Communications and Networking Conference* WCNC 2011.
- A. Laufer and Y. Bar-Ness, "Full Rate Space Time Codes for Large Number of Transmitting Antennas with Linear Complexity Decoding," *Wireless Personal Communications*, Springer, Vol. 57, pp. 465-480, Apr. 2011.

In the second part of this work, the problem of imperfect channel estimation is addressed. This topic is covered in Chapter 3 and include an analysis of the effect of STBC decoding using a mismatched filter due to the use of erroneous channel coefficients. It is shown that this type of decoding results in the introduction of

coupling or symbols interference at the output of the decoding filter. The bootstrap algorithm is adopted as an adaptive method for suppressing the interference levels. This method is thoroughly analyzed using the Alamouti 2×2 codeword as a case study. An alternate method for the bootstrap's weights calculation is emerged from this analysis, which employs the use of orthogonal sequences. The concept of weights calculation via orthogonal sequences is further expanded to general STBC to overcome some major practical issues regarding the implementation of the bootstrap decoding. To mitigate the rate loss due to the use of orthogonal data sequences, a novel method for extracting orthogonal segments out of the data is introduced. The results of this chapter have been extended from

- A. Laufer and Y. Bar-Ness, "Adaptive Decoding for Space Time Codes with Imperfect Channel Estimation, Using the Bootstrap Algorithm," *IEEE Broadband Wireless Access (BWA) Workshop*, GLOBECOM 2010.

- A. Laufer and Y. Bar-Ness, "Improved Bootstrap Decoding Scheme for Space Time Codes with Imperfect Channel Estimation," *Conference on Information Science and Systems* CISS 2011.

- A. Laufer and Y. Bar-Ness, "Bootstrap Decoding for the Alamouti Space-Time Scheme with Imperfect Channel Estimation," *Wireless and Optical Communication Conference* WOCC 2011.

Finally, the contributions and results of this work are summarized in Chapter 4.

Part I

High Rate Space Time Block Codes with Linear Decoding Complexity

Preface

STC designs that can achieve full/high transmit diversity and high rate, but requiring only moderate fixed decoding complexity are highly desirable for practical applications. The term fixed decoding complexity is viewed as linear complexity in the scope of this work. The goal is to have a linear decoding complexity in both number of transmit antennas and the constellation size. There are three factors that contribute to the total decoding complexity, the filter calculations, filtering the received samples and de-mapping the filtered vector back to the constellation points. The linearity in the number of constellation size (i.e., the cardinality of the modulation) of the last factor is guaranteed by having a symbol by symbol de-mapping. This ensures that finding the nearest constellation point can be found with a search size equal to the constellation size for each transmitted symbol. The filtering linearity is defined as processing each data symbol with a linear combination of the received samples. In addition, it is much desired that the overhead for calculating the filter used in the decoding process will also be of linear order, i.e., doesn't require matrix inversion, etc.

In this part of the work, a new approaches is considered that achieve full rate along with the constraint for linear complexity decoding. This approach is a modification of the OSTBC, which involves a method for transmitting a full rate STBC generated from an OSTBC combined with a novel decoding scheme which comply with the linear complexity decoding restriction.

CHAPTER 2

SEMI ORTHOGONAL SPACE TIME CODES

In this chapter, OSTBC are used as the basis of a new transmission and decoding schemes, which will be shown to achieve the desired code properties. Since OSTBC come with inherent simple, symbol by symbol decoding, the modifications which are needed include the rate increase while maintaining the simple decoding. In a nutshell, the idea is to take advantage of the OSTBC extra redundancy, and transmit only part of the codeword. Obviously, by doing that, the regular OSTBC decoding scheme cannot be used directly due to the missing, non-transmitted, part of the codeword. To that end, a modified decoding scheme is proposed. The proposed decoding scheme comprises of two steps. At the first step, the missing part of the codeword is estimated from the transmitted part. At the second step, the received vector, which corresponds to the transmitted part of the codeword, combined with the estimated part are jointly processed through a regular OSTBC decoder. Given that the estimation at the first step of the decoder can be performed in linear computational complexity, the overall complexity order of the decoder is linear. Thus, a rate one code with linear decoding complexity is achieved.

2.1 Transmission Method

The core idea for achieving higher transmission rate, is to have a smaller transmitted codeword [23]. That is, by omitting row/s from the codeword, the total number of channel uses is reduced. The new overall transmission time is smaller $T_{new} < T$, hance, the achieved rate is higher,

$$R_{new} = \frac{k}{T_{new}} > \frac{k}{T} = R \tag{2.1}$$

This technique is referred as the *Row Elimination Method* (REM).

2.1.1 Row Elimination Method

The row elimination method, that is used to increase the rate of any OSTBC, can simply be explained with an example. Consider the following 4 Tx OSTBC codeword [24],

$$\mathbf{X}_{4,3/4} = \begin{pmatrix} s_1 & s_2 & s_3 & 0 \\ -s_2^* & s_1^* & 0 & s_3 \\ -s_3^* & 0 & s_1^* & -s_2 \\ 0 & -s_3^* & s_2^* & s_1 \end{pmatrix} \tag{2.2}$$

where the first index of $\mathbf{X}$ represents the number of transmitting antennas, and the second stands for the code's rate. Since there are three different data symbols ($k = 3$) transmitted over four time slot (the codeword rows dimension), the rate of this code is 3/4. Note that the three data symbols appear in each row of the codeword. If one of the rows will be omitted, there will still be three channel outputs, and the three data symbols could be recovered from the received vector. In other word, from mathematic point of view, when using the original codeword given in Equation (2.2), the decoding process is equivalent to solving an equations system with three variables that need to be found using four equations. An equation system like this can be generally solved using only three equations, which leads to the possibility of omitting one of the rows in the codeword.

Assuming, without lost of generality, that the the last row is deleted. This results in the following rate one code

$$\tilde{\mathbf{X}}_{4,1} = \begin{pmatrix} s_1 & s_2 & s_3 & 0 \\ -s_2^* & s_1^* & 0 & s_3 \\ -s_3^* & 0 & s_1^* & -s_2 \end{pmatrix} \tag{2.3}$$

This type of codeword will be referred as *Semi-Orthogonal* STBC (SSTBC), to reflect the fact that they were generated from an OSTBC. Obviously, this codes cannot be categorized as orthogonal since

$$\tilde{\mathbf{X}}_{4,1}^{H}\tilde{\mathbf{X}}_{4,1} \neq \mathbf{D} \tag{2.4}$$

with $\mathbf{D}$ being diagonal matrix. Hence, the decoding of this code should be performed in a different way than the decoding of an OSTBC. It will be shown that, by a proper choice of the original OSTBC that is used in the REM, as well as the appropriate selection of the row/s to eliminate, the decoding complexity of the new, rate one, code remains linear.

2.2 Decoding Scheme

The new code is of full rate but obviously not orthogonal. As a result, it requires different decoding scheme than that for the regular OSTBC. Given a SSTBC codeword, one can rearrange the rows of the original OSTBC codeword and write

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}_r \\ \mathbf{X}_d \end{pmatrix} \tag{2.5}$$

Where $\mathbf{X}_r$ stands for the remaining codeword (the new truncated codeword), and $\mathbf{X}_d$ contains the deleted rows. Using the same notations the EVC matrix is now

$$\mathbf{H} = \begin{pmatrix} \mathbf{H}_r \\ \mathbf{H}_d \end{pmatrix} \tag{2.6}$$

The received vector can be written as

$$\tilde{\mathbf{y}}_r = \mathbf{H}_r\mathbf{s} + \tilde{\mathbf{z}}_r \tag{2.7}$$

The simple ML decoder of the OSTBC cannot be used here since $\mathbf{H}_r^H\mathbf{H}_r$ is not diagonal, and as a result the symbol by symbol decoding cannot be maintained. In order to force a symbol by symbol decoding, the sub-optimal Zero Forcing (ZF) decoder can be used, results in

$$\mathbf{r} = (\mathbf{H}_r^H\mathbf{H}_r)^{-1}\mathbf{H}_r^H\tilde{\mathbf{y}}_r \tag{2.8}$$

or

$$\mathbf{r} = \mathbf{H}_r^{-1}\tilde{\mathbf{y}}_r \tag{2.9}$$

if $\mathbf{H}_r$ is invertible.

The main problem with the suggested ZF decoder is its high computational complexity. An order of N^3 operations should be performed to calculate the corresponding filter, which is severely inefficient in comparison to the linear complexity of the OSTBC decoder. To tackle this problem, a different way to handle the decoding is presented. Taking advantage of the fact that the transmitted codeword was generated from an OSTBC, the proposed decoding scheme comprises of two steps. At the first step, the missing part of the codeword is estimated to reconstruct the original codeword. At the second step, the regular OSTBC decoder is applied. Following the same notations as before, the received vector of the original OSTC can be written as

$$\tilde{\mathbf{y}} = \begin{pmatrix} \tilde{\mathbf{y}}_r \\ \tilde{\mathbf{y}}_d \end{pmatrix} \tag{2.10}$$

The first step is to estimate the non-transmitted part $\tilde{\mathbf{y}}_d$. Using Equation (1.43), while disregarding the noise term, one have

$$\tilde{\mathbf{y}} = \mathbf{H}\mathbf{s} \tag{2.11}$$

Since

$$\mathbf{H}^H\mathbf{H} = (|h_1|^2 + |h_2|^2 + \cdots + |h_N|^2)\mathbf{I}_k \triangleq \alpha\mathbf{I}_k \tag{2.12}$$

hence, one can write $\mathbf{s}$ as

$$\mathbf{s} = \frac{1}{\alpha}\mathbf{H}^H\tilde{\mathbf{y}} \tag{2.13}$$

and $\tilde{\mathbf{y}}_d$ can be written as

$$\tilde{\mathbf{y}}_d = \mathbf{H}_d\mathbf{s} \tag{2.14}$$

Substituting (2.13) in (2.14) results in

$$\tilde{\mathbf{y}}_d = \frac{1}{\alpha}\mathbf{H}_d\mathbf{H}^H\tilde{\mathbf{y}} \tag{2.15}$$

or after some basic algebraic manipulations

$$\tilde{\mathbf{y}}_d = (\alpha\mathbf{I} - \mathbf{H}_d\mathbf{H}_d^H)^{-1}\mathbf{H}_d\mathbf{H}_r^H\tilde{\mathbf{y}}_r \tag{2.16}$$

which gives an expression of estimating $\tilde{\mathbf{y}}_d$ out of $\tilde{\mathbf{y}}_r$. Reconstructing $\tilde{\mathbf{y}}$ from $\tilde{\mathbf{y}}_r$ and $\tilde{\mathbf{y}}_d$, one can continue to the second step of the decoder and decode the data symbols vector $\mathbf{s}$ out of $\tilde{\mathbf{y}}$ using the regular OSTBC decoder given in (2.13).

A fundamental relationship between the proposed decoding method and the ZF decoder is summarized in the following claim;

Claim 2.2.1 *If $\mathbf{H}_r$ is nonsingular, the proposed two steps decoder is identical to the ZF decoder.*

Proof The filtered vector at the output of the second decoding step is given by

$$\mathbf{r} = \frac{1}{\alpha}\mathbf{H}^H\tilde{\mathbf{y}} = \frac{1}{\alpha}\begin{pmatrix} \mathbf{H}_r^H & \mathbf{H}_d^H \end{pmatrix}\begin{pmatrix} \tilde{\mathbf{y}}_r \\ \\ \tilde{\mathbf{y}}_d \end{pmatrix} \tag{2.17}$$

Substituting $\tilde{\mathbf{y}}_d$ from (2.16), one obtain

$$\begin{aligned}\mathbf{r} &= \tfrac{1}{\alpha}\begin{pmatrix} \mathbf{H}_r^H & \mathbf{H}_d^H \end{pmatrix}\begin{pmatrix} \tilde{\mathbf{y}}_r \\ (\alpha\mathbf{I} - \mathbf{H}_d\mathbf{H}_d^H)^{-1}\mathbf{H}_d\mathbf{H}_r^H\tilde{\mathbf{y}}_r \end{pmatrix} \\ &= \tfrac{1}{\alpha}\begin{pmatrix} \mathbf{H}_r^H & \mathbf{H}_d^H \end{pmatrix}\begin{pmatrix} \mathbf{I} \\ (\alpha\mathbf{I} - \mathbf{H}_d\mathbf{H}_d^H)^{-1}\mathbf{H}_d\mathbf{H}_r^H \end{pmatrix}\tilde{\mathbf{y}}_r\end{aligned} \tag{2.18}$$

which results in

$$\mathbf{r} = \frac{1}{\alpha}\left(\mathbf{H}_r^H + \mathbf{H}_d^H(\alpha\mathbf{I} - \mathbf{H}_d\mathbf{H}_d^H)^{-1}\mathbf{H}_d\mathbf{H}_r^H\right)\tilde{\mathbf{y}}_r \tag{2.19}$$

For nonsingular $\mathbf{H}_r$ the ZF decoder has the form

$$\mathbf{r}_{ZF} = \mathbf{H}_r^{-1}\tilde{\mathbf{y}}_r \tag{2.20}$$

Hence, one need to show that the following holds

$$\mathbf{H}_r^{-1} = \frac{1}{\alpha}\left(\mathbf{H}_r^H + \mathbf{H}_d^H(\alpha\mathbf{I} - \mathbf{H}_d\mathbf{H}_d^H)^{-1}\mathbf{H}_d\mathbf{H}_r^H\right) \tag{2.21}$$

Or equivalently, that

$$\mathbf{H}_r^{-1}\left(\mathbf{H}_r^H\right)^{-1} = \frac{1}{\alpha}\left(I + \mathbf{H}_d^H(\alpha\mathbf{I} - \mathbf{H}_d\mathbf{H}_d^H)^{-1}\mathbf{H}_d\right) \tag{2.22}$$

is satisfied. Since $\mathbf{H}^H\mathbf{H} = \alpha\mathbf{I}$ one can write

$$\begin{aligned}\mathbf{H}^H\mathbf{H} = \begin{pmatrix} \mathbf{H}_r^H & \mathbf{H}_d^H \end{pmatrix}\begin{pmatrix} \mathbf{H}_r \\ \mathbf{H}_d \end{pmatrix} = \mathbf{H}_r^H\mathbf{H}_r + \mathbf{H}_d^H\mathbf{H}_d = \alpha\mathbf{I} \\ \rightarrow \quad \mathbf{H}_r^H\mathbf{H}_r = \alpha\mathbf{I} - \mathbf{H}_d^H\mathbf{H}_d\end{aligned} \tag{2.23}$$

But $\left(\mathbf{H}_r^H\mathbf{H}_r\right)^{-1} = \mathbf{H}_r^{-1}\left(\mathbf{H}_r^H\right)^{-1}$, hence, combining (2.22) and (2.23) the following equation holds

$$\left(\alpha\mathbf{I} - \mathbf{H}_d^H\mathbf{H}_d\right)^{-1} = \frac{1}{\alpha}\left(\mathbf{I} + \mathbf{H}_d^H(\alpha\mathbf{I} - \mathbf{H}_d\mathbf{H}_d^H)^{-1}\mathbf{H}_d\right) \tag{2.24}$$

Multiplying by α and applying some algebra on the left hand side of the equation results in

$$\left(\mathbf{I} - \frac{1}{\alpha}\mathbf{H}_d^H\mathbf{H}_d\right)^{-1} = \mathbf{I} + \mathbf{H}_d^H(\alpha\mathbf{I} - \mathbf{H}_d\mathbf{H}_d^H)^{-1}\mathbf{H}_d \tag{2.25}$$

The last equation is the well known Woodbury matrix identity [25]

$$(\mathbf{A} + \mathbf{UCV})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{U}\left(\mathbf{C}^{-1} + \mathbf{VA}^{-1}\mathbf{U}\right)^{-1}\mathbf{VA}^{-1} \tag{2.26}$$

with $\mathbf{A} = \mathbf{I}$, $\mathbf{U} = -\mathbf{H}_d^H$, $\mathbf{V} = \mathbf{H}_d$ and $\mathbf{C} = \frac{1}{\alpha}\mathbf{I}$. Hence, (2.22) holds, which concludes the proof. □

Note that the demand on $\mathbf{H}_r$ to be invertible is crucial for the correctness of the claim. In wider sense, it relates to the REM process. If at the REM stage of generating the SSTBC out of an OSTBC codeword, the deleted rows were chosen such that the corresponding $\mathbf{H}_r$ is non-invertible, the SSTBC is poorly constructed, and the transmitted data cannot be recovered.

The performance of the new decoder is identical to the performance of the ZF decoder since their mathematical representation is the same. The incentive to use the two steps decoding scheme comes from its reduces computational complexity. The second step of the decoding is done in linear complexity as it is the regular OSTBC decoder. It is yet to be shown that the first step of estimating the missing part of the codeword can also be executed with linear complexity. A linear estimation of the missing part of the codeword, using the received vector can generally be written as

$$\tilde{\mathbf{y}}_d = \mathbf{B}\tilde{\mathbf{y}}_r \tag{2.27}$$

using Equation (2.16), the expression for the matrix $\mathbf{B}$ is given by

$$\mathbf{B} = (\alpha\mathbf{I} - \mathbf{H}_d\mathbf{H}_d^H)^{-1}\mathbf{H}_d\mathbf{H}_r^H \tag{2.28}$$

To force the required linear computational complexity, the calculation of the matrix $\mathbf{B}$ should also be restricted to be linear in its dimensions. For general OSTBC the calculation of the left hand side in Equation (2.28) does require more than linear order of operations. This is due to the matrices product and the matrix inversion. Hence, it is requires to hand pick the OSTBC from which the SSTBC is generated such that the left hand side expression in Equation (2.28) will converges to a more simple expression that can be calculate with a linear number of operations. In the next section, a systematic method for generating OSTBC which comply with this condition is presented. This method can generate OSTBC for any number of transmitting antennas. An REM rule, tailored for this new OSTBC family, is also provided such that the invertibility of $\mathbf{H}_r$ is guaranteed, and the expression of $\mathbf{B}$ becomes very simple to calculate.

2.3 New OSTBC

Applying the new transmission and decoding schemes to any OSTBC results in a full rate code which can be decoded by first, using (2.16), to estimate the missing part of the original code's received vector, followed by the regular OSTBC decoder. In order to benefit the most out of the two steps decoder, one should find an OSTBC as well as a rule for the REM such that the calculation of the matrix expression in (2.28) will be of low complexity. Although many of the known in the literature OSTBCs do comply with the codeword conditions of the proposed transmission / decoding schemes, it is desirable to come with a systematic way for generating such codes. This leads to a new OSTBC that addresses this requirement [26].

The new OSTBC can be used with any number of transmit antennas, and is

generated in a recursive way, i.e., $\mathbf{X}_{n+1}$ depends upon $\mathbf{X}_n$, where the subscript index represents both the number of Tx antennas and the number of the different data symbols. The algorithm for the code generation is summarized below.

Algorithm 1 Generate codeword $\mathbf{X}_{n+1}$ from $\mathbf{X}_n$

Given $\mathbf{X}_n$, then

$$\mathbf{X}_{n+1} = \begin{pmatrix} \mathbf{X}_n & \mathbf{b} \\ \mathbf{C} & -\mathbf{d} \end{pmatrix}$$

Where $[t, n] = \text{size}(\mathbf{X}_n)$

$$\mathbf{b}_{t\times 1} = (s_{n+1} \;\; \mathbf{0}_{1\times(t-1)})^T$$

$$\mathbf{C}_{n\times n} = s^*_{n+1} \cdot \mathbf{I}_n$$

$$\mathbf{d}_{n\times 1} = (s^*_1 \;\; s^*_2 \;\; ... \;\; s^*_n)^T$$

The starting point of the algorithm is simply the "codeword" for 1 Tx antenna $\mathbf{X}_1 = s_1$. Applying the algorithm on $\mathbf{X}_1$ results in

$$\mathbf{X}_2 = \begin{pmatrix} s_1 & s_2 \\ s^*_2 & -s^*_1 \end{pmatrix} \tag{2.29}$$

which is the famous Alamouti code. Making another recursive step produces $\mathbf{X}_3$,

$$\begin{aligned}\mathbf{X}_3 &= \begin{pmatrix} \mathbf{X}_2 & \begin{pmatrix} s_3 \\ 0 \end{pmatrix} \\ \begin{pmatrix} s_3^* & 0 \\ 0 & s_3^* \end{pmatrix} & -\begin{pmatrix} s_1^* \\ s_2^* \end{pmatrix} \end{pmatrix} \\ &= \begin{pmatrix} s_1 & s_2 & s_3 \\ s_2^* & -s_1^* & 0 \\ s_3^* & 0 & -s_1^* \\ 0 & s_3^* & -s_2^* \end{pmatrix}\end{aligned} \tag{2.30}$$

The $\mathbf{X}_3$ code as well as its successor $\mathbf{X}_4$ can be found in [27].

The code's orthogonality property follows directly from the code generation method (the full proof is given in Appendix A.). The rate of this code is the ratio between the number of rows and the number of columns of the code matrix and is given by

$$R_n = \frac{2n}{n^2 - n + 2} \tag{2.31}$$

Obviously this rate is very bad since it decreases fast with the number of Tx antennas. In order to transform this code to a rate 1 code through the row elimination method one need to delete $\frac{n^2-3n+2}{2}$ rows. This number is exactly the number of null entries in each column, which leads to the following REM rule. One needs to choose a column of the codeword and delete the rows that correspond to the null entries of the selected column. Recall that the codeword columns represent the different antennas (channels), hence this row elimination rule corresponds to the selection of the channel that will be utilized the most.

Applying the aforementioned REM rule, it can be shown that the resulted $\mathbf{H}_r$

is indeed invertible and that the expression for the matrix $\mathbf{B}$ can be simplified to (derived in details in Appendix B.1)

$$(\alpha\mathbf{I} - \mathbf{H}_d\mathbf{H}_d^H)^{-1}\mathbf{H}_d\mathbf{H}_r^H = \frac{1}{|h_i|^2}\mathbf{H}_d\mathbf{H}_r^H \tag{2.32}$$

where i is the index of the 'selected' column. Furthermore, it can be shown that (Appendix B.2)

$$\frac{1}{|h_i|^2}\mathbf{H}_d\mathbf{H}_r^H = \frac{1}{h_i^*}\tilde{\mathbf{H}}_{d,i} \tag{2.33}$$

where $\tilde{\mathbf{H}}_{d,i}$ is equal to $\mathbf{H}_d$ up to a possibly columns position exchange and sign inverse, hence, can be derived with no computational overhead. Combining Equations (2.32) and (2.33) results in

$$(\alpha\mathbf{I} - \mathbf{H}_d\mathbf{H}_d^H)^{-1}\mathbf{H}_d\mathbf{H}_r^H == \frac{1}{h_i^*}\tilde{\mathbf{H}}_{d,i} \tag{2.34}$$

This property greatly reduces the computational complexity of the decoding process, since the first step of the decoding process, given in (2.16), reduces to

$$\tilde{\mathbf{y}}_d = \frac{1}{h_i^*}\tilde{\mathbf{H}}_{d,i}\tilde{\mathbf{y}}_r \tag{2.35}$$

which has linear computational complexity. The second decoding step is the regular OSTBC decoder which is known to have linear complexity. Thus, the complexity order of the two steps decoder, applied on the new OSTBC with the proposed REM rule, is linear.

2.4 4 Tx Example

To better illustrates the new transmission/decoding schemes and the REM rule, consider the 4 Tx codeword generated using the new algorithm for OSTBC

construction [28]. The new code for in this case is given by

$$\mathbf{X}_{4,4/7} = \begin{pmatrix} s_1 & s_2 & s_3 & s_4 \\ s_2^* & -s_1^* & 0 & 0 \\ s_3^* & 0 & -s_1^* & 0 \\ 0 & s_3^* & -s_2^* & 0 \\ s_4^* & 0 & 0 & -s_1^* \\ 0 & s_4^* & 0 & -s_2^* \\ 0 & 0 & s_4^* & -s_3^* \end{pmatrix} \tag{2.36}$$

its rate is $\frac{4}{7}$ and one need to delete 3 rows to get a full rate code.

Following the proposed rule for the rows elimination, assuming without loss of generality, that the first column is chosen as the 'selected' one. Since the first column contains zeros in the 4th, 6th and 7th entries, these rows will be deleted from the code word, resulting in

$$\tilde{\mathbf{X}}_{4,1} = \begin{pmatrix} s_1 & s_2 & s_3 & s_4 \\ s_2^* & -s_1^* & 0 & 0 \\ s_3^* & 0 & -s_1^* & 0 \\ s_4^* & 0 & 0 & -s_1^* \end{pmatrix} \tag{2.37}$$

To ease the presentation, the 'rate' index will be removed in the sequel. The equivalent virtual channel (EVC) form for the full OSTC code ($\mathbf{X}_4$) is

$$\begin{pmatrix} y_1 \\ y_2^* \\ y_3^* \\ y_4^* \\ y_5^* \\ y_6^* \\ y_7^* \end{pmatrix} = \begin{pmatrix} h_1 & h_2 & h_3 & h_4 \\ -h_2^* & h_1^* & 0 & 0 \\ -h_3^* & 0 & h_1^* & 0 \\ 0 & -h_3^* & h_2^* & 0 \\ -h_4^* & 0 & 0 & h_1^* \\ 0 & -h_4^* & 0 & h_2^* \\ 0 & 0 & -h_4^* & h_3^* \end{pmatrix} \begin{pmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \end{pmatrix} \tag{2.38}$$

Rearranging the rows in the matrix results in the following EVC terms

$$\tilde{\mathbf{y}}_r = \begin{pmatrix} y_1 \\ y_2^* \\ y_3^* \\ y_5^* \end{pmatrix} \quad ; \quad \mathbf{H}_r = \begin{pmatrix} h_1 & h_2 & h_3 & h_4 \\ -h_2^* & h_1^* & 0 & 0 \\ -h_3^* & 0 & h_1^* & 0 \\ -h_4^* & 0 & 0 & h_1^* \end{pmatrix}$$
$$\tilde{\mathbf{y}}_d = \begin{pmatrix} y_4^* \\ y_6^* \\ y_7^* \end{pmatrix} \quad ; \quad \mathbf{H}_d = \begin{pmatrix} 0 & -h_3^* & h_2^* & 0 \\ 0 & -h_4^* & 0 & h_2^* \\ 0 & 0 & -h_4^* & h_3^* \end{pmatrix} \tag{2.39}$$

In this case $\tilde{\mathbf{H}}_{d,1}$ is equal to $\mathbf{H}_d$ and the estimation of $\tilde{\mathbf{y}}_d$ is simply given by

$$\begin{pmatrix} \hat{y_4}^* \\ \hat{y_6}^* \\ \hat{y_7}^* \end{pmatrix} = \frac{1}{h_1^*} \begin{pmatrix} 0 & -h_3^* & h_2^* & 0 \\ 0 & -h_4^* & 0 & h_2^* \\ 0 & 0 & -h_4^* & h_3^* \end{pmatrix} \begin{pmatrix} y_1 \\ y_2^* \\ y_3^* \\ y_5^* \end{pmatrix} \tag{2.40}$$

Once $\tilde{\mathbf{y}}_d$ is estimated, $\tilde{\mathbf{y}}$ is reconstructed and plugged in to the regular OSTBC decoder

$$\mathbf{r} = \frac{1}{\alpha} \mathbf{H}^H \tilde{\mathbf{y}} \tag{2.41}$$

If different than the first column was chosen, $\tilde{\mathbf{H}}_{d,i}$ will not be equal to $\mathbf{H}_d$ but could be derived in a very simple way. Consider the case were the second column is selected ($i = 2$). The second column contains zeros in its 3rd, 5th and 7th entries, these rows will be deleted from the $\mathbf{X}_4$, resulting in

$$\tilde{\mathbf{X}}_4 = \begin{pmatrix} s_1 & s_2 & s_3 & s_4 \\ s_2^* & -s_1^* & 0 & 0 \\ 0 & s_3^* & -s_2^* & 0 \\ 0 & s_4^* & 0 & -s_2^* \end{pmatrix} \tag{2.42}$$

Rearranging the rows in the $\mathbf{H}$ matrix results in the following EVC terms

$$\tilde{\mathbf{y}}_r = \begin{pmatrix} y_1 \\ y_2^* \\ y_4^* \\ y_6^* \end{pmatrix} \quad ; \quad \mathbf{H}_r = \begin{pmatrix} h_1 & h_2 & h_3 & h_4 \\ -h_2^* & h_1^* & 0 & 0 \\ 0 & -h_3^* & h_2^* & 0 \\ 0 & -h_4^* & 0 & h_2^* \end{pmatrix}$$
$$\tilde{\mathbf{y}}_d = \begin{pmatrix} y_3^* \\ y_5^* \\ y_7^* \end{pmatrix} \quad ; \quad \mathbf{H}_d = \begin{pmatrix} -h_3^* & 0 & h_1^* & 0 \\ -h_4^* & 0 & 0 & h_1^* \\ 0 & 0 & -h_4^* & h_3^* \end{pmatrix} \tag{2.43}$$

In this case $\tilde{\mathbf{H}}_{d,2}$ is given by

$$\tilde{\mathbf{H}}_{d,2} = \begin{pmatrix} 0 & h_3^* & h_1^* & 0 \\ 0 & h_4^* & 0 & h_1^* \\ 0 & 0 & -h_4^* & h_3^* \end{pmatrix} \tag{2.44}$$

which can be simply derived from $\mathbf{H}_d$ by inverting the sign of the first column and flipping the position of the first and second columns. The estimation of $\tilde{\mathbf{y}}_d$ is given in this case by

$$\begin{pmatrix} \hat{y_3}^* \\ \hat{y_5}^* \\ \hat{y_7}^* \end{pmatrix} = \frac{1}{h_2^*} \begin{pmatrix} 0 & h_3^* & h_1^* & 0 \\ 0 & h_4^* & 0 & h_1^* \\ 0 & 0 & -h_4^* & h_3^* \end{pmatrix} \begin{pmatrix} y_1 \\ y_2^* \\ y_4^* \\ y_6^* \end{pmatrix} \tag{2.45}$$

2.5 5 Tx Example

The proposed transmission (using REM) and decoding schemes work well with any number of transmit antennas. In this section, an 5 Tx OSTBC is presented with its REM and decoding schemes. For a system with five transmitting antennas, one can

use the following OSTBC codeword[1] [29],

$$\mathbf{X}_{5,2/3} = \begin{pmatrix} s_1 & s_2^* & s_3^* & s_4^* & 0 \\ s_2 & -s_1^* & 0 & 0 & s_5^* \\ s_3 & 0 & -s_1^* & 0 & -s_6^* \\ 0 & s_3 & -s_2 & 0 & s_7 \\ s_4 & 0 & 0 & -s_1^* & s_8^* \\ 0 & -s_4 & 0 & s_2 & -s_9 \\ 0 & 0 & -s_4 & s_3 & s_{10} \\ s_5 & 0 & -s_7^* & -s_9^* & -s_2^* \\ 0 & s_5 & -s_6 & s_8 & s_1 \\ s_6 & -s_7^* & 0 & -s_{10}^* & s_3^* \\ s_7 & s_6^* & s_5^* & 0 & 0 \\ s_8 & s_9^* & -s_{10}^* & 0 & -s_4^* \\ s_9 & -s_8^* & 0 & s_5^* & 0 \\ s_{10} & 0 & s_8^* & s_6^* & 0 \\ 0 & -s_{10} & -s_9 & s_7 & 0 \end{pmatrix} \tag{2.46}$$

which is orthogonal (i.e., $X^H X = (|s_1|^2 + |s_2|^2 + \cdots + |s_{10}|^2)\mathbf{I}_5$) with rate 2/3 since 10 data symbols are transmitted over 15 time slots. To get a full rate code, one need to delete five rows from the original codeword. As in the OSTBC proposed in Section 2.3, the number of null entries in each column of the codeword matrix is equal to number or rows that need to be eliminated from the codeword. This leads to a similar REM rule. After a column is chosen (in an arbitrary fashion or with some side information as will be explained later) the rows that coresponds to the null entries of this column are deleted from the codeword. Table 2.1 summarize the REM rule

[1]This code is not from the new OSTBC family presented earlier, but rather, off the shelf code. It was chosen as an example for the new transmission and decoding schemes to emphasis the fact that most of the OSTBC known in literature can be handle with the proposed schemes to result in rate one code with linear decoding complexity.

Table 2.1 REM Rule for 5 Tx OSTBC

Chosen Column	Rows to Delete
1	{4,6,7,9,15}
2	{3,5,7,8,14}
3	{2,5,6,10,13}
4	{2,3,4,11,12}
5	{1,11,13,14,15}

for the 5 Tx codeword given by (2.46). Choosing the non-transmitted rows according to Table 2.1, results in $\mathbf{H}_r$ and $\mathbf{H}_d$ which satisfies Equation (2.34), i.e., the following equation holds

$$(\alpha\mathbf{I} - \mathbf{H}_d\mathbf{H}_d^H)^{-1}\mathbf{H}_d\mathbf{H}_r^H = \frac{1}{h_i^*}\tilde{\mathbf{H}}_{d,i} \tag{2.47}$$

where the index i corresponds to the chosen column. The matrix $\tilde{\mathbf{H}}_{d,i}$ is equal to $\mathbf{H}_d$ up to a column sign inversion and position change. This means that the calculation of the filter (matrix B) for the first step of the decoding have no calculation overhead, hence, the total decoding complexity remains linear.

2.6 Improved Transmission / Decoding

The new transmission / decoding schemes combined with the new code and REM rule enable having a full rate code with linear complexity implementation of the ZF decoder. The performance of the proposed decoding method is highly depend on the quality of the estimation in the first decoding step. If the estimated missing part of the codeword will be severely non-accurate the resulted error rate will be high and intolerable. Having a closer look at the expression for the estimation of the missing part, given by Equation (2.34), reveals that the estimation is inverse proportional to h_i which is the channel coefficient that corresponds to the selected

column i at the REM stage. Since h_i can be relatively small, the estimated part can suffer a substantial noise enhancement, which in turn, effects the recovering of the data symbols. This problem can be shown by calculating the power of the noise terms at the output of the filter. After proving that the proposed decoding scheme is equivalent to the ZF decoding scheme one can simply analyze the performance of the ZF decoder. Revisiting the 4 Tx codeword given in Section 2.4, where the first column was chosen for the REM, the filtered vector is given by

$$\mathbf{r} = \mathbf{H}_r^{-1}\tilde{\mathbf{y}}_r = \mathbf{H}_r^{-1}\left(\mathbf{H}_r\mathbf{s} + \tilde{\mathbf{z}}\right) = \mathbf{s} + \mathbf{H}_r^{-1}\tilde{\mathbf{z}} = \mathbf{s} + \mathbf{v} \tag{2.48}$$

where $\tilde{\mathbf{z}} \sim \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I})$. The covariance matrix of $\mathbf{v}$ is given by

$$\mathbf{K}_v = \mathbf{H}_r^{-1}\sigma^2\mathbf{I}(\mathbf{H}_r^{-1})^H = \sigma^2(\mathbf{H}_r^H\mathbf{H}_r)^{-1} \tag{2.49}$$

Evaluating this matrix for the four antennas example results in a diagonal (the filtered noise variances) of the form (derived in Appendix C.)

$$\mathrm{diag}(\mathbf{K}_v) = \frac{\sigma^2}{\alpha}\begin{pmatrix} 1 \\ 1 + \left(\frac{|h_3|}{|h_1|}\right)^2 + \left(\frac{|h_4|}{|h_1|}\right)^2 \\ 1 + \left(\frac{|h_2|}{|h_1|}\right)^2 + \left(\frac{|h_4|}{|h_1|}\right)^2 \\ 1 + \left(\frac{|h_2|}{|h_1|}\right)^2 + \left(\frac{|h_3|}{|h_1|}\right)^2 \end{pmatrix} \tag{2.50}$$

It is clear from these expressions that given a channel realization where $|h_1|$ is very small, the noise term can be very large. This may happen if $|h_1|$ goes to zero or even for a moderate value of $|h_1|$ but accompanied with large $|h_i|$ terms ($i \neq 1$). In the general case, where are rows are deleted from the original OSTBC codeword (following a different selected column), the term in the denominators of Equation (2.50) will be $|h_i|^2$ where i corresponds to the index of the selected column. Never the less, the problem of having small h_i remains the same and can result in severe performance loss.

To tackle this problem, modified transmission and decoding schemes are presented [30]. In the following, the more compact 4 Tx OSTBC (given in Equation (2.2)) will be thoroughly analyzed and modified under various different setups. Recalling the rate $\frac{3}{4}$, 4 Tx codeword

$$\mathbf{X}_{4,3/4} = \begin{pmatrix} s_1 & s_2 & s_3 & 0 \\ -s_2^* & s_1^* & 0 & s_3 \\ -s_3^* & 0 & s_1^* & -s_2 \\ 0 & -s_3^* & s_2^* & s_1 \end{pmatrix} \tag{2.51}$$

Eliminating (without loss of generality) the last row, results in a new semi-orthogonal, rate 1 codeword

$$\mathbf{X}_{4,1} = \begin{pmatrix} s_1 & s_2 & s_3 & 0 \\ -s_2^* & s_1^* & 0 & s_3 \\ -s_3^* & 0 & s_1^* & -s_2 \end{pmatrix} \tag{2.52}$$

The two steps decoding scheme is demonstrated for this code. Having one receiving antenna, the received vector $\mathbf{y}$ entries are given by

$$\begin{aligned} y_1 &= h_1 s_1 + h_2 s_2 + h_3 s_3 + n_1 \\ y_2 &= -h_1 s_2^* + h_2 s_1^* + h_4 s_3 + n_2 \\ y_3 &= -h_1 s_3^* + h_3 s_1^* - h_4 s_2 + n_3 \end{aligned} \tag{2.53}$$

In order to recover the missing last row of the codeword $\mathbf{X}_{4,1}$ from $\mathbf{y}$, one need to find a linear combination of $\mathbf{y}$ entries which will results in the desired y_4 given by ([31]),

$$y_4 = h_4 s_1 + h_3 s_2^* - h_2 s_3^* \tag{2.54}$$

It is simple to show that the following combination results in (2.54) up to the noise term

$$\hat{y}_4 = \frac{h_4}{h_1} y_1 - \frac{h_3}{h_1} y_2 + \frac{h_2}{h_1} y_3 \tag{2.55}$$

Substituting (2.53) in the above equation one get

$$\hat{y}_4 = h_4 s_1 + h_3 s_2^* - h_2 s_3^* + \tilde{n}_4 \tag{2.56}$$

where $\tilde{n}_4$, the noise term of the estimated channel output (or the estimation error) is given by

$$\tilde{n}_4 = \frac{h_4}{h_1} n_1 - \frac{h_3}{h_1} n_2 + \frac{h_3}{h_1} n_3 \tag{2.57}$$

With the recovered y_4, the regular OSTBC decoder is applied. The "matched" filtering is given by the following combining equations

$$\begin{aligned} r_1 &= \tfrac{1}{\alpha}\left(h_1^* y_1 + h_2 y_2^* + h_3 y_3^* + h_4^* \hat{y}_4\right) = s_1 + \xi_1 \\ r_2 &= \tfrac{1}{\alpha}\left(h_2^* y_1 - h_1 y_2^* - h_4^* y_3 + h_3 \hat{y}_4^*\right) = s_2 + \xi_2 \\ r_3 &= \tfrac{1}{\alpha}\left(h_3^* y_1 + h_4^* y_2 - h_1 y_3^* - h_2 \hat{y}_4^*\right) = s_3 + \xi_3 \end{aligned} \tag{2.58}$$

where $\alpha = |h_1|^2 + |h_2|^2 + |h_3|^2 + |h_4|^2$ and the filtered noise terms are a linear combination of the channel noise terms $n_i,\ i = 1, 2, 3$, given by

$$\begin{aligned} \xi_1 &= \tfrac{1}{\alpha}\left(h_1^* n_1 + h_2 n_2^* + h_3 n_3^* + h_4^* \tilde{n}_4\right) \\ \xi_2 &= \tfrac{1}{\alpha}\left(h_2^* n_1 - h_1 n_2^* - h_4^* n_3 + h_3 \tilde{n}_4^*\right) \\ \xi_3 &= \tfrac{1}{\alpha}\left(h_3^* n_1 + h_4^* n_2 - h_1 n_3^* - h_2 \tilde{n}_4^*\right) \end{aligned} \tag{2.59}$$

The filtered noise powers are given by

$$E(\xi_1\xi_1^*) = \frac{\sigma^2}{\alpha}\left(1+\frac{|h_4|^2}{|h_1|^2}\right)$$

$$E(\xi_2\xi_2^*) = \frac{\sigma^2}{\alpha}\left(1+\frac{|h_3|^2}{|h_1|^2}\right) \tag{2.60}$$

$$E(\xi_3\xi_3^*) = \frac{\sigma^2}{\alpha}\left(1+\frac{|h_2|^2}{|h_1|^2}\right)$$

or in general by

$$E(\xi_i\xi_i^*) = \frac{\sigma^2}{\alpha}\left(1+\frac{|h_j|^2}{|h_1|^2}\right) \quad , \quad i \in \{1,2,3\},\ j \in \{2,3,4\} \tag{2.61}$$

where i is the symbol index and j is the channel or the transmit antenna index. Once again, it is clear from the expression in Equation (2.61) that the noise terms are inverse proportional to the channel coefficient that corresponds to the selected column at the REM stage. This may results in a poor performance in the case where this channel coefficient is very small.

2.6.1 Sequential Decoding

A simple improvement, yet performance boosting, can be made to this decoding scheme if considering sequential decoding. The motivation for the sequential decoding is the understanding that with high probability the largest term among the filtered noise powers can be very large which results in poor decoding performance for that symbol. In order to prevent this, one can first decode the two symbols which associated with the lesser two noise powers using (2.58). The next step is to subtract the decoded symbols from the received signals (2.53) and using the modified signals for the decoding of the last symbol without the use of the estimated $\hat{y}_4$. This method can be illustrated using the 4 Tx example. Consider, without loss of generality, that $|h_2| > |h_3| > |h_4|$. In this case s_1 and s_2 enjoys better SNR and will be decoded first

(i.e., the noise terms associated with r_1 and r_2 have less power than the one associated with r_3). Subtracting them from (2.53) one gets

$$\begin{aligned}\tilde{y}_1 &= y_1 - (h_1\hat{s}_1 + h_2\hat{s}_2) = h_3 s_3 + n_1 + e_1 \\ \tilde{y}_2 &= y_2 - (-h_1\hat{s}_2^* + h_2\hat{s}_1^*) = h_4 s_3 + n_2 + e_2 \\ \tilde{y}_3 &= y_3 - (h_3\hat{s}_1^* - h_4\hat{s}_2) = -h_1 s_3^* + n_3 + e_3\end{aligned} \tag{2.62}$$

where e_i are terms which correspond to the estimation errors of s_1 and s_2. The regular combining results in

$$\tilde{r}_3 = \frac{1}{\alpha_s}\left(h_3^*\tilde{y}_1 + h_4^*\tilde{y}_2 - h_1\tilde{y}_3^*\right) = s_3 + \psi_3 \tag{2.63}$$

where $\alpha_s = |h_1|^2 + |h_3|^2 + |h_4|^2$. Assuming low error rate for the first two decoded symbols, this scheme greatly decrease the filtered noise power of the last symbol. This will be further elaborated in the performance analysis section. Note that this addition for the decoding scheme doesn't change its complexity order which remains optimal, i.e., linear in the number of antennas and in the constellation size.

The main problem of this scheme, as can be viewed from (2.61), is that it depends greatly on the value of h_1 (note that the channel index is due to the choice of the last row as the one be eliminated, given a deletion of another row, the channel index in the denominator of (2.61) would have been different). If h_1 is close to zero, or more generally, if $|h_1| << |h_j|$, $j \in \{2, 3, 4\}$, the filtered noise power is very large resulting in poorly decoded symbols even for the two symbols that associated with the lesser noise power terms. To overcome this, some modifications for the proposed transmission and decoding schemes are presented under different system settings.

Table 2.2 Dependency of the Channel Coefficient at the Denominator of the Noise Power on the Deleted Row

Deleted Row	Denominator Channel Index
1	4
2	3
3	2
4	1

2.6.2 Limited Feedback

The simplest way to avoid a weak channel at the denominator of (2.61), is to have some feedback to the transmitter. A very limited feedback of 1 and 2 bits is considered and the modification of the transmission scheme is explained. The feedback is used at the transmitter for the choice of the row to delete which corresponds to the term in the noise power denominator. Table 2.2 summarizes this dependency which can be simply demonstrated.

2.6.2.1 2 Bits Feedback. With 2 feedback bits, or more generally, $\lceil log_2(M_t) \rceil$ feedback bits (where M_t is the number of transmit antennas), the receiver can inform the transmitter which channel is the strongest. With this information, the transmitter can choose the row to delete from the codeword matrix (2.51) such that the strongest channel will appear in the denominator of (2.61) so that

$$E(\xi_i \xi_i^*) \leq 2\frac{\sigma^2}{\alpha} \tag{2.64}$$

This is a simple and robust way to prevent the noise power from divergence.

2.6.2.2 1 Bit Feedback. With only 1 feedback bit, it can be pre-agreed on only two optional rows (instead of four), from which one row will be chosen as the deleted

row. For example, the transmitter deletes either the third or the last row. This corresponds to having either h_2 or h_1 at the denominator of (2.61) respectively (see Table 2.2). The feedback bit is determined according to the following rule

$$\begin{array}{ll} 0 & \text{if } |h_1| > |h_2| \\ 1 & \text{otherwise} \end{array} \tag{2.65}$$

The transmitter chooses the row to delete (among the pre-agreed two rows) according to the feedback bit value. By this method the probability of having very small denominator decreases due to the smaller probability of having both values of $|h_1|$ and $|h_2|$ close to zero at the same time.

2.6.3 Multiple Receiving Antennas

While the transmission scheme is modified when having a limited feedback, it remains the same for multiple receive antennas. Rather, the decoding scheme is modified to exploit the benefits of the additional antennas used to mitigate the "small denominator" problem presented. Consider the case of 2 Rx, the filtered noise powers of the combined signals (2.58) at each receive antenna are given now by

$$\begin{gathered} E(\xi_{i,k}\xi_{i,k}^*) = \frac{\sigma^2}{\alpha_k}\left(1 + \frac{|h_{j,k}|^2}{|h_{1,k}|^2}\right) \\ i \in \{1,2,3\},\ j \in \{2,3,4\},\ k \in \{1,2\} \end{gathered} \tag{2.66}$$

where k is the receive antenna index. At the receiver, a maximum ratio combining (MRC) approach can be used to maximize the output SNR by properly weighting the combined signals at each antenna output and combine them accordingly. Thus, the inputs to the de-mapper will be of the form

$$z_i = \frac{b_i}{a_i + b_i} r_{i,1} + \frac{a_i}{a_i + b_i} r_{i,2} \tag{2.67}$$

where $r_{i,k}$ are the combined outputs (2.58) associated with each receiver antenna. The weights a_i and b_i are given by

$$a_i = E(\xi_{i,1}\xi_{i,1}^*)$$
$$b_i = E(\xi_{i,2}\xi_{i,2}^*) \tag{2.68}$$

It is easy to see from (2.67) and (2.68) that in the case of very small $h_{1,k}$ at one of the receiving antennas the MRC will weight out this antenna output enabling a better decoding performance.

All three schemes (1 bit feedback, 2 bits feedback and multiple receiving antennas), can be further improved when combined with the proposed sequential decoding. This is done by following the proposed modified decoding schemes only for the two symbols having the two small noise power terms. The third symbol is decoded after the previous symbols are subtracted from the received signals. It is important to note that all the modification for the basic transmission/decoding schemes of the SSTBC do not change the complexity order which remains linear in the number of antennas and the constellation size of the modulation used.

2.7 Performance Analysis

In order to evaluate the performance gain of the different decoding schemes the foundation for the probability of error calculations is established. With the different SSTBC transmission and decoding scheme the instantaneous SNR is usually given by

$$SNR = \frac{P}{\sigma^2}x = \gamma x \tag{2.69}$$

where P is the antenna transmission power (per time slot), σ^2 is the channel additive noise power and x, which is a function of channel coefficients $\mathbf{h}$, is an r.v. which varies according to the applied transmission and decoding schemes. γ can be viewed

as SNR_0 or the SNR in the case of single input single output channel. The general expression for the instantaneous error probability is proportional to

$$P_{e,inst}(SNR) \propto Q(\sqrt{SNR}) \tag{2.70}$$

To evaluate the probability of error one need to average (2.70) over the r.v. x, namely

$$P_e = \int_0^\infty Q(\sqrt{SNR}) f_x(x) dx = \int_0^\infty Q(\sqrt{\gamma x}) f_x(x) dx \tag{2.71}$$

To solve this integral, two approximation are used. The first is the famous upper bound of the Q function

$$Q(x) \leq \frac{1}{2} e^{-\frac{x^2}{2}} \tag{2.72}$$

The p.d.f of x should be found or approximated such that this integral can be evaluated. For simple cases, such as the regular OSTBC, the p.d.f of x can be found analytically and it is said to follow the Chi Square distribution, χ^2, with $2 \cdot M_t$ degrees of freedom. It can be simply shown that for this type of distribution, the achieved diversity order is equal to half of the degrees of freedom. Hence, the OSTBC achieves diversity order of M_t which is full diversity in the simple case of MISO with M_t transmitting antennas and 1 receiving antenna.

The problem starts with more involved transmission / decoding schemes which involve feedback or multiple receiving antennas. It is much harder to derive an analytical expression for x in these cases. The common practice to tackle this problem is to approximate its empirical distribution with a known p.d.f function. To that end, the use of the Gamma distribution is suggested. Initially, the achieved diversity order for a general system where x's p.d.f is approximated to a Gamma distributed is derived. The p.d.f of the Gamma distributed r.v. is given by

$$f_x(x) = \frac{\theta^k}{\Gamma(k)} x^{k-1} e^{-\theta x} \quad , \quad x \geq 0 \tag{2.73}$$

where k and θ are the distribution parameters. Rewriting (2.71) one have

$$P_e \leq \int_0^\infty \frac{1}{2} e^{-\frac{\gamma x}{2}} \frac{\theta^k}{\Gamma(k)} x^{k-1} e^{-\theta x} dx = \frac{\theta^k}{2\Gamma(k)} \int_0^\infty x^{k-1} e^{-\frac{x}{2}(\gamma+2\theta)} dx \tag{2.74}$$

Using the following identity

$$\int_0^\infty x^a e^{-bx} dx = \frac{\Gamma(a+1)}{b^{a+1}} \tag{2.75}$$

results in

$$P_e \leq \frac{\theta^k}{2\Gamma(k)} \frac{\Gamma(k)}{\left(\frac{1}{2}(\gamma+2\theta)\right)^k} = 2^{k-1}\theta^k \frac{1}{(\gamma+2\theta)^k} \tag{2.76}$$

For high SNR (i.e., $\gamma >> 2\theta$) one can approximate and write

$$P_e \leq 2^{k-1}\theta^k \frac{1}{\gamma^k} \tag{2.77}$$

which indicates that a diversity order of k is achieved. Hence, the behavior of the probability of error in the high SNR region is a function of the distribution parameters. A diversity order of k is achieved, while θ determines how large γ should be to achieve this diversity order. These two parameters will be the basis for comparison of the various transmission / decoding schemes that were presented. Note that the χ^2 distribution is a private case of the Gamma distribution with $\theta = \frac{1}{2}$.

This approximation is preferred due to the fact that it is more general than the traditionally used Chi-square distribution, i.e., the additional parameter θ allows better fitting to the given empirical distribution while the χ^2 distribution consist of only one parameter n which corresponds to the number of degrees of freedom and is equivalent to k of the Gamma distribution. In addition, the Gamma distribution has an important scaling property, namely, if

$$x \sim Gamma(k, \theta) \tag{2.78}$$

then for any $a > 0$

$$ax \sim Gamma(k, \frac{\theta}{a}) \tag{2.79}$$

This implies the following; any scaling change of x will not result in a different achieved diversity order but rather will change only the value of γ for which this diversity order is achieved. This result is expected, yet since the χ^2 p.d.f function have no similar property, i.e., scaling a χ^2 distributed r.v. results in a non χ^2 distribution. This property is important when comparing the performance of different transmission / decoding methods which, in many cases, results in different effective transmission power. This difference in γ can be embedded into x resulting in a scaling of the random variable. Using the Gamma p.d.f as the approximation for the various distributions enables an easier comparison since the p.d.f function remains the same (up to a scaling parameter) and there is no need to compare different p.d.f functions. In addition, x can be written as the product of two terms $\alpha(h) \cdot f(h)$, where $\alpha(h)$ is a sum of $|h_i|^2$ terms and $f(h)$ is usually some rational function involving $|h_i|^2$ terms. This eases the performance comparison of different schemes with common $\alpha(h)$ since the comparison is done only on the $f(h)$ part of x.

2.7.1 Basic SSTBC

From Equations (2.61) and (2.69) one can write the resulted SNR of the basic SSTBC as

$$SNR_i = P\frac{\alpha}{\sigma^2}\left(\frac{|h_1|^2}{|h_j|^2 + |h_1|^2}\right) = P\frac{\alpha}{\sigma^2}\nu \tag{2.80}$$

where it can be shown (Appendix D.1) that the r.v. ν has uniform distribution between 0 and 1. This degrades the performance significantly since the probability of having low SNR at the receiver is substantial.

When ordering the decoded symbols by their filtered noise powers their

associated $\nu_{(i)}$ (i.e., the r.v. that corresponds to their instantaneous SNR) are given by

$$\nu_i = \frac{|h_1|^2}{h_{(i)} + |h_1|^2} \tag{2.81}$$

where $h_{(i)}$ are the ordered statistics of $\{|h_2|^2\ |h_3|^2\ |h_4|^2\}$. Their p.d.fs calculated in Appendix D.2 and are given by

$$\begin{aligned}
f_{\nu_{(1)}}(\nu_{(1)}) &= \tfrac{3}{(3-2\nu_{(1)})^2} \quad , \quad 0 \le \nu_{(1)} \le 1 \\
f_{\nu_{(2)}}(\nu_{(2)}) &= \tfrac{3}{2}\left[\tfrac{1}{(1-\frac{1}{2}\nu_{(2)})^2} - \tfrac{1}{(\frac{3}{2}-\nu_{(2)})^2}\right] \quad , \quad 0 \le \nu_{(2)} \le 1 \\
f_{\nu_{(3)}}(\nu_{(3)}) &= \tfrac{3}{4}\left[4 - \tfrac{2}{(1-\frac{1}{2}\nu_{(3)})^2} + \tfrac{1}{(\frac{3}{2}-\nu_{(3)})^2}\right] \quad , \quad 0 \le \nu_{(3)} \le 1
\end{aligned} \tag{2.82}$$

Figure 2.1 shows the p.d.fs of $\nu_{(i)}$ and it can easily viewed that for the symbol with the smallest SNR (associated with $\nu_{(3)}$) the probability for having close to zero SNR is substantial resulting in very poor performance. To that end, the sequential decoding was suggested.

2.7.2 Sequential Decoding

By sequential decoding, the symbol with the largest noise term is handled differently than the basic decoding. After having two decoded symbols they subtracted from the received signals and then simply combined to form the third symbol input to the de-mapper. Since there is no use of the noisy estimation of $\hat{y}_4$ the total noise power via this method is reduced. The calculation of the resulted noise power is done by revisiting (2.63). The noise term ψ_3 is given by

$$\psi_3 = \frac{1}{\alpha_s}\left(h_3^*(n_1 + e_1) + h_4^*(n_2 + e_2) - h_1(n_3^* + e_3)\right) \tag{2.83}$$

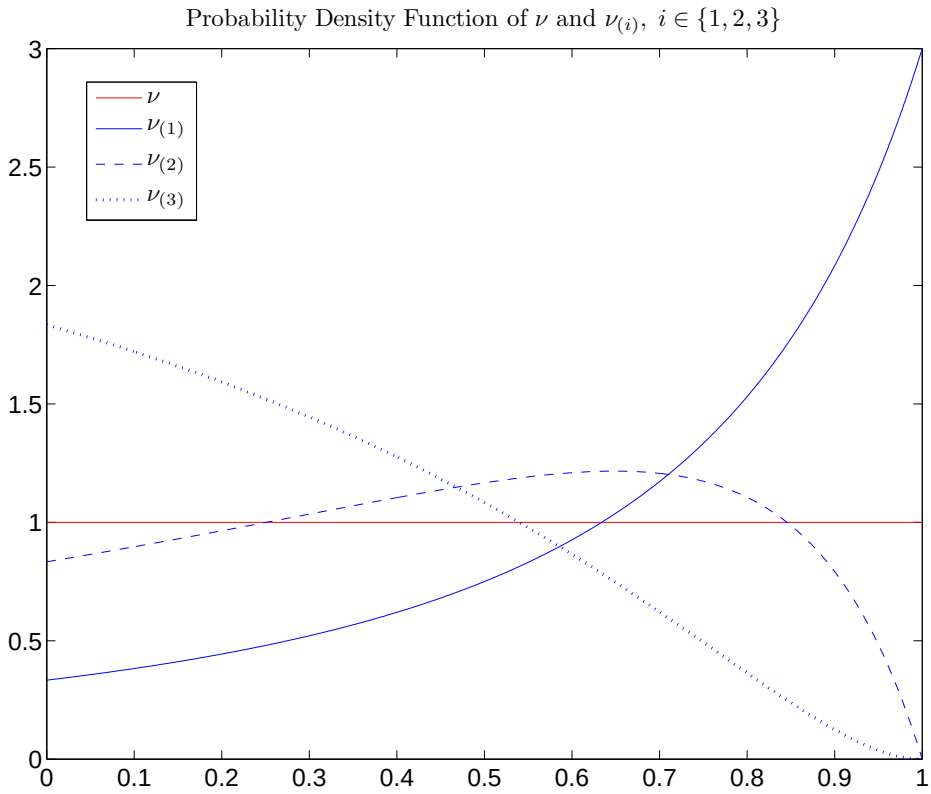

Figure 2.1 Probability density function of ν and $\nu_{(i)}$.

Assuming perfect estimation of the first two symbols (i.e., $e_i = 0$) the power of the filtered noise is given by

$$E(\psi_3\psi_3^*) = \frac{\sigma^2}{(|h_1|^2 + |h_3|^2 + |h_4|^2)} \tag{2.84}$$

It can be simply shown that this noise power is always smaller than the power of ξ_3, hence, the total performance is better than decoding with the basic decoding scheme. In general the noise power can be written as

$$E(\psi_3\psi_3^*) = \frac{\sigma^2}{(\alpha - h_{(3)})} = \frac{\sigma^2}{\upsilon} \tag{2.85}$$

where $h_{(3)}$ is the maximum between $\{|h_2|^2,\ |h_3|^2,\ |h_4|^2\}$. The p.d.f of υ is derived in Appendix D.3 and given by

$$f_\upsilon(\upsilon) = 3\left(e^{-\upsilon} + e^{-\frac{1}{2}\upsilon} - 2e^{-\frac{3}{4}\upsilon}\right) \tag{2.86}$$

Figure 2.2 illustrates the instantaneous SNR p.d.fs of the decoding of the last symbol and compares between the regular ($\alpha \cdot \nu_{(3)}$) and the sequential decoding (υ). It is clear how dramatic is the SNR improvement when using the sequential decoding scheme with no substantial addition to the computational complexity of the decoder.

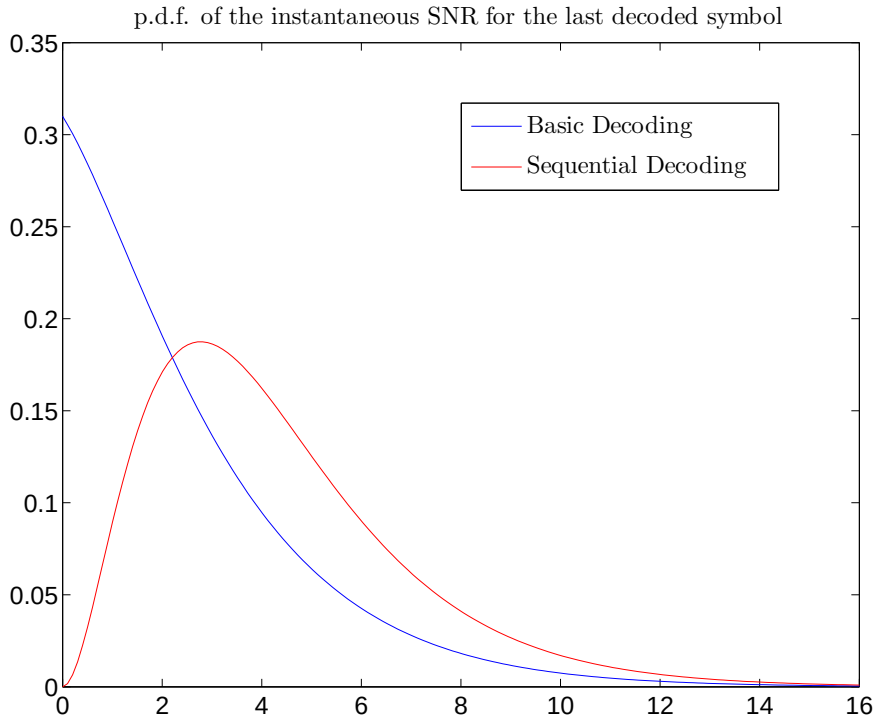

Figure 2.2 Probability density function of the instantaneous SNR of the last decoded symbol for the basic and sequential decoding.

2.7.3 Modified SSTBC

The p.d.fs of the instantaneous SNR terms of the modified SSTBC for systems with limited feedback or multiple receiving antennas is more involved to calculate analytically. The histograms of these terms are presented and compared in Figure 2.3 to illustrate how these modifications can boost the performance and mitigate the problem presented with the regular scheme. As mentioned before, the sequential decoding can be incorporated with the decoding of the modified SSTBC for further performance enhancement.

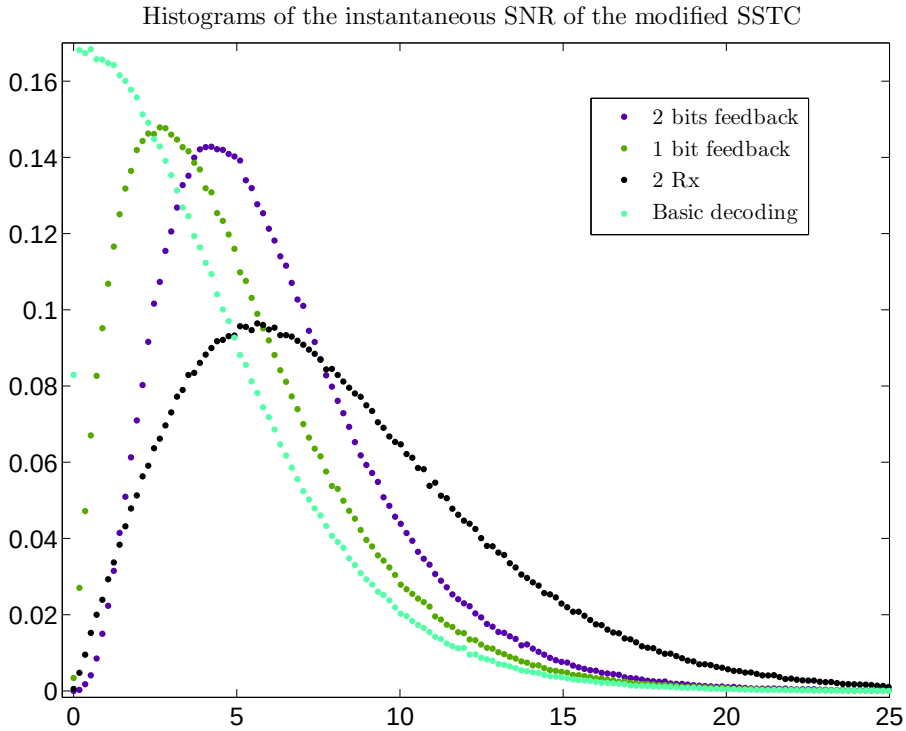

Figure 2.3 Histograms of the instantaneous SNR of the modified SSTBC.

Figure 2.4 shows the performance curves for an extensive simulations that have been performed to compare a 4 Tx system using SSTBC and QSTBC under various settings[2]. The black curve is the regular SSTBC decoding, with poor performance. The sequential decoding scheme is performed on top of the other modified schemes (based on the system settings). It is clear that the performance is similar to the one achieved by the QSTBC for these setting, but with lower decoding complexity.

[2]For the QSTBC encoding with limited feedback and 2 receiving antennas the decoding schemes that been used in the simulation were adopted from [32],[33] and [34]

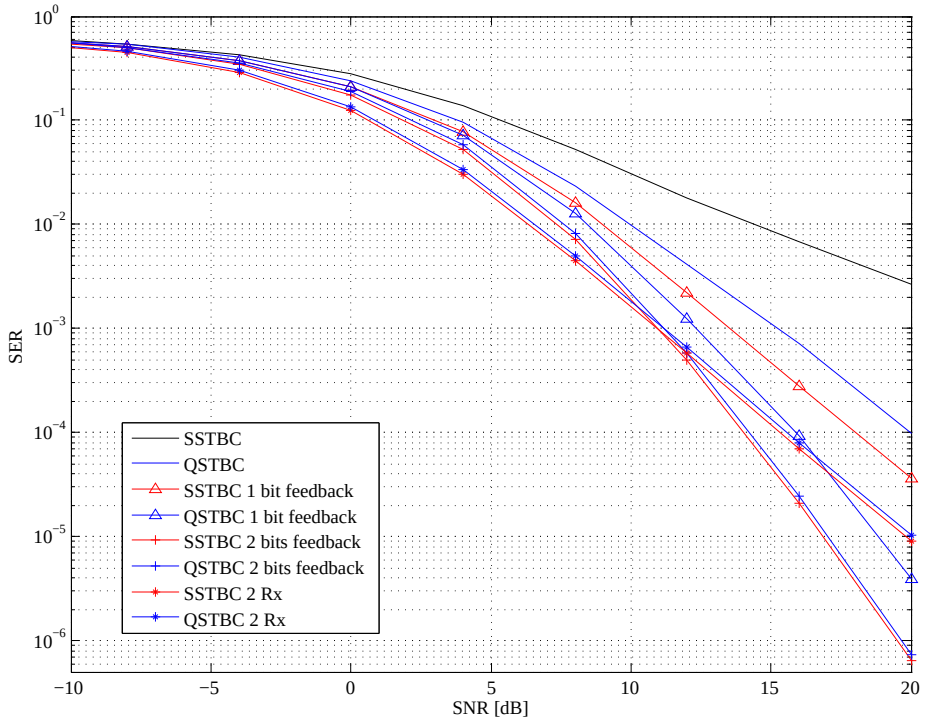

Figure 2.4 SSTBC Vs. QSTBC for different system settings.

Part II

Adaptive Decoding for STBC with Imperfect Channel Estimation

Preface

In every communication system that requires some channel state information at the receiver, a portion of the transmission is dedicated to training sequences. The quality of the channel estimation is a function of the total energy / length of the training sequence. It was shown that in space time codes (STC) systems it is crucial to have a very accurate channel estimation in order to have low error rate [35]. This implies that a substantial portion of the transmission will be "wasted" on training to obtain the required accurate channel estimation, resulting in significant effective rate loss. This problem is even more acute in high data rate systems where large constellation size is used and the error probability increase dramatically for inaccurate channel coefficient at the receiver.

Substantial work has been done in this area, covering many aspects relating to imperfect channel estimation (ICE). Zheng and Tse [36] addressed the capacity of the MIMO channel with imperfect side information caused by ICE from an information-theoretic point of view . The design criteria of STC in the presence of ICE is discussed in [7]. A performance analysis of various decoders for STC with ICE can be found in [37], where the maximum likelihood (ML) decoding scheme is used. The problem with this approach is its computational complexity overload, which becomes non-tractable for large constellation size and / or for large number of transmit antennas.

This part of the dissertation focuses on a method for handling the ICE scenario while maintaining low decoding complexity. This can be achieved by forcing the regular symbol by symbol decoding followed by a reduction of the estimation errors effect. The effect of using the mismatch decoder, i.e., decoding regularly with an erroneous filter, is the introduction of signals interference at the receiver. Hence, it is suggested to use an adaptive method , namely the bootstrap algorithm [38], for signals separation on the output of the mismatched filter.

CHAPTER 3

SPACE TIME BLOCK CODES WITH IMPERFECT CHANNEL ESTIMATION

The presence of channel estimation errors degraded the performance of STBC dramatically. One approach to mitigate the affect of imperfect channel estimation (ICE) is to enhance the estimation quality. This, inevitably, comes with the penalty of the resources that are dedicated for the estimation process, namely, the time and power of the training sequence. To get better estimation, one need to increase the length of the training sequence and / or to invest more power to its transmission. Both harm the transmission of the data itself since, for achieving the same data bit rate, one need to use larger constellation size to compensate for the shorter transmission time. In addition, less power is available per data block due to the excess power that was invested in the training portion of the transmitted block. Another approach is a complex ML decoding that involves both the channel coefficients estimation and the data symbols.

This approach was analyzed in [37] and shown to be very effective in combating the ICE effect on STBC decoding. Nether the less, the computational complexity of the implementation of such ML decoding scheme is very high. For system with large number of transmit antennas and high constellation size, this scheme is practically non-tractable. Hence, it is highly desirable to have a method for minimize the ICE effect on the performance of STBC while maintaining the low complexity of the decoding process. Moreover, the dedicated portion of the transmitted block for estimating the channel coefficients, should remain untouched in terms of its length and total power, in order to maintain the system data rate and performance.

With these restrictions in mind, this chapter addresses the ICE problem in a

different way. It is suggested to start with the regular STBC decoding, where simple decoding is guaranteed, which resulting in inter symbol interference due to the use of erroneous channel coefficients. The next decoding phase is dedicated for interference reduction where a simple, adaptive scheme, is proposed as a mean for refining the output of the regular decoder by lowering the interference levels.

3.1 System Model

The effect of ICE on the decoding of a STBC system is demonstrated for both the OSTBC and QSTBC families. Consider the received signal of a MISO channel with M_t transmitting antennas [1]

$$\mathbf{y} = \mathbf{X}\mathbf{h} + \mathbf{n} \tag{3.1}$$

where $\mathbf{h} = [h_1 \ \dots \ h_{M_t}]^T$ are the channel coefficients.

3.1.1 OSTBC

To present the ICE effect on the decoding, the basics properties of the OSTBC decoding scheme are reviewed. The OSTBC codeword $\mathbf{X}$ satisfies

$$\mathbf{X}^H\mathbf{X} = \gamma\mathbf{I} \tag{3.2}$$

where $\gamma = \frac{1}{k}||\mathbf{X}||^2$, where k is the number of different data symbols and $||\cdot||$ is the Forbenius norm.

$\mathbf{y}$ can be rewritten as the output of a MIMO channel (in the EVC form)

$$\tilde{\mathbf{y}} = \mathbf{H}\mathbf{s} + \tilde{\mathbf{n}} \tag{3.3}$$

[1] While this work can be applied to a general STBC system operating over a MIMO channel with M_t transmitting antennas and M_r receiving antennas, the emphasis will be on the multiple input single output (MISO) channel. This is due to the more intuitive aspects of the decoding process for system with only one receive antenna.

where $\mathbf{s}$ is the data symbols vector. $\tilde{\mathbf{y}}$ and $\tilde{\mathbf{n}}$ are equal to $\mathbf{y}$ and $\mathbf{n}$ respectively, up to the conjugation of some of the vectors entries. The equivalent channel matrix $\mathbf{H}$ has similar properties as $\mathbf{X}$ and it also satisfies

$$\mathbf{H}^H\mathbf{H} = \alpha\mathbf{I} \tag{3.4}$$

where $\alpha = \frac{1}{M_t}||\mathbf{H}||^2$. This attribute enables the simple maximum likelihood (ML) decoding of the OSTBC given by

$$\mathbf{r} = \frac{1}{\alpha}\mathbf{H}^H\tilde{\mathbf{y}} = \mathbf{s} + \frac{1}{\alpha}\mathbf{H}^H\tilde{\mathbf{n}} \tag{3.5}$$

In the presence of ICE, an erroneous version of the channel coefficients vector $\hat{\mathbf{h}} = \mathbf{h} + \mathbf{h}_e$ is available at the receiver. This transforms to an erroneous equivalent channel matrix $\hat{\mathbf{H}}$, written as

$$\hat{\mathbf{H}} = \mathbf{H} + \mathbf{H}_e \tag{3.6}$$

This results in the mismatch decoder

$$\hat{\mathbf{r}} = \frac{1}{\hat{\alpha}}\hat{\mathbf{H}}^H\tilde{\mathbf{y}} = \frac{1}{\hat{\alpha}}\left(\alpha\mathbf{I} + \mathbf{H}_e^H\mathbf{H}\right)\mathbf{s} + \frac{1}{\hat{\alpha}}\hat{\mathbf{H}}^H\tilde{\mathbf{n}} \tag{3.7}$$

where $\hat{\alpha} = \frac{1}{M_t}||\hat{\mathbf{H}}||^2$. Assuming small estimation errors one can approximate $\hat{\alpha} \simeq \alpha$ and write

$$\hat{\mathbf{r}} = \left(\mathbf{I} + \frac{1}{\hat{\alpha}}\mathbf{H}_e^H\mathbf{H}\right)\mathbf{s} + \frac{1}{\hat{\alpha}}\hat{\mathbf{H}}^H\tilde{\mathbf{n}} \tag{3.8}$$

$\hat{\mathbf{r}}$ can be viewed as the output of a system with signals cross interference, where the off diagonal elements of the matrix $\left(\mathbf{I} + \frac{1}{\hat{\alpha}}\mathbf{H}_e^H\mathbf{H}\right)$ create the coupling between the different signals.

3.1.2 QSTBC

Quasi OSTBC codes have, in general, higher rate than the OSTBC codes. This rate increase comes with a penalty of an inherent coupling among the data symbols which create interference. The presence of ICE at the receiver adds further interference in the decoding process. To illustrate this fact, consider the 4 Tx Extended Alamouti (EA) QSTBC which is presented [19], and whose codeword is given by

$$\mathbf{X}_{EA} = \begin{pmatrix} s_1 & s_2 & s_3 & s_4 \\ s_2^* & -s_1^* & s_4^* & -s_3^* \\ s_3^* & s_4^* & -s_1^* & -s_2^* \\ s_4 & -s_3 & -s_2 & s_1 \end{pmatrix} \tag{3.9}$$

The received vector can be written as

$$\mathbf{y} = \mathbf{X}_{EA}\mathbf{h} + \mathbf{n} \tag{3.10}$$

Applying the EVC model, this can, equivalently, be written as an output of a MIMO channel

$$\tilde{\mathbf{y}} = \mathbf{H}\mathbf{s} + \tilde{\mathbf{n}} \tag{3.11}$$

where $\tilde{\mathbf{y}} = [y_1\ y_2^*\ y_3^*\ y_4]^T$, $\tilde{\mathbf{n}} = [n_1\ n_2^*\ n_3^*\ n_4]^T$ and

$$\mathbf{H} = \begin{pmatrix} h_1 & h_2 & h_3 & h_4 \\ -h_2^* & h_1^* & -h_4^* & h_3^* \\ -h_3^* & -h_4^* & h_1^* & h_2^* \\ h_4 & -h_3 & -h_2 & h_1 \end{pmatrix} \tag{3.12}$$

In order to enable simple symbol by symbol decoding the ZF decoder is used. It is basically computes the following

$$\mathbf{r}_{ZF} = \mathbf{H}^{-1}\tilde{\mathbf{y}} = \mathbf{s} + \mathbf{H}^{-1}\tilde{\mathbf{n}} \tag{3.13}$$

Having ICE, the equivalent channel matrix available at the receiver is $\hat{\mathbf{H}}$ as in (3.6). Applying the regular ZF decoding with $\hat{\mathbf{H}}$ results in

$$\hat{\mathbf{r}}_{ZF} = \hat{\mathbf{H}}^{-1}\tilde{\mathbf{y}} = \hat{\mathbf{H}}^{-1}\mathbf{H}\mathbf{s} + \hat{\mathbf{H}}^{-1}\tilde{\mathbf{n}} \tag{3.14}$$

After some matrix algebra manipulations this can be written as

$$\hat{\mathbf{r}}_{ZF} = \left(\mathbf{I} - \left(\mathbf{H} + \mathbf{H}_e\right)^{-1}\mathbf{H}_e\right)\mathbf{s} + \hat{\mathbf{H}}^{-1}\tilde{\mathbf{n}} \tag{3.15}$$

As in the OSTBC case, the term $\left(\mathbf{H} + \mathbf{H}_e\right)^{-1}\mathbf{H}_e$ is the symbols coupling term which results in the inter symbol interference caused due to the ICE.

3.2 Adaptive Decoding

The output of the STBC decoder, when using the mismatched filter, can be viewed as the output of a system with symbol cross interference ((3.8) and (3.15)). Since the coupling matrix is not known to the receiver, a blind method is required to handle the interference. An adaptive scheme is in favor due to its simple implementation which usually requires minimal hardware and small number of computations per iteration. For the proposed ICE mitigation scheme, the bootstrap algorithm is adapted as a powerful, blind, adaptive method which can reduce the interference level caused by ICE [39].

3.2.1 The Bootstrap Algorithm

The bootstrap algorithm, first presented in [38], is an adaptive method for signals separation. Consider a multi signal system with coupling matrix $\mathbf{P}$, where the noise is currently neglected for the presentation clarity,

$$\mathbf{x} = \mathbf{P}\mathbf{s} \tag{3.16}$$

where $\mathbf{x}$ is the system's output, $\mathbf{s}$ is the signals vector both of dimension $K \times 1$ and $\mathbf{P}$ is given by

$$P = \begin{pmatrix} 1 & \rho_{12} & \cdots & \rho_{1K} \\ \rho_{21} & \ddots & & \rho_{2K} \\ \vdots & & \ddots & \vdots \\ \rho_{K1} & \rho_{K2} & \cdots & 1 \end{pmatrix} \tag{3.17}$$

where ρ_{ij} is the coupling coefficient between the symbols i and j.

Different from zero forcing, which applies $\mathbf{P}^{-1}$, the bootstrap algorithm calculates

$$\mathbf{z} = \mathbf{V}\mathbf{x} = \mathbf{V}\mathbf{P}\mathbf{s} \tag{3.18}$$

where $\mathbf{V}$ is chosen such that $\mathbf{VP}$ is a diagonal matrix but not necessarily $\mathbf{V} = \mathbf{P}^{-1}$. The suggestion is to take $\mathbf{V} = \mathbf{I} - \mathbf{W}$ where $\mathbf{W}$ is given by

$$\mathbf{W} = \begin{pmatrix} 0 & w_{12} & \cdots & w_{1K} \\ w_{21} & \ddots & & w_{2K} \\ \vdots & & \ddots & \vdots \\ w_{K1} & w_{K2} & \cdots & 0 \end{pmatrix} \tag{3.19}$$

The weights w_{ij} are chosen so that

$$E[z_k \text{sgn}\{\mathbf{z}_k\}] = 0 \tag{3.20}$$

where $\mathbf{z}_k$ is the vector $\mathbf{z}$ without z_k, its kth element. The recursion for calculating the weights is given by

$$\mathbf{w}_k(n+1) = \mathbf{w}_k(n) - \mu z_k(n) \text{sgn}\{\mathbf{z}_k(n)\} \tag{3.21}$$

where $\mathbf{w}_k$ is the kth column of $\mathbf{W}$ without the kth element (i.e., the kkth element of $\mathbf{W}$).

The bootstrap algorithm was first implemented for real signals with real coupling matrix P (i.e., the coupling coefficients ρ_{ij} are also real). In [40], a complex implementation of the bootstrap algorithm was presented, where the signals, the coupling coefficients and the weights, can be complex. In this more general case, the complex weights are chosen such that

$$E[z_k^* \text{sgnc}\{\mathbf{z}_k\}] = 0 \tag{3.22}$$

and the recursion for calculating the weights is given by

$$\mathbf{w}_k(n+1) = \mathbf{w}_k(n) - \mu z_k^*(n) \text{sgnc}\{\mathbf{z}_k(n)\} \tag{3.23}$$

The signum function for complex numbers, sgnc, is defined by

$$\text{sgnc}(\cdot) = \text{sgn}\left(\Re(\cdot)\right) + j\ \text{sgn}\left(\Im(\cdot)\right) \tag{3.24}$$

At the steady state of the algorithm (3.22) holds, i.e., the correlation between the elements of $\mathbf{z}$ goes to zero which implies that symbols separation is achieved. Figure 3.1 depicts the general structure of a 2 rails complex bootstrap algorithm.

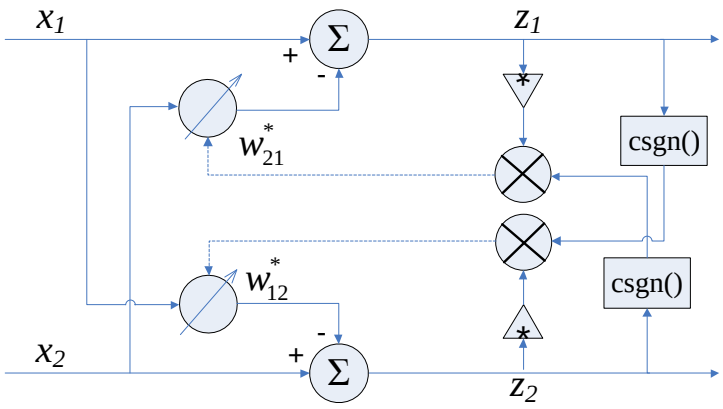

Figure 3.1 Schematics of a two users, complex implementation of the bootstrap algorithm.

3.2.2 Reduced Complexity QSTBC Decoder

Before getting into the details of the bootstrap algorithm implementation, a beneficial by product of the use of the bootstrap is presented. In Section 3.1.2, the effect of ICE on a QSTBC system which utilizes a mismatched ZF decoding scheme is derived. The reason for using the ZF decoding is to avoid the QSTBC inherent intefrence, thus, enabling a symbol by symbol decoding. Nether the less, due to ICE, inter symbol interference does appear at the output of the ZF decoder as shown in Equation (3.15). Since the bootstrap algorithm is used to reduce the interference levels, one might suggested to apply the simpler OSTBC filter. Using the OSTBC filter, $\mathbf{H}^H$, will inevitably result in the introduction of more symbol interference at its output, but since the bootstrap decoder is already applied, it will reduce this interference also. Using $\mathbf{H}^H$ instead of $\mathbf{H}^{-1}$ results in reduced decoding complexity as the need for matrix inversion is avoided.

The QSTBC got its name from the following basic property of the QSTBC [34], which contains extra off diagonal terms

$$\mathbf{G} = \mathbf{H}^H\mathbf{H} = \alpha \begin{pmatrix} 1 & 0 & 0 & \frac{\beta}{\alpha} \\ 0 & 1 & -\frac{\beta}{\alpha} & 0 \\ 0 & -\frac{\beta}{\alpha} & 1 & 0 \\ \frac{\beta}{\alpha} & 0 & 0 & 1 \end{pmatrix} \tag{3.25}$$

where $\alpha = \sum_{i=1}^{4} |h_i|^2$ and $\beta = 2\Re\{h_1 h_4^* - h_2 h_3^*\}$. The off diagonal elements in the matrix $\mathbf{G}$ define the code as quasi orthogonal and represents the inherent symbols coupling of the QSTBC codes. Having in mind that ICE adds interference at the decoder it is clear that any adaptive interference cancelation scheme might be able to handle simultaneously both types of interference in the system. Hence, the following decoding scheme is suggested.

Initially, the output vector is multiplied by $\frac{1}{\hat{\alpha}}\hat{\mathbf{H}}^H$, the normalized hermitian

conjugate of the erroneous channel matrix (instead of its inverse)

$$\begin{aligned}\tilde{\mathbf{r}} &= \tfrac{1}{\hat{\alpha}}\hat{\mathbf{H}}^H\tilde{\mathbf{y}} = \tfrac{1}{\hat{\alpha}}\left(\mathbf{H}+\mathbf{H}_e\right)^H\mathbf{H}\mathbf{s} + \tfrac{1}{\hat{\alpha}}\hat{\mathbf{H}}^H\tilde{\mathbf{n}} \\ &= \tfrac{1}{\hat{\alpha}}\left(\mathbf{G}+\mathbf{H}_e^H\mathbf{H}\right)\mathbf{s} + \tfrac{1}{\hat{\alpha}}\hat{\mathbf{H}}^H\tilde{\mathbf{n}}\end{aligned} \tag{3.26}$$

$\mathbf{G}$ can be rewritten as

$$\mathbf{G} = \alpha\left(\mathbf{I}+\frac{\beta}{\alpha}\mathbf{J}\right) \tag{3.27}$$

where

$$\mathbf{J} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \tag{3.28}$$

Hence, one have

$$\tilde{\mathbf{r}} = \frac{1}{\hat{\alpha}}\left(\alpha\left(\mathbf{I}+\frac{\beta}{\alpha}\mathbf{J}\right)+\mathbf{H}_e^H\mathbf{H}\right)\mathbf{s} + \frac{1}{\hat{\alpha}}\hat{\mathbf{H}}^H\tilde{\mathbf{n}} \tag{3.29}$$

Under the assumption of small estimation errors, α can be approximated by $\hat{\alpha}$, resulting in

$$\tilde{\mathbf{r}} = \left(\mathbf{I}+\frac{\beta}{\hat{\alpha}}\mathbf{J}+\frac{1}{\hat{\alpha}}\mathbf{H}_e^H\mathbf{H}\right)\mathbf{s} + \frac{1}{\hat{\alpha}}\hat{\mathbf{H}}^H\tilde{\mathbf{n}} \tag{3.30}$$

The term $\frac{\beta}{\hat{\alpha}}\mathbf{J} + \frac{1}{\hat{\alpha}}\mathbf{H}_e^H\mathbf{H}$ can be viewed as the coupling matrix which induced the interference from the other symbols. The first term in the coupling matrix is generated by the codeword itself while the second term caused by ICE. Once again this form is similar to the one in (3.16) and the bootstrap decoder can be used to iteratively decrease both types of interference simultaneously. The fact that the bootstrap can handle the regular decoding (even without ICE) of QSTBC is very appealing. It basically implies that ones the ICE effect is dealt by applying the bootstrap decoding technique, the regular ZF decoder ($\hat{\mathbf{H}}^{-1}$) can be replaced with the less computational

demanding filter, $\hat{\mathbf{H}}^H$, and the bootstrap will overcome the inherent inference induced by this filter. As a result, applying the bootstrap algorithm in this case reduces the decoding complexity by omitting the need for the matrix inversion of the QSTBC decoder.

3.2.3 Implementation

Various implementation modes of the bootstrap algorithm as cross coupling reduction method for STBC with ICE, are detailed in this section. A real versus complex implementation will be discussed as well as the type of limiter used at the weights updating formula. Two examples will be given to highlight the different considerations that lead to the different implementations of the algorithm.

A. 4 Tx OSTBC Consider the following 4 Tx OSTBC given in Equation (2.51),

$$\mathbf{X}_4 = \begin{pmatrix} s_1 & s_2 & s_3 & 0 \\ -s_2^* & s_1^* & 0 & s_3 \\ -s_3^* & 0 & s_1^* & -s_2 \\ 0 & -s_3^* & s_2^* & s_1 \end{pmatrix} \tag{3.31}$$

Due to the fact that the rows of the codeword contains both conjugate and non-conjugate versions of the data symbols, the EVC model for this codeword is a different than the regular EVC model presented in Section 1.4.5. To be able to express the channel output $\mathbf{y}$, given by

$$\mathbf{y} = \mathbf{X}_4\mathbf{h} + \mathbf{n} \tag{3.32}$$

as the output of a linear system of the form

$$\tilde{\mathbf{y}} = \mathbf{H}\mathbf{s} + \tilde{\mathbf{n}} \tag{3.33}$$

one must convert the data symbols vector $\mathbf{s}$ to a real one, i.e., considering each complex data symbol as two real symbols,

$$\mathbf{s} = \begin{pmatrix} s_1 \\ s_2 \\ s_3 \end{pmatrix} \Rightarrow \tilde{\mathbf{s}} = \begin{pmatrix} Re(s_1) \\ Im(s_1) \\ Re(s_2) \\ Im(s_2) \\ Re(s_3) \\ Im(s_3) \end{pmatrix} \tag{3.34}$$

Now one can rewrite (3.32) as

$$\tilde{\mathbf{y}} = \mathbf{H}\tilde{\mathbf{s}} + \tilde{\mathbf{n}} \tag{3.35}$$

where the EVC matrix $\mathbf{H}$ have only real entries and given by

$$H = \begin{pmatrix} Re(h_1) & -Im(h_1) & Re(h_2) & -Im(h_2) & Re(h_3) & -Im(h_3) \\ Im(h_1) & Re(h_1) & Im(h_2) & Re(h_2) & Im(h_3) & Re(h_3) \\ Re(h_2) & Im(h_2) & -Re(h_1) & -Im(h_1) & Re(h_4) & -Im(h_4) \\ Im(h_2) & -Re(h_2) & -Im(h_1) & Re(h_1) & Im(h_4) & Re(h_4) \\ Re(h_3) & Im(h_3) & -Re(h_4) & Im(h_4) & -Re(h_1) & -Im(h_1) \\ Im(h_3) & -Re(h_3) & -Im(h_4) & -Re(h_4) & -Im(h_1) & Re(h_1) \end{pmatrix} \tag{3.36}$$

$\tilde{\mathbf{y}}$ and $\tilde{\mathbf{n}}$ are also real and have similar structure as $\tilde{\mathbf{s}}$.

B. 4 Tx QSTBC For the second example of STBC, consider the 4 Tx QSTBC given in Equation (3.9). The H matrix for this code is complex and given by

$$\mathbf{H} = \begin{pmatrix} h_1 & h_2 & h_3 & h_4 \\ -h_2^* & h_1^* & -h_4^* & h_3^* \\ -h_3^* & -h_4^* & h_1^* & h_2^* \\ h_4 & -h_3 & -h_2 & h_1 \end{pmatrix} \tag{3.37}$$

Using the regular OSTBC decoding for this code (as suggested in Section 3.2.2) results in the following coupling matrix

$$\frac{\beta}{\hat{\alpha}}\mathbf{J} + \frac{1}{\hat{\alpha}}\mathbf{H}_e^H\mathbf{H} \tag{3.38}$$

Since β, which given by $\beta = 2\Re\{h_1 h_4^* - h_2 h_3^*\}$ can be close to ± 1, the interference levels at the output of this decoder can be very high in comparison to the interference caused by the ICE.

The different implementations modes are now presented in light of the above examples.

3.2.3.1 Real / Complex. Originally the bootstrap algorithm was presented for real signals [38],[41]. Later on it was modified to handle complex signal as well [40]. The benefit of implementing the complex mode is mainly from hardware saving perspective. For STBCs, the decision on whether to implement the real or the complex version of the algorithm is mainly based on the EVC model. If the codeword structure enforce a real EVC model, it will be more natural to implement the real bootstrap algorithm. Such codeword is the 4 Tx OSTBC given in example A above. If the codeword allows complex EVC channel matrix one may choose if to remain in the complex space and implement the 'complex bootstrap' or to implement the 'real bootstrap' and double the size of the system due to the conversion from complex to real.

3.2.3.2 Hard / Soft Limiter. In the bootstrap algorithm, the 'limiter' is referred as the function used in the weight calculation process. More specifically, the limiter is the function that applied on each signal of the bootstrap's output before calculating the correlation between that signal and the rest of the signals. Consider the weights

calculation formula given in Equation (3.23) in more general form

$$\mathbf{w}_k(n+1) = \mathbf{w}_k(n) - \mu z_k^*(n) f\left(\mathbf{z}_k(n)\right) \tag{3.39}$$

The simplest implementation is the 'hard' limiter where the function f is simply the signum function. This works well for low level of interference and for small constellation size where a symbol can be approximated by its sign. For the system where the interference level maybe in the the same order as the signal itself (as the 4 Tx QSTBC in example B.), the simple hard limiter may not work and the algorithm will not converge in a fast and accurate manner. This is due to the fact that having high interference levels means that the coupling matrix may be close to singular. Instead, a 'soft' limiter was presented in [42]. The soft limiter is given by

$$f(x) = \begin{cases} -1 & , \ x \leq -\lambda \\ x & , \ -\lambda < x < \lambda \\ 1 & , \ x \geq \lambda \end{cases} \tag{3.40}$$

Although the implementation of this limiter is more involved than the simple hard limiter, it ensure robust convergence of the bootstrap algorithm even when for high interference levels.

3.3 Analytical Analysis

Due to its non-linear nature, it is not trivial to analytically analyze to performance of the bootstrap algorithm in terms of its converges rate and error performance. Nether the less, for small systems, some size analysis can be made, namely, one can calculate the optimal weights of the diagonalize matrix V, which sets an upper bound on the performance of the bootstrap algorithm for that case. To that end, the simple 2×2 Alamouti codeword is analyzed thoroughly.

3.3.1 Alamouti's Code with ICE

The use of the bootstrap algorithm to reduce the interference level caused by the use of the mismatched filter in the presence of ICE is demonstrated and analyzed in the section for the 2×2 Alamouti code (which was introduced in details in Section 1.4.2) [43]. The Almaouti's codeword is given by

$$\mathbf{X} = \begin{pmatrix} s_1 & s_2 \\ -s_2^* & s_1^* \end{pmatrix} \tag{3.41}$$

where s_i are the data symbols drawn from arbitrary complex modulation. The channel output is then given by

$$\mathbf{y} = \mathbf{X}\mathbf{h} + \mathbf{n} \tag{3.42}$$

where $\mathbf{h}$ and $\mathbf{n}$ are 2×1 vectors whose entries h_i and n_i are the channel coefficients and the additive noise respectively. Both h_i and n_i are zero mean complex Gaussian random variables, i.e., the channel is modeled as Rayleigh fading channel. The channel output can be written as the the output of a linear system (EVC model)

$$\tilde{\mathbf{y}} = \mathbf{H}\mathbf{s} + \tilde{\mathbf{n}} \tag{3.43}$$

where

$$\mathbf{H} = \begin{pmatrix} h_1 & h_2 \\ h_2^* & -h_1^* \end{pmatrix} \quad , \quad \mathbf{s} = \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} \tag{3.44}$$

and the relation between $\tilde{\mathbf{y}}$ and $\mathbf{y}$ is the following $\tilde{y}_1 = y_1$, $\tilde{y}_2 = y_2^*$. Similar relations hold for $\tilde{\mathbf{n}}$ and $\mathbf{n}$.

The regular decoding scheme for the Alamouti code is the use of $\frac{1}{\alpha}\mathbf{H}^H$ as the filter at the receiver, resulting in

$$\mathbf{r} = \frac{1}{\alpha}\mathbf{H}^H\tilde{\mathbf{y}} = \frac{1}{\alpha}\mathbf{H}^H\mathbf{H}\mathbf{s} + \tilde{\mathbf{n}} = \mathbf{s} + \frac{1}{\alpha}\mathbf{H}^H\tilde{\mathbf{n}} \tag{3.45}$$

where $\alpha = |h_1|^2 + |h_2|^2$. In the presence of ICE, an erroneous version of the channel coefficients is used at the receiver, resulting in the mismatched filter

$$\mathbf{r}_{mf} = \frac{1}{\alpha_e}\mathbf{H}_e^H\mathbf{y} \tag{3.46}$$

where

$$\mathbf{H}_e = \begin{pmatrix} h_1 + h_{1,e} & h_2 + h_{2,e} \\ (h_2 + h_{2,e})^* & -(h_1 + h_{1,e})^* \end{pmatrix} \tag{3.47}$$

and $\alpha_e = |h_1 + h_{1,e}|^2 + |h_2 + h_{2,e}|^2$. $h_{i,e}$ are the errors associate with each channel coefficient estimation. The output of the mismatched filter (3.46) can be written as

$$\begin{aligned}
\mathbf{r}_{mf} &= \tfrac{1}{\alpha_e}\begin{pmatrix} (h_1 + h_{1,e})^* & h_2 + h_{2,e} \\ (h_2 + h_{2,e})^* & -(h_1 + h_{1,e}) \end{pmatrix} \cdot \begin{pmatrix} h_1 & h_2 \\ h_2^* & -h_1^* \end{pmatrix} \mathbf{s} + \tfrac{1}{\alpha_e}H_e^H\mathbf{n} \\
&= \tfrac{1}{\alpha_e}\left(\alpha I + \begin{pmatrix} h_{1,e}^* & h_{2,e} \\ h_{2,e}^* & -h_{1,e} \end{pmatrix}\begin{pmatrix} h_1 & h_2 \\ h_2^* & -h_1^* \end{pmatrix}\right)\mathbf{s} + \tfrac{1}{\alpha_e}H_e^H\mathbf{n} \\
&= \tfrac{1}{\alpha_e}\left(\alpha I + \begin{pmatrix} \gamma & \delta \\ -\delta^* & \gamma^* \end{pmatrix}\right)\mathbf{s} + \tfrac{1}{\alpha_e}H_e^H\mathbf{n} \\
&= \tfrac{1}{\alpha_e}\begin{pmatrix} \alpha + \gamma & \delta \\ -\delta^* & \alpha + \gamma^* \end{pmatrix}\mathbf{s} + \underbrace{\frac{1}{\alpha_e}H_e^H\mathbf{n}}_{\psi}
\end{aligned} \tag{3.48}$$

where the following definition were used

$$\begin{aligned}
\gamma &= h_1 h_{1,e}^* + h_2^* h_{2,e} \\
\delta &= -h_1^* h_{2,e} + h_2 h_{1,e}^*
\end{aligned} \tag{3.49}$$

The filtered noise variance is given by

$$\mathrm{cov}\{\psi\} = E\{\frac{1}{\alpha_e^2}\mathbf{H}_e^H\mathbf{n}\mathbf{n}^H\mathbf{H}_e\} = \frac{1}{\alpha_e}\sigma^2\mathbf{I} \tag{3.50}$$

and the resulted SINR is given by

$$SINR = \frac{|\alpha+\gamma|^2}{|\delta|^2 + \alpha_e\frac{\sigma^2}{E_s}} \tag{3.51}$$

Applying the bootstrap algorithm on the output of the mismatched filter (as in Figure 3.1), results in the addition of the following to the decoding scheme

$$\mathbf{z} = \mathbf{V}\mathbf{r}_{mf} = (\mathbf{I}-\mathbf{W})\mathbf{r}_{mf} \tag{3.52}$$

In the 2×2 case this is simply

$$\begin{aligned}\mathbf{z} &= \tfrac{1}{\alpha_e}\begin{pmatrix} 1 & -w_{12} \\ -w_{21} & 1 \end{pmatrix}\begin{pmatrix} \alpha+\gamma & \delta \\ -\delta^* & \alpha+\gamma^* \end{pmatrix}\mathbf{s} + \tfrac{1}{\alpha_e}(I-W)H_e^H\mathbf{n} \\ &= \tfrac{1}{\alpha_e}\begin{pmatrix} \alpha+\gamma+w_{12}\delta^* & \delta - w_{12}(\alpha+\gamma^*) \\ -\delta^* - w_{21}(\alpha+\gamma) & \alpha+\gamma^* - w_{21}\delta \end{pmatrix}\mathbf{s} + \underbrace{\frac{1}{\alpha_e}(I-W)H_e^H\mathbf{n}}_{\psi_{bs}}\end{aligned} \tag{3.53}$$

The filtered noise variance at the output of the bootstrap decoder is given by

$$\begin{aligned}\mathrm{cov}\{\psi_{bs}\} &= E\{\tfrac{1}{\alpha_e^2}(\mathbf{I}-\mathbf{W})\mathbf{H}_e^H\mathbf{n}\mathbf{n}^H\mathbf{H}_e(\mathbf{I}-\mathbf{W})^H\} \\ &= \tfrac{1}{\alpha_e}\sigma^2\mathbf{I}(\mathbf{I}-\mathbf{W})(\mathbf{I}-\mathbf{W})^H \\ &= \tfrac{1}{\alpha_e}\sigma^2\begin{pmatrix} 1+|w_{12}|^2 & -w_{12}-w_{21}^* \\ -w_{21}-w_{12}^* & 1+|w_{21}|^2 \end{pmatrix}\end{aligned} \tag{3.54}$$

and the SINRs are given by

$$SINR_{bs}^{1} = \frac{|\alpha+\gamma+w_{12}\delta^{*}|^{2}}{|\delta-w_{12}(\alpha+\gamma^{*})|^{2}+\alpha_{e}\frac{\sigma^{2}}{E_{s}}(1+|w_{12}|^{2})}$$
$$SINR_{bs}^{2} = \frac{|\alpha+\gamma^{*}-w_{21}\delta|^{2}}{|-\delta^{*}-w_{21}(\alpha+\gamma)|^{2}+\alpha_{e}\frac{\sigma^{2}}{E_{s}}(1+|w_{21}|^{2})} \tag{3.55}$$

3.3.2 Weights Calculation

The bootstrap algorithm iteratively calculates the weights w_{12} and w_{21} resulting in a maximization of the SINR terms in Equation (3.55). In the 2×2 case, the optimal weights can be calculated analytically. The optimal weights can be used as an upper bound on the performance of the bootstrap decoder. In this particular example of the Alamouti code, it will also be shown that the bootstrap's iterative weights calculation method cannot converge to the optimal weights. This invokes the use of different method for the bootstrap's weights calculation.

3.3.2.1 Optimal Weights. The optimal weights can be defined as the weights that will maximize the achieved $SINR$. Consider $SINR_{bs}^{1}$, its derivative in accordance to w_{12} is given by

$$\frac{d}{dw_{12}}SINR_{bs}^{1} = \frac{d}{dw_{12}}\left\{\frac{(\alpha+\gamma)(\alpha+\gamma^{*})+w_{12}^{*}\delta(\alpha+\gamma)+w_{12}\delta^{*}(\alpha+\gamma^{*})+w_{12}w_{12}^{*}\delta\delta^{*}}{\delta\delta^{*}-w_{12}^{*}\delta(\alpha+\gamma)-w_{12}\delta^{*}(\alpha+\gamma^{*})+w_{12}w_{12}^{*}(\alpha+\gamma)(\alpha+\gamma^{*})+\alpha_{e}\frac{\sigma^{2}}{E_{s}}(1+w_{12}w_{12}^{*})}\right\} \tag{3.56}$$

Equating (3.56) to zero results in the following quadratic equation (in w_{12}^{*})

$$a_{1}(w_{12}^{*})^{2} + b_{1}w_{12}^{*} + c_{1} = 0 \tag{3.57}$$

where

$$\begin{aligned} a_{1} &= -\delta(\alpha+\gamma) \\ b_{1} &= \delta^{*}\delta - (\alpha+\gamma)(\alpha+\gamma^{*}) \\ c_{1} &= \delta^{*}(\alpha+\gamma^{*}) \end{aligned} \tag{3.58}$$

The two roots of this equation are given by

$$w_{12}^{1} = -\frac{\alpha+\gamma}{\delta^*}$$
$$w_{12}^{2} = \frac{\delta}{\alpha+\gamma^*} \tag{3.59}$$

It is easy to verify that the first root takes $SINR_{bs}^{1}$ to its minimum (namely zero) since it zeros the nominator while the second root is the maximum where the interference term is nulled out. Hence, the optimal w_{12} is given by

$$w_{12}^{opt} = \frac{\delta}{\alpha+\gamma^*} \tag{3.60}$$

Similar handling for w_{21} results in the value that maximizes $SINR_{bs}^{2}$, which is given by

$$w_{21}^{opt} = -\frac{\delta^*}{\alpha+\gamma} \tag{3.61}$$

It is worth noting that the optimal weights values are not influence by the nominal SNR value $\frac{\sigma^2}{E_s}$. In addition, since

$$w_{12}^{opt} = -(w_{21}^{opt})^* \tag{3.62}$$

holds, the filtered noise at the output of the bootstrap block remains white due to the fact that the off diagonal elements in (3.54) are zeros. Moreover, in the next section it will be shown that forcing (3.62) on the bootstrap's iterative weights control algorithm results in a process deadlock which basically makes it non applicable to this system.

3.3.2.2 Bootstrap's Weights Calculation. The bootstrap's weights control block adjusts the weights such that in the steady state the following holds

$$E\left\{z_1^* \mathrm{csgn}\{z_2\}\right\} = 0 \tag{3.63}$$

where the function csgn$\{\cdot\}$ is defined as

$$\text{csgn}\{x\} = \text{sgn}\{\Re(x)\} + i \cdot \text{sgn}\{\Im(x)\} \tag{3.64}$$

In practice, the weights are calculated iteratively using

$$w_{ij}^{(n+1)} = w_{ij}^{(n)} + \mu \left(z_i^* \text{csgn}\{z_j\}\right) \tag{3.65}$$

The problem with this method, when applied to the 2×2 Alamouti code, is that due to the code's symmetry, the correlation between z_1 and z_2 is always zero for any set of weights w_{ij} (this is true when forcing $w_{21} = -w_{12}^*$ which holds for the optimal weights),

$$\begin{aligned}
E\left\{z_1^* z_2\right\} = & \\
& = E\left\{([(\alpha+\gamma) + w_{12}\delta^*]s_1 + [\delta - w_{12}(\alpha+\gamma^*)]s_2)^* \right. \\
& \quad \left. \cdot([-\delta^* - w_{21}(\alpha+\gamma)]s_1 + [(\alpha+\gamma^*) - w_{21}\delta]s_2)\right\} \\
& = E\left\{([(\alpha+\gamma^*) + w_{12}^*\delta]s_1^* + [\delta^* - w_{12}^*(\alpha+\gamma)]s_2^*) \right. \\
& \quad \left. \cdot([-\delta^* - w_{21}(\alpha+\gamma)]s_1 + [(\alpha+\gamma^*) - w_{21}\delta]s_2)\right\} \\
& = [(\alpha+\gamma^*) + w_{12}^*\delta][-\delta^* - w_{21}(\alpha+\gamma)]E\{s_1^* s_1\} \\
& \quad + [\delta^* - w_{12}^*(\alpha+\gamma)][(\alpha+\gamma^*) - w_{21}\delta]E\{s_2^* s_2\} \\
& = E_s\left(-\delta^*(\alpha+\gamma^*) - w_{12}^*\delta^*\delta - w_{21}(\alpha+\gamma)(\alpha+\gamma^*)\right. \\
& \quad -w_{12}^* w_{21}\delta(\alpha+\gamma) + \delta^*(\alpha+\gamma^*) - w_{12}^*(\alpha+\gamma)(\alpha+\gamma^*) \\
& \quad \left. -w_{21}\delta^* + w_{12}^* w_{21}\delta(\alpha+\gamma)\right) \\
& = E_s(-w_{21} - w_{12}^*)\left[\delta^*\delta + (\alpha+\gamma)(\alpha+\gamma^*)\right] = 0
\end{aligned} \tag{3.66}$$

where in the derivation $E\{s_1^* s_2\} = E\{s_2^* s_1\} = 0$ and $w_{21} = -w_{12}^*$ were used.

This imposes a significant problem over the implementation of the bootstrap's iterative weights calculation method since theoretically the algorithm will stop after the first iteration because the update element $(z_i^* \text{csgn}\{z_j\})$ is zero for any initial weights values. In order to be able to implement the bootstrap algorithm for this case one need to find an alternative method for the weights calculation, that will be more robust in converging to the optimal weights shown above. Such a method is presented in the next section.

3.3.3 Orthogonal Training Sequences

The core idea beyond this method is to use the training sequences for the weights calculation. Usually an orthogonal block is used for the channel coefficients estimation. Consider, for example the use of the Alamouti codeword itself as the training block

$$\mathbf{X}_t = \begin{pmatrix} s_1 & s_2 \\ -s_2^* & s_1^* \end{pmatrix} \tag{3.67}$$

where s_i are the training symbols drawn from a given complex modulation with symbol power P. The receiver gets

$$\mathbf{y}_t = \mathbf{X}_t \mathbf{h} + \mathbf{n} \tag{3.68}$$

The matrix $\mathbf{X}_t$ is known to the receiver which in turn calculates

$$\mathbf{r}_t = \frac{1}{2P} \mathbf{X}_t^H \mathbf{y}_t = \mathbf{h} + \frac{1}{2P} \mathbf{X}_t^H \mathbf{n} \tag{3.69}$$

Thus one have $\hat{\mathbf{h}} = \mathbf{r}_t$ where

$$\begin{pmatrix} h_{1,e} \\ h_{2,e} \end{pmatrix} = \frac{1}{2P} \mathbf{X}_t^H \mathbf{n} \tag{3.70}$$

are the errors terms of the estimated channel coefficients. Now consider that s_is are vectors of length L and are chosen such that they are orthogonal to each other, i.e., their inner product is zero

$$\langle \mathbf{s}_1, \mathbf{s}_2 \rangle = 0 \tag{3.71}$$

After the first use of the training sequences to estimate $\hat{\mathbf{h}}$, one can build $\mathbf{H}_e$ and try to "decode" these sequences

$$\mathbf{r}_0 = \tfrac{1}{\alpha_e}\mathbf{H}_e^H \tilde{\mathbf{y}}_t = \tfrac{1}{\alpha_e} \begin{pmatrix} \alpha+\gamma & \delta \\ -\delta^* & \alpha+\gamma^* \end{pmatrix} \mathbf{s} + \underbrace{\frac{1}{\alpha_e}\mathbf{H}_e^H \tilde{\mathbf{n}}}_{\varphi} \tag{3.72}$$

where $\tilde{\mathbf{y}}_t$ and $\tilde{\mathbf{n}}$ are given by

$$\tilde{\mathbf{y}}_t = \begin{pmatrix} y_{t,1} \\ y_{t,2}^* \end{pmatrix} \quad ; \quad \tilde{\mathbf{n}} = \begin{pmatrix} n_1 \\ n_2^* \end{pmatrix} \tag{3.73}$$

The transformation from $\mathbf{y}_t$ to $\tilde{\mathbf{y}}_t$ enables the channel output to be viewed as the output of a linear system and to be handled accordingly. The noise term φ can be written as

$$\begin{aligned} \varphi &= \begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix} = \tfrac{1}{\alpha_e}\mathbf{H}_e^H \tilde{\mathbf{n}} \\ &= \tfrac{1}{\alpha_e} \begin{pmatrix} h_1 + h_{1,e} & h_2 + h_{2,e} \\ (h_2 + h_{2,e})^* & -(h_1 + h_{1,e})^* \end{pmatrix}^H \begin{pmatrix} n_1 \\ n_2^* \end{pmatrix} \end{aligned} \tag{3.74}$$

Consider $\mathbf{r}_{0,1}$;

$$\mathbf{r}_{0,1} = \frac{1}{\alpha_e}\left((\alpha+\gamma)\mathbf{s}_1 + \delta \mathbf{s}_2\right) + \varphi_1 \tag{3.75}$$

The optimal weights can be derived by calculating the following inner products

$$\begin{aligned}
\langle \mathbf{s}_1, \mathbf{r}_{0,1} \rangle &= \tfrac{1}{\alpha_e} \left((\alpha + \gamma) \langle \mathbf{s}_1, \mathbf{s}_1 \rangle + \delta \langle \mathbf{s}_1, \mathbf{s}_2 \rangle \right) + \langle \mathbf{s}_1, \varphi_1 \rangle \\
&= \tfrac{1}{\alpha_e} (\alpha + \gamma) LP \\
\langle \mathbf{s}_2, \mathbf{r}_{0,1} \rangle &= \tfrac{1}{\alpha_e} \left((\alpha + \gamma) \langle \mathbf{s}_2, \mathbf{s}_1 \rangle + \delta \langle \mathbf{s}_2, \mathbf{s}_2 \rangle \right) + \langle \mathbf{s}_2, \varphi_1 \rangle \\
&= \tfrac{1}{\alpha_e} \delta LP
\end{aligned} \tag{3.76}$$

where the following identities were used

$$\begin{aligned}
\langle \mathbf{s}_i, \mathbf{s}_j \rangle &= 0 \\
\langle \mathbf{s}_i, \mathbf{s}_i \rangle &= LP \\
\langle \mathbf{s}_i, \varphi_i \rangle &= 0
\end{aligned} \tag{3.77}$$

Similarly, using $\mathbf{r}_{0,2}$;

$$\mathbf{r}_{0,2} = \frac{1}{\alpha_e} \left(-\delta^* \mathbf{s}_1 + (\alpha + \gamma^*) \mathbf{s}_2 \right) + \varphi_2 \tag{3.78}$$

The inner products will yield

$$\begin{aligned}
\langle \mathbf{s}_1, \mathbf{r}_{0,2} \rangle &= \tfrac{1}{\alpha_e} \left((\alpha + \gamma^*) \langle \mathbf{s}_1, \mathbf{s}_2 \rangle - \delta^* \langle \mathbf{s}_1, \mathbf{s}_1 \rangle \right) + \langle \mathbf{s}_1, \varphi_2 \rangle \\
&= -\tfrac{1}{\alpha_e} \delta^* LP \\
\langle \mathbf{s}_2, \mathbf{r}_{0,2} \rangle &= \tfrac{1}{\alpha_e} \left((\alpha + \gamma^*) \langle \mathbf{s}_2, \mathbf{s}_2 \rangle - \delta^* \langle \mathbf{s}_2, \mathbf{s}_1 \rangle \right) + \langle \mathbf{s}_2, \varphi_1 \rangle \\
&= \tfrac{1}{\alpha_e} (\alpha + \gamma^*) LP
\end{aligned} \tag{3.79}$$

Hence, one can estimate $\alpha+\gamma$ and δ with

$$\begin{aligned}\widehat{\alpha+\gamma} &= \tfrac{\alpha_e}{2LP}\left(\langle \mathbf{s}_1, \mathbf{r}_{0,1}\rangle + \langle \mathbf{s}_2, \mathbf{r}_{0,2}\rangle^*\right)\\ \hat{\delta} &= \tfrac{\alpha_e}{2LP}\left(\langle \mathbf{s}_2, \mathbf{r}_{0,1}\rangle - \langle \mathbf{s}_1, \mathbf{r}_{0,2}\rangle^*\right)\end{aligned} \tag{3.80}$$

The optimal weights can be simply constructed with $\widehat{\alpha+\gamma}$ and $\hat{\delta}$ according to (3.60) and (3.61),

$$\begin{aligned}\tilde{w}_{12}^{opt} &= \frac{\hat{\delta}}{\widehat{\alpha+\gamma}^*}\\ \tilde{w}_{21}^{opt} &= -\frac{\hat{\delta}^*}{\widehat{\alpha+\gamma}}\end{aligned} \tag{3.81}$$

Figure 3.2 shows the promising performance when using the new weights calculation method. In the simulation, the Alamouti code was used with 16-QAM modulation. The total transmission length is 512 Alamouti blocks while 16 blocks were used as pilots. The solid black curve represents the performance of a system with perfect knowledge of the channel coefficients at the receiver and is given as a reference. The broken black curve represents the performance of system using only the mismatched filter without any additional processing. The red curve is the best theoretical performance of the bootstrap decoding using the analytically calculated optimal weights. The blue curve is the achieved performance of the bootstrap algorithm with the new weights calculation method. It can be clearly shown how the bootstrap algorithm improves the performance of a system with ICE as well as how the performance of the new method is close to the analytical optimal weights performance.

3.4 Advanced Bootstrap Implementation

Even though the bootstrap decoding method was shown to improve the error rate of system with ICE, it has one major drawback. For proper operation, the decoding

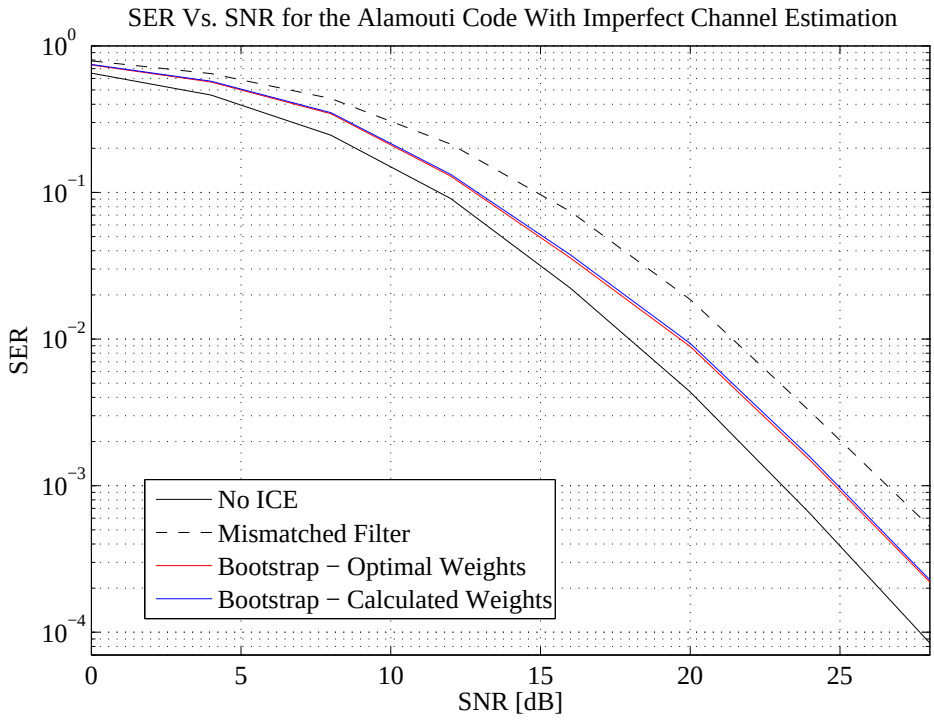

Figure 3.2 SER Vs. SNR for the Alamouti code, 16-QAM with length of 512 blocks and pilot length of 16 blocks.

algorithm assumes that the data sequences are statistically independent. In practice, to have an empirical correlation (as in (3.23)) resembles the theoretical correlation (3.22), one needs to use a very long vectors. The requirement for large data vectors cannot be satisfied given a short coherence time of the channel. Since the underlying assumption for any STBC system is that the channel coefficients do not change faster than the transmission time of a code block, having fast changing channel put limits on the block length. Moreover, even if the channel allows transmission of lengthy data sequence, the computation load required to handle these long vectors becomes significant. In this section, an improve scheme for better use of the bootstrap algorithm is presented. Inspired by the use of orthogonal training sequences with the Alamouti scheme (Section 3.3.3) it is suggested to transmit orthogonal data sequences [44]. These sequences can be used for the bootstrap's weights calculation, which in

turn enables the use of the bootstrap algorithm even for very short data sequences. This also results in a substantial computational burden reduction.

3.4.1 Orthogonal Data Vectors

In order to ensure fast and accurate converges of the bootstrap algorithm one may suggest the use of orthogonal data vectors. Forcing the transmitted data vectors to be orthogonal to each other, eases the requirement of long vector which become orthogonal due to statistic properties. Due to the obvious rate loss by using orthogonal vectors, it is suggested not to force the whole transmitted vector to be orthogonal. Rather, only short section of each data vector satisfies the orthogonality requirement hence, the rate penalty is minimized.

The improved bootstrap decoding scheme is modified in both the transmitter and receiver ends. The transmitted block is adjusted to include a short portion of orthogonal data sequence (Figure (3.3)). The decoding comprise of the mismatched filter followed by the bootstrap algorithm where the iterative weights calculation is performed only with the short orthogonal sequence. The calculated weights are then applied to the rest (the non-orthogonal part) of the output vector.

3.4.1.1 Transmission. Consider a data vector s to be transmitted using a STBC in a system with N transmit antennas. The data is split to k vectors s_i, where $k = R \cdot N$ and R is the code's rate (for example with 4 Tx and Orthogonal STBC with rate 3/4, $k = 3$). In general, there is no requirement on the relation between the different s_is and they are assumed independent of each other. The main idea is to require that a small portion of each vector s_i will be orthogonal to the corresponding portion in the other vectors. Defining

$$s_i = [s_i^o \quad s_i^d] \tag{3.82}$$

where s_i^o is the orthogonal portion and $\bar{s}_i$ is the rest of the data vector. The requirement is that

$$(s_i^o)^* \cdot s_j^o = 0 \quad , \quad i \neq j \tag{3.83}$$

Note that this is not a statistic attribute but rather a deterministic uncorrelated sequences.

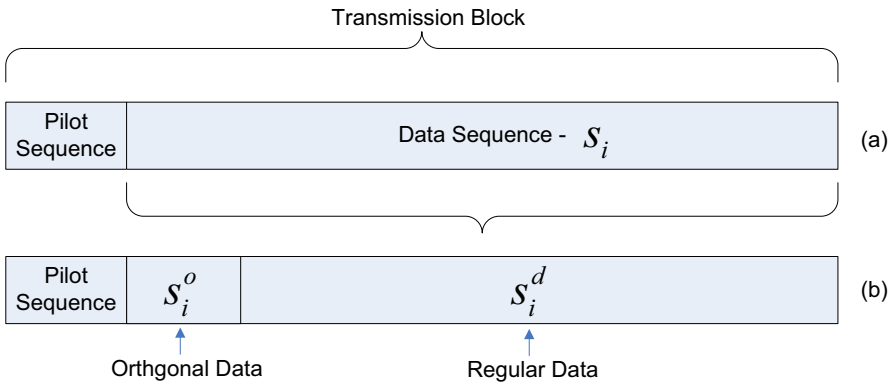

Figure 3.3 Transmission block, (a) old structure, (b) proposed new structure containing a portion of orthogonal sequence.

3.4.1.2 Decoding. With the orthogonal vectors the bootstrap decoder is implemented as follows. The weights control part (3.23) of the bootstrap algorithm is applied only to the orthogonal portion of the data vectors. This enables fast and robust convergence due to the enforced orthogonality of the inputs. In addition the computational load of the algorithm is significantly reduced since the orthogonal portions of the data vectors are much shorter relative to the total data vector length. The algorithm produces V, the diagonalize matrix which in turn applied to the rest of the data vector for the interference reduction.

This works well since V is the matrix that diagonalize the matrix $\left(I + \frac{1}{\hat{\alpha}} H_e^H H\right)$ which entries are function of the channel coefficients. Even though the calculation of the weights w_{ij} is done on the orthogonal portions, the channel coefficients aren't

changing throughout the whole block transmission hence it they can be applied later to the rest of the data vector. It will be shown later that for data vectors with no orthogonal requirement, the bootstrap algorithm performs bad since the assumption of orthogonality doesn't hold. The concept of computing the matrix V using only the orthogonal portion can be viewed as if 'pilot' sequences are transmitted and dedicated for the calculation of V. Even though the receiver doesn't know the the transmitted symbols of the orthogonal sequences, it takes advantage of their orthogonal property to compute the matrix V.

3.4.1.3 Rate loss. Forcing a portion of the data vectors to be orthogonal inevitably results in some rate loss. Plainly speaking, the rate is dropping by a factor of k since instead of having k independent sequences, one have only one independent sequence and $k-1$ sequences that are dependent. With more sophisticated encoding, one might save some of lost rate. This can be done, for example, by assigning date values to the different orthogonal sequences. Let L^o be the length of the orthogonal sequences. Defining U to be the cardinality of the group of distinguish orthogonal vector of length L^o that one can generate with symbols over some given constellation. By assigning a value for each vector of the orthogonal vector group, one can 'save' an additional $log_2(U)$ bits. In addition, if there are several groups, one use the group selection as another method for data delivery. Yet, even with these techniques, the rate loss cannot be ignored and it is desirable to minimize the length of the orthogonal portion of the data vector such that the rate loss is also minimized.

3.4.1.4 Simulations. While involving with iterative algorithms can be sometimes hard to analyze rigorously, simulations show the potential of the presented method. The settings are 4 Tx, 1 Rx system with a rate 3/4 OSTBC and 16-QAM modulation. Each data vector (s_i) is of length 264. The black and blue lines in Figure (3.4) represent the performance of a system without ICE, i.e., system with perfect channel

state information at the receiver, and of the mismatched decoding respectively. The performance of the regular bootstrap decoder (broken red line) clearly implies that even a vector length of 264 is not long enough for the independent data sequences assumption to hold. Hence, not only the decoder has a computational load involving a vector of length 264, it also not converging fast (i.e., large number of iterations that adds to the computation burden) and not to the right weights even for large SNR values.

With the new method, an orthogonal portion only occupies a length of 8 symbols. The performance of the new decoder (solid red line) shows that even this short sequence is enough for the weights calculation algorithm to converge such that signals separation is achieved. Thus, not only enabling the use of the bootstrap algorithm in a scenario of short data vector in term of acceptable performance, a significant computation load reduction is also achieved.

The total rate loss for this settings is limited to about 2% of the total symbol rate. While for the old method the symbol rate transmission is $\frac{3}{4} \cdot 264 = 198$, the new transmission rate is $\frac{3}{4}(\frac{1}{3} \cdot 8 + 256) = 194$. The rate loss $4/198 = 0.202$ is acceptable in light of the gains achieved both in the error rate and the computational load of the decoder.

3.4.2 Zero Rate-loss Implementation

The main draw back of the use of orthogonal data sequences for the bootstrap's weights calculation is the rate loss. This is due to the inherent data redundancy of the orthogonal structure. In this section, a innovative approach is presented, wherein, no rate loss is incurred in the system while the weights calculation is still performed using orthogonal sequences. The idea is to extract orthogonal sequences out of the data itself and use it for calculate the bootstrap's weights. since no measure is taken at

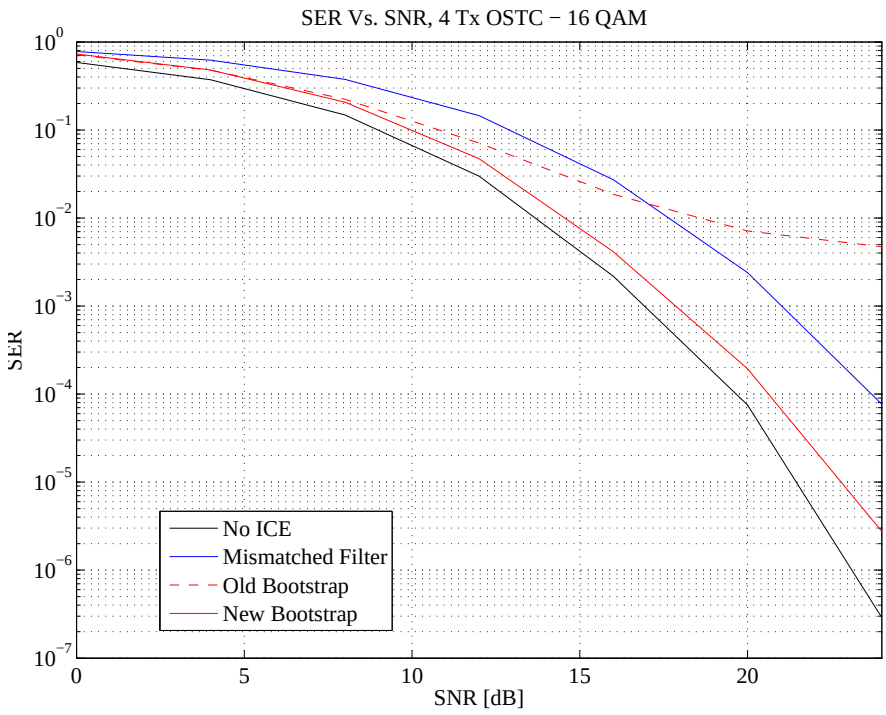

Figure 3.4 SER Vs. SNR for 4 Tx system with OSTBC encoding, 16-QAM modulation. The data vector size is 264 comprises of an orthogonal portion (8) and regular data (256).

the transmitter end to ensure data orthogonality, there is no rate loss. The proposed decoding scheme comprise of the following steps -

1. Mismatched filter.
2. Orthogonal sequences extraction.
3. Weight's calculation.
4. Bootstrap algorithm.

At the first decoding phase, the mismatched filter is used to get an initial data recovery. Out of the recovered data, orthogonal data is extracted to form a set of orthogonal sequences which are then used for the bootstrap's weights calculation. Once the weights are calculated, the bootstrap diagonalize matrix is applied to the output of mismatched filter to improve the performance by reducing the interference

Table 3.1 Example for Orthogonal Data Extraction

Time Index	1	2	3	4	5	6	7	8
Data Stream 1	-1	1	1	-1	-1	1	-1	1
Data Stream 2	1	-1	1	-1	1	1	1	-1

levels.

In order to understand the concept of orthogonal data extraction, consider to following example. Consider two data streams with BPSK symbols shown in Table 3.1, For orthogonal sequences of length four, one may simply choose the symbols with time indices 2,3,6 and 8 resulting in the two orthogonal sequences

$$\begin{aligned} \mathbf{s}_1^o = &\quad [1 \quad 1 \quad 1 \quad 1] \\ \mathbf{s}_2^o = &\quad [-1 \quad 1 \quad 1 \quad -1] \end{aligned} \tag{3.84}$$

This can be simply expand to larger constellation size and to more than two data streams. Obviously, as the number of data sequences increase, it is harder to find a set of indices that will form an orthogonal set for all the different sequences. The strength of this method is that it suffices to find orthogonal sequences only in a pairwise fashion and calculate the weights. In other words, the weights calculation based on orthogonal sequences can work well even if the data sets which apply to it is only pairwise orthogonal and not orthogonal to all other sequences. This may cause slower converges, but enables the implementation of the bootstrap algorithm with any rate loss.

3.4.2.1 Simulations. The same setup as in Section 3.4.1 was used for simulating the new, zero rate loss scheme. A data block size of 512 symbols was used and an orthogonal sequences of length eight were extracted from it. The difference from the last simulation shown in Figure 3.4 is that in this simulation the new method for

extracting orthogonal sequences from the data is added. This shown in Figure 3.5 by the red curve with star marks. Although the new method performs worse than the one which contains transmitted orthogonal sequences (solid red curve), it still performs much better than the old bootstrap implementation with less computational complexity and with no rate loss compared to the method that transmits orthogonal data. The new method enhances the performance of the mismatched filter by an order of magnitude in the high SNR region. It is worth noting that the new method for extracting orthogonal sequences out the decoded data performs poorly in the low SNR region. This is due to the fact that in this region the initial decoded data has high error rate resulting in erroneous orthogonal sequences, i.e., data that is not really orthogonal, which in turn results in erroneous weights.

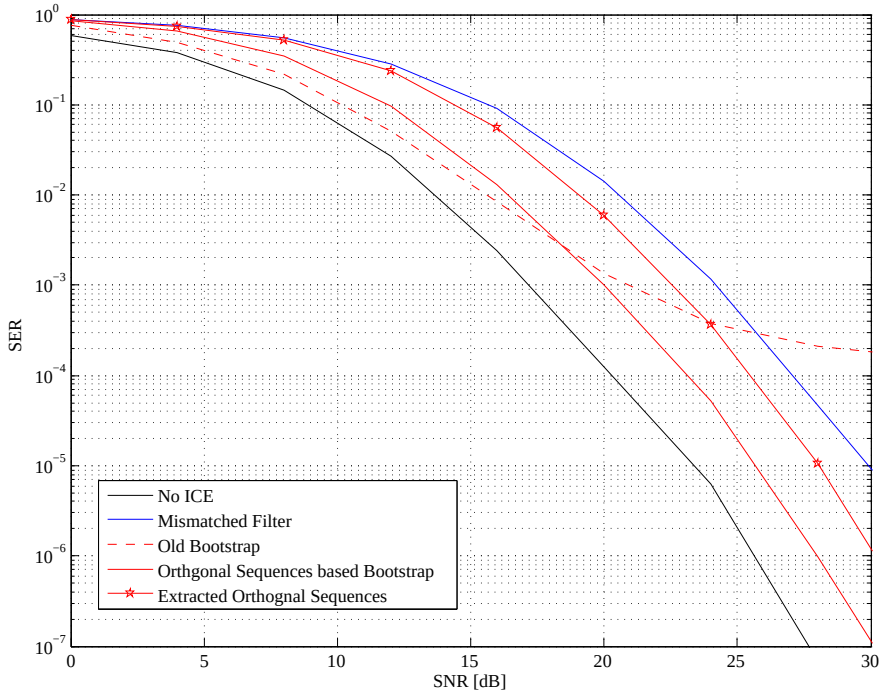

Figure 3.5 SER Vs. SNR for different bootstrap implementations, 4 Tx system with OSTBC encoding, 16-QAM modulation. The data vector size is 512 with an orthogonal portion length of 8.

CHAPTER 4

CONCLUSIONS

In this work the semi-orthogonal space time codes (SSTBC) was presented and showed to enjoy full rate along with linear computational decoding complexity. The transmission and decoding schemes was presented and analyzed and further improvements were suggested involving sequential decoding as well as modified schemes for system with limited feedback or multiple transmit antennas. An iterative method for generating OSTBC that complies with the requirement of the new decoding / transmission schemes was also presented. This new OSTBC is suitable for any number of transmit antennas and can be used easily as the basis for the SSTBC. Comparing the performance to a system with full rate but with non-linear decoding complexity, namely the QSTBC with ZF decoding, it was shown that the achieved error rate matched and even exceeds the QTSBC's error rate performance for most of the setting that were simulated. Thus, making the proposed SSTBC very appealing as it enjoys full rate transmission, linear computational decoding complexity and high performance as well.

In addition, the bootstrap algorithm was adopted as an adaptive method for suppressing the interference levels caused by the use of mismatched filter due to imperfect channel estimation. This method was thoroughly analyzed using the Alamouti 2×2 codeword as a case study. It was shown that for the Alamouti code, the bootstrap's weight calculation scheme will not converges to the optimal solution, hence, an alternate method for the bootstrap's weights was presented which employs the use of orthogonal sequences. Inspired by the use of orthogonal sequences as a mean for the weights calculation, this concept was further expanded to general STBC to overcome two major practical issues regarding the implementation of the bootstrap

decoding. Both issues emerge from the need of relatively long data sequences for the bootstrap algorithm to converges. For channels with short coherence time long data blocks cannot be transmitted. Even, when long data blocks are available, the computational burden of the weights calculation using the regular bootstrap's scheme over long data vectors is too high. To that end, using very short orthogonal data segments for the weights calculation was proposed which dramatically reduces the computational burden and even more importantly, enables to apply the bootstrap algorithm for short data block. To further mitigate the rate loss due to the use of orthogonal data sequences, a novel method for extracting orthogonal segments out of the transmitted data was introduced, thus, eliminating the need of dedicated orthogonal data segments which reduces the rate loss to zero.

To conclude, the field of STC in general and STBC in particular have been widely and thoroughly studied. Although the general framework and the theoretical boundaries have been defined, there is always room for improvement and new ideas that enables both improved STBC systems and better understanding towards the implementation of such systems. This work introduces these type of new ideas which not only were shown to enhance existing STBC system but are also simple to implement and can be incorporated into current STBC systems with no major adjustments.

APPENDIX A

ORTHOGONALITY OF THE NEW OSTBC

To show that the proposed code is orthogonal, one need to show that

$$\mathbf{X}_n^H \mathbf{X}_n = \alpha_n \mathbf{I}_n \qquad \forall n \tag{A.1}$$

where $\alpha_n = \sum_{i=1}^{n} |s_i|^2$.

Assuming that $\mathbf{X}_n$ is indeed an orthogonal code it will be shown by induction that $\mathbf{X}_{n+1}$ is also orthogonal.

By the code formulation, the expression for $\mathbf{X}_{n+1}$ is given by

$$\mathbf{X}_{n+1} = \begin{pmatrix} \mathbf{X}_n & \mathbf{b} \\ \mathbf{C} & -\mathbf{d} \end{pmatrix} \tag{A.2}$$

where $\mathbf{b}_{t\times 1} = (s_{n+1} \;\; \mathbf{0}_{1\times(t-1)})^T$, $\mathbf{C}_{n\times n} = s_{n+1}^* \cdot \mathbf{I}_n$ and $\mathbf{d}_{n\times 1} = (s_1^* \;\; s_2^* \;\; ... \;\; s_n^*)^T$. To verify that X_{n+1} is also orthogonal the product $\mathbf{X}_{n+1}^H \mathbf{X}_{n+1}$ is calculated

$$\begin{aligned} \mathbf{X}_{n+1}^H \mathbf{X}_{n+1} &= \begin{pmatrix} \mathbf{X}_n & \mathbf{b} \\ \mathbf{C} & -\mathbf{d} \end{pmatrix}^H \begin{pmatrix} \mathbf{X}_n & \mathbf{b} \\ \mathbf{C} & -\mathbf{d} \end{pmatrix} = \begin{pmatrix} \mathbf{X}_n^H & \mathbf{C}^H \\ \mathbf{b}^H & -\mathbf{d}^H \end{pmatrix} \begin{pmatrix} \mathbf{X}_n & \mathbf{b} \\ \mathbf{C} & -\mathbf{d} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{X}_n^H \mathbf{X}_n + \mathbf{C}^H \mathbf{C} & \mathbf{X}_n^H \mathbf{b} - \mathbf{C}^H \mathbf{d} \\ \mathbf{b}^H \mathbf{X}_n - \mathbf{d}^H \mathbf{C} & \mathbf{b}^H \mathbf{b} + \mathbf{d}^H \mathbf{d} \end{pmatrix} \end{aligned} \tag{A.3}$$

By the definition of $\mathbf{C}, \mathbf{b}$ and $\mathbf{d}$ the elements of the above matrix are given by

$$
\begin{aligned}
&(i)\ \mathbf{X}_n^H \mathbf{X}_n + \mathbf{C}^H \mathbf{C} = \alpha_n \mathbf{I}_n + |s_{n+1}|^2 \mathbf{I}_n = \alpha_{n+1} \mathbf{I}_n \\
&(ii)\ \mathbf{b}^H \mathbf{X}_n - \mathbf{d}^H \mathbf{C} = (s_{n+1}^* \ \ \mathbf{0}) \mathbf{X}_n - s_{n+1}^* \mathbf{d}^H = s_{n+1}^* (s_1 \ \ s_2 \ \ ... \ \ s_n) - s_{n+1}^* \mathbf{d}^H = \mathbf{0}^H \\
&(iii)\ \mathbf{X}_n^H \mathbf{b} - \mathbf{C}^H \mathbf{d} = (\mathbf{b}^H \mathbf{X}_n - \mathbf{d}^H \mathbf{C})^H = \mathbf{0} \\
&(iv)\ \mathbf{b}^H \mathbf{b} + \mathbf{d}^H \mathbf{d} = |s_{n+1}|^2 + \textstyle\sum_{i=1}^{n} |s_i|^2 = \alpha_{n+1}
\end{aligned}
\tag{A.4}
$$

Putting it all together results in

$$
\mathbf{X}_{n+1}^H \mathbf{X}_{n+1} = \begin{pmatrix} \alpha_{n+1} \mathbf{I}_n & \mathbf{0} \\ \mathbf{0}^H & \alpha_{n+1} \end{pmatrix} = \alpha_{n+1} \mathbf{I}_{n+1} \tag{A.5}
$$

Hence, given an orthogonal codeword $\mathbf{X}_n$, that was generated through the proposed algorithm, the next generated codeword $\mathbf{X}_{n+1}$ will also be orthogonal. A starting point for this induction can be $\mathbf{X}_2$ which is the famous Alamouti code and is known to be orthogonal. This concludes the proof that for any n this algorithm generate an orthogonal codeword $\mathbf{X}_n$.

APPENDIX B

DERIVATION OF THE SIMPLIFIED B MATRIX

B.1

The derivation of (2.32) is as follows. It will be shown for the 4 Tx codeword (presented in Eq. (2.36)) but can be simply generalized to any number of transmit antennas. The key feature is that for the new code the following holds

$$\alpha\mathbf{I} - \mathbf{H}_d\mathbf{H}_d^H = |h_1|^2\mathbf{I} + \mathbf{w}\mathbf{w}^H \tag{B.1}$$

where $\mathbf{w} = [h_4 \;\; -h_3 \;\; h_2]^T$. This is due to the way the codeword $\mathbf{X}_4$ is constructed and the similar structure of $\mathbf{X}_4$ and the equivalent channel matrix $\mathbf{H}$ (see (2.36) and (2.38)). One can use (B.1) to calculate the inverse of $\alpha\mathbf{I} - \mathbf{H}_d\mathbf{H}_d^H$.

$$(\alpha\mathbf{I} - \mathbf{H}_d\mathbf{H}_d^H)^{-1} = (|h_1|^2\mathbf{I} + \mathbf{w}\mathbf{w}^H)^{-1} \tag{B.2}$$

Applying the following matrix identity

$$(\mathbf{A}^{-1} + \mathbf{u}\mathbf{v}^H)^{-1} = \mathbf{A} - \mathbf{A}\mathbf{u}\mathbf{v}^H\mathbf{A}/(1 + \mathbf{v}^H\mathbf{A}\mathbf{u}) \tag{B.3}$$

with $\mathbf{A} = \frac{1}{|h_1|^2}\mathbf{I}$ and $\mathbf{u} = \mathbf{v} = \mathbf{w}$ results in

$$(|h_1|^2\mathbf{I} + \mathbf{w}\mathbf{w}^H)^{-1} = \frac{1}{|h_1|^2}\mathbf{I} - \frac{\frac{1}{|h_1|^2}\mathbf{w}\mathbf{w}^H\frac{1}{|h_1|^2}}{1 + \mathbf{w}^H\frac{1}{|h_1|^2}\mathbf{I}\mathbf{w}} \tag{B.4}$$

since $\mathbf{w}^H\frac{1}{|h_1|^2}\mathbf{I}\mathbf{w} = \frac{1}{|h_1|^2}\mathbf{w}^H\mathbf{w} = \frac{|h_2|^2+|h_3|^2+|h_4|^2}{|h_1|^2} = \frac{\alpha-|h_1|^2}{|h_1|^2}$, one can write

$$\begin{aligned}(|h_1|^2\mathbf{I} + \mathbf{w}\mathbf{w}^H)^{-1} &= \frac{1}{|h_1|^2}\mathbf{I} - \frac{\frac{1}{|h_1|^2}\mathbf{w}\mathbf{w}^H\frac{1}{|h_1|^2}}{\frac{\alpha}{|h_1|^2}} = \frac{1}{|h_1|^2}\mathbf{I} - \frac{1}{\alpha|h_1|^2}\mathbf{w}\mathbf{w}^H \\ &= \frac{1}{\alpha|h_1|^2}\left(\alpha\mathbf{I} - \mathbf{w}\mathbf{w}^H\right)\end{aligned} \tag{B.5}$$

substituting $\mathbf{w}\mathbf{w}^H$ with $\alpha\mathbf{I} - \mathbf{H}_d\mathbf{H}_d^H - |h_1|^2\mathbf{I}$ (using (B.1)) results in

$$(|h_1|^2\mathbf{I} + \mathbf{w}\mathbf{w}^H)^{-1} = \frac{1}{\alpha|h_1|^2}\left(|h_1|^2\mathbf{I} + \mathbf{H}_d\mathbf{H}_d^H\right) \tag{B.6}$$

Returning to (B.2), one gets

$$(\alpha\mathbf{I} - \mathbf{H}_d\mathbf{H}_d^H)^{-1} = \frac{1}{\alpha|h_1|^2}\left(|h_1|^2\mathbf{I} + \mathbf{H}_d\mathbf{H}_d^H\right) \tag{B.7}$$

Plugging (B.7) into the the left hand side of (2.32), results in

$$\begin{aligned}(\alpha\mathbf{I} - \mathbf{H}_d\mathbf{H}_d^H)^{-1}\mathbf{H}_d\mathbf{H}_r^H &= \tfrac{1}{\alpha|h_1|^2}\left(|h_1|^2\mathbf{I} + \mathbf{H}_d\mathbf{H}_d^H\right)\mathbf{H}_d\mathbf{H}_r^H \\ &= \tfrac{1}{\alpha}\mathbf{H}_d\mathbf{H}_r^H + \tfrac{1}{\alpha|h_1|^2}\mathbf{H}_d\mathbf{H}_d^H\mathbf{H}_d\mathbf{H}_r^H \\ &= \tfrac{1}{\alpha}\mathbf{H}_d\mathbf{H}_r^H + \tfrac{1}{\alpha|h_1|^2}(\alpha - |h_1|^2)\mathbf{H}_d\mathbf{H}_r^H\end{aligned} \tag{B.8}$$

The last expression holds since $\mathbf{H}_d\mathbf{H}_d^H\mathbf{H}_d = (\alpha - |h_1|^2)\mathbf{H}_d$. This can be verified by substituting $\mathbf{H}_d\mathbf{H}_d^H$ with $(\alpha - |h_1|^2)\mathbf{I} - \mathbf{w}\mathbf{w}^H$ (using (B.1)) and since $\mathbf{w}$ is orthogonal to $\mathbf{H}_d$ (i.e. $\mathbf{w}^H\mathbf{H}_d = \mathbf{0}$). simplifying the last term, results in the desired form of

$$(\alpha\mathbf{I} - \mathbf{H}_d\mathbf{H}_d^H)^{-1}\mathbf{H}_d\mathbf{H}_r^H = \frac{1}{|h_1|^2}\mathbf{H}_d\mathbf{H}_r^H \tag{B.9}$$

which conclude the derivations of (2.32).

B.2

The last step is to verify (2.33). This can simply done by noting the following properties of $\mathbf{H}, \mathbf{H}_r$ and $\mathbf{H}_d$.

1. All the rows of the matrix $\mathbf{H}$ are orthogonal to its first row.
2. $\mathbf{H}_r$ will always contain the first row of $\mathbf{H}$.
3. $\mathbf{H}_d$ will always contain a zero column.

The first property is due to the way $\mathbf{X}_n$ is generated and the similar structure of $\mathbf{X}_n$ and $\mathbf{H}$. The second and third properties are simply derived from the proposed REM rule. Since the rule is to delete the rows which contain null at the entries which corresponds to the chosen column, the resulted column in $\mathbf{H}_d$ will contain only zeros. Moreover, since the first row of $\mathbf{H}$ doesn't contain any null entries, this row will never be deleted i.e. will be in $\mathbf{H}_r$.

Rewriting $\mathbf{H}_r$ as

$$\mathbf{H}_r = \begin{pmatrix} h_1 & h_2 & h_3 & h_4 \\ -h_2^* & h_1^* & 0 & 0 \\ -h_3^* & 0 & h_1^* & 0 \\ -h_4^* & 0 & 0 & h_1^* \end{pmatrix} = \begin{pmatrix} \mathbf{h}_1 \\ -\mathbf{k} \;\; h_1^* \mathbf{I} \end{pmatrix} \tag{B.10}$$

where $\mathbf{h}_1$ is the first row of $\mathbf{H}$ and $\mathbf{k} = [h_2 \;\; h_3 \;\; h_4]^H$. One can rewrite $\mathbf{H}_d$ as

$$\mathbf{H}_d = \begin{pmatrix} 0 & -h_3^* & h_2^* & 0 \\ 0 & -h_4^* & 0 & h_2^* \\ 0 & 0 & -h_4^* & h_3^* \end{pmatrix} = \begin{pmatrix} \mathbf{0} & \bar{\mathbf{H}}_d \end{pmatrix} \tag{B.11}$$

where $\bar{\mathbf{H}}_d$ is the remaining part of $\mathbf{H}_d$ after removing the null column. Referring to the left hand side of (2.33), one gets

$$\begin{aligned}
\tfrac{1}{|h_1|^2}\mathbf{H}_d\mathbf{H}_r^H &= \tfrac{1}{|h_1|^2}\mathbf{H}_d \begin{pmatrix} \mathbf{h}_1 \\ -\mathbf{k} \quad h_1^*\mathbf{I} \end{pmatrix}^H = \tfrac{1}{|h_1|^2}\mathbf{H}_d \begin{pmatrix} \mathbf{h}_1^H & -\mathbf{k}^H \\ & h_1\mathbf{I} \end{pmatrix} \\
&= \tfrac{1}{|h_1|^2}\left(\mathbf{H}_d\mathbf{h}_1^H \quad \mathbf{H}_d \begin{pmatrix} -\mathbf{k}^H \\ h_1\mathbf{I} \end{pmatrix}\right) = \tfrac{1}{|h_1|^2}\left(\mathbf{H}_d\mathbf{h}_1^H \quad \begin{pmatrix} \mathbf{0} & \bar{\mathbf{H}}_d \end{pmatrix} \begin{pmatrix} -\mathbf{k}^H \\ h_1\mathbf{I} \end{pmatrix}\right) \\
&= \tfrac{1}{|h_1|^2}\left(\mathbf{H}_d\mathbf{h}_1^H \quad -\mathbf{0}\cdot\mathbf{k}^H + \bar{\mathbf{H}}_d h_1\mathbf{I}\right) = \tfrac{1}{|h_1|^2}\left(\mathbf{0} \quad \mathbf{0} + h_1\bar{\mathbf{H}}_d\right) \\
&= \tfrac{h_1}{|h_1|^2}\left(\mathbf{0} \quad \bar{\mathbf{H}}_d\right) = \tfrac{1}{h_1^*}\mathbf{H}_d
\end{aligned} \tag{B.12}$$

where using the fact that $\mathbf{H}_d$ rows were originated from $\mathbf{H}$, hence, they are orthogonal to the its first row (property 1) resulting in $\mathbf{H}_d\mathbf{h}_1^H = \mathbf{0}$. This concludes the derivations of (2.32) and (2.33) which shows how the the first step of the decoder can be implemented with linear computational complexity. For other choices of $\mathbf{H}_r$ and $\mathbf{H}_d$ (i.e choosing different columns for the REM rule) these derivations will remain the same up to some columns displacement in the corresponding matrices.

APPENDIX C

FILTERED NOISE VARIANCE CALCULATION

To evaluate the diagonal of the filtered noise covariance matrix one need to calculate the inverse of the matrix $\mathbf{H}_r^H \mathbf{H}_r$. To simplify it one cab rewrite the inverse as

$$(\mathbf{H}_r^H \mathbf{H}_r)^{-1} = \mathbf{H}_r^{-1} (\mathbf{H}_r^H)^{-1} = \mathbf{H}_r^{-1} (\mathbf{H}_r^{-1})^H \tag{C.1}$$

hence, only the calculation of the inverse of $\mathbf{H}_r$ is needed. To that end, one can rewrite $\mathbf{H}_r$ as

$$\mathbf{H}_r = \begin{pmatrix} h_1 & h_2 & h_3 & h_4 \\ -h_2^* & h_1^* & 0 & 0 \\ -h_3^* & 0 & h_1^* & 0 \\ -h_4^* & 0 & 0 & h_1^* \end{pmatrix} = \begin{pmatrix} h_1 & \mathbf{h}^H \\ -\mathbf{h} & h_1^* I \end{pmatrix} \tag{C.2}$$

where $\mathbf{h} = [h_2 \;\; h_3 \;\; h_4]^H$. Using a known block matrix inversion formula one gets

$$\mathbf{H}_r^{-1} = \frac{1}{\alpha} \begin{pmatrix} h_1^* & -\mathbf{h}^H \\ \mathbf{h} & \frac{1}{h_1^*}(\alpha \mathbf{I} - \mathbf{h}\mathbf{h}^H) \end{pmatrix} \tag{C.3}$$

and

$$(\mathbf{H}_r^{-1})^H = \frac{1}{\alpha} \begin{pmatrix} h_1 & \mathbf{h}^H \\ -\mathbf{h} & \frac{1}{h_1}(\alpha \mathbf{I} - \mathbf{h}\mathbf{h}^H) \end{pmatrix} \tag{C.4}$$

The product of these two matrices is given by

$$\begin{aligned}\mathbf{H}_r^{-1}(\mathbf{H}_r^{-1})^H &= \tfrac{1}{\alpha}\begin{pmatrix} h_1^* & -\mathbf{h}^H \\ \mathbf{h} & \frac{1}{h_1^*}(\alpha\mathbf{I}-\mathbf{h}\mathbf{h}^H)\end{pmatrix}\tfrac{1}{\alpha}\begin{pmatrix} h_1 & \mathbf{h}^H \\ -\mathbf{h} & \frac{1}{h_1}(\alpha I-\mathbf{h}\mathbf{h}^H)\end{pmatrix} \\ &= \tfrac{1}{\alpha^2}\begin{pmatrix} |h_1|^2+\mathbf{h}^H\mathbf{h} & h_1^*\mathbf{h}^H-\mathbf{h}^H\frac{1}{h_1}(\alpha\mathbf{I}-\mathbf{h}\mathbf{h}^H) \\ \mathbf{h}h_1-\frac{1}{h_1^*}(\alpha\mathbf{I}-\mathbf{h}\mathbf{h}^H)\mathbf{h} & \mathbf{h}\mathbf{h}^H+\frac{1}{h_1^*}(\alpha\mathbf{I}-\mathbf{h}\mathbf{h}^H)\frac{1}{h_1}(\alpha\mathbf{I}-\mathbf{h}\mathbf{h}^H)\end{pmatrix}\end{aligned} \tag{C.5}$$

Focusing in the diagonal of this matrix (since only the noise powers are of interest), the first element of the diagonal is

$$\frac{1}{\alpha^2}(|h_1|^2+\mathbf{h}^H\mathbf{h}) = \frac{1}{\alpha^2}(|h_1|^2+\sum_{i=2}^{4}|h_i|^2) = \frac{\alpha}{\alpha^2} = \frac{1}{\alpha} \tag{C.6}$$

The other three element of the diagonal are the diagonal of the matrix given by

$$\frac{1}{\alpha^2}(\mathbf{h}\mathbf{h}^H+\frac{1}{|h_1|^2}(\alpha\mathbf{I}-\mathbf{h}\mathbf{h}^H)^2) = \frac{1}{\alpha^2}(\mathbf{h}\mathbf{h}^H+\frac{1}{|h_1|^2}(\alpha^2\mathbf{I}-2\alpha\mathbf{h}\mathbf{h}^H+(\mathbf{h}\mathbf{h}^H)^2)) \tag{C.7}$$

The diagonals of $\mathbf{h}\mathbf{h}^H$ and $(\mathbf{h}\mathbf{h}^H)^2$ are given by

$$\begin{aligned}\text{diag}(\mathbf{h}\mathbf{h}^H) &= [|h_2|^2 \;\; |h_3|^2 \;\; |h_4|^2]^T \\ \text{diag}((\mathbf{h}\mathbf{h}^H)^2) &= diag(\mathbf{h}(\mathbf{h}^H\mathbf{h})\mathbf{h}^H) = \text{diag}(\mathbf{h}(\textstyle\sum_{i=2}^{4}|h_i|^2)\mathbf{h}^H) = \\ &= \textstyle\sum_{i=2}^{4}|h_i|^2 diag(\mathbf{h}\mathbf{h}^H) = \sum_{i=2}^{4}|h_i|^2[|h_2|^2 \;\; |h_3|^2 \;\; |h_4|^2]^T\end{aligned} \tag{C.8}$$

which results in

$$
\begin{aligned}
&diag(\tfrac{1}{\alpha^2}(\mathbf{h}\mathbf{h}^H + \tfrac{1}{|h_1|^2}(\alpha I - \mathbf{h}\mathbf{h}^H)^2)) = \\
&= \tfrac{1}{\alpha^2}\left(\begin{bmatrix} |h_2|^2 \\ |h_3|^2 \\ |h_4|^2 \end{bmatrix} + \tfrac{1}{|h_1|^2}\left(\alpha^2\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - 2\alpha\begin{bmatrix} |h_2|^2 \\ |h_3|^2 \\ |h_4|^2 \end{bmatrix} + \textstyle\sum_{i=2}^{4}|h_i|^2\begin{bmatrix} |h_2|^2 \\ |h_3|^2 \\ |h_4|^2 \end{bmatrix}\right)\right) = \\
&= \tfrac{1}{\alpha^2|h_1|^2}\left(\alpha^2\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \alpha\begin{bmatrix} |h_2|^2 \\ |h_3|^2 \\ |h_4|^2 \end{bmatrix}\right) = \tfrac{1}{\alpha|h_1|^2}\left(\alpha\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} |h_2|^2 \\ |h_3|^2 \\ |h_4|^2 \end{bmatrix}\right) = \\
&= \tfrac{1}{\alpha|h_1|^2}\begin{pmatrix} |h_1|^2 + |h_3|^2 + |h_4|^2 \\ |h_1|^2 + |h_2|^2 + |h_4|^2 \\ |h_1|^2 + |h_2|^2 + |h_3|^2 \end{pmatrix} = \tfrac{1}{\alpha}\begin{pmatrix} 1 + \left(\frac{|h_3|}{|h_1|}\right)^2 + \left(\frac{|h_4|}{|h_1|}\right)^2 \\ 1 + \left(\frac{|h_2|}{|h_1|}\right)^2 + \left(\frac{|h_4|}{|h_1|}\right)^2 \\ 1 + \left(\frac{|h_2|}{|h_1|}\right)^2 + \left(\frac{|h_3|}{|h_1|}\right)^2 \end{pmatrix}
\end{aligned}
\tag{C.9}
$$

Combining these results with the one derived in (C.6), results in the diagonal of covariance matrix $\mathbf{K}_v$

$$
\mathrm{diag}(\mathbf{K}_v) = \mathrm{diag}(\sigma^2(\mathbf{H}_r^H\mathbf{H}_r)^{-1}) = \frac{\sigma^2}{\alpha}\begin{pmatrix} 1 \\ 1 + \left(\frac{|h_3|}{|h_1|}\right)^2 + \left(\frac{|h_4|}{|h_1|}\right)^2 \\ 1 + \left(\frac{|h_2|}{|h_1|}\right)^2 + \left(\frac{|h_4|}{|h_1|}\right)^2 \\ 1 + \left(\frac{|h_2|}{|h_1|}\right)^2 + \left(\frac{|h_3|}{|h_1|}\right)^2 \end{pmatrix} \tag{C.10}
$$

which concludes the derivation of the filtered noise variances.

APPENDIX D

PROBABILITY DENSITY FUNCTIONS EVALUATION

D.1

The p.d.f of $x = |h_i|^2$ is a Chi square with two degrees of freedom which is given by $\chi_2^2(x) = \frac{1}{2}e^{-\frac{1}{2}x}$. The p.d.f of the ratio of two χ_2^2 r.vs can be calculated as ($x_i = |h_i|^2$, $x_j = |h_j|^2$)

$$y = \frac{x_j}{x_i} \tag{D.1}$$

$$f_y(y) = \frac{1}{4}\int_0^\infty \left(x_i e^{-\frac{1}{2}yx_i} e^{-\frac{1}{2}x_i}\right) dx_i = \frac{1}{(y+1)^2} \quad , \quad y \geq 0$$

The r.v. $z = 1 + y$ is distributed

$$f_z(z) = f_y(z-1) = \frac{1}{(z)^2} \quad , \quad z \geq 1 \tag{D.2}$$

Finally the p.d.f of the r.v. $v = z^{-1}$ is given by

$$f_v(v) = \frac{1}{v^2} f_z(\frac{1}{v}) = \frac{1}{v^2}\frac{1}{(\frac{1}{v})^2} = 1 \quad , \quad 0 \geq v \geq 1 \tag{D.3}$$

which is the uniform distribution.

D.2

Starting with the calculation of the p.d.fs of the ith orders statistic ($i = 1, 2, 3$) of $\{|h_2|^2\ |h_3|^2\ |h_4|^2\}$, one have

$$f_{h_{(1)}}(x) = \frac{3!}{0!2!}\left(F(x)\right)^0 \left(1 - F(x)\right)^2 f(x) \tag{D.4}$$

where $F(x) = 1 - e^{-\frac{1}{2}x}$ and $f(x) = \frac{1}{2}e^{-\frac{1}{2}x}$. This reduces to

$$f_{h_{(1)}}(x) = 3\left(e^{-\frac{1}{2}x}\right)^2 \frac{1}{2}e^{-\frac{1}{2}x} = \frac{3}{2}e^{-\frac{3}{2}x} \tag{D.5}$$

Similarly, the p.d.fs of $h_{(2)}$ and $h_{(3)}$ are

$$\begin{aligned} f_{h_{(2)}}(x) &= \tfrac{3!}{1!1!}\left(F(x)\right)^1\left(1-F(x)\right)^1 f(x) \\ &= 6\left(1-e^{-\frac{1}{2}x}\right)e^{-\frac{1}{2}x}\tfrac{1}{2}e^{-\frac{1}{2}x} \\ &= 3\left(e^{-x}-e^{-\frac{3}{2}x}\right) \\ f_{h_{(3)}}(x) &= \tfrac{3!}{2!0!}\left(F(x)\right)^2\left(1-F(x)\right)^0 f(x) \\ &= 3\left(1-e^{-\frac{1}{2}x}\right)^2\tfrac{1}{2}e^{-\frac{1}{2}x} \\ &= \tfrac{3}{2}\left(e^{-\frac{1}{2}x}-2e^{-x}+e^{-\frac{3}{2}x}\right) \end{aligned} \tag{D.6}$$

The first decoded symbol is the one associated with the largest SNR which in turn is a function of the following r.v.

$$\nu_{(1)} = \frac{|h_1|^2}{|h_1|^2+h_{(1)}} = \left(1+\frac{h_{(1)}}{|h_1|^2}\right)^{-1} \tag{D.7}$$

To find its p.d.f consider the ratio $z = \frac{h_{(1)}}{|h_1|^2} = \frac{x}{y}$;

$$\begin{aligned} f_z(z) &= \textstyle\int_0^\infty y f_{xy}(yz,y)dy = \int_0^\infty y f_x(yz)f_y(y)dy \\ &= \textstyle\int_0^\infty \frac{3}{2}ye^{-\frac{3}{2}yz}\frac{1}{2}e^{-\frac{1}{2}y}dy = \frac{3}{4}\int_0^\infty ye^{-y(\frac{3}{2}z+\frac{1}{2})}dy \\ &= \frac{\frac{3}{4}}{\left(\frac{3}{2}z+\frac{1}{2}\right)^2} = \frac{3}{(3z+1)^2} \quad , \quad z \geq 0 \end{aligned} \tag{D.8}$$

The r.v. $w = 1 + z$ is distributed

$$\begin{aligned} f_w(w) &= f_z(w-1) = \tfrac{3}{(3(w-1)+1)^2} \\ &= \tfrac{3}{(3w-2)^2} \quad , \quad w \geq 1 \end{aligned} \tag{D.9}$$

finally $\nu_{(1)} = w^{-1}$ is distributed

$$\begin{aligned} f_{\nu_{(1)}}(\nu_{(1)}) &= \tfrac{1}{\nu_{(1)}^2} f_w(\tfrac{1}{\nu_{(1)}}) = \tfrac{1}{\nu_{(1)}^2} \tfrac{3}{\left(3\frac{1}{\nu_{(1)}}-2\right)^2} \\ &= \tfrac{3}{\left(3-2\nu_{(1)}\right)^2} \quad , \quad 0 \geq \nu_{(1)} \geq 1 \end{aligned} \tag{D.10}$$

Similar steps can be taken to calculate the p.d.f of the r.v. associated with the second large SNR -

$$\nu_{(2)} = \frac{|h_1|^2}{|h_1|^2 + x_{(2)}} = \left(1 + \frac{x_{(2)}}{|h_1|^2}\right)^{-1} \tag{D.11}$$

Once again starting with the ratio $z = \frac{x_{(2)}}{|h_1|^2} = \frac{x}{y}$;

$$\begin{aligned} f_z(z) &= \textstyle\int_0^\infty y f_{xy}(yz, y)dy = \int_0^\infty y f_x(yz) f_y(y)dy \\ &= \textstyle\int_0^\infty 3y\left(e^{-yz} - e^{-\frac{3}{2}yz}\right)\frac{1}{2}e^{-\frac{1}{2}y}dy = \frac{3}{2}\int_0^\infty ye^{-y(z+\frac{1}{2})}dy - \frac{3}{2}\int_0^\infty ye^{-y(\frac{3}{2}z+\frac{1}{2})}dy \\ &= \tfrac{3}{2}\left(\tfrac{1}{\left(z+\frac{1}{2}\right)^2} - \tfrac{1}{\left(\frac{3}{2}z+\frac{1}{2}\right)^2}\right) \quad , \quad z \geq 0 \end{aligned} \tag{D.12}$$

The r.v. $w = 1 + z$ is distributed

$$f_w(w) = f_z(w-1) = \frac{3}{2}\left(\frac{1}{\left(w-\frac{1}{2}\right)^2} - \frac{1}{\left(\frac{3}{2}w-1\right)^2}\right) \quad , \quad w \geq 1 \tag{D.13}$$

and $\nu_{(2)} = w^{-1}$ is distributed

$$f_{\nu_{(2)}}(\nu_{(2)}) = \frac{1}{\nu_{(2)}^2} f_w(\frac{1}{\nu_{(2)}}) = \frac{3}{2}\left(\frac{1}{\left(1-\frac{1}{2}\nu_{(2)}\right)^2} - \frac{1}{\left(\frac{3}{2}-\nu_{(2)}\right)^2}\right) \quad , \quad 0 \geq \nu_{(2)} \geq 1 \tag{D.14}$$

For the last decoded symbol the p.d.f of $\nu_{(3)}$ needs to be derived,

$$\nu_{(3)} = |h_1|^2 + |h_3|^2 + |h_4|^2 = |h_1|^2 + x_{(1)} + x_{(2)} \tag{D.15}$$

To that end, the joint p.d.f of $x_{(1)}$ and $x_{(2)}$ needs to be found, where, to ease the notation burden y and z are defined as $y = x_{(1)}$ and $z = x_{(2)}$.

Using the joint order statistic p.d.f formula given by

$$f_{n,j,k}(y,z) = \left(\begin{array}{c} n \\ \\ j-1, 1, k-j-1, 1, n-k \end{array} \right) F(y)^{j-1} f(y) \left(F(z) - F(y)\right)^{k-j-1} f(z) \left(1 - F(z)\right)^{n-k} \tag{D.16}$$

for $1 \leq j \leq k \leq n$, $y \leq z$, one obtain

$$f_{3,1,2}(y,z) = \frac{3!}{0!1!0!1!1!} \frac{1}{2} e^{-\frac{1}{2}y} \frac{1}{2} e^{-\frac{1}{2}z} e^{-\frac{1}{2}z} = \frac{3}{2} e^{-\frac{1}{2}y} e^{-z} \quad , \quad y \leq z \tag{D.17}$$

The p.d.f of $w = y + z$ is given by

$$\begin{aligned} f_w(w) &= \int_0^{\frac{w}{2}} f_{y,z}(y, w-y) dy = \tfrac{3}{2} \int_0^{\frac{w}{2}} e^{-\frac{1}{2}y} e^{-(w-y)} dy \\ &= \tfrac{3}{2} e^{-w} \int_0^{\frac{w}{2}} e^{\frac{1}{2}y} = 3e^{-w} \left(e^{\frac{w}{4}} - 1\right) = 3\left(e^{-\frac{3}{4}w} - e^{-w}\right) \end{aligned} \tag{D.18}$$

The p.d.f of $\nu_{(3)} = |h_1|^2 + w = x + w$ can be calculated by

$$\begin{aligned} f_{\nu_{(3)}}(\nu_{(3)}) &= \int_0^{\nu_{(3)}} f_x(x) f_w(\nu_{(3)} - x) dx = \tfrac{3}{2} \int_0^{\nu_{(3)}} e^{-\frac{1}{2}x} \left(e^{-\frac{3}{4}(\nu_{(3)}-x)} - e^{-(\nu_{(3)}-x)}\right) dx \\ &= \tfrac{3}{2} \left[e^{-\frac{3}{4}\nu_{(3)}} \int_0^{\nu_{(3)}} e^{\frac{1}{4}x} dx - e^{-\nu_{(3)}} \int_0^{\nu_{(3)}} e^{\frac{1}{2}x} dx \right] \\ &= \tfrac{3}{2} \left[e^{-\frac{3}{4}\nu_{(3)}} 4 \left(e^{\frac{1}{4}\nu_{(3)}} - 1\right) - 2e^{-\nu_{(3)}} \left(e^{\frac{1}{2}\nu_{(3)}} - 1\right) \right] \\ &= 3\left(e^{-\nu_{(3)}} + e^{-\frac{1}{2}\nu_{(3)}} - 2e^{-\frac{3}{4}\nu_{(3)}}\right) \quad , \quad 0 \geq \nu_{(3)} \geq 1 \end{aligned} \tag{D.19}$$

D.3

υ can be written as

$$\upsilon = \alpha - h_{(3)} = |h_1|^2 + h_{(1)} + h_{(2)} \tag{D.20}$$

For the p.d.f calculation, one initially needs to find the joint p.d.f of $h_{(1)}$ and $h_{(2)}$. To ease the notation burden y and z are defined as $y = h_{(1)}$ and $z = h_{(2)}$.

Using the joint order statistic p.d.f formula given by

$$f_{n,j,k}(y,z) = \left(\begin{array}{c} n \\ \\ j-1, 1, k-j-1, 1, n-k \end{array} \right) \cdot \tag{D.21}$$

$$\cdot F(y)^{j-1} f(y) \left(F(z) - F(y)\right)^{k-j-1} f(z) \left(1 - F(z)\right)^{n-k}$$

for $1 \leq j \leq k \leq n$, $y \leq z$, this results in

$$\begin{aligned} f_{3,1,2}(y,z) &= \tfrac{3!}{0!1!0!1!1!} \tfrac{1}{2} e^{-\frac{1}{2}y} \tfrac{1}{2} e^{-\frac{1}{2}z} e^{-\frac{1}{2}z} \\ &= \tfrac{3}{2} e^{-\frac{1}{2}y} e^{-z} \quad , \quad y \leq z \end{aligned} \tag{D.22}$$

The p.d.f of $w = y + z$ is given by

$$\begin{aligned} f_w(w) &= \int_0^{\frac{w}{2}} f_{y,z}(y, w-y) dy = \tfrac{3}{2} \int_0^{\frac{w}{2}} e^{-\frac{1}{2}y} e^{-(w-y)} dy \\ &= \tfrac{3}{2} e^{-w} \int_0^{\frac{w}{2}} e^{\frac{1}{2}y} = 3e^{-w} \left(e^{\frac{w}{4}} - 1\right) \\ &= 3 \left(e^{-\frac{3}{4}w} - e^{-w}\right) \end{aligned} \tag{D.23}$$

The p.d.f of $\upsilon = |h_1|^2 + w = x + w$ can be calculated by

$$\begin{aligned} f_\upsilon(\upsilon) &= \int_0^\upsilon f_x(x) f_w(\upsilon - x) dx \\ &= \tfrac{3}{2} \int_0^\upsilon e^{-\frac{1}{2}x} \left(e^{-\frac{3}{4}(\upsilon - x)} - e^{-(\upsilon - x)} \right) dx \\ &= \tfrac{3}{2} \left[e^{-\frac{3}{4}\upsilon} \int_0^\upsilon e^{\frac{1}{4}x} dx - e^{-\upsilon} \int_0^\upsilon e^{\frac{1}{2}x} dx \right] \\ &= \tfrac{3}{2} \left[e^{-\frac{3}{4}\upsilon} 4 \left(e^{\frac{1}{4}\upsilon} - 1 \right) - 2e^{-\upsilon} \left(e^{\frac{1}{2}\upsilon} - 1 \right) \right] \\ &= 3 \left(e^{-\upsilon} + e^{-\frac{1}{2}\upsilon} - 2e^{-\frac{3}{4}\upsilon} \right) \end{aligned} \tag{D.24}$$

REFERENCES

[1] J. P. Proakis, *Digital Communications*, 4th ed. New York: McGraw Hill, 2001.

[2] A. Papoulis, *Probability, Random Variables and Stochastic Processes*, 4th ed. New York: McGraw Hill, 2002.

[3] W. C. Y. Lee, *Mobile Communications Engineering*. New York: McGraw Hill, 1982.

[4] W. C. Jakes, *Microwave Mobile Communications*. New York: Wiley, 1974.

[5] J. Guey, M. Fitz, M. Bell, and W. Kuo, "Signal design for transmitter diversity wireless communication systems over Rayleigh fading channels," *Proc. IEEE VTC*, pp. 136-140, 1996.

[6] N. Seshadri and J. H. Winters, "Two signaling schemes for improving the error performance of frequency-division-duplex (FDD) transmission systems using transmitter antenna diversity," *Vehicular Technology Conference*, pp. 508-511, 1993.

[7] V. Tarokh, , N. Seshadri, and A. R. Calderbank, "Spacetime codes for high data rate wireless communication: Performance criteria and code construction," *IEEE Trans. Information Theory*, vol. 44, pp. 744-765, Mar. 1998.

[8] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Information Theory*, vol. 45, pp. 1456-1467, Jul. 1999.

[9] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multielement antennas," *Bell Labs Tech. J.*, pp. 41-59, Autumn 1996.

[10] I. E. Telatar, "Capacity of multi antenna gaussian channels," AT&T Bell Labs., Tech. Rep., Jun. 1995.

[11] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, pp. 311-335, Mar. 1998.

[12] S. M. Alamouti, "A simple transmitter diversity scheme for wireless communications," *IEEE J. Selected Areas in Communications.*, vol. 16, pp. 1451-1458, Oct. 1998.

[13] A. Wittneben, "A new bandwidth efficient transmit antenna modulation diversity scheme for linear digital modulation," in *IEEE International Conference on Communications (ICC)*, vol. 3, May. 1993, pp. 1630-1634.

[14] X. B. Liang, "Orthogonal designs with maximum rates," *IEEE Trans. Information Theory*, vol. 49, pp. 2468-2503, Oct. 2003.

[15] K. Lu, S. Fu., and X.-G. Xia, "Closed form design of complex orthogonal space-time block codes of rates (k+1)/2k for 2k-1 and 2k transmit antennas," in *Proc. Int. Symp. Information Theory (ISIT 2004)*, Jun/Jul 2004, p. 307.

[16] H. Jafarkhani, "A quasi orthogonal space-time block code," *IEEE Trans. Communications.*, vol. 49, pp. 1-4, 2001.

[17] O. Tirkkonen, A. Boariu, and A. Hottinen, "Minimal nonorthogonality rate one space-time block codes for 3+ tx antennas," *IEE International Symposium on Spread Spectrum Techniques and Applications (ISSSTA)*, vol. 2, pp. 429-432, 2000.

[18] C. Papadias and G. Foschini, "Capacity-approaching space-time codes for system employing four transmitter antennas," *IEEE Trans. Information Theory*, vol. 49, no. 3, pp. 726-733, Mar. 2003.

[19] C. Mecklenbrauker and M. Rupp, "On extended Alamouti schemes for space-time coding," *WPMC02, Honolulu*, pp. 115-119, Oct. 2002.

[20] E. Agrell, T. Eriksson, A. Vardy, and K. Zeger, "Closest point search in lattices," *IEEE Trans. Information Theory*, vol. 48, no. 8, pp. 2201-2214, 2002.

[21] J. Jaldén and B. Ottersten, "On the complexity of sphere decoding in digital communications," *IEEE Transactions on Signal Processing*, vol. 53, no. 4, pp. 1474-1485, 2005.

[22] G. D. Golden, G. J. Foschini, R. A. Valenzuela, and P. W. Wolniansky, "Detection algorithm and initial laboratory results using the V-BLAST space-time communication architecture," *Electronics Letters*, vol. 35, no. 1, pp. 14-15, Jan. 1999.

[23] A. Laufer and Y. Bar-Ness, "Improved transmission scheme for orthogonal space time codes," in *Information Sciences and Systems, 2008. CISS 2008. 42nd Annual Conference on*, Mar. 2008, pp. 1108-1113.

[24] G. Ganesan and P. Stoica, "Spacetime block codes: A maximum SNR approach," *IEEE Trans. Information Theory*, vol. 47, no. 4, pp. 1650-1656, May 2001.

[25] M. A. Woodbury, "Inverting modified matrices," Memorandum Rept. 42, Statistical Research Group, Princeton University, Princeton, NJ, 1950.

[26] A. Laufer and Y. Bar-Ness, "Full rate space time codes for large number of transmitting antennas with linear complexity decoding," *Wireless Personal Communications*, vol. 57, pp. 465-480, 2011.

[27] X. Changlong, G. Gong, and K. Ben-Letaief, "High-rate complex orthogonal space-time block codes for high number of transmit antennas," in *IEEE International Conference on Communications (ICC)*, vol. 2, Jun. 2004, pp. 823-826.

[28] A. Laufer and Y. Bar-Ness, "Full rate space time codes for large number of transmitting antennas with linear complexity decoding and high performance," in *Information Theory Workshop, 2009. ITW 2009. IEEE*, Oct. 2009, pp. 416-420.

[29] X. B. Liang, "A high-rate orthogonal space-time block code," *IEEE Communications Letters*, vol. 7, pp. 222-223, May. 2003.

[30] A. Laufer and Y. Bar-Ness, "Linear computational complexity decoding for semi orthogonal full rate space time codes," in *Wireless Communications and Networking Conference (WCNC), 2011 IEEE*, Mar. 2011, pp. 1534-1539.

[31] A. Jain, A. Laufer, and Y. Bar-Ness, "On converting ostc scheme from non-full rate to full-rate with better error performance," in *Wireless Communication and Sensor Networks, 2008. WCSN 2008. Fourth International Conference on*, Dec. 2008, pp. 230-235.

[32] B. Badic, M. Rupp, and H. Weinrichter, "Quasi-orthogonal space-time block codes for data transmission over four and eight transmit antennas with very low feedback rate," in *5th International ITG Conference on Source and Channel Coding (SCC)*, Jan. 2004, pp. 157-164.

[33] ——, "High diversity with simple space time block-codes and linear receivers," in *Global Telecommunications Conference, 2003. GLOBECOM '03. IEEE*, vol. 7, Dec. 2003.

[34] ——, "Quasi-orthogonal space-time block codes: approaching optimality," in *13. European Signal Processing Conference*, Antalya, Turkey, Sep. 2005.

[35] D. Gu and C. Leung, "Performance analysis of transmit diversity scheme with imperfect channel estimation," *"Electronics Letters"*, vol. 39, pp. 402-403, Feb 2003.

[36] L. Zheng and D. N. C. Tse, "Communication on the grassman manifold: A geometric approach to the noncoherent multiple-antenna channel," *IEEE Trans. Information Theory*, vol. 48, pp. 359-383, Feb 2002.

[37] G. Taricco and E. Biglieri, "Space time decoding with imperfect channel estimation," *IEEE Trans. Wireless Commun*, vol. 4, pp. 1874-1888, Jul 2005.

[38] Y. Bar-Ness and J. Rokah, "Cross-coupled bootstrapped interference canceler," *Antenna and Propagation Symposium, Conf. Proc.*, pp. 292-295, 1981.

[39] A. Laufer and Y. Bar-Ness, "Adaptive decoding for space time codes with imperfect channel estimation, using the bootstrap algorithm," in *GLOBECOM Workshops, 2010 IEEE*, Dec. 2010, pp. 855-859.

[40] Y. Bar-Ness and N. VanWaes, "The complex bootstrap algorithm for blind separation of QAM multiuser CDMA signals," *Wireless Personal Commun.*, vol. 12, pp. 1-14, 2000.

[41] Y. Bar-Ness and J. Punt, "Adaptive 'bootstrap' CDMA multi-user detector," *Wireless Personal Communications: An International Journal, special issue on Signal Separation and Interference Cancellation for PIMRC*, vol. 3, pp. 55-71, 1996.

[42] N. Sezgin and Y. Bar-Ness, "Adaptive soft limiter bootstrap separator for one-shot asynchronous CDMA channel with singular partial cross-correlation matrix," in *IEEE International Conference on Communications (ICC)*, Dallas Tx, USA, Jun 1996, pp. 73-77.

[43] A. Laufer and Y. Bar-Ness, "Bootstrap decoding for the alamouti space-time scheme with imperfect channel estimation," in *Wireless and Optical Communications Conference (WOCC), 2011 20th Annual*, Apr. 2011, pp. 1-5.

[44] ——, "Improved bootstrap decoding scheme for space time codes with imperfect channel estimation," in *Information Sciences and Systems (CISS), 2011 45th Annual Conference on*, Mar. 2011, pp. 1-5.