

Gallager's Random Coding Exponent for Differential Space Time Modulation

by

Siddharth K. Srinivasan

B.E, Government College of Engineering, Pune (2001)

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF APPLIED SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

(Electrical and Computer Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA

March 2006

© Siddharth Srinivasan, 2006

Abstract

In this thesis, we study differential space-time modulation with multiple transmit and receive antennas over a frequency non-selective block fading channel in the absence of channel state information (CSI) at both the transmitter and the receiver. We focus on the analysis of the random coding exponent proposed by Gallager for such space-time channels employing differential modulation with finite signal sets, and consider *multiple symbol differential detection* (MSDD) with an observation window size N . The underlying principle of MSDD is to utilize an increased *observation window* of $N > 2$ consecutively received space-time samples to yield decision variables on $N - 1$ space-time data symbols. We thus take into account the channel memory, and can improve power efficiency over *conventional differential detection*, which employs $N = 2$.

We extend previous work for a similar setup, in that we consider channels where the fading coherence interval is different from the MSDD *observation window* N , and can be arbitrary with L independent fading realizations per coding frame. We analyze the effect of arbitrary fading coherence intervals on the random coding exponent, and therefore analyze the achievable performance of coded transmission over block fading channels with low to moderate code lengths, i.e. with decoding delay constraints. In this context we also analyze space-time transmission systems with spatial correlation between antennas. Such an analysis allows for a fair comparison of DSTM with MSDD where the *window size* may vary but the *coded diversity* remains fixed.

We also devise upper and lower bounds and approximations for the random coding exponent, which allows us to not only bound the achievable performance of such space-time systems but also allows for efficient numerical evaluation of the random coding exponent by limiting the search space for the metric calculation. To this end, we make use of tree-search based sphere decoding algorithms for efficient decoding, and along with novel *stopping criteria* to control the accuracy of the approximation and correspondingly the computational complexity, we apply these sphere decoders for a reduction in complexity cost upto many orders of magnitude.

The presented numerical results provide useful information on the performance of coded differential space-time transmission for short to moderate code lengths.

Contents

Abstract	ii
Contents	iv
List of Figures	vii
Acknowledgements	x
1 Introduction	1
2 Channel Models and Differential Encoding	6
2.1 Multiple Transmit and Receive Antenna Systems with MSDD	7
2.1.1 Discrete Time Channel Model for Multiple Transmit and Receive Antenna Transmission	8
2.1.2 Differential Space Time Modulation: Vector Channel Model . .	9
2.1.3 Single Antenna Transmission: A Special Case	13
2.2 Signal Constellations for Differential Space Time Modulation	15

2.2.1	Constellations from Group Codes: Cyclic Codes or Diagonal Signals	16
2.2.2	Constellations from Non-Group Codes	17
3	Gallager Random Coding Exponent for Differential Transmission and Noncoherent Reception	19
3.1	Introduction	19
3.2	Multiple Antenna Systems	20
3.2.1	Noncoherent Probability Density Functions	20
3.2.2	Random Coding Exponent and Channel Capacity	22
3.2.3	Reduced Complexity Metric to Calculate the Random Coding Exponent	28
3.2.4	Results and Discussion	30
3.3	Single Antenna Systems	40
3.3.1	Noncoherent SISO Channel PDF	40
3.3.2	Random Coding Exponent and Capacity Analysis	41
3.3.3	Performance Analysis: Results and Discussion	42
4	Modified Tree-Search and Sphere Decoding MSDD Algorithms	49
4.1	Single Antenna Systems: Multiple Symbol Differential Sphere Decoding (MSDSD)	51
4.2	Multiple Antenna Systems: Tree Search based Sphere Decoding Algorithms	53

4.2.1	Algorithms for Orthogonal Space-Time Codes	54
4.2.2	Algorithms for Diagonal Constellations	54
4.3	Bounds Using Tree-Search Algorithms	55
4.3.1	$D = 1$ and Spatially Uncorrelated Fading	57
4.3.2	$D > 1$ and Spatially Uncorrelated Fading	59
4.3.3	Upper Bound to the Random Coding Exponent	60
4.4	Performance Analysis and Results	60
4.4.1	Single Antenna Systems	61
4.4.2	Multiple Antenna Systems	65
5	Conclusions and Future Work	76
A	Derivation of Gallager's Function with Coherent metric	80
B	List of Symbols and Acronyms	84

List of Figures

2.1	<i>Overlapping of consecutive $2N_T \times N_T$ matrix symbols $\Phi[k]$. After right multiplication with the $N_T \times N_T$ submatrix $\mathbf{S}[k-1] = \prod_{\zeta=1}^{\infty} \mathbf{V}[k-\zeta]$ of the previous matrix symbol $\Phi[k-1]$, only the bracketed regions i.e. $N_T \times N_T$ matrices $\mathbf{S}[k]$ are transmitted.</i>	10
2.2	<i>Block diagrams for Differential Space-Time Modulation: Discrete Time Model and Vector Channel Model</i>	11
3.1	<i>Gallager Exponent vs Rate for DSTM, Orthogonal codes with CDD ($N = 2$), target rate $R_t = 1$ bit/(channel use): Effect of coding diversity</i>	31
3.2	<i>Required SNR ($10 \log_{10}(E_s/N_0)$) vs t_d in PSK samples, D4PSK and $C_{OD}(16)$. Dashed lines: DSTM, Solid lines: D4PSK</i>	32
3.3	<i>Required $10 \log_{10}(E_s/N_0)$ vs Decoding delay t_d in PSK samples. Target rate $R_t=1$ bit/(channel use), CDD ($N = 2$) and MSDD with $N = 3$. . .</i>	33
3.4	<i>Gallager Exponent results: Required SNR ($10 \log_{10}(E_s/N_0)$) vs Decoding delay t_d in PSK samples for DSTM with Orthogonal Codes $C_{OD}(16)$ showing the effect of increasing D while keeping N constant: Reduced diversity. Solid horizontal line: channel capacity for DSTM, $N = 3$. . .</i>	35

3.5	<i>Comparison between D4PSK and DSTM: Diversity Analysis</i>	36
3.6	<i>Capacity $C_{DSTM}(N)$ results for DSTM, Orthogonal Codes with modulation rate $R_m = 2$ bits/(channel use)</i>	38
3.7	<i>Capacity $C_{DSTM}(N)$ for DSTM</i>	39
3.8	<i>Capacity and Gallager functions for Single Antenna D4PSK</i>	43
3.9	<i>Gallager exponent $E(R, N, D)$ vs Rate R, effect of coding diversity . . .</i>	44
3.10	<i>Gallager random coding exponent results: Single Antenna D4PSK, delay-SNR analysis</i>	45
3.11	<i>Required $10 \log_{10}(E_s/N_0)$ vs coding delay t_d for single antenna D8PSK, CDD($N = 2$), fading coherence interval increasing with D</i>	47
3.12	<i>Gallager Exponent $E(R, N, D)$ vs R for D4PSK ($N_T = 1$) and $10 \log_{10}(E_s/N_0) = 10$ dB</i>	48
4.1	<i>Stack algorithm bounds for D4PSK: MSDD with 99% accuracy</i>	62
4.2	<i>Random Coding Exponent for D4PSK: Lower bounds using Stack Algorithm with different accuracies. CDD ($N = 2$) and MSDD ($N = 6$) with 90% and 99% accuracy</i>	63
4.3	<i>Complexity and random coding analysis of Stack Algorithm, D4PSK and D8PSK with 99,90 and 85% accuracy</i>	64
4.4	<i>$E(R, N, D)$ vs R for DSTM, different ST codes with $R_t = 1$ (bit/ch. use), spatially correlated antennas.</i>	66
4.5	<i>Required E_s/N_0 vs decoding delay t_d for DSTM with CDD, $C_{OD}(16)$ with $R_t = 1$ (bit/ch. use), spatially correlated antennas</i>	67

4.6	<i>Random coding exponent for DSTM:$\mathbf{C}_{OD}(16)$ and $\mathbf{C}_{DI}(16)$ codes . . .</i>	69
4.7	<i>.</i>	71
4.8	<i>$E(R, N, D)$ vs. R: D4PSK and $10 \log_{10}(E_s/N_0) = 10$ dB. (a): MSDD with different N. (b): $N = 6$. Exact expression and approximation for $E_r(R, N, D)$. $n_c/L = 40$.</i>	73
4.9	<i>Required $10 \log_{10}(E_b/N_0)$ for $P_e \leq 10^{-3}$ vs. n_c. 4DPSK ($N_T = 1$) and 16OD ($N_T = 2$). $n_c/L = 100$, $R = 1$ bit/(scalar channel use): upper bound approximation</i>	75

Acknowledgements

It gives me great pleasure to convey my gratitude to Dr. Lutz Lampe for his able guidance throughout the course of this work. I have learnt a lot from him and consider it a privilege to have worked with him. His constant encouragement and support helped me in overcoming the challenges I faced while pursuing this project. The numerous discussions we had were very fruitful and the team work was instrumental in allowing me to achieve the goals that I set out for. I would also like to thank all the professors of the Department of Electrical and Computer Engineering, specially Dr. Robert Schober, who taught me various courses and helped me enhance my knowledge.

I thank all my colleagues in the Communication Theory Group, Dept. of ECE, UBC for creating a friendly and stimulating environment. I would also like to thank my family for being very patient and for believing in me during all these years. Without their love and support I would have not been the person I am today.

Chapter 1

Introduction

The field of digital transmission systems has been long driven by the need for higher data rates, lower probabilities of error and reduced power consumption. Especially in the case of modern wireless communication systems, the quest for low power, bandwidth efficient digital communication systems has been the major area of research and development over the past years. The goal of this thesis is to contribute to the design and analysis of noncoherent wireless communication systems, and present a detailed analysis of the random coding exponent proposed by Gallager (see [10]) and its properties.

In digital communication systems, the overall system model of the transmitter, the channel under consideration and the receiver is given by the equation

$$r[k] = e^{j\theta[k]}h[k]s[k] + n[k] \quad (1.1)$$

relating the discrete time, complex valued data or transmit symbol $s[k]$ to the received symbol $r[k]$. A time variant phase shift is taken into account by the term $e^{j\theta[k]}$, which is introduced generally by oscillator instabilities or by Doppler effects. The additive noise component $n[k]$ is generally modeled as additive white Gaussian noise (AWGN). The techniques for detecting $r[k]$ from $s[k]$ can be broadly classified into two categories,

based on the knowledge about the phase shift $\theta[k]$. In *coherent detection*, $\theta[k]$ is assumed to be perfectly known, implying the existence of a phase-synchronization mechanism, often involving the transmission of pilot tones or symbols (cf. e.g. [31, 30, 33, 3]). On the other hand, *noncoherent detection* assumes no knowledge about the phase, but treats the estimation of the transmitted data as well as the channel phase as one operation. The signal fading, which is one of the major disturbances in mobile communications is represented by $h[k]$, and this enables us to distinguish between coherent detection with perfect channel state estimation (CSI) at the receiver, i.e., $e^{j\theta[k]}h[k]$ is presumed to be perfectly known by means of explicit channel estimation, and noncoherent detection without CSI where neither $\theta[k]$ nor $h[k]$ are explicitly estimated for the purpose of data selection. One of the common ways to implement a noncoherent receiver is to base the estimation of the data vector on blocks of $N > 2$ received symbols (see [43]), also known as multiple-symbol differential detection (MSDD). This form of noncoherent reception assumes that the receiver is aware of the statistical characterization of the channel.

The concept of noncoherent reception where explicit channel estimation is not required becomes even more attractive for the case of transmission systems employing multiple antennas at the transmit and receive points, where transmit diversity is achieved through the use of $N_T > 1$ antennas at the transmitter. Such space-time (ST) communication systems employing N_R receive antennas require extensive training intervals to estimate the $N_T \cdot N_R$ fading coefficients, and due to this restriction, space-time modulation formats which do not require the receiver to know or estimate the channel coefficients have been proposed (see [28, 14]). We use an approach known as *differential space-time modulation* [16, 15], which generalizes the notion of differential phase-shift keying (DPSK) to multiple antenna transmission, and is a low complexity scheme providing full-antenna diversity.

In this thesis, we study the application of bandwidth efficient noncoherent coded mod-

ulation to differential space-time coding for multiple transmit and receive antenna systems. We focus our research on the error exponent analysis of such systems, and consider a *block fading channel* of arbitrary fading coherence interval, i.e. where the channel remains constant over an arbitrary interval. This is an extension to previous work [24] on single antenna DPSK where the channel coherence interval was assumed to be equal to the MSDD observation interval N . Error exponent analysis is a powerful tool to study the achievable performance of coded transmission over fading channels with short to moderate code lengths due to e.g. delay constraints. In particular, random coding exponents have been used to bound error and/or outage probability for certain types of channels and different degrees of channel state information being available. In this context the block-fading channel model is often employed to (approximately) represent realistic fading channel scenarios with possibly non-ideal interleaving. More recently, random coding exponents for space-time (ST) modulation over block fading channels have been evaluated in [4, 18]. We consider Gallager's random coding exponent [10, Ch. 5] for ST transmission over block fading channels with no CSI at both the transmitter and the receiver. We consider bandwidth-efficient differential space-time modulation with practical signal constellations at the transmitter and power-efficient MSDD at the receiver.

Chapter 2 introduces the discrete time, frequency non-selective block fading channel model under consideration, and also introduces the *differential encoding* technique used to resolve the ambiguities in the data detection. The concept of MSDD is described in detail, where blocks of N consecutively received symbols are processed for joint detection of the corresponding data symbols. In particular, we consider spectrally efficient DPSK based encoding strategies for single antenna transmission and multiple antenna DSTM based on unitary signal elements. We also introduce the different signal constellations for DSTM that will be used throughout the thesis for analysis purposes.

Chapter 3 is dedicated to the random coding exponent and its properties as a tool for

error exponent analysis. We focus on absolute performance limits of coded transmission over block fading channels, and to this end we consider Gallager's random coding exponent and, to a lesser degree, channel capacity as appropriate information theoretical parameters. We derive expressions for the random coding exponent for DSTM, and widen the range of block fading channels that can be analyzed by extending the definition of the random coding exponent to channels with arbitrary fading coherence intervals. For the sake of completeness, we present results corresponding to the case of single antenna transmission as well as DSTM, and show the effect of arbitrary fading coherence intervals on the random coding exponent.

In Chapter 4, we present efficient decoding algorithms for DSTM with orthogonal and diagonal constellations. We describe in detail the need for such algorithms based on the analysis in Chapter 3, and describe how efficient tree-search based sphere decoding algorithms considerably reduce the computational complexity of the derived metrics, where the search space for the decoding algorithm increases exponentially with the length of the fading interval. In order to reduce this exponential dependency on the fading coherence interval, we devise modified sphere decoding algorithms (cf. e.g. [22]) to limit the complexity cost for ST transmission systems with arbitrarily large fading intervals. We also attempt to use a combination of the sphere decoding algorithms and a metric to upper and lower bound the random coding exponent and hence the probability of error. We devise a novel *performance metric* using the best maximum likelihood (ML) metrics obtained for the sphere decoding algorithms along with a *stopping criteria*, which not only allows us to bound the random coding exponent to a user-defined accuracy, but correspondingly also reduces the complexity cost for the decoder by several orders of magnitude, depending on the required accuracy. Finally in Appendix A, we derive in detail the expression for Gallager's random coding exponent used in this thesis, where the effect of transmit diversity is taken into account.

We present results and analysis separately for multiple transmit and receive antenna

systems and single antenna systems, and present our findings at the end of each chapter. We make a note of the fact that the focus of this thesis is on block fading frequency non-selective channels with arbitrary fading coherence intervals, so the class of broadband communication systems entailing a frequency-selective or time-dispersive channel are not directly addressed by this thesis, neither are channels with continuous fading. Nevertheless, by means of widely accepted multicarrier transmission technologies [2], which decompose a frequency-selective channel into a number of parallel frequency non-selective channels, the concepts proposed throughout the thesis can be applied to such broadband systems.

Chapter 2

Channel Models and Differential Encoding

This chapter briefly introduces the discrete time channel models and signal constellations used throughout the thesis for the single antenna systems considered as well as space-time transmission. In accordance with the proposed work, we distinguish between the single antenna channel with one transmit and one receive antenna and the multiple antenna channel with more than one transmit and/or receive antennas. The corresponding channels are thus referred to as single-input-single-output (SISO) and multiple-input-multiple-output (MIMO) channels respectively [23]. In section 2.1, we introduce the multiple transmit and receive antenna channel model that we will use throughout this thesis. We introduce the vector channel model of the system, and explain in detail the process of differential space-time modulation (DSTM) and multiple symbol differential decoding (MSDD) for noncoherent reception without channel state information (CSI). The block fading channel model for our MIMO channel is also described, along with the transmit diversity effect of the channel. We also describe, for reference, the single antenna channel model used in literature in section 2.1.3, and

introduce the basic channel model and differential encoding for such a SISO channel model. Finally in section 2.2, we introduce the space-time signal constellations that we use throughout the thesis for analysis purposes, and mention in brief the design of orthogonal and diagonal space-time block codes for DSTM.

2.1 Multiple Transmit and Receive Antenna Systems with MSDD

Multiple symbol differential decoding works on blocks of N consecutively received matrix symbols (in the case of DSTM, these symbols are $N_T \times N_T$ matrices) such that decision variables on data symbols are based on independent evaluations of these blocks cf. e.g. [42, 8, 25, 9]. Since, in general, computational complexity grows exponentially with N , a judicious choice of N is required to balance performance and complexity. Since in MSDD, a constant phase offset does not reflect in decision metrics, differential encoding techniques are required to resolve phase ambiguities. This process of differential encoding can be achieved by multiplying the data carrying *differential symbol* by the previously transmitted symbol. Let $\mathbf{S}[k]$ be the reference symbol, and $\mathbf{V}[k]$ be the data carrying differential symbol drawn from a signal set \mathcal{V} . The transmitted symbol at the current instant is then given by

$$\mathbf{S}[k] = \mathbf{V}[k] \cdot \mathbf{S}[k - 1] \quad (2.1)$$

This process of differential encoding is applied to blocks of N symbols for MSDD, and at the receiver, N consecutively received symbols are grouped together to yield a decision on $N - 1$ differential data symbols.

2.1.1 Discrete Time Channel Model for Multiple Transmit and Receive Antenna Transmission

In the previous section we introduced the concept of MSDD, and we use this concept to formulate our system model with multiple transmit and receive antennas. We consider space-time transmission with N_T transmit antennas and N_R receive antennas over a frequency non-selective block fading channel. We consider transmitting blocks of N space-time symbols over a particular fading subchannel, and D such blocks are transmitted over one fading realization. Since the total number of fading realizations is L , one codeword spans

$$n_v \triangleq LD(N - 1) \quad (2.2)$$

differential symbols, i.e. the code length is $n_c = n_v \log_2(M)$. This is an extension to work done in literature thus far, where the channel coherence interval and the MSDD observation window N were equal. Here, the channel remains constant for DN consecutive matrix symbols, and L such independent fading realizations are seen in one coding frame, thereby introducing a transmit *diversity* which is then used to parameterize our channel model. The transmitted symbols are $N_T \times N_T$ matrices denoted by $\mathbf{S}_{\ell,d,n}[k]$, where subscripts ℓ , d , and n , $1 \leq \ell \leq L$, $1 \leq d \leq D$, $1 \leq n \leq N$, specify the fading subchannel, the block transmitted over this subchannel, and the symbol position within this block, respectively, and $k \in \mathbb{Z}$ is the time-index with respect to frames of

$$n_s \triangleq LDN \quad (2.3)$$

symbols. The ℓ th subchannel is represented by the $N_T \times N_R$ channel matrix $\mathbf{H}_\ell[k]$. For a compact notation, it is convenient to introduce the matrix notations

$$\begin{aligned} \mathbf{S}_{\ell,d}[k] &\triangleq [\mathbf{S}_{\ell,d,1}^T[k] \dots \mathbf{S}_{\ell,d,N}^T[k]]^T, \\ \mathbf{S}_\ell[k] &\triangleq [\mathbf{S}_{\ell,1}^T[k] \dots \mathbf{S}_{\ell,D}^T[k]]^T, \end{aligned}$$

to represent the transmitted symbols,

$$\begin{aligned}\mathbf{R}_{\ell,d}[k] &\triangleq [\mathbf{R}_{\ell,d,1}^T[k] \dots \mathbf{R}_{\ell,d,N}^T[k]]^T, \\ \mathbf{R}_{\ell}[k] &\triangleq [\mathbf{R}_{\ell,1}^T[k] \dots \mathbf{R}_{\ell,D}^T[k]]^T,\end{aligned}$$

to represent the received symbols, and

$$\begin{aligned}\mathbf{N}_{\ell,d}[k] &\triangleq [\mathbf{N}_{\ell,d,1}^T[k] \dots \mathbf{N}_{\ell,d,N}^T[k]]^T, \\ \mathbf{N}_{\ell}[k] &\triangleq [\mathbf{N}_{\ell,1}^T[k] \dots \mathbf{N}_{\ell,D}^T[k]]^T,\end{aligned}$$

to represent the noise samples, corresponding to blocks of N and DN matrix-symbol transmissions respectively.

2.1.2 Differential Space Time Modulation: Vector Channel Model

We consider the channel model from section 2.1.1, and apply differential space-time modulation at the transmitter to achieve the necessary transmit diversity. DSTM can be described by considering transmission of $B \times N_T$ matrices ([28],[14]), if channel state information is not available at the receiver. Hence during each modulation interval T , one row vector is transmitted, corresponding to using all N_T antennas, assuming the channel remains constant for all B intervals. This approach is known as unitary space-time (ST) modulation. It is possible to show ([28]) that right multiplication of these $B \times N_T$ matrices by arbitrary $N_T \times N_T$ matrices does not affect the pairwise error probability, hence consecutive $B \times N_T$ matrices can overlap by such submatrices, thereby increasing power and bandwidth efficiency. In the considered case, $B = 2N_T$, the $N_T \times N_T$ unitary submatrices are chosen, without loss of generality to be the identity matrix I_{N_T} . Hence the overlapping is achieved by successively right multiplying the current matrix symbol with the corresponding $N_T \times N_T$ unitary submatrices of all

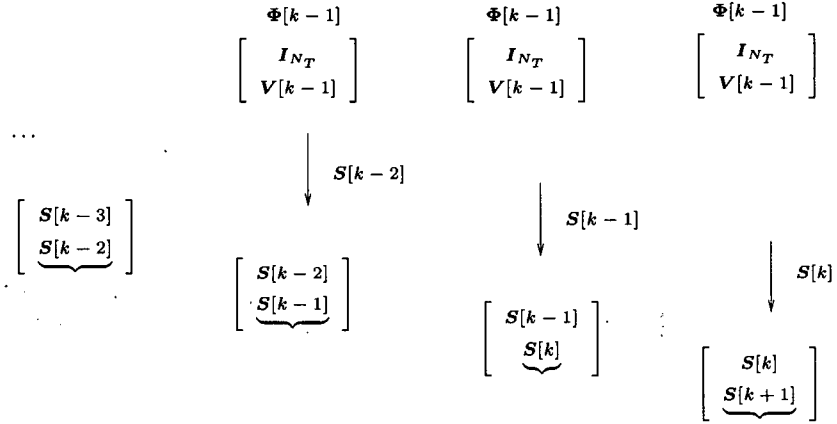


Figure 2.1: *Overlapping of consecutive $2N_T \times N_T$ matrix symbols $\Phi[k]$. After right multiplication with the $N_T \times N_T$ submatrix $S[k-1] = \prod_{\zeta=1}^{\infty} V[k-\zeta]$ of the previous matrix symbol $\Phi[k-1]$, only the bracketed regions i.e. $N_T \times N_T$ matrices $S[k]$ are transmitted.*

previous matrix symbols. As an illustration, let $V[k]$ be the chosen matrix for the current transmission. Hence the $2N_T \times N_T$ transmit matrix $\Phi[k] = [I_{N_T} V[k]^T]^T$. At a given time instant k , $\Phi[k]$ is right multiplied by the unitary matrix

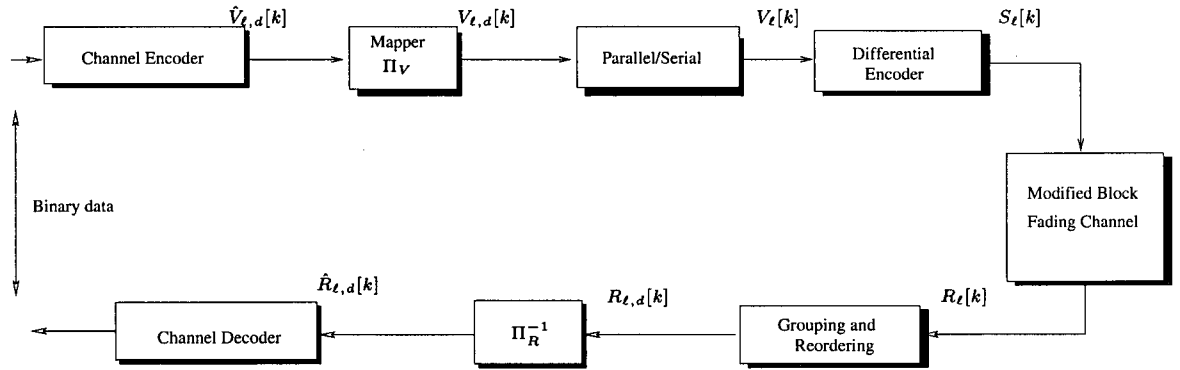
$$S[k-1] = \prod_{\zeta=1}^{\infty} V[k-\zeta] \quad (2.4)$$

This is illustrated in fig. 2.1.

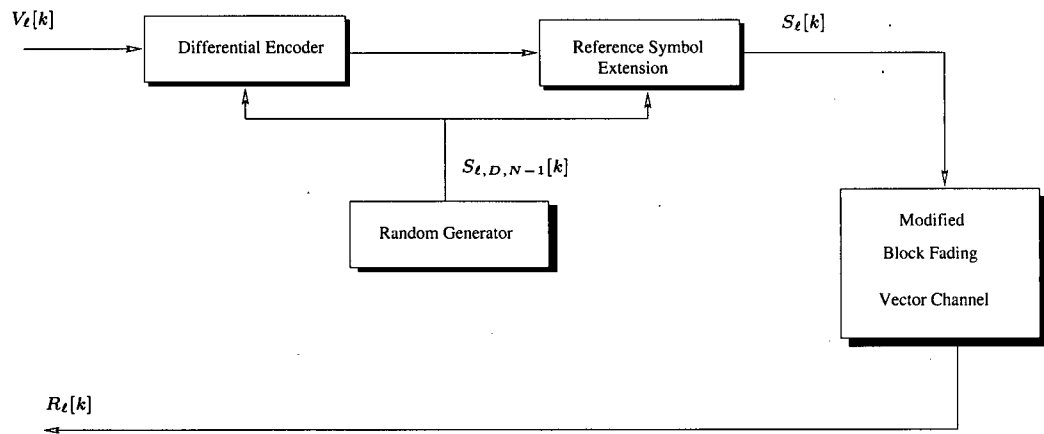
The block diagram of the equivalent system is shown in fig. 2.2. We map binary coded symbols to a $D(N-1)N_T \times N_T$ dimensional matrix

$$V_{\ell}[k] \triangleq [V_{\ell,d,1}[k]^T V_{\ell,d,2}[k]^T \dots V_{\ell,d,N-1}[k]^T]^T, 1 \leq d \leq D \quad (2.5)$$

of $D(N-1)$ space-time $N_T \times N_T$ data matrix symbols $V_{\ell,d,n}[k]$, $1 \leq n \leq N-1, 1 \leq d \leq D$. The matrix symbols are taken from a signal set \mathcal{V} as described in section 2.2. This signal set forms a space-time code with $M = 2^{N_T R_m}$ elements, where R_m is the modulation rate in bits/(channel use), and one channel use corresponds to the use of all N_T transmit antennas simultaneously. In our case of DSTM with unconstrained



(a) Discrete-time channel model for DSTM with MSDD



(b) Equivalent vector channel model for DSTM with MSDD

Figure 2.2: Block diagrams for Differential Space-Time Modulation: Discrete Time Model and Vector Channel Model

block fading, DSTM can best be described by generating a *block transmit matrix*

$$\mathbf{S}_\ell[k] \triangleq [\mathbf{S}_{\ell,d,1}[k]^T \mathbf{S}_{\ell,d,2}[k]^T \dots \mathbf{S}_{\ell,d,N}[k]^T]^T, 1 \leq d \leq D \quad (2.6)$$

from the data block matrix $\mathbf{V}_\ell[k]$ via

$$\mathbf{S}_{\ell,d,n-j}[k] = \mathbf{V}_{\ell,d,n-j}[k] \mathbf{S}_{\ell,d,n-j-1}[k], \quad 1 \leq j \leq N-1 \quad (2.7)$$

where the reference symbol $\mathbf{S}_{\ell,D,N-1}[k]$, which is an independent u.i.i.d random variable is taken arbitrarily from the signal set \mathcal{S} . Since we are primarily interested in performance analysis using the random coding exponent, we can conveniently assume interleaving to be limited to blocks of LDN space-time symbols without imposing any more restrictions.

Taking the expressions above and defining block matrix symbols

$$\mathbf{S}_\ell[k] \triangleq \begin{bmatrix} \mathbf{S}_{\ell,d,1}[k] & \mathbf{0}_{N_T} & \dots & \mathbf{0}_{N_T} \\ \mathbf{0}_{N_T} & \mathbf{S}_{\ell,d,2}[k] & \dots & \mathbf{0}_{N_T} \\ \vdots & & \ddots & \vdots \\ \mathbf{0}_{N_T} & & \mathbf{0}_{N_T} & \mathbf{S}_{\ell,d,DN-1}[k] \end{bmatrix} \quad (2.8)$$

$$\mathbf{H}_\ell[k] \triangleq [\mathbf{H}_{\ell,d,1}[k]^T \mathbf{H}_{\ell,d,2}[k]^T \dots \mathbf{H}_{\ell,d,N}[k]^T]^T, 1 \leq d \leq D \quad (2.9)$$

$$\mathbf{N}_\ell[k] \triangleq [\mathbf{N}_{\ell,d,1}[k]^T \mathbf{N}_{\ell,d,2}[k-1]^T \dots \mathbf{N}_{\ell,d,N}[k-(DN-1)]^T]^T, 1 \leq d \leq D \quad (2.10)$$

we obtain the $DNN_T \times N_R$ matrix $\mathbf{R}_\ell[k]$ of DN received samples as

$$\mathbf{R}_\ell[k] = \mathbf{S}_\ell[k] \mathbf{H}_\ell[k] + \mathbf{N}_\ell[k], \quad 1 \leq \ell \leq L, \quad (2.11)$$

where $\mathbf{N}_\ell[k]$ denotes additive spatially and temporally white circularly-symmetric Gaussian noise (AWGN) with variance σ_n^2 per complex scalar component. We also

assume that the elements of $\mathbf{H}_\ell[k]$ are circularly-symmetric Gaussian distributed, i.e., block Rayleigh fading, with common variance σ_h^2 and possible spatial correlation,

$$\psi_{\mathbf{H}}[\mu_1, \mu_2, \nu_1, \nu_2] \triangleq \mathcal{E}\{[\mathbf{H}_\ell[k]]_{\mu_1, \nu_1}([\mathbf{H}_\ell[k]]_{\mu_2, \nu_2})^*\}, \quad (2.12)$$

which is assumed identical for all subchannels ℓ .

At the receiver, MSDD with an observation window size N is applied. This means that the decoder input is based on independent processing of blocks of N received samples collected in $\mathbf{R}_{\ell,d}[k]$ corresponding to $N - 1$ differential symbols

$$\mathbf{V}_{\ell,d}[k] \triangleq [\mathbf{V}_{\ell,d,1}^T[k] \dots \mathbf{V}_{\ell,d,N-1}^T[k]]^T.$$

We use the notations

$$\begin{aligned} \mathbf{R} &\triangleq [\mathbf{R}_1^T[k] \dots \mathbf{R}_L^T[k]]^T \\ \mathbf{V} &\triangleq [\mathbf{V}_1^T[k] \dots \mathbf{V}_L^T[k]]^T \triangleq [\mathbf{V}_{1,1}^T[k] \mathbf{V}_{1,2}^T[k] \dots \mathbf{V}_{L,D}^T[k]]^T \end{aligned} \quad (2.13)$$

to represent the collection of all received samples and differential symbols corresponding to the transmission of one codeword. Throughout this thesis, for the sake of simplicity, we omit the frame index k , and use the representations $\mathbf{X}_{\ell,d}[k] = \mathbf{X}_{\ell,d}$, $\mathbf{X}_\ell[k] = \mathbf{X}_\ell$, $\mathbf{X} \in \{\mathbf{S}, \mathbf{R}, \mathbf{N}\}$.

2.1.3 Single Antenna Transmission: A Special Case

For the special case of $N_T = N_R = 1$, we revert back to the traditional single antenna transmission scenario with multiple symbol differential detection at the receiver. For a compact notation, it is convenient to introduce the vector notation

$$\begin{aligned} \mathbf{x}_{\ell,d}[k] &\triangleq [\mathbf{x}_{\ell,d,1}^T[k] \dots \mathbf{x}_{\ell,d,N}^T[k]]^T, \\ \mathbf{x}_\ell[k] &\triangleq [\mathbf{x}_{\ell,1}^T[k] \dots \mathbf{x}_{\ell,D}^T[k]]^T, \end{aligned}$$

corresponding to blocks of N and DN vector-symbol transmissions, respectively, and $\mathbf{x} \in \{\mathbf{s}, \mathbf{r}, \mathbf{n}\}$ corresponding to the transmitted, received and noise vector block symbols respectively.

Here, instead of transmitting $DN N_T \times N_R$ matrix symbols as in the ST case, we transmit differentially encoded vectors of length DN corresponding to the fading coherence interval, and we can represent the received vector as

$$\mathbf{r}_\ell[k] = \mathbf{s}_\ell[k] \mathbf{h}_\ell[k] + \mathbf{n}_\ell[k], \quad 1 \leq \ell \leq L, \quad (2.14)$$

where

$$\mathbf{x}_{\ell,d}[k] \triangleq [\mathbf{x}_{\ell,d,1}^T[k] \dots \mathbf{x}_{\ell,d,N}^T[k]]^T,$$

$$\mathbf{x}_\ell[k] \triangleq [\mathbf{x}_{\ell,1}^T[k] \dots \mathbf{x}_{\ell,D}^T[k]]^T,$$

corresponding to blocks of N and DN vector symbols transmissions respectively.

Vector Channel Model for Single Antenna DPSK

Similar to the DSTM case, we consider transmission of blocks of DN symbols over a frequency non-selective block fading channel, with L independent fading realizations per coding frame (see [19, 20, 27]). A sequence of information bits is encoded at the encoder, and mapped to data carrying *differential symbols*

$$\mathbf{v}_\ell[k] = [v_{\ell,d,1}[k] v_{\ell,d,2}[k-1] \dots v_{\ell,d,N-1}[k]]^T \quad (2.15)$$

consisting of $D(N-1)$ scalar differential symbols $v_{\ell,d,j}[k] \in \mathcal{V}$, $1 \leq j \leq N, 1 \leq d \leq D$. Here, since we consider interleaving and deinterleaving as an integral part of the channel, a *vector channel* is considered between the transmitter and the receiver. Assuming interleaving to be restricted to frames of LDN symbols, as in the DSTM case, at the transmitter, the process of differential encoding can be described by generating the *transmit vector* of dimension DN

$$\mathbf{s}_\ell[k] = [s_{\ell,d,1}[k] s_{\ell,d,2}[k-1] \dots s_{\ell,d,N}[k]]^T, \quad 1 \leq d \leq D \quad (2.16)$$

from the $D(N-1)$ dimensional data vector $\mathbf{v}_{\ell,d,N}[k]$ via

$$\mathbf{s}_{\ell,d,n-j}[k] = \mathbf{v}_{\ell,d,n-j}[k] \cdot \mathbf{s}_{\ell,d,n-j-1}[k], \quad 1 \leq j \leq N, 1 \leq d \leq D \quad (2.17)$$

where the reference symbol is chosen arbitrarily from the signal set \mathcal{S} due to noncoherent detection at the receiver. Here, the first (reference) symbol $s_{\ell,D,N}[k]$ of the block $\mathbf{s}_\ell[k]$ is the last symbol of the previously transmitted block $\mathbf{s}_{\ell-1}[k]$. As [21] and [35] note, this overlapping is advantageous with respect to spectral and power efficiency. Defining the DN -dimensional vectors of channel gains and noise samples corresponding to $\mathbf{s}_\ell[k]$ by

$$\mathbf{h}_\ell[k] = [h_{\ell,d,1}[k] h_{\ell,d,2}[k-1] \dots h_{\ell,d,N}[k]]^T, 1 \leq d \leq D \quad (2.18)$$

$$\mathbf{n}_\ell[k] = [n_{\ell,d,1}[k] n_{\ell,d,2}[k-1] \dots n_{\ell,d,N}[k]]^T, 1 \leq d \leq D \quad (2.19)$$

respectively, the input-output relation of the vector channel is represented as

$$\mathbf{r}_\ell[k] = \mathbf{s}_\ell[k] \mathbf{h}_\ell[k] + \mathbf{n}_\ell[k] \quad (2.20)$$

where the fading gains $\mathbf{h}_\ell[k]$ and the noise $\mathbf{n}_\ell[k]$ are complex valued Gaussian random variables with variances $1/\sqrt{2}$ and σ_n^2 respectively per complex component. For the single antenna case too, throughout this thesis we omit the frame index k for the sake of simplicity.

2.2 Signal Constellations for Differential Space Time Modulation

The signal constellations for DSTM that are supported in this thesis have been widely studied in literature ([16, 15, 36, 12, 17]). Signal Constellation design for DSTM basically involves design of constellations $\mathcal{V} = \{\mathbf{V}_0, \mathbf{V}_1, \dots, \mathbf{V}_{M-1}\}$ of $M = 2^{N_T R_m}$ unitary matrices \mathbf{V}_m , where R_m is the maximum supported bit rate in bits per channel use, and one channel use corresponds to all N_T antennas being used simultaneously. This ensures that throughout the course of this thesis, the rate R_m is independent of

the number of transmit antennas. We consider two basic kinds of signal constellations, namely constellations from Group Codes and Non-Group Codes ([36], Table 4).

2.2.1 Constellations from Group Codes: Cyclic Codes or Diagonal Signals

To simplify design, group codes have been developed [16, 15, 36, 17], which form a group under matrix multiplication. A set of group codes implies $\mathcal{V} = \mathcal{S}$, i.e. the differential symbols are the same set as that of the transmitted symbols. Cyclic group codes or diagonal signals are defined as

$$C_{Cy}(M, u_1, \dots, u_{N_T}) = \left\{ \begin{bmatrix} e^{j2\pi u_1/M} & 0 & \dots & 0 \\ 0 & e^{j2\pi u_2/M} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e^{j2\pi u_{N_T}/M} \end{bmatrix}^m \mid m \in \{0, 1, \dots, M-1\} \right\}$$

where $u_l \in \{0, \dots, M-1\}$, $1 \leq l \leq N_T$ are coefficients optimized with respect to minimum distance. Some particular diagonal signal constellations that we use later for analysis purposes are

$$C_{Cy}(16, 1, 3) = \left\{ \begin{bmatrix} e^{j2\pi/16} & 0 \\ 0 & e^{j2\pi \cdot 3/16} \end{bmatrix}^m \mid m \in \{0, 1 \dots 15\} \right\} \quad (2.21)$$

for $N_T = 2$ transmit antennas and modulation rate $R_m = 2$ bits per channel use and

$$C_{Cy}(9, 1, 2, 5) = \left\{ \begin{bmatrix} e^{j2\pi/9} & 0 & 0 \\ 0 & e^{j2\pi \cdot 2/9} & 0 \\ 0 & 0 & e^{j2\pi \cdot 5/9} \end{bmatrix}^m \mid m \in \{0, 1 \dots 8\} \right\} \quad (2.22)$$

for $N_T = 3$ transmit antennas and $R_m = 1$ bits/(ch. use).

2.2.2 Constellations from Non-Group Codes

Based on group structures, in the case of group codes considered by [36], we consider another set of space-time constellations where $\mathcal{V} \neq \mathcal{S}$.

Orthogonal Codes

Codes from orthogonal designs ([39]) can be applied with good performance results for DSTM. For the special and widely considered case in literature of $N_T = 2$ transmit antennas, DSTM based on Alamouti's space-time block code is considered and applied. These codes can be represented by

$$\mathbf{C}_{OD}(M) = \left\{ \frac{1}{\sqrt{2}} \begin{bmatrix} x & -y^* \\ y & x^* \end{bmatrix} \mid x, y \in \{1, e^{j\frac{2\pi}{M}}, \dots, e^{j\frac{2\pi(\sqrt{M}-1)}{\sqrt{M}}}\}, M = 2^{2l} \right\} \quad (2.23)$$

where the two symbols x and y are taken from \sqrt{M} PSK constellations. Such codes yield good results and are used for analysis purposes, and also allow particularly simple detection [1, 40].

Cayley Codes

Another class of non-group unitary constellations $C_{Ca}(N_T, M)$ has been proposed by Hassibi and Hochwald [12]. The differential matrix symbols are obtained via the Cayley Transform of $j\mathbf{A}$, where $\mathbf{A} = \sum_{q=1}^Q \mathbf{A}_Q a_q$, and $\mathbf{A}_1, \dots, \mathbf{A}_Q$ are preselected $N_T \times N_T$ complex Hermitian matrices and a_1, \dots, a_Q are chosen from a set A with P real values. Similar to previous work [23], we focus on the advantageous effect of transmit diversity considering low and moderate modulation rates of $R_m = 1, 2, 3$ bits/(channel use). One such Cayley constellation giving good performance and complexity results is the

Cayley Code $\mathbf{C}_{Ca}(2, 4)$ given by

$$\left\{ \begin{array}{l} (\mathbf{I}_2 + j\mathbf{A})^{-1}(\mathbf{I}_2 - j\mathbf{A}) \\ \mathbf{A} = \mathbf{A}a_1 + \mathbf{A}a_2 \\ \mathbf{A}_1 = \begin{bmatrix} -0.1853 & 0.8218 + 0.2694j \\ 0.8218 - 0.2694j & 0.1359 \end{bmatrix} \\ \mathbf{A}_2 = \begin{bmatrix} -0.2935 & -0.5885 - 0.5704j \\ -0.5885 + 0.5704j & 0.3452 \end{bmatrix} \\ a_1, a_2 \in \{-1, 1\} \end{array} \right\} \quad (2.24)$$

for a modulation rate of $R_m = 1$ bit/(ch. use). We will use this particular diagonal code for the random coding analysis in later chapters.

Chapter 3

Gallager Random Coding Exponent for Differential Transmission and Noncoherent Reception

3.1 Introduction

In this chapter, performance limits of coded transmission from an information theoretic perspective are discussed. We discuss the performance of differential space-time modulation with MSDD using Gallager's random coding exponent [10] as a performance metric for the general block fading channel already described, and also use channel capacity to characterize the transmission system. The first part of the chapter focuses on multiple transmit and receive antenna systems, where we first introduce the required noncoherent channel probability density functions (PDFs) to characterize the block fading channel in section 3.2.1 and then analyze the random coding exponent for space-time modulation in section 3.2.2, and also introduce the channel capacity for our transmission scheme. We provide detailed derivations of the random coding exponent,

and explain in detail the effect of transmit diversity introduced due to our general block fading model on the random coding exponent. In section 3.2.3 we introduce a reduced complexity metric to calculate the error exponent, which eliminates the complexity drawback where the calculation of the channel metrics involved a high dimensional search space corresponding to the length DN of the channel coherence interval. We present results illustrating our analysis for DSTM in section 3.2.4, and to complete our analysis of the random coding exponent, we present the corresponding single antenna performance metrics and results in section 3.3.

3.2 Multiple Antenna Systems

In this section, we introduce the noncoherent PDFs required to characterize the space-time channel and introduce Gallager's random coding exponent for DSTM, along with a detailed derivation of the exponent for the block fading channel model under consideration.

3.2.1 Noncoherent Probability Density Functions

For noncoherent transmission, the channel described in section 2.1.1 is completely characterized by the DN dimensional probability density function $p(\mathbf{R}_\ell | \mathbf{V}_\ell)$ for a given differential vector \mathbf{V}_ℓ . Under our block fading assumption, in order to characterize our vector channel model, we express the conditional noncoherent pdf (see [37]) as

$$p(\mathbf{R}_\ell | \mathbf{V}_\ell) = \frac{\exp(-\text{tr}\{\mathbf{R}_\ell^H \boldsymbol{\Psi}_R^{-1} \mathbf{R}_\ell\})}{(\pi^{NN_T} \det\{\boldsymbol{\Psi}_R\})^{N_R}}, \quad (3.1)$$

where

$$\boldsymbol{\Psi}_R = \mathcal{E}_{\mathbf{R}_\ell} \{\mathbf{R}_\ell \mathbf{R}_\ell^H | \mathbf{V}_\ell\} = N_R (\sigma_h^2 \mathbf{S}_\ell \mathbf{S}_\ell^H + \sigma_n^2 \mathbf{I}_{NN_T}) \quad (3.2)$$

is the conditional autocorrelation matrix of the received vector \mathbf{R}_ℓ given the transmitted vector \mathbf{V}_ℓ , with some arbitrarily chosen $\mathbf{S}_{\ell,1}$. We note that $\det\{\Psi_{\mathbf{R}}\}$ is independent of \mathbf{V}_ℓ . Equation 3.1 is the expression for the conditional noncoherent pdf of the output vector given the input vector each of length DN , and we differentiate from previous work where the channel coherence interval was equal to N . We represent the overall maximum-likelihood (ML) decoding metric as

$$\Lambda(\mathbf{R}|\mathbf{V}) = \sum_{\ell=1}^L \sum_{d=1}^D \lambda(\mathbf{R}_{\ell,d}|\mathbf{V}_{\ell,d}), \quad (3.3)$$

where the metric increments $\lambda(\mathbf{R}_{\ell,d}|\mathbf{V}_{\ell,d})$ are log-likelihood expressions derived from the probability density function (pdf) $p(\mathbf{R}_{\ell,d}|\mathbf{V}_{\ell,d})$. The rationale for (3.3) is that since the size of the effective signal constellation and hence the complexity of soft-output MSDD increases exponentially with the window size N , the value for N has often to be chosen (much) smaller than the channel coherence interval of length DN , i.e., $D > 1$ (cf. e.g. [8]).

The pdf for the received sequence \mathbf{R} given the sequence of differential symbols \mathbf{V} assuming coherent detection with CSI can be decomposed into the product ($\mathbf{H} \triangleq [\mathbf{H}_1^T \dots \mathbf{H}_L^T]^T$)

$$p(\mathbf{R}|\mathbf{V}, \mathbf{H}) = \prod_{\ell=1}^L \prod_{d=1}^D p(\mathbf{R}_{\ell,d}|\mathbf{V}_{\ell,d}, \mathbf{H}_\ell), \quad (3.4)$$

with the NN_TN_R -dimensional pdf

$$p(\mathbf{R}_{\ell,d}|\mathbf{V}_{\ell,d}, \mathbf{H}_\ell) = \mathcal{E}_{\mathbf{S}_{\ell,d,1}} \left\{ \frac{\exp\left(-\frac{\|\mathbf{R}_{\ell,d} - \mathbf{S}_{\ell,d}\mathbf{H}_\ell\|^2}{\sigma_n^2}\right)}{(\pi\sigma_n^2)^{NN_TN_R}} \right\}. \quad (3.5)$$

We will use the pdf in (3.5) to derive certain metrics to bound the random coding exponent in section 3.2.3.

3.2.2 Random Coding Exponent and Channel Capacity

We consider the random coding exponent to be the definitive information theoretic parameter to analyze the maximum data rate supported by the channel. Since channel capacity assumes infinite code length and zero error probability, for most practical applications we require an information theoretic performance measure that incorporates code length and/or decoding delay and error probability to give an accurate performance analysis. The random coding exponent proposed by Gallager relates code length or decoding delay to the reliability of the decoded data, namely the word error rate (WER) P_w . The random coding exponent is given by [10]

$$E(R_V, N, D) = \max_{0 \leq \rho \leq 1} \{E_0(\rho, N, D) - \rho R_V\} \quad (3.6)$$

where R_V is the rate in bits per vector symbol, and $E_0(\rho, N, D)$ is Gallager's function, derived in the next section. From the random coding exponent, given the number n_v of vector symbols v_k per code word, the word error rate P_w is bounded by

$$P_w \leq 2^{-n_v E(R_V, N)} \quad (3.7)$$

Derivation of Gallager's Random Coding Exponent:

We refer to Gallager's original derivation for the random coding exponent in [10], and modify it to reflect the block fading channel under consideration. Since we consider noncoherent transmission with L vector symbols $\mathbf{V}_{\ell,d}$, each consisting of $N - 1$ components drawn from a signal set \mathcal{V} , we denote the codeword of length L as $\bar{\mathbf{V}}_\ell$ i.e.

$$\bar{\mathbf{V}}_\ell = [\mathbf{V}_{\ell,1}^T \mathbf{V}_{\ell,2}^T \dots \mathbf{V}_{\ell,D}^T]^T \quad (3.8)$$

and the corresponding received vector $\bar{\mathbf{R}}$ consists of L overlapping received blocks $\mathbf{R}_{\ell,d}$, i.e. $\bar{\mathbf{R}}_\ell = [\mathbf{R}_{\ell,1}^T \mathbf{R}_{\ell,2}^T \dots \mathbf{R}_{\ell,D}^T]^T$. We assume the transmission rate is R_V bits per vector symbol.

At the receiver, we assume decoding metrics of the form

$$\lambda_{n_w}(\bar{\mathbf{R}}_\ell|\bar{\mathbf{V}}_\ell) = \sum_{d=1}^D \sum_{\ell=1}^L \lambda_{\ell,d}(\mathbf{R}_{\ell,d}|\mathbf{V}_{\ell,d}) \quad (3.9)$$

corresponding to (3.3), where $n_w = DNL$ denotes the dimension of $\bar{\mathbf{R}}$ i.e. we assume a memoryless channel. Let us assume that a particular codeword $\bar{\mathbf{V}}_t$ is transmitted, and if $\lambda_{n_w}(\bar{\mathbf{R}}|\bar{\mathbf{V}}_t) \leq \lambda_{n_w}(\bar{\mathbf{R}}|\bar{\mathbf{V}}_d)$ for some $d \neq t$ the receiver decides in favor of another codeword. Then, given $\bar{\mathbf{V}}_t$ and a set of $2^{R_v L}$ codewords, the word error rate (WER) reads as

$$P_{w,t} = \int_{\bar{\mathbf{R}} \in \mathcal{C}^{n_w}} p_{n_w}(\bar{\mathbf{R}}|\bar{\mathbf{V}}_t) \tau_t(\bar{\mathbf{R}}) d\bar{\mathbf{R}} \quad (3.10)$$

where $p_{n_w}(\bar{\mathbf{R}}|\bar{\mathbf{V}}_t)$ is the n_w dimensional complex pdf of $\bar{\mathbf{R}}$ given $\bar{\mathbf{V}}_t$ and the indicator function $\tau_t(\bar{\mathbf{R}})$ is given as

$$\tau(\bar{\mathbf{R}}) \triangleq \begin{cases} 1, & \text{if } \lambda_{n_w}(\bar{\mathbf{R}}|\bar{\mathbf{V}}_t) \leq \lambda_{n_w}(\bar{\mathbf{R}}|\bar{\mathbf{V}}_d) \text{ for some } d \neq t \\ 0, & \text{otherwise} \end{cases} \quad (3.11)$$

Upper bounding $\tau_t(\bar{\mathbf{R}})$ by

$$\tau_t(\bar{\mathbf{R}}) \leq \left[\sum_{d=1, d \neq t}^{2^{R_v L}} \left[\frac{\lambda_{n_w}(\bar{\mathbf{R}}|\bar{\mathbf{V}}_d)}{\lambda_{n_w}(\bar{\mathbf{R}}|\bar{\mathbf{V}}_t)} \right]^{\frac{1}{1+\rho}} \right]^\rho, \quad \rho > 0 \quad (3.12)$$

(3.10) can be upper bounded by

$$P_{w,t} \leq \int_{\bar{\mathbf{R}} \in \mathcal{C}^{n_w}} p_{n_w}(\bar{\mathbf{R}}|\bar{\mathbf{V}}_t) \left[\sum_{d=1, d \neq t}^{2^{R_v L}} \left[\frac{\lambda_{n_w}(\bar{\mathbf{R}}|\bar{\mathbf{V}}_d)}{\lambda_{n_w}(\bar{\mathbf{R}}|\bar{\mathbf{V}}_t)} \right]^{\frac{1}{1+\rho}} \right]^\rho d\bar{\mathbf{R}}, \quad (3.13)$$

Now, additionally imposing $\rho \leq 1$, the random coding upper bound on the average WER P_w follows from averaging the right hand side of 3.12 with respect to $\bar{\mathbf{V}}_d$ and

$\bar{\mathbf{V}}_t$. Assuming the codeword symbols $\mathbf{V}_{\ell,d}$ are chosen independently and identically distributed with a fixed distribution $Pr\{\mathbf{V}_{\ell,d}\}$ not subject to optimization, we have

$$P_w \leq 2^{R_v n_v \rho} \int_{\bar{\mathbf{R}} \in \mathbf{C}^{n_w}} \left(\sum_{\bar{\mathbf{V}} \in \mathbf{V}^{(n_w-L)}} \prod_{d=1}^D \prod_{\ell=1}^L Pr\{\mathbf{V}_{\ell,d}\} \lambda_N(\mathbf{R}_{\ell,d} | \mathbf{V}_{\ell,d})^{\frac{-\rho}{1+\rho}} p_{n_w}(\bar{\mathbf{R}} | \bar{\mathbf{V}}) \right) \left(\prod_{d=1}^D \prod_{\ell=1}^L \sum_{\mathbf{V}_{\ell,d} \in \mathbf{V}^{N-1}} Pr\{\mathbf{V}_{\ell,d}\} \lambda_N(\mathbf{R}_{\ell,d} | \mathbf{V}_{\ell,d})^{\frac{1}{1+\rho}} \right)^\rho d\bar{\mathbf{R}} \quad (3.14)$$

Hence we have

$$P_w = 2^{-L(E_0(\rho, N, D) - \rho R_v)}, \quad 0 \leq \rho \leq 1 \quad (3.15)$$

where Gallager's function $E_0(\rho, N, D)$ is defined as

$$E_0(\rho, N, D) \triangleq \frac{-1}{L} \log_2 \left(\int_{\bar{\mathbf{R}} \in \mathbf{C}^{n_w}} \left(\sum_{\bar{\mathbf{V}} \in \mathbf{V}^{(n_w-L)}} \prod_{d=1}^D \prod_{\ell=1}^L Pr\{\mathbf{V}_{\ell,d}\} \lambda_N(\mathbf{R}_{\ell,d} | \mathbf{V}_{\ell,d})^{\frac{-\rho}{1+\rho}} p_{n_w}(\bar{\mathbf{R}} | \bar{\mathbf{V}}) \right) \left(\prod_{d=1}^D \prod_{\ell=1}^L \sum_{\mathbf{V}_{\ell,d} \in \mathbf{V}^{N-1}} Pr\{\mathbf{V}_{\ell,d}\} \lambda_N(\mathbf{R}_{\ell,d} | \mathbf{V}_{\ell,d})^{\frac{1}{1+\rho}} \right)^\rho d\bar{\mathbf{R}} \right) \quad (3.16)$$

Here, we reiterate the fact that the considered channel has L independent fading realizations per coding frame, and we can rewrite the expression for Gallager's function as

$$E_0(\rho, N, D) = -\frac{1}{n_v} \log_2 \int \dots \int_{\mathbf{R}_1} \dots \int_{\mathbf{R}_L} \sum_{\mathbf{V}_1} \dots \sum_{\mathbf{V}_L} \left(\prod_{\ell=1}^L \prod_{d=1}^D Pr\{\mathbf{V}_{\ell,d}\} (\lambda_N(\mathbf{R}_{\ell,d} | \mathbf{V}_{\ell,d}))^{\frac{-\rho}{1+\rho}} \right) \left(\prod_{\ell=1}^L p(\mathbf{R}_\ell | \mathbf{V}_\ell) \right) \left(\prod_{\ell=1}^L \prod_{d=1}^D \left(\sum_{\mathbf{V}_1} \dots \sum_{\mathbf{V}_L} Pr\{\mathbf{V}_{\ell,d}\} (\lambda_N(\mathbf{R}_{\ell,d} | \mathbf{V}_{\ell,d}))^{\frac{1}{1+\rho}} \right)^\rho \right) d\mathbf{R}_1 \dots d\mathbf{R}_L. \quad (3.17)$$

We expand the above products to read

$$E_0(\rho, N, D) = -\frac{1}{n_v} \log_2 \int_{\mathbf{r}_1} \dots \int_{\mathbf{R}_L} \sum_{\mathbf{V}_1} \dots \sum_{\mathbf{V}_L} \left(\prod_{\ell=1}^L \prod_{d=1}^D Pr\{\mathbf{V}_{\ell,d}\} (\lambda_N(\mathbf{R}_{\ell,d}|\mathbf{V}_{\ell,d}))^{\frac{-\rho}{1+\rho}} \right)_{e_1} \cdot p(\mathbf{R}_\ell|\mathbf{V}_\ell) \cdot \left(\prod_{\ell=1}^L \prod_{d=1}^D \left(\sum_{\mathbf{V}_{\ell,d}} Pr\{\mathbf{V}_{\ell,d}\} (\lambda_N(\mathbf{R}_{\ell,d}|\mathbf{V}_{\ell,d}))^{\frac{1}{1+\rho}} \right)^\rho \right) d\mathbf{R}_1 \dots d\mathbf{R}_L. \quad (3.18)$$

Combining the product terms and interchanging the order of integration and summation, we have

$$E_0(\rho, N, D) = -\frac{1}{n_v} \log_2 \int_{\mathbf{R}_1} \dots \int_{\mathbf{R}_L} \prod_{\ell=1}^L \left[\sum_{\mathbf{V}_\ell} \left(Pr\{\mathbf{V}_{\ell,d}\} (\lambda_N(\mathbf{R}_{\ell,d}|\mathbf{V}_{\ell,d}))^{\frac{-\rho}{1+\rho}} \cdot p(\mathbf{R}_\ell|\mathbf{V}_\ell) \right) \right] \cdot \left[\left(\sum_{\mathbf{V}_\ell} \prod_{d=1}^D Pr\{\mathbf{V}_{\ell,d}\} \lambda_N(\mathbf{R}_{\ell,d}|\mathbf{V}_{\ell,d})^{\frac{1}{1+\rho}} \right)^\rho \right] d\mathbf{R}_1 \dots d\mathbf{R}_L. \quad (3.19)$$

Since each block of DN samples experiences independent fading and is independently processed at the receiver, we represent $\mathbf{R}_1 \dots \mathbf{R}_L$ as $\mathbf{R}_i, 1 \leq i \leq L$ to give

$$E_0(\rho, N, D) = -\frac{1}{n_v} \log_2 \prod_{\ell=1}^L \int_{\mathbf{R}_\ell} \left(\sum_{\mathbf{V}_\ell} \prod_{d=1}^D Pr\{\mathbf{V}_{\ell,d}\} \lambda_N(\mathbf{R}_{\ell,d}|\mathbf{V}_{\ell,d})^{\frac{-\rho}{1+\rho}} \cdot p(\mathbf{R}_\ell|\mathbf{V}_\ell) \right) \cdot \left(\sum_{\mathbf{V}_\ell} \prod_{d=1}^D Pr\{\mathbf{V}_{\ell,d}\} \lambda_N(\mathbf{R}_{\ell,d}|\mathbf{V}_{\ell,d})^{\frac{1}{1+\rho}} \right)^\rho d\mathbf{R}_\ell \quad (3.20)$$

Finally, we obtain

$$E_0(\rho, N, D) = -\frac{L}{n_v} \log_2 \int_{\mathbf{R}_\ell} \left(\sum_{\mathbf{V}_\ell} \prod_{d=1}^D Pr\{\mathbf{V}_{\ell,d}\} \lambda_N(\mathbf{R}_{\ell,d}|\mathbf{V}_{\ell,d})^{\frac{-\rho}{1+\rho}} \cdot p(\mathbf{R}_\ell|\mathbf{V}_\ell) \right) \cdot \left(\sum_{\mathbf{V}_\ell} \prod_{d=1}^D Pr\{\mathbf{V}_{\ell,d}\} \lambda_N(\mathbf{R}_{\ell,d}|\mathbf{V}_{\ell,d})^{\frac{1}{1+\rho}} \right)^\rho d\mathbf{R}_\ell \quad (3.21)$$

Here, the product over all transmitted vectors L is equivalent to multiplying the results of each vector by L , since each vector represents an independent fading coherence interval. This expression 3.21 represents Gallager's function for any arbitrary channel with a fading coherence interval independent from the MSDD observation interval, and simulation results shown in section 3.2.4 demonstrate the effect of various fading coherence intervals on the random coding exponent. This is an invaluable tool to accurately demonstrate the effect of varying MSDD observation intervals with the *coding diversity* kept constant, and allows for analysis of a far wider range of channels that until now were restricted to having fading coherence intervals tied to N . Because of the varying fading coherence interval, for the noncoherent case we need to take a product over all possible N -dimensional (or $N - 1$ -dimensional differential symbols) transmitted vectors $\mathbf{V}_{\ell,d}$ within a transmitted vector of dimension DN and sum over all possible transmitted vectors \mathbf{V}_ℓ to obtain the DN -dimensional pdf

$$p_{n_w}(\mathbf{r}|\mathbf{v}) = \prod_{i=1}^L p(\mathbf{R}_{\ell,i}|\mathbf{V}_{\ell,i}) \quad (3.22)$$

where $n_w = n_v \cdot N$ is the codeword length for the block fading channel under consideration having DN consecutive transmit vectors seeing the same fading coefficients.

For the special case of $D = 1$, we refer to 3.14, and can express it compactly by optimizing with respect to ρ . Hence the random coding exponent $E(R_v, N, D)$ is given by

$$E(R_v, N, D) \triangleq \max_{0 \leq \rho \leq 1} \{E_0(\rho, N, D) - \rho R_v\} \quad (3.23)$$

and the expression for Gallager's function

$$E_0(\rho, N, 1) = -\log_2 \left(\mathcal{E}_R \left\{ \mathcal{E}_V \left\{ \left(\frac{p_N(\mathbf{R}_{\ell,1}|\mathbf{V}_{\ell,1})}{p_N(\mathbf{R}_{\ell,1})} \right)^{\frac{1}{1+\rho}} \right\}^{1+\rho} \right\} \right) \quad (3.24)$$

with the WER bounded by

$$P_w \leq 2^{R_v L \rho} \left[\int_{\mathbf{R}_\ell \in \mathcal{C}^N} \left(\sum_{\mathbf{V}_\ell \in \mathcal{V}^{N-1}} Pr\{\mathbf{V}_{\ell,1}\} p_N(\mathbf{R}_{\ell,1}|\mathbf{V}_{\ell,1})^{\frac{1}{1+\rho}} \right)^{1+\rho} d\mathbf{R}_\ell \right]^L \quad (3.25)$$

This is exactly the expression seen in [10] to calculate the random coding exponent for various channels and constellations. From information theory, we know that Gallager's function $E_0(\rho, N, D)$ is upper bounded by the coherent channel with perfect channel state information (CSI) as $E_{0,CSI}(\rho, N, D)$. The performance of noncoherent transmission without CSI is clearly expected to improve as the MSDD observation length N increases, as the dependencies amongst received symbols are compromised by the increasing N , and the channel memory is more completely taken into account. However, when we look at the performance taking decoding delay constraints into account, we observe an interesting effect. We see that for noncoherent transmission over a block fading channel, where the channel coherence interval is equal to the MSDD observation window, the error exponent and the cutoff rate are decreasing functions of N , and decreases substantially as the fading memory becomes large [29, 41]. This behavior illustrates the reduced diversity achievable by coding when the MSDD interval N increases and the delay $(N - 1)n_v$ is fixed, resulting in a higher WER as shown by 3.7.

Channel Capacity

From information theory [10], the average *mutual information* $I(\cdot; \cdot)$ in bits per vector symbol is given by

$$I(\mathbf{R}_{\ell,d}; \mathbf{V}_{\ell,d}) = \mathcal{E}_{\mathbf{R}_{\ell,d}, \mathbf{V}_{\ell,d}} \left\{ \log_2 \left(\frac{p(\mathbf{R}_{\ell,d} | \mathbf{V}_{\ell,d})}{p(\mathbf{R}_{\ell,d})} \right) \right\} \quad (3.26)$$

where

$$p(\mathbf{R}_{\ell,d}) = \mathcal{E}_{\mathbf{V}_{\ell,d}}(p(\mathbf{R}_{\ell,d} | \mathbf{V}_{\ell,d})) \quad (3.27)$$

is the average pdf of the channel output.

Another information theoretic parameter to analyze the performance of block fading channels is the channel capacity $C(N)$ where we will consider the input distribution to be uniformly distributed, and hence refer to "capacity" meaning *constellation con-*

strained capacity [23]. For the conventional case of channel coherence interval equal to the MSDD observation window N , the average mutual information

$$I(\mathbf{R}_{\ell,d}; \mathbf{V}_{\ell,d}) = \mathcal{E}_{\mathbf{R}_{\ell,d}, \mathbf{V}_{\ell,d}} \left\{ \log_2 \left(\frac{p_{NN_T N_R}(\mathbf{R}_{\ell,d} | \mathbf{V}_{\ell,d})}{p_{NN_T N_R}(\mathbf{R}_{\ell,d})} \right) \right\} \quad (3.28)$$

measured in bits per matrix symbol \mathbf{V} is considered, where

$$p_{NN_T N_R}(\mathbf{R}_{\ell,d}) = \mathcal{E}_{\mathbf{V}_{\ell,d}} \{ p_{NN_T N_R}(\mathbf{R}_{\ell,d} | \mathbf{V}_{\ell,d}) \} \quad (3.29)$$

is the average pdf of the channel output. We normalize the channel capacity with respect to $DN - 1$ components per data vector symbol $\mathbf{V}_{\ell,d}$, giving

$$C(N) \triangleq \frac{1}{DN - 1} \cdot I(\mathbf{R}_{\ell,d}; \mathbf{V}_{\ell,d}) = \frac{1}{DN - 1} \cdot \mathcal{E}_{\mathbf{R}_{\ell,d}, \mathbf{V}_{\ell,d}} \left\{ \log_2 \left(\frac{p(\mathbf{R}_{\ell,d} | \mathbf{V}_{\ell,d})}{p(\mathbf{R}_{\ell,d})} \right) \right\} \quad (3.30)$$

3.2.3 Reduced Complexity Metric to Calculate the Random Coding Exponent

In order to efficiently compute the random coding exponent for higher dimensions (i.e $D, N \gg 1$), we introduce a reduced complexity derivation of the random coding exponent in this section. We begin with the expression for the WER bound, obtained from Gallager's function given by 3.21. The average probability of decoding error P_e is bounded by

$$P_e \leq 2^{-\rho N_T R n_v} \int_{\mathbf{R}_\ell} \left(\sum_{\mathbf{V}_\ell} \prod_{d=1}^D Pr\{\mathbf{V}_{\ell,d}\} \lambda(\mathbf{R}_{\ell,d} | \mathbf{V}_{\ell,d})^{\frac{-\rho}{1+\rho}} p(\mathbf{R}_\ell | \mathbf{V}_\ell) \right) \left(\sum_{\mathbf{V}_\ell} \prod_{d=1}^D Pr\{\mathbf{V}_{\ell,d}\} \lambda(\mathbf{R}_{\ell,d} | \mathbf{V}_{\ell,d})^{\frac{1}{1+\rho}} \right)^\rho d\mathbf{R}_\ell \quad (3.31)$$

We note the simplification

$$\begin{aligned} p(\mathbf{R}_\ell | \mathbf{V}_\ell) &\triangleq \int_{\mathbf{H}_\ell} p(\mathbf{H}_\ell) p(\mathbf{R}_\ell | \mathbf{V}_\ell, \mathbf{H}_\ell) d\mathbf{H}_\ell \\ &\triangleq \int_{\mathbf{H}_\ell} \prod_{d=1}^D p(\mathbf{R}_{\ell,d} | \mathbf{V}_{\ell,d}, \mathbf{H}_\ell) d\mathbf{H}_\ell \end{aligned} \quad (3.32)$$

where $p(\mathbf{R}_\ell | \mathbf{V}_\ell, \mathbf{H}_\ell)$ is the pdf for the received sequence \mathbf{R}_ℓ given the sequence of differential symbols \mathbf{V}_ℓ assuming coherent detection with CSI. Using the pdf's from section 3.2.1, 3.31 and 3.32, we can rewrite the expression for the probability of decoding as

$$\begin{aligned} P_e &\leq 2^{-\rho N_T R n_v} \int \frac{1}{M^{n_v}} \sum_{\substack{\mathbf{H}_\ell \in \mathbb{C}^{LN_T \times NR} \\ \mathbf{R}_\ell \in \mathbb{C}^{n_s N_T \times NR}}} p(\mathbf{R} | \mathbf{V}, \mathbf{H}) [\exp(\Lambda(\mathbf{R} | \mathbf{V}))]^{-\frac{\rho}{1+\rho}} \\ &\quad \left(\frac{1}{M^{n_v}} \sum_{\mathbf{V} \in \mathcal{V}^{n_v}} [\exp(\Lambda(\mathbf{R} | \mathbf{V}))]^{\frac{1}{1+\rho}} \right)^\rho d(\mathbf{H}_\ell, \mathbf{R}_\ell) \end{aligned} \quad (3.33)$$

for $0 \leq \rho \leq 1$, where the PDFs $p(\mathbf{R} | \mathbf{V}, \mathbf{H})$ and $\Lambda(\mathbf{R} | \mathbf{V})$ have been described in section 3.2.1. Optimizing the parameter ρ such that the upper bound (3.33) becomes tightest, we can write

$$P_e \leq 2^{-n_v E_r(R, N, D)}, \quad (3.34)$$

where $E_r(R, N, D)$ is the random coding exponent given by

$$E_r(R, N, D) = \max_{0 \leq \rho \leq 1} \{E_0(\rho, N, D) - \rho N_T R\}. \quad (3.35)$$

After manipulations using (3.3), (3.4), and (3.33), the Gallager function $E_0(\rho, N, D)$ is obtained as

$$\begin{aligned} E_0(\rho, N, D) &= -\frac{1}{(N-1)D} \log_2 \left[\mathcal{E}_{\mathbf{R}_\ell, \mathbf{H}_\ell} \left\{ \prod_{d=1}^D \right. \right. \\ &\quad \mathcal{E}_{\mathbf{V}_{\ell,d}} \left\{ \exp(\lambda(\mathbf{R}_{\ell,d} | \mathbf{V}_{\ell,d}))^{-\frac{\rho}{1+\rho}} p(\mathbf{R}_{\ell,d} | \mathbf{V}_{\ell,d}, \mathbf{H}_\ell) \right\} \\ &\quad \left. \left(\mathcal{E}_{\mathbf{V}_{\ell,d}} \left\{ \exp(\lambda(\mathbf{R}_{\ell,d} | \mathbf{V}_{\ell,d}))^{\frac{1}{1+\rho}} \right\} \right)^\rho \right. \\ &\quad \left. \left. \left(\mathcal{E}_{\mathbf{V}_{\ell,d}} \{p(\mathbf{R}_{\ell,d} | \mathbf{V}_{\ell,d}, \mathbf{H}_\ell)\} \right)^{-1} \right\} \right]. \end{aligned} \quad (3.36)$$

Here, using the Monte-Carlo technique we rewrite the integrations as expected values $\mathcal{E}(\cdot)$ and again make use of 3.3 and 3.4 to rewrite $\mathbf{R}_{\ell,d}$ and $\mathbf{V}_{\ell,d}$ in terms of \mathbf{R}_{ℓ} and \mathbf{V}_{ℓ} . In Appendix A, we look into this derivation of Gallager's function with reduced complexity in greater detail, and present only the result here for coherence.

3.2.4 Results and Discussion

In this section, we discuss the performance of differential space time modulation based on channel capacity and the random coding exponent. The random coding exponent analysis from section 3.2.3 is an invaluable tool to calculate the random coding exponent for channels not restricted by previous analysis, where the channel fading coherence interval was tied to N . Here we first analyze the performance of DSTM, and the advantage over single antenna systems for the purpose of random coding exponent analysis, and show that orthogonal constellations perform better than any other kind of space-time code such as diagonal constellations.

A) Random Coding Exponent Results

In this section, we demonstrate the random coding exponent for differential space-time modulation for the channel under considerations with fading coherence interval varying and not constrained to N . We use the results from section 3.2.3 to obtain the simplified expression for Gallager's function and correspondingly the random coding exponent. In order to compare and analyze the random coding exponent for DSTM, we plot Gallager's exponent against the rate R as in the single antenna case to justify our claim that the above analysis enables us to analyze channels where the coding diversity is kept constant though the MSDD observation interval N may change. We also perform SNR-delay analysis for the exponent and compare it with the single antenna case, to fairly compare the potential advantages of DSTM with multiple transmit antennas over

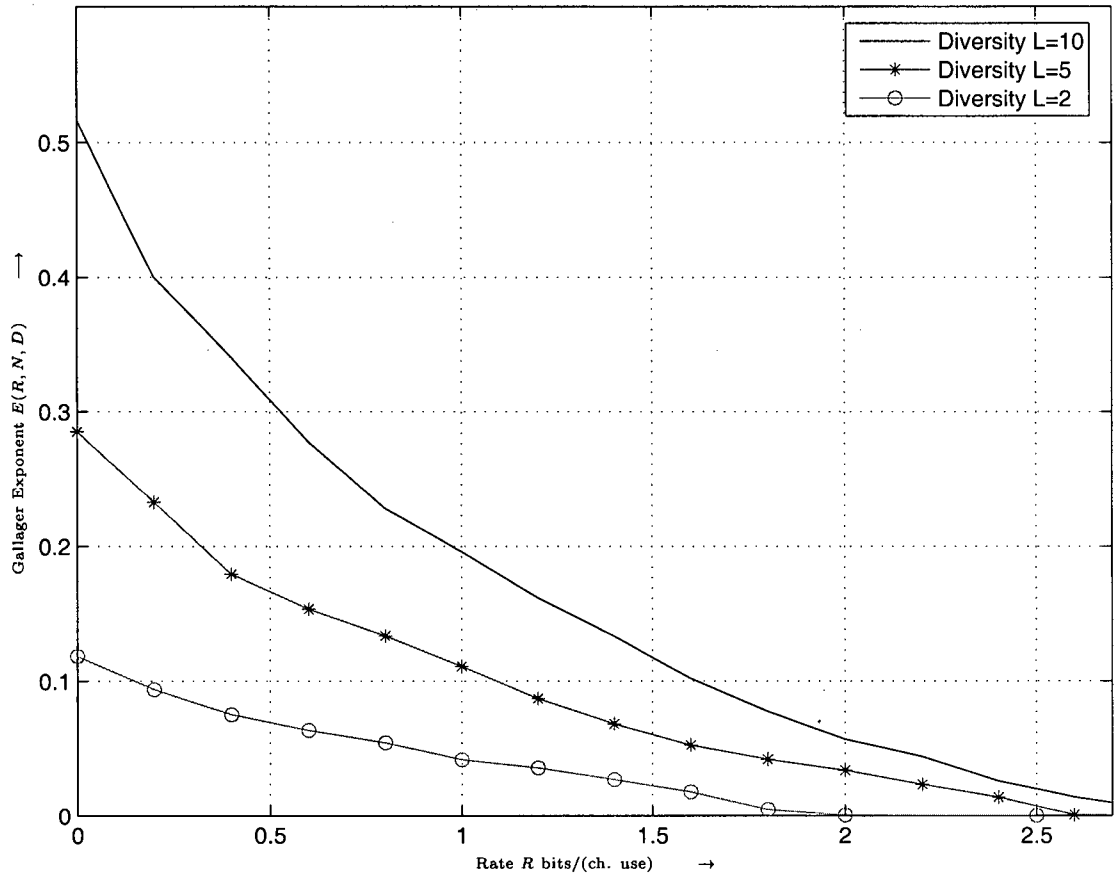


Figure 3.1: *Gallager Exponent vs Rate for DSTM, Orthogonal codes with CDD ($N = 2$), target rate $R_t = 1$ bit/(channel use): Effect of coding diversity*

single antenna traditional DPSK. The decoding delay t_d is measured as the number of scalar transmitted symbols required for decoding, i.e in this case

$$t_d = (DN - 1)n_d = (DN - 1)L \quad (3.37)$$

where $n_d = L$ is the number of vector symbols per code word.

1) $E(R, N, D)$ analysis for DSTM: In fig. 3.1, we illustrate the advantage of the derivation in section 3.2.3. When we compare schemes with different values for N and/or N_T , we use the normalized coherence length $n_c/L = (N - 1)D \log_2(M)$ as parameter

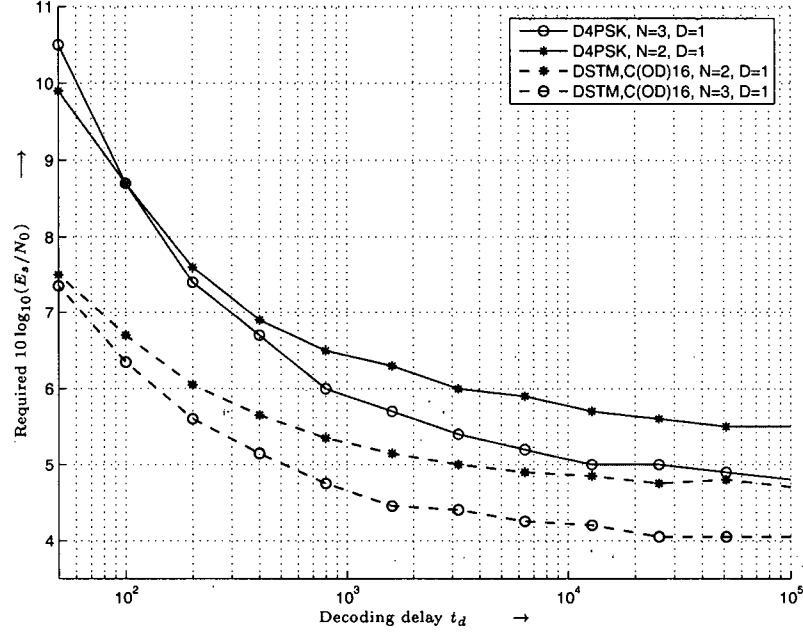


Figure 3.2: Required SNR ($10\log_{10}(E_s/N_0)$) vs t_d in PSK samples, D4PSK and $C_{OD}(16)$. Dashed lines: DSTM, Solid lines: D4PSK

to specify the amount of coded diversity (either temporally or spectrally) available. As a counterpart to the advantage of being able to compare different MSDD observation intervals N for the same coding diversity, we show the effect of different coding diversities $n_c/L = 20, 40, 100$ i.e coding diversities 10, 5 and 2 respectively for DSTM with orthogonal designs and target rate $R_t = 2$ bits/(channel use). We observe that keeping the observation window N constant while increasing the diversity allows the transmitter to see more independent fading realizations, thereby increasing the coding diversity and correspondingly a better random coding exponent performance.

2) SNR-delay analysis for DSTM: Advantage over Single Antenna Systems: Here, we demonstrate the advantage, taking decoding delay constraints into account, of DSTM with multiple transmit antennas over single antenna modulation. Fig. 3.2 shows the required $E_s/N_0/N_R$ per received antenna to achieve a WER $P_w \leq 0.001$ as a function of the decoding delay. We plot the random coding exponent for single antenna DPSK with

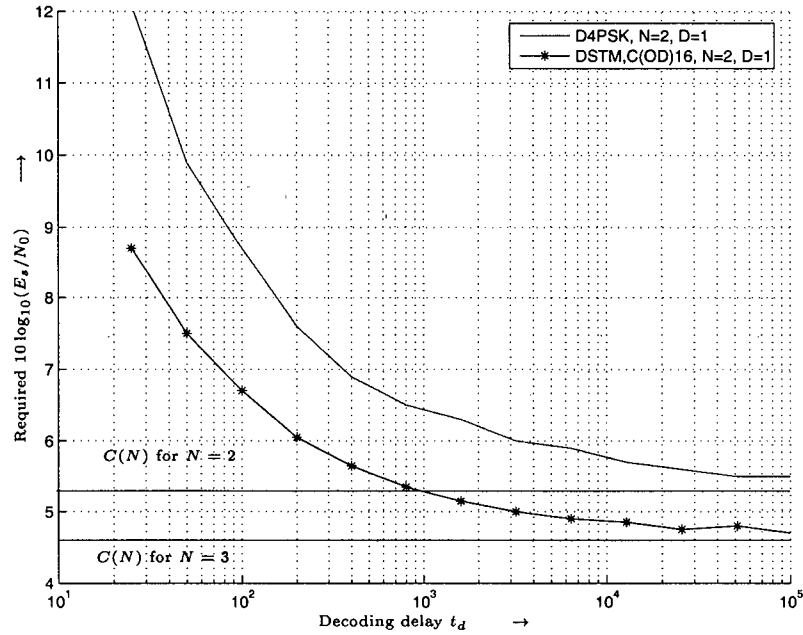
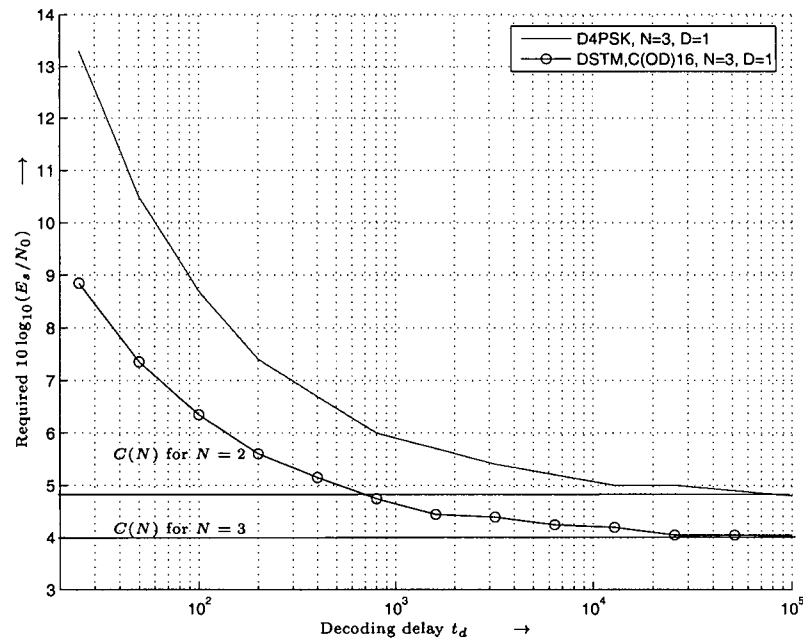
(a) CDD ($N = 2$): DSTM with $C_{OD}(16)$ (b) MSDD ($N = 3$): DSTM with $C_{OD}(16)$

Figure 3.3: Required $10 \log_{10}(E_s/N_0)$ vs Decoding delay t_d in PSK samples. Target rate $R_t=1$ bit/(channel use), CDD ($N = 2$) and MSDD with $N = 3$

CDD and MSDD, and compare it with the corresponding DSTM results. Interestingly, we note that not only do we get an added SNR boost by switching for $N_T = 1$ to $N_T = 2$, but for smaller delays $t_d \leq 10^2 \dots 10^3$ DSTM performs significantly better than DPSK. For the single antenna case we noted that for smaller delays, the gain obtained from MSDD is countered by the reduced diversity of not seeing enough independent fading realizations, hence the curves for different N 's intersected. However, for DSTM we see that the potential advantage of DSTM remains even for small delays, effectively providing enough diversity via space-time coding to retain the advantage of MSDD. In fig. 3.3 we look more closely into the advantage of DSTM over single antenna systems. For this purpose, we compare CDD ($N = 2$) and MSDD with $N = 3$ for both schemes. We note of interest that while the advantage of DSTM over single antenna DPSK is apparent at high decoding delays where the situation closely resembles the capacity case with potentially infinite delay, the advantage of DSTM is more apparent for low delays. This shows that with reduced diversity available for low delays and a specific window size N , the DSTM advantage far exceeds the single antenna case due to the space-time coding and overlapping process. This result shows the beneficial advantages of transmit diversity with $N_T > 1$ over standard single antenna differential modulation.

3) *Effect of D* : In order to characterize and analyze our proposed modified channel, we observe the effect increasing the fading coherence interval has on the random coding exponent. In fig 3.4, we show the effect increasing D has on the coding exponent. The curve for standard DSTM as in section 2.1.2 with $D = 1$ is shown as a reference. We observe that increasing D reduces the diversity observed on the channel by reducing the number of independent fading observations seen by the channel, hence requiring a higher SNR to achieve the same WER. This results in a degraded random coding exponent performance as seen by the increasing curves in the figure. Here we consider transmission of orthogonal codes with MSDD observation window $N = 3$, and change D to reflect the increased fading coherence interval.

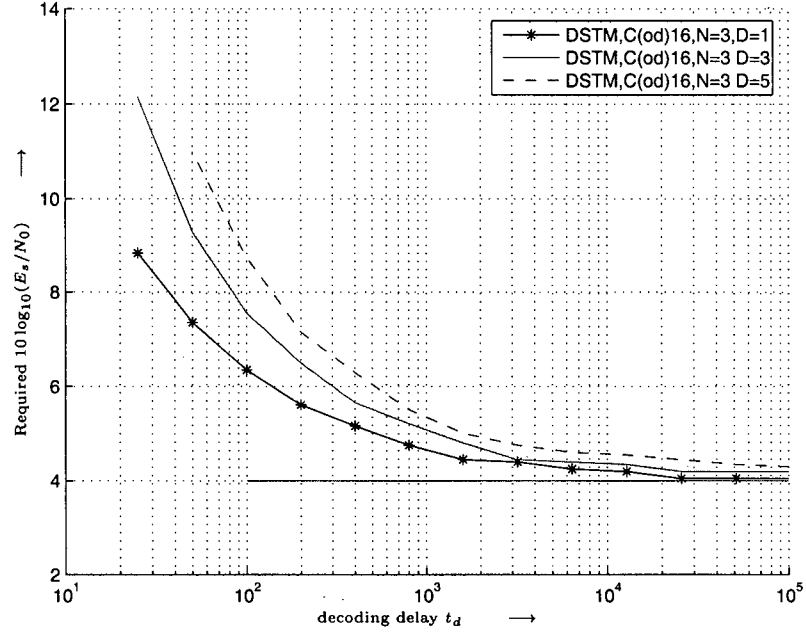
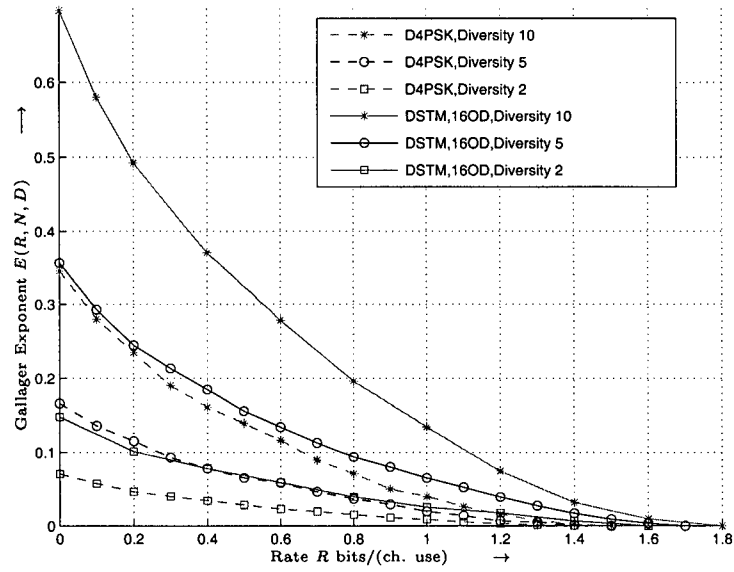


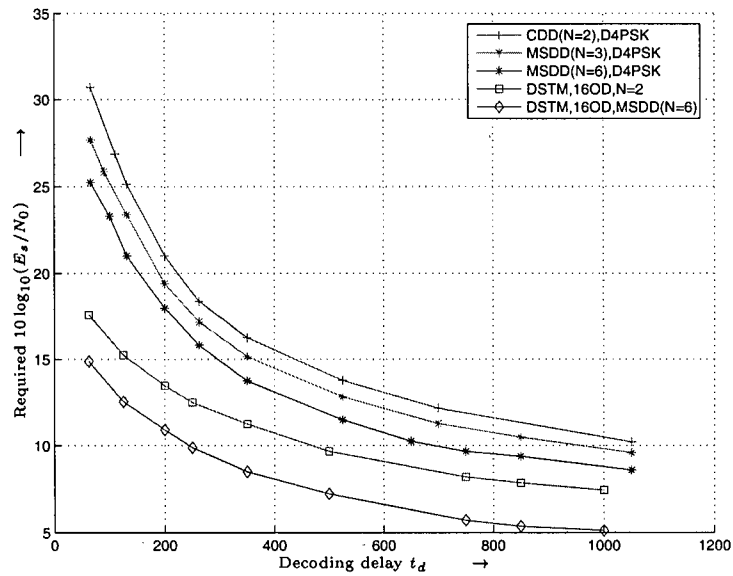
Figure 3.4: *Gallager Exponent results: Required SNR ($10 \log_{10}(E_s/N_0)$) vs Decoding delay t_d in PSK samples for DSTM with Orthogonal Codes $C_{OD}(16)$ showing the effect of increasing D while keeping N constant: Reduced diversity. Solid horizontal line: channel capacity for DSTM, $N = 3$*

Finally, in order to complete our introduction to the random coding exponent for DSTM, we compare $E(R, N, D)$ for single antenna systems and DSTM from a coding diversity point of view. Fig. 3.5(a) compares the random coding exponent $E(R, N, D)$ for single and multiple antenna systems where the coding diversity is the same in both cases. For a fair comparison, the normalized fading coherence length n_c/L is taken in PSK samples, hence an accurate comparison between single and multiple antenna systems can be made. In each case, for diversities 10, 5, 2 i.e $n_c/L = 20, 40, 100$ respectively, the random coding exponent obtained from DSTM offers a significant advantage over the single antenna case. We keep in mind the fact that since the WER bound is inversely proportional to the random coding exponent

$$P_w \leq 2^{-n_v E(R, N, D)} \quad (3.38)$$



(a) $E(R, N, D)$ vs R comparison between D4PSK and DSTM: Overall Diversity $L = 10, 5, 2$, DSTM with $C_{OD}(16)$, $R_{target} = 1$ bit/(ch. use) vs Single Antenna D4PSK



(b) Required SNR vs decoding delay: DSTM vs D4PSK and fading coherence length $n_c/L = 100$ fixed: Advantage of DSTM over single antenna modulation

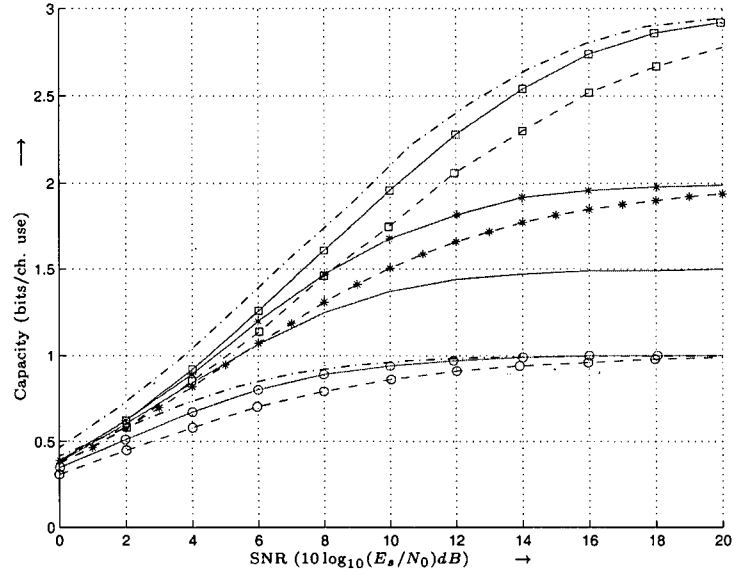
Figure 3.5: Comparison between D4PSK and DSTM: Diversity Analysis

a higher random coding exponent implies better performance since less E_s/N_0 is required to achieve the same WER. In fig. 3.5(b), the random coding exponent is shown as a function of the total decoding delay t_d for the case when the transmitted vector length DN is maintained constant. We plot the SNR $10\log_{10}(E_b/N_0)$ required to achieve $P_e \leq 10^{-3}$ according to (3.33) as function of code length n_c for transmission with rate $R = 1$ bit per scalar channel use. The normalized coherence length is $n_c/L = 100$, i.e. L degrees of freedom, and thus coded diversity increases linearly with code length n_c . Two things are interesting to note. First, increasing N consistently improves power efficiency regardless of n_c . This is decidedly different from [24], where coded diversity was decreasing with N . Second, the gain due to transmit diversity increases with decreasing n_c . This result reiterates the benefits of spatial transmit diversity if coded diversity is limited, e.g., due to delay constraints.

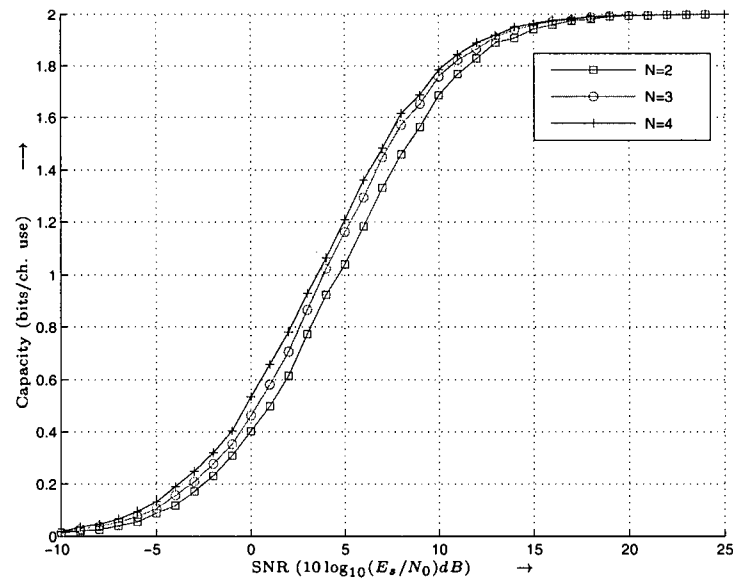
B) Capacity Results

In fig. 3.6 we compare the channel capacity obtained with DSTM with that from single antenna DPSK, and see the potential advantages of coding over multiple antennas i.e $N_T > 1$. In 3.6(a), we plot the normalized capacity $C(N)$ over the SNR i.e $10\log_{10}(E_s/N_0/N_R)$ for CDD ($N = 2$). We can see that increasing the number of transmit antennas N_T from 1 to 2 consistently increases the gain by about 2 dB in power efficiency for $N_R = 1$. We take constellations with the best performance [[16, 15, 36, 17]], and refer to section 2.2 to obtain the best constellations. As a summary, the rates and the corresponding DSTM constellations are given as follows

Modulation Rate R_m in bits/(channel use)	$N_T = 2$
1.0	$C_{CA}(2, 4)$
2.0	$C_{OD}(16)$
3.0	$C_{OD}(64)$

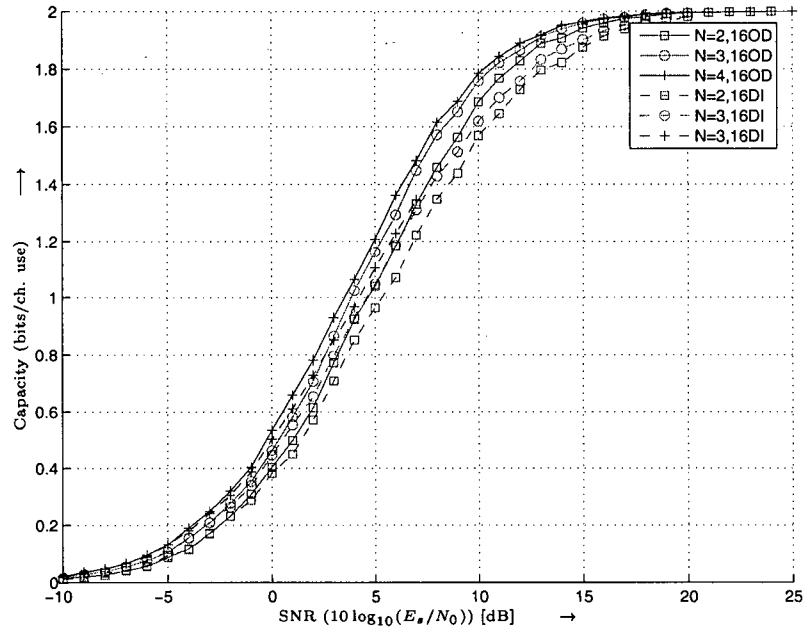


(a) Capacity $C(N)$ for DSTM: Advantage over DPSK. Solid lines: DSTM, $N = 2$; Dashed lines: D_4PSK , $N = 2$; Dash-dotted lines: DSTM, MSDD ($N = 3$)

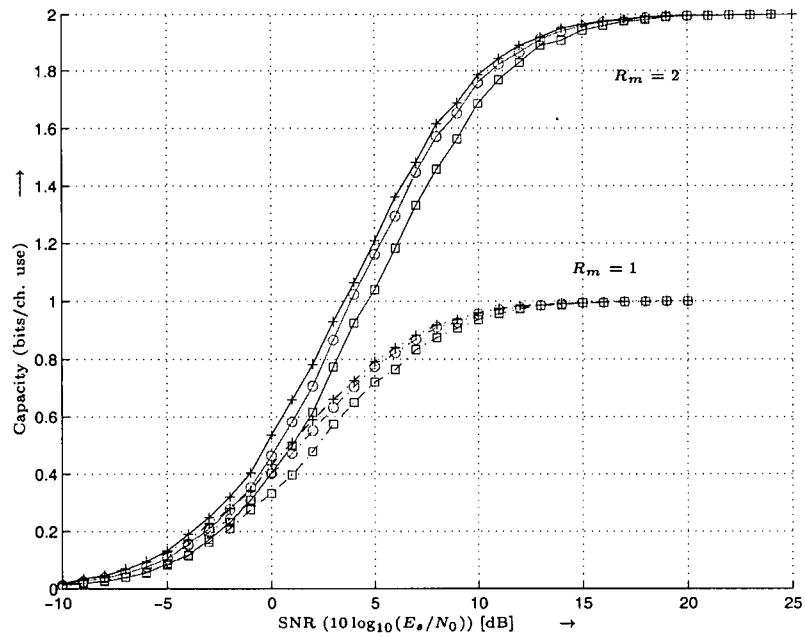


(b) Capacity $C(N)$ for DSTM: Orthogonal Codes with modulation rate $R_m = 2$ bits/(channel use), MSDD with $N \in \{2, 3, 4\}$

Figure 3.6: Capacity $C_{DSTM}(N)$ results for DSTM, Orthogonal Codes with modulation rate $R_m = 2$ bits/(channel use)



(a) $C(N)$ for $R_m = 2$ bits/(ch. use), orthogonal and diagonal constellations for $N = 2, 3, 4$



(b) $C(N)$ for $C_{OD}(\cdot)$, rates $R_m = 1, 2$ bits/(ch. use)

Figure 3.7: Capacity $C_{DSTM}(N)$ for DSTM

The advantage of MSDD with DSTM is shown in fig. 3.6(b), where we show the effect of increasing N on the channel capacity. As expected, while operating under unconstrained decoding delay scenarios for channel capacity, the increasing MSDD observation window N provides an increase in channel capacity for any given SNR of interest. Noting from [23] that in a perfectly interleaved fading channel, time diversity is utilized through channel coding, and additional transmit diversity is limited, we take into consideration $N_T = 2$ DSTM analysis. As mentioned, the best DSTM constellations are taken according to table 3.2.4, and this can be seen from fig. 3.7(a). A comparison is shown between orthogonal constellations and diagonal constellations for $N = 2, 3, 4$, and we can see that orthogonal constellations perform best for modulation rate $R_m = 2$ bits/(channel use). We also show in fig. 3.7(b) the capacities for 2 different modulation rates i.e $R_m = 1, 2$ bits/(channel use), and $N = 2, 3, 4$. The constellations are taken from table 3.2.4.

3.3 Single Antenna Systems

For the case of $N_T = N_R = 1$, traditional single antenna differential modulation with MSDD results. The analysis of the random coding exponent exactly follows along the lines of 3.2.3, and we forego the analysis and present only the results, which we apply in section 3.3.3 for performance analysis using the random coding exponent.

3.3.1 Noncoherent SISO Channel PDF

For noncoherent transmission over a single transmit and receive antenna system, the channel described in 2.1.3 is completely characterized by the N dimensional pdf $p(\mathbf{r}_{\ell,d}[k]|\mathbf{v}_{\ell,d}[k])$ of the received vector $\mathbf{r}_{\ell,d}[k]$ given the transmitted vector $\mathbf{v}_{\ell,d}[k]$. Since $\mathbf{r}_{\ell,d}[k]$ conditioned on $\mathbf{s}_{\ell,d}[k]$ is the sum of zero-mean complex Gaussian random variables, the

conditional pdf according to [32] reads as

$$p(\mathbf{r}_{\ell,d}[k]|\mathbf{s}_{\ell,d}[k]) = \frac{\exp(-\mathbf{r}_{\ell,d}^H[k]\mathbf{s}_{\ell,d}[k]\psi_{hh}\mathbf{s}_{\ell,d}[k]^H + \sigma_n^2\mathbf{I}_N)^{-1}\mathbf{r}_{\ell,d}[k])}{\pi^N \det\{\mathbf{s}_{\ell,d}[k]\psi_{hh}\mathbf{s}_{\ell,d}[k]^H + \sigma_n^2\mathbf{I}_N\}} \quad (3.39)$$

where ψ_{hh} is the autocorrelation matrix. From this equation, the conditional probability $p(\mathbf{r}_{\ell,d}[k]|\mathbf{v}_{\ell,d}[k])$ is obtained by averaging the above probability $p(\mathbf{r}_{\ell,d}[k]|\mathbf{s}_{\ell,d}[k])$ with respect to the reference symbol $\mathbf{s}_{\ell,d}[k]$. For the special case of M -ary DPSK, the averaging mentioned above can be omitted, and we can further simplify the pdf 3.39 as

$$p(\mathbf{r}_{\ell,d}[k]|\mathbf{v}_{\ell,d}[k]) = \frac{\exp(-\mathbf{r}_{\ell,d}^H[k]\mathbf{V}_{\ell,d}[k](\psi_{hh} + \sigma_n^2)^{-1}(\mathbf{r}_{\ell,d}[k]\mathbf{V}_{\ell,d}[k]^H))}{\pi^N \det\{\psi_{hh} + \sigma_n^2\mathbf{I}_N\}} \quad (3.40)$$

analogous to the multiple transmit and receive antenna metric.

3.3.2 Random Coding Exponent and Capacity Analysis

As previously mentioned, we disassociate the channel coherence interval from the MSDD observation window, making the channel fading remain constant for DN symbols, where $DN - 1$ transmitted differential symbols are differentially encoded to produce transmission vectors of length DN . The analysis for the random coding exponent closely follows from the analysis done in section 3.2.3, so only the result will be reproduced here. Following the theoretical analysis, the random coding exponent for single antenna differential modulation can be written compactly as

$$E(\rho, N, D) = \max_{0 \leq \rho \leq 1} \{E_0(\rho, N, D) - \rho R_v\} \quad (3.41)$$

with Gallager's function $E_0(\rho, N, D)$ given by

$$E_0(\rho, N, D) = -\frac{-L}{n_v} \log_2 \int_{\mathbf{r}_\ell} \left(\sum_{\mathbf{v}} \prod_{d=1}^D Pr\{\mathbf{v}_{\ell,d}\} \lambda(\mathbf{r}_{\ell,d}|\mathbf{v}_{\ell,d})^{\frac{-\rho}{1+\rho}} \cdot p(\mathbf{r}_\ell|\mathbf{v})_\ell \right) \cdot \left(\sum_{\mathbf{v}} \prod_{d=1}^D Pr\{\mathbf{v}_{\ell,d}\} \lambda(\mathbf{r}_{\ell,d}|\mathbf{v}_{\ell,d})^{\frac{1}{1+\rho}} \right)^\rho d\mathbf{r}_\ell \quad (3.42)$$

Once again, we apply the metric simplification from section 3.2.3 to reduce the number of candidate vectors for maximum likelihood decoding, by rewriting the expression for Gallager's function as

$$E_0(\rho, N, D) = -\frac{1}{(N-1)D} \log_2 \left[\mathcal{E}_{\mathbf{r}_\ell, \mathbf{h}_\ell} \left\{ \prod_{d=1}^D \mathcal{E}_{\mathbf{v}_{\ell,d}} \left\{ \exp(\lambda(\mathbf{r}_{\ell,d}|\mathbf{v}_{\ell,d}))^{-\frac{\rho}{1+\rho}} p(\mathbf{r}_{\ell,d}|\mathbf{v}_{\ell,d}, \mathbf{h}_\ell) \right\} \left(\mathcal{E}_{\mathbf{v}_{\ell,d}} \left\{ \exp(\lambda(\mathbf{r}_{\ell,d}|\mathbf{v}_{\ell,d}))^{\frac{1}{1+\rho}} \right\} \right)^\rho (\mathcal{E}_{\mathbf{v}_{\ell,d}} \{p(\mathbf{r}_{\ell,d}|\mathbf{v}_{\ell,d}, \mathbf{h}_\ell)\})^{-1} \right\} \right] . \quad (3.43)$$

Channel Capacity

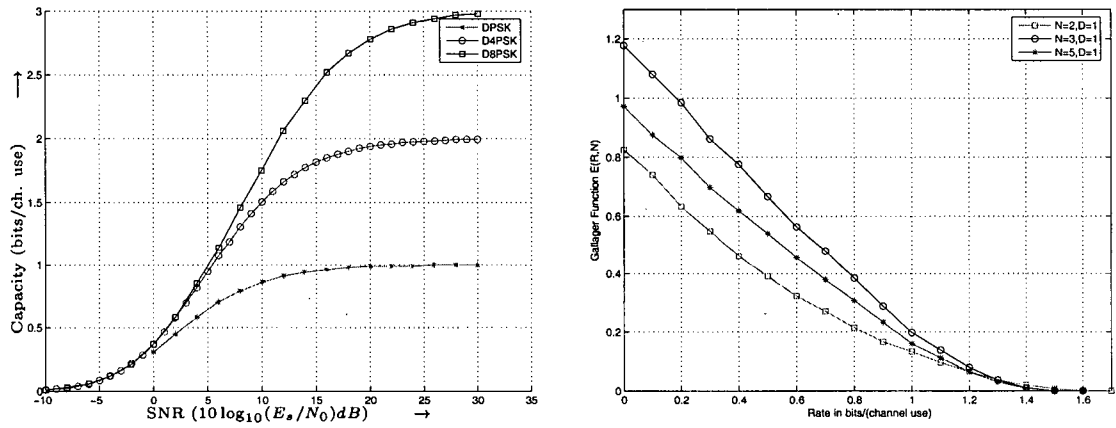
As discussed in section 3.2.2, channel capacity is another essential information theoretic tool to classify the memoryless block matrix channel between the transmitted symbols $\mathbf{v}_{\ell,d}$ and the received symbols $\mathbf{r}_{\ell,d}$. Similar to the single antenna case, "capacity" in this case refers to the associated mutual information normalized with the respect to each channel use i.e $N-1$ as

$$C(N) \triangleq \frac{1}{N-1} \cdot I(\mathbf{r}_{\ell,d}; \mathbf{v}_{\ell,d}) = \frac{1}{N-1} \mathcal{E}_{\mathbf{r}_{\ell,d}, \mathbf{v}_{\ell,d}} \left\{ \log_2 \left(\frac{p(\mathbf{r}_{\ell,d}|\mathbf{v}_{\ell,d})}{p(\mathbf{r}_{\ell,d})} \right) \right\} \quad (3.44)$$

Fig. 3.8(a) demonstrates the channel capacity for single antenna DPSK for different rates $R = 1, 2, 3$ bits/(channel use), and fig. 3.8(b) illustrates the random coding exponent for different MSDD observation intervals N for single antenna DPSK. Both channel capacity and the random coding exponent are demonstrated and explained in section 3.3.3.

3.3.3 Performance Analysis: Results and Discussion

In this section, the presented analysis for the random coding exponent for single antenna DPSK as well as channel capacity are discussed. Correspondingly, channel ca-



(a) Channel capacities $C(N)$ for single antenna DPSK, Rates $R_v = 1, 2, 3$ bits/(channel use)

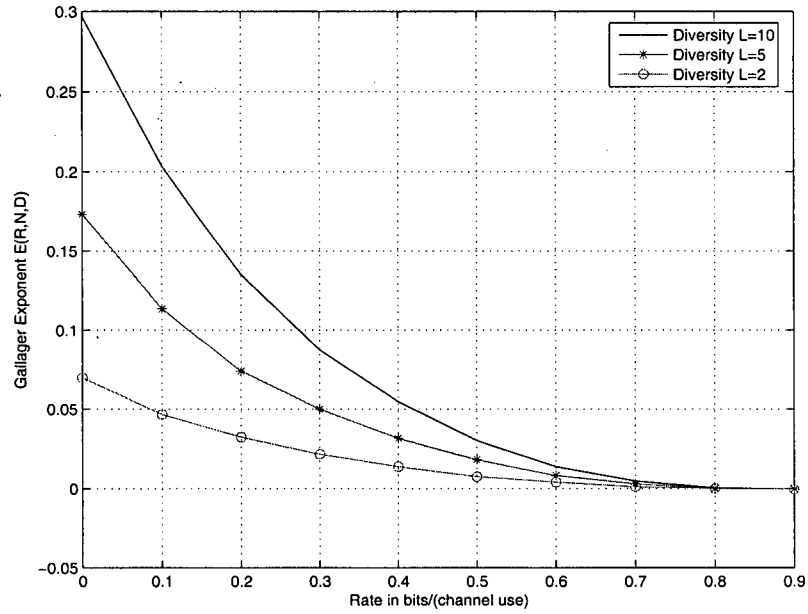
(b) Random coding exponent $E(R, N, 1)$ for single antenna D4PSK

Figure 3.8: Capacity and Gallager functions for Single Antenna D4PSK

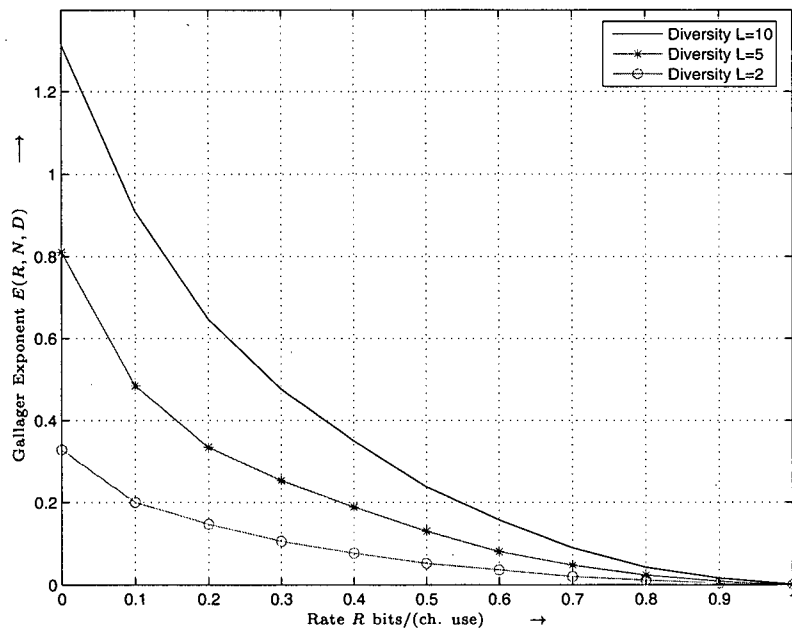
capacity and the word error rate obtained from the random coding exponent are analyzed as a function of the required signal to noise ratio $10 \log_{10}(E_s/N_0)$, and decoding delay in PSK samples t_d respectively. As general and interesting cases, for single antenna modulation we consider 4-ary and 8-ary modulation supporting 2 and 3 bits/(channel use) respectively. Specifically, we consider D4PSK and D8PSK as suitable signal constellations for transmission. Throughout the thesis we consider block fading channels as already discussed in section 2.1.3.

Random Coding Exponent Results

In this analysis, unlike channel capacity which has been discussed at length in literature, we put our transmission system under decoding delay constraints. As mentioned, we consider the random coding exponent to be a necessary and adequate information theoretic tool to analyze the transmission system, and to this end we discuss the average

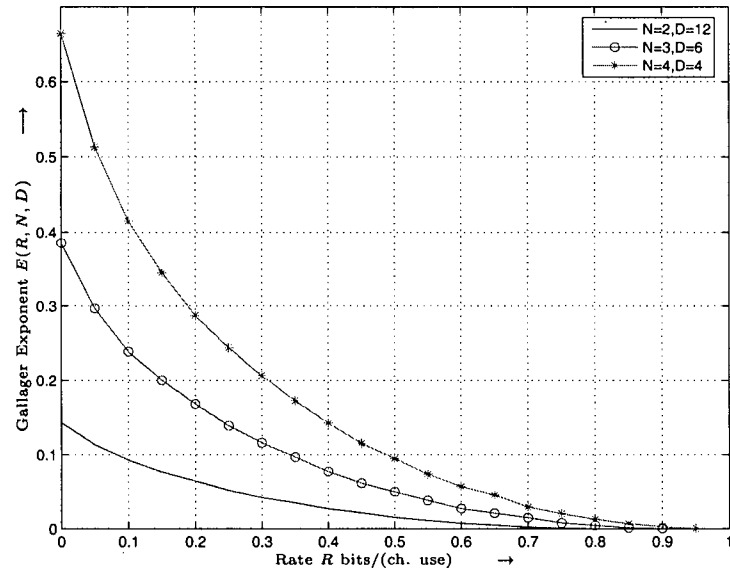


(a) $E(R, N, D)$ vs Rate R bits/(channel use) for single antenna D_4PSK , CDD ($N = 2$)

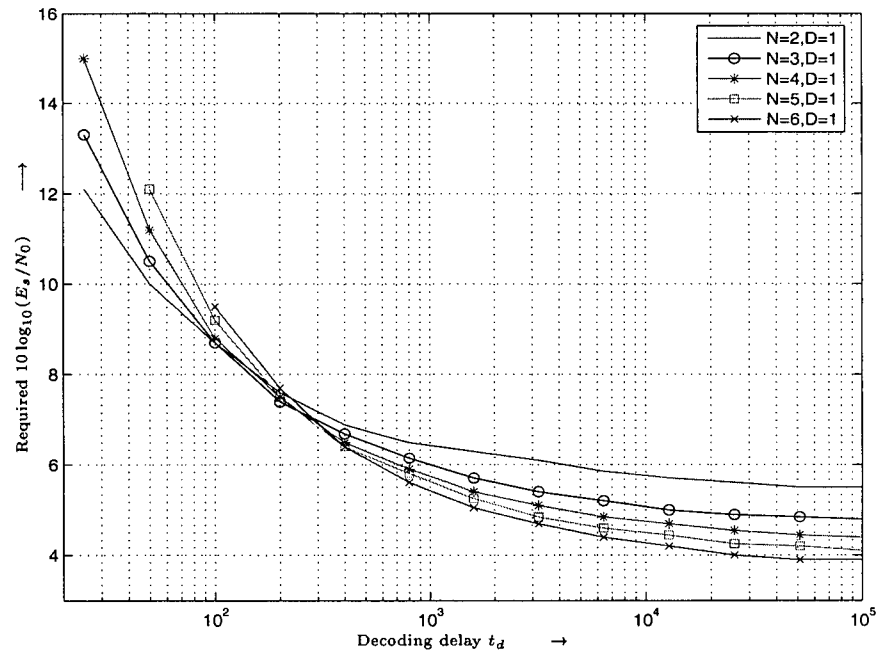


(b) $E(R, N, D)$ vs Rate R bits/(channel use) for single antenna D_4PSK , MSDD ($N = 6$), decoding delay in PSK symbols=100

Figure 3.9: Gallager exponent $E(R, N, D)$ vs Rate R , effect of coding diversity



(a) Gallager exponent $E(R, N, D)$ vs Rate for single antenna D4PSK, coding diversity=12, $N = 2, 3, 4$



(b) Required $10 \log_{10}(E_s/N_0)$ vs Coding Delay t_d in PSK symbols, standard MSDD vector channel with channel coherence interval= N i.e $D = 1$

Figure 3.10: Gallager random coding exponent results: Single Antenna D4PSK, delay-SNR analysis

signal to noise ratio (SNR) yielding a random coding exponent, which in turn leads to a word error rate (WER) not greater than a given threshold P_w . Also, in order to clearly emphasize the effect of various block fading scenarios on the random coding exponent, we show the Gallager exponent $E(R, N, D)$ as a function of rate R in addition to the Gallager exponent-delay-SNR examples, to clearly illustrate the effect of the channel coherence interval on the random coding exponent. In all cases, we consider either D8PSK or D4PSK with supported rate 3 and 2 bits/(channel use) respectively and a block fading channel with channel coherence interval $DN - 1$ differential symbols. The target rate considered is $R = R_v/(D.(N - 1))$ where R_v is the rate in bits/(vector symbol).

In figure 3.9, we show the random coding exponent $E(R, N, D)$ as a function of rate R for D4PSK, where the decoding delay t_d is kept fixed at 100 PSK samples, and the transmission diversity obtained by the varying fading coherence interval is changed. From figure 3.9(a) we note that as the fading diversity L is increased from 2 to 5 to 10, the random coding exponent increases, agreeing with our discussing wherein for the same decoding delay, the channel sees more independent fading realizations, thereby increasing the diversity and correspondingly the random coding exponent. The same argument is employed for figure 3.9(b), where instead of CDD as in the previous figure, MSDD with $N = 6$ is employed.

In figure 3.10, the effect of MSDD is shown. For a fair comparison, the normalized coherence length DN is kept constant at 12, and the MSDD observation window N is varied. From this, we can clearly see the beneficial effect of increasing N while keeping the overall diversity constant, as the differential encoding effect becomes more apparent. Fig. 3.10(b) shows the required $10 \log_{10}(E_s/N_0)$ required to obtain a WER not greater than 0.001 i.e $P_w \leq 0.001$. For long delays, the situation resembles the capacity case, where the decoding delay does not affect the performance significantly, and the effect of MSDD is apparent in the reduced SNR required to achieve a particular

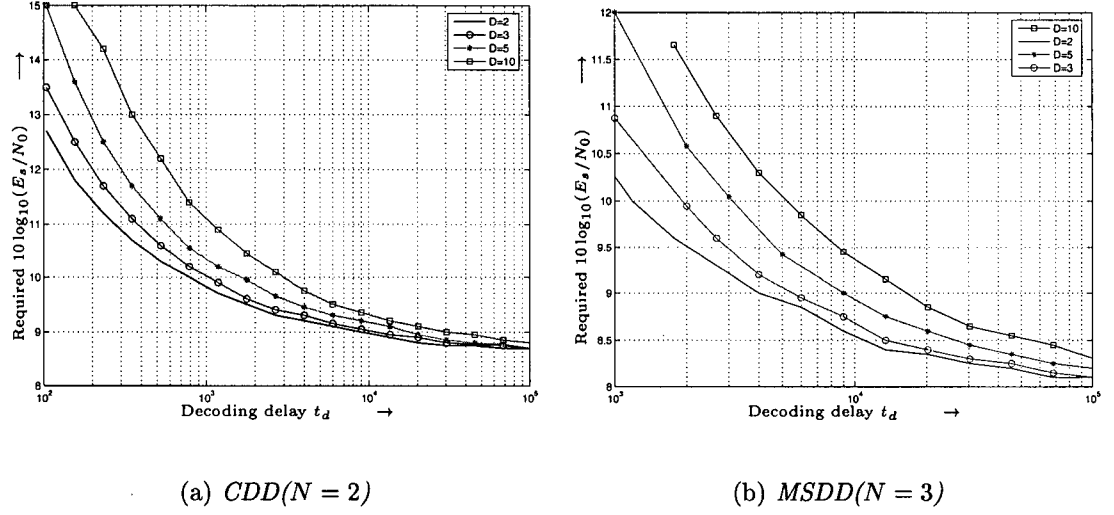


Figure 3.11: Required $10 \log_{10}(E_s/N_0)$ vs coding delay t_d for single antenna D8PSK, $CDD(N=2)$, fading coherence interval increasing with D

WER. However for shorter delays, the effect of shorter component codes or the reduced diversity seen from the fewer independent fading realizations outweighs the effect of MSDD, and the curves intersect where increasing the observation window results in a deterioration in performance for $t_d < 10^2 \dots 10^3$.

In order to demonstrate the effect of the channel with fading coherence interval not equal to N , we observe from figure 3.11 the effect of increasing D while keeping N the same. Increasing D effectively reduces the number of independent fading realizations seen for a given decoding delay t_d , hence reducing the effective coding diversity. The effect of this reduction can be seen in the increased SNR required to achieve a particular WER as D increases. Finally, fig. 3.12 shows $E(R, N, D)$ as function of R for D4PSK transmission, SNR $10 \log_{10}(E_s/N_0) = 10$ dB, and fixed normalized coherence length $n_c/L = 40$. As can be seen from fig. 3.12, and as expected from MSDD performance results in e.g. [8], the random coding exponent for MSDD approaches that of coherent detection with CSI, which is included as reference curve, with increasing N .

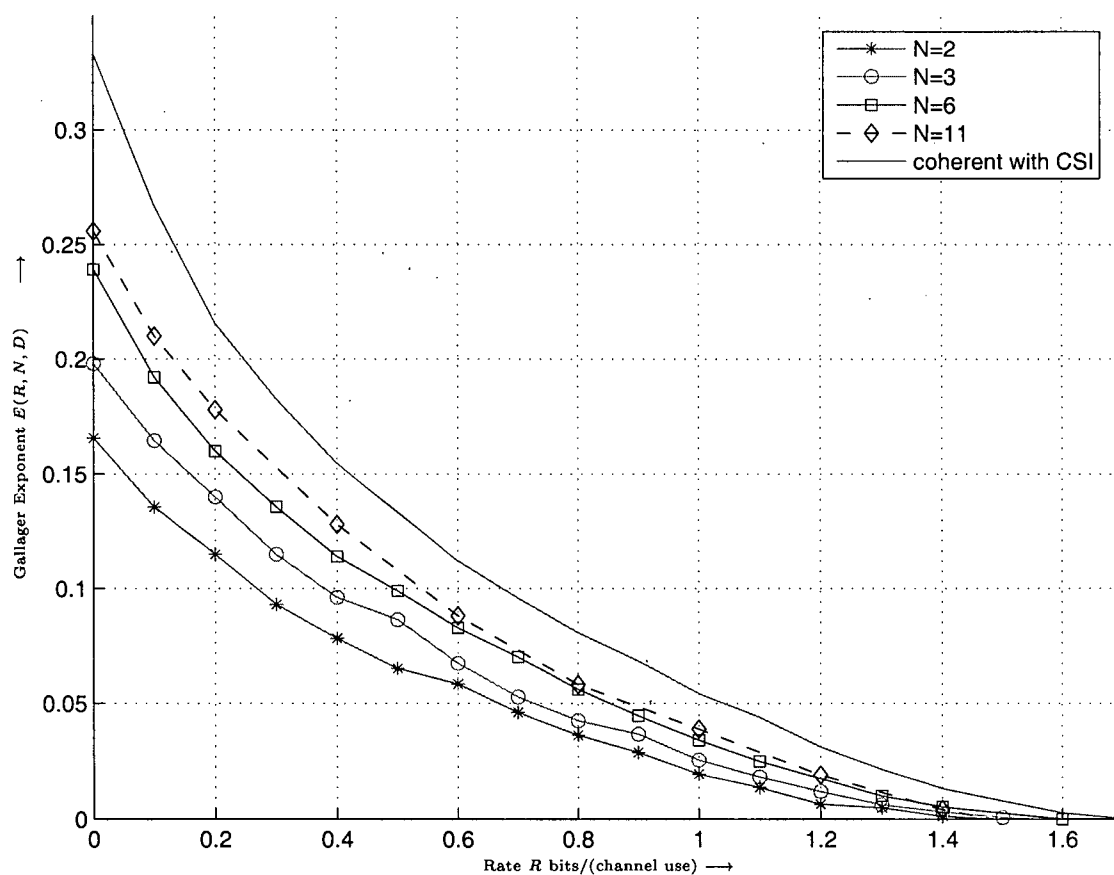


Figure 3.12: Gallager Exponent $E(R, N, D)$ vs R for $D4PSK$ ($N_T = 1$) and $10 \log_{10}(E_s/N_0) = 10$ dB

Chapter 4

Modified Tree-Search and Sphere Decoding MSDD Algorithms

In the previous chapter, we derived expressions for the random coding exponent for both single antenna systems and DSTM where the channel coherence interval was independent from the MSDD *observation window*. This helped us analyze a wide range of channels without any restrictions on the channel coherence interval, for example if the *coding diversity* was kept fixed but the MSDD observation window N was allowed to vary for analysis of different MSDD strategies. However, one drawback of this derivation was that the expression for Gallager's function $E_0(\rho, D, N)$ given by

$$E_0(\rho, N, D) = -\frac{1}{(N-1)D} \log_2 \left[\mathcal{E}_{\mathbf{R}_\ell, \mathbf{H}_\ell} \left\{ \prod_{d=1}^D \mathcal{E}_{\mathbf{V}_{\ell,d}} \left\{ \exp(\lambda(\mathbf{R}_{\ell,d} | \mathbf{V}_{\ell,d}))^{-\frac{\rho}{1+\rho}} p(\mathbf{R}_{\ell,d} | \mathbf{V}_{\ell,d}, \mathbf{H}_\ell) \right\} \right. \right. \\ \left. \left. \left(\mathcal{E}_{\mathbf{V}_{\ell,d}} \left\{ \exp(\lambda(\mathbf{R}_{\ell,d} | \mathbf{V}_{\ell,d}))^{\frac{1}{1+\rho}} \right\} \right)^\rho \right. \right. \\ \left. \left. \left(\mathcal{E}_{\mathbf{V}_{\ell,d}} \{ p(\mathbf{R}_{\ell,d} | \mathbf{V}_{\ell,d}, \mathbf{H}_\ell) \} \right)^{-1} \right\} \right] . \quad (4.1)$$

involved summing over all possible vectors $\mathbf{V}_{\ell,d}$ of length N , where the value for N could involve an extremely high complexity cost as the complexity for MSDD increases exponentially with not only N but also with the number N_T of transmit antennas and the modulation rate R_m . In order to reduce the complexity cost for analysis of such block fading channels, we make use of efficient tree-search algorithms to overcome the complexity limitations of MSDD for differential space-time modulation([34]), and with novel modifications to this tree-search approach we obtain an efficient implementation of Gallager's random coding exponent for our block fading channel under consideration. Furthermore, we propose a novel "modified metric" using this stack algorithm to provide accurate bounds for the random coding exponent and time delay analysis, which reduces the computational cost for analysis of our modified block fading channel by many orders of magnitude.

In section 4.1 we consider the tree-search algorithm for single antenna systems, and explain the working of the tree-search stack algorithm. This algorithm will later be extended to the DSTM case in section 4.2 for Alamouti's $N_T \times N_T$ for $N_T = 2$ code as well as diagonal DSTM constellations, where we use a modified sphere decoding algorithm to obtain computational efficiency during our search. We then achieve accurate bounds for the random coding exponent as well as the time-delay analysis of the random coding exponent in section 4.3 by using a combination of efficient modified tree-search algorithms, the modified performance metric for calculation of the random coding exponent and a novel sorting approach that enables us to formulate an efficient "stopping criteria" to terminate our MSDD tree-search for an increase in computational efficiency. We also analyze the effect of this modified metric on the random coding exponent and formulate a tradeoff between the efficiency or complexity of the tree-search algorithm and the tightness of the corresponding bound on the random coding exponent. Since our work is focused primarily on space-time coding, the analysis done for single-antenna systems serves as a reference for the DSTM case,

4.1 Single Antenna Systems: Multiple Symbol Differential Sphere Decoding (MSDSD)

where we include some more detail to the analysis. We present results for the various approximations and bounds in section 4.4 for multiple antenna transmission as well as single antenna DPSK for reference.

4.1 Single Antenna Systems: Multiple Symbol Differential Sphere Decoding (MSDSD)

In this section, we review some efficient algorithms for reducing the complexity cost of MSDD, where the complexity for decoding increases exponentially with the MSDD observation window N . We mainly focus on tree-search algorithms for MSDD for single antenna systems, and apply sphere decoding to Maximum Likelihood-MSDD (ML-MSDD) also known as multiple symbol differential sphere decoding (MSDSD) (see [22]).

In this section, we devise efficient sphere decoding algorithms for MSDD corresponding to [22], where we reproduce the results here to make the analysis in further sections clearer. For maximum likelihood MSDD (ML-MSDD), we obtain the estimate $\hat{\mathbf{s}}_{\ell,d}$ based on the observation of the N -dimensional received MSDD vector \mathbf{r} . For the case of Rayleigh fading, the maximum likelihood decision rule is [13]

$$\hat{\mathbf{s}}_{\ell,d} = \underset{\mathbf{s}_{\ell,d}}{\operatorname{argmin}} \{ \mathbf{r}_{\ell,d}^H \mathbf{R}_{\mathbf{r}\mathbf{r}}^{-1} \mathbf{r}_{\ell,d} \} \quad (4.2)$$

with the correlation matrix

$$\mathbf{R}_{\mathbf{r}\mathbf{r}} \triangleq \mathcal{E} \{ \mathbf{r}_{\ell,d} \mathbf{r}_{\ell,d}^H | \mathbf{s}_{\ell,d} \} \quad (4.3)$$

Applying the Cholesky factorization of the inverse matrix $\mathbf{C}^{-1} = \mathbf{L}\mathbf{L}^H$, and further defining $\mathbf{U} \triangleq (\mathbf{L}^H \operatorname{diag}\{\mathbf{r}_{\ell,d}\})^*$, where \mathbf{L}, \mathbf{U} are upper and lower triangular matrices

respectively, we obtain the ML decision metric

$$\hat{\mathbf{s}}_{\ell,d} = \underset{\mathbf{s}_{\ell,d}}{\operatorname{argmin}} \{ \|\mathbf{U}\mathbf{s}_{\ell,d}\|^2 \} \quad (4.4)$$

This is generally regarded as a shortest vector problem that can be solved using sphere decoding. The sphere decoder examines only those candidate vectors $\mathbf{s}_{\ell,d}$ that lie within a sphere of radius (distance) R

$$\|\mathbf{U}\mathbf{s}_{\ell,d}\|^2 \leq R \quad (4.5)$$

As has been discussed, due to the upper triangular form of \mathbf{U} , the metric for ML detection (see 4.4) can be checked component-wise i.e. having found initial candidates $\hat{\mathbf{s}}_l$ for the $N - i$ previous components $\mathbf{s}_l, i + 1 \leq l \leq N$, where we eliminate the indices ℓ, d for simplicity, we can obtain a condition for the i th component as follows. If the squared length,

$$d_{i+1}^2 = \prod_{l=i+1}^N \left| \prod_{j=l}^N u_{lj} \hat{\mathbf{s}}_j \right|^2 \quad (4.6)$$

where u_{ij} is the \mathbf{U} entry in row i and column j , $1 \leq i, j \leq N$, then the candidates \mathbf{s}_i have to satisfy the criteria

$$d_i^2 = \left| u_{ii}\mathbf{s}_i + \prod_{j=i+1}^N u_{ij} \hat{\mathbf{s}}_j \right|^2 + d_{i+1}^2 \leq R^2 \quad (4.7)$$

Once a valid vector $\hat{\mathbf{s}}$ is found ($i = 1$), then the radius is dynamically updated by $R = \|\mathbf{U}\hat{\mathbf{s}}\|$ and the sphere decoding is repeated with $i = 2$ and the updated radius. If no candidate vector $\hat{\mathbf{s}}$ is found within the updated radius R , MSDSD stops.

The stack algorithm used in this thesis basically eliminates one drawback of the sphere decoding algorithm, where searching needs to start from the beginning each time the decoded is called to provide the next best candidate. In the stack algorithm, we calculate an initial metric (or sphere decoding distance) from the upper triangular matrix \mathbf{U} , and each time replace the path at the top of the stack with all its extensions. This process is repeated, and the first path of MSDD length N at the top of the stack

is the optimal path corresponding to the best transmitted vector $\mathbf{V}_{\ell,d}$. We then store the stack and continue with this stack when the next best candidate is required (see section 4.4.1).

4.2 Multiple Antenna Systems: Tree Search based Sphere Decoding Algorithms

As mentioned previously, similar to the single antenna case, the complexity of MSDD for differential space time modulation too increases exponentially with the MSDD observation interval N . A number of sub-optimum decoders have been proposed for DSTM specific to certain Space-Time constellations and/or MSDD observation intervals ([11],[38],[6]). Lattice reduction based techniques ([5]) and sphere decoding techniques ([7]) amongst others have been proposed in literature, and to this end we apply tree-search based sphere decoding algorithms for efficient decoding for our purpose. From [34] the maximum likelihood (ML) metric can be rewritten as a sum of N non-negative scalar terms

$$\delta_n^2 \triangleq \|\mathbf{V}_{\ell,d} \mathbf{R}_{n,n,(\ell,d)} + \mathbf{X}_{\ell,d}\|^2 \quad (4.8)$$

where

$$\mathbf{X}_{\ell,d} \triangleq \mathbf{S}_{n+1,\ell,d} \sum_{j=n+1}^N \mathbf{S}_{j,\ell,d}^H \mathbf{R}_{n,j,(\ell,d)} \quad (4.9)$$

and $\mathbf{R}_{n,j,(\ell,d)} \triangleq \frac{p_{j-n}^{N-n}}{\sigma_e^{N-n}} \mathbf{R}_{j,\ell,d}$, $1 \leq n \leq N-1$, $n \leq j \leq N$. $p_j^{(n)}$ and $(\sigma_e^{(n)})^2$ denote the j th coefficient of the n th order linear backward minimum mean squared error (MMSE) predictor for the discrete time random process $\mathbf{H}_{\ell,d} + \mathbf{N}_{\ell,d}$ and the corresponding variance respectively, and $p_0^{(n)} = -1$. Defining

$$d_n^2 \triangleq \sum_{i=n}^{N-1} \|\mathbf{V}_i \mathbf{R}_{i,i} + \mathbf{X}_i\|^2 = d_{n+1}^2 + \delta_n^2, \quad 1 \leq n \leq N-1, \quad (4.10)$$

where $d_N^2 = 0$ and d_1^2 is the ML metric in 4.9, we simplify the notation by eliminating ℓ, d . We begin the sphere decoding algorithm at $n = N - 1$, and select the candidate transmitted vector \mathbf{V}_n based on tentative decisions $\mathbf{V}_j = \hat{\mathbf{V}}_j, n + 1 \leq j \leq N - 1$ and we decrement n until the current metric d_n does not exceed a certain maximum metric ρ i.e $d_n \leq \rho$. Once we reach the top of the "stack" i.e. have traversed through until $n = 1$, we use the metric of the currently examined best candidate \mathbf{V}_n to reduce the search space by updating $\rho = d_1$. If, for any n , d_n exceeds ρ , n is incremented and a new candidate transmitted vector $\hat{\mathbf{V}}_{\ell,d}$ is examined. Once the sphere decoder reaches $n = N$ meaning that there are no candidate vectors inside the current search sphere, the algorithm stops.

4.2.1 Algorithms for Orthogonal Space-Time Codes

For orthogonal designs, given by

$$\mathbf{V}_o \triangleq \left\{ \frac{1}{\sqrt{2}} \begin{bmatrix} a & -b^* \\ b & a^* \end{bmatrix} \quad a, b \in \sqrt{M}PSK \right\} \quad (4.11)$$

it is relatively easy to show that

$$\delta_n^2 = \gamma_n + \text{Re}\{a_n \alpha_n\} + \text{Re}\{b_n \beta_n\}, 1 \leq n \leq N - 1 \quad (4.12)$$

with α_n, β_n being functions of the elements in $\mathbf{R}_{n,n}, \mathbf{X}_{n,n}$ respectively. It turns out that the parameters a_n, b_n minimizing the distance δ_n^2 are independent of each other, hence the sphere decoding algorithm can be regarded as an extension of single antenna MSDSD as in section 4.1 to differential space-time modulation.

4.2.2 Algorithms for Diagonal Constellations

For diagonal constellations, given by the set

$$\mathbf{V}_D \triangleq \left\{ \text{diag}\{e^{j\frac{2\pi}{L}c_1}, \dots, e^{j\frac{2\pi}{L}c_{N_T}}\}^l \mid l \in \{0, \dots, L - 1\} \right\} \quad (4.13)$$

using the cosine approximation, the problem for minimizing δ_n^2 can be turned into a N_T -dimensional lattice decoding [5, 26] of the form

$$\hat{\mathbf{x}} = \underset{\mathbf{x} \in \mathbf{Z}^{N_T}}{\operatorname{argmin}} \left\{ \left\| \begin{bmatrix} b_{1,1} & & & \mathbf{0} \\ & b_{2,1} & b_{2,2} & \\ & \vdots & & \ddots \\ b_{N_T,1} & \mathbf{0} & & b_{N_T,N_T} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{N_T} \end{bmatrix} - \begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_{N_T} \end{bmatrix} \right\| \right\} \quad (4.14)$$

To obtain $\hat{\mathbf{x}}$, we apply the sphere decoding algorithm and take advantage of the lower triangular structure of 4.14. We find possible candidates for $X_i \in \mathbf{Z}$ and increase i , $2 \leq i \leq N_T$, as long as

$$\mu_i^2 \triangleq |b_{1,1}x_1 - t_1|^2 + \sum_{j=2}^i |b_{j,1}x_1 + b_{j,j}\hat{x}_j - t_j|^2 < r^2 \quad (4.15)$$

with $\hat{x}_j \triangleq [(t_j - b_{j,1}x_1)/b_{j,j}]$. If $\mu_{N_T} < r$, x_1 is considered to be the temporary decoder result, and the search continues with the updated radius $r = \mu_{N_T}$. The sphere decoder terminates when $|b_{1,1}x_1 - t_1| > r$. Hence the overall search is limited to a single dimension rather than the N_T dimensional search space of ML-MSDD.

4.3 Bounds Using Tree-Search Algorithms

In this section, we aim to present accurate bounds to the random coding exponent using the results from section 3.3.2 and the above mentioned tree-search or Stack based decoding algorithms. Here, we reintroduce the notations for the modified channel of coherence interval DN , and differentiate between the corresponding metrics for the block fading channel where the fading coherence interval is equal to the MSDD observation window N . We begin by revisiting the ML noncoherent metric $p_N(\mathbf{R}_{\ell,d}|\mathbf{V}_{\ell,d})$ given by

$$p(\mathbf{R}_{\ell,d}|\mathbf{V}_{\ell,d}) = \frac{\exp(-\operatorname{tr}\{\mathbf{R}_{\ell,d}^H \mathbf{\Psi}_R^{-1} \mathbf{R}_{\ell,d}\})}{(\pi^{NN_T} \det\{\mathbf{\Psi}_R\})^{N_R}}, \quad (4.16)$$

where

$$\Psi_{\mathbf{R}} = \mathcal{E}_{\mathbf{R}_{\ell,d}} \{ \mathbf{R}_{\ell,d} \mathbf{R}_{\ell,d}^H | \mathbf{V}_{\ell,d} \} = N_R (\sigma_h^2 \mathbf{S}_{\ell,d} \mathbf{S}_{\ell,d}^H + \sigma_n^2 \mathbf{I}_{NN_T}) \quad (4.17)$$

with some arbitrarily chosen $\mathbf{S}_{\ell,d,1}$. We note that $\det\{\Psi_{\mathbf{R}}\}$ is independent of $\mathbf{V}_{\ell,d}$. Inserting (4.17) into (4.16) the decoding metric follows as

$$\begin{aligned} \lambda_N(\mathbf{R}_{\ell,d} | \mathbf{V}_{\ell,d}) &= \log(p(\mathbf{R}_{\ell,d} | \mathbf{V}_{\ell,d})) \\ &= -\frac{\|\mathbf{R}_{\ell,d}\|^2}{N_R \sigma_n^2} + \frac{\sigma_h^2 \|\mathbf{R}_{\ell,d}^H \mathbf{S}_{\ell,d}\|^2}{N_R \sigma_n^2 (\sigma_n^2 + N \sigma_h^2)} \\ &\quad - N_R \log(\pi^{NN_T} \det\{\Psi_{\mathbf{R}}\}). \end{aligned} \quad (4.18)$$

Since we would like to provide accurate bounds on the random coding exponent $E(\rho, N, D)$, we begin with the already derived expression for Gallager's function for differential space-time modulation for the modified channel, with channel coherence interval DN given by

$$\begin{aligned} E_0(\rho, N, D) &= -\frac{n_D}{n_v} \log_2 \int \left(\sum_{\mathbf{R}_{\ell}} \prod_{\mathbf{V}_{\ell}} \prod_{d=1}^D Pr\{\mathbf{V}_d\} \lambda_N(\mathbf{R}_{\ell,d} | \mathbf{V}_{\ell,d})^{\frac{\rho}{1+\rho}} \cdot p_{DN}(\mathbf{R}_{\ell} | \mathbf{V}_{\ell}) \right) \\ &\quad \left(\sum_{\mathbf{V}_{\ell}} \prod_{k=1}^D Pr\{\mathbf{V}_{\ell,k}\} \lambda_N(\mathbf{R}_{\ell,k} | \mathbf{V}_{\ell,k})^{\frac{1}{1+\rho}} \right)^{\rho} d\mathbf{R}_{\ell} \end{aligned} \quad (4.19)$$

Different from the case of coherent detection with CSI, the MSDD metric $\lambda_N(\mathbf{R}_{\ell,d} | \mathbf{V}_{\ell,d})$ cannot be split into individual terms depending only on $\mathbf{V}_{\ell,d,n}$, $1 \leq n \leq N-1$. Thus, averaging over the set \mathbf{V}^{N-1} instead of \mathbf{V} is necessary in (4.19), which becomes computationally expensive for large values of N . We therefore consider an approximation of $E_0(\rho, N, D)$ using only a subset $\mathbf{V}_s \subset \mathbf{V}^{N-1}$ of differential symbols to reduce computational complexity when numerically evaluating $E_0(\rho, N, D)$. We begin our analysis with the special case of $D = 1$ and spatially uncorrelated fading, and based on the results obtained we look at the block fading channel under consideration, where the fading coherence interval is not fixed.

4.3.1 $D = 1$ and Spatially Uncorrelated Fading

For the interesting special case of $D = 1$ and spatially uncorrelated channels, i.e., each block of N transmitted space-time symbols experiences independent fading and 4.16 is the correct ML decoding metric, the Gallager function can be simplified to

$$E_0(\rho, N, D = 1) = -\frac{1}{(N-1)} \cdot \log_2 \left[\mathcal{E}_{\mathbf{R}_{\ell,1}} \left\{ \left(\frac{\mathcal{E}_{\mathbf{V}_{\ell,1}} \left\{ p(\mathbf{R}_{\ell,1} | \mathbf{V}_{\ell,1})^{\frac{1}{1+\rho}} \right\}}{(\mathcal{E}_{\mathbf{V}_{\ell,1}} \{ p(\mathbf{R}_{\ell,1} | \mathbf{V}_{\ell,1}) \})^{\frac{1}{1+\rho}}} \right)^{1+\rho} \right\} \right] \quad (4.20)$$

Considering (4.20) it is reasonable to choose the subset \mathbf{V}_s of the ML search space such that the ratio

$$r(C_s) \triangleq \sigma_s / \sigma < 1 \quad (4.21)$$

with

$$\sigma \triangleq \mathcal{E}_{\mathbf{V}_{\ell,1}} \left\{ p(\mathbf{R}_{\ell,1} | \mathbf{V}_{\ell,1})^{\frac{1}{1+\rho}} \right\} \quad (4.22)$$

$$\sigma_s \triangleq \sum_{\mathbf{V}_{\ell,1} \in \mathbf{V}_s} \frac{1}{M^{N-1}} (p(\mathbf{R}_{\ell,1} | \mathbf{V}_{\ell,1}))^{\frac{1}{1+\rho}} \quad (4.23)$$

is maximized for a given cardinality

$$C_s \triangleq |\mathbf{V}_s| < M^{N-1}. \quad (4.24)$$

This means we only consider the dominating terms when averaging over $\mathbf{V}_{\ell,1}$ and to do so, we need to sort matrices $\mathbf{V}_{\ell,1} \in \mathbf{V}^{N-1}$ according to decreasing values of $p(\mathbf{R}_{\ell,1} | \mathbf{V}_{\ell,1})$.

A) Sorting:

An efficient way to accomplish sorting without actually calculating $p(\mathbf{R}_{\ell,1} | \mathbf{V}_{\ell,1})$ for all $\mathbf{V}_{\ell,1} \in \mathbf{V}^{N-1}$ is to run a detection algorithm which returns the best (ML) and successively the next best estimates for $\mathbf{V}_{\ell,1}$ according to $p(\mathbf{R}_{\ell,1} | \mathbf{V}_{\ell,1})$ and given $\mathbf{R}_{\ell,1}$.

As this detection requires a tree search and the search metric is additive, we use a modified version of the stack algorithm as described in section 4.2, which continuously finds the next best signal point. We calculate the maximum likelihood metric using the tree-search based stack algorithm as described in the previous section, each time saving the stack and corresponding the path through the tree as we proceed. We then rerun the algorithm starting from the previously saved locations, thereby giving us the next best metric to the ML metric. We continue this process until sufficient accuracy and/or complexity has been reached according to 4.21.

B) Choice of C_s :

For a given C_s the ratio $r(C_s)$ in (4.21) significantly varies depending on the realization $\mathbf{R}_{\ell,1}$. It is therefore advisable not to fix C_s but to choose it adaptively when Monte-Carlo integrating (4.20). To this end, let $\mathbf{V}_{\ell,1}^{(C_s)}$ be the C_s th best term found by the stack algorithm. The upper bound

$$\sigma_u \triangleq \sigma_s + (1 - C_s/M^{N-1}) \left(p(\mathbf{R}_{\ell,1} | \mathbf{V}_{\ell,1}^{(C_s)}) \right)^{\frac{1}{1+\rho}} \geq \sigma \quad (4.25)$$

for σ allows to lower bound $r(C_s)$ by

$$r(C_s) \geq \frac{\sigma_s}{\sigma_u}. \quad (4.26)$$

We then can choose C_s such that the lower bound exceeds a certain threshold value of, e.g., 0.9 for 90% accuracy of the Stack Algorithm. We can then control the accuracy of our approximation, by modifying the parameter $r(C_s)$. A higher value of $r(C_s)$ closer to 1 indicates a higher accuracy, however it would also involve a higher complexity. In this case, rather than computing the additions or complex multiplications to gauge complexity, we define complexity as the number of candidate vectors in the signal space of dimension M^{DN} that the stack algorithm tests in order to achieve a certain accuracy for the random coding exponent (RCE) bound.

C) Bounds for $E_0(\rho, N, D)$:

Let us define

$$\eta \triangleq (\mathcal{E}_{\mathbf{V}_{\ell,1}} \{p(\mathbf{R}_{\ell,1}|\mathbf{V}_{\ell,1})\})^{\frac{1}{1+\rho}} \quad (4.27)$$

$$\eta_s \triangleq \left(\sum_{\mathbf{V}_{\ell,1} \in \mathbf{V}_s} \frac{1}{M^{N-1}} p(\mathbf{R}_{\ell,1}|\mathbf{V}_{\ell,1}) \right)^{\frac{1}{1+\rho}} \quad (4.28)$$

$$\eta_u \triangleq \left(\sum_{\mathbf{V}_{\ell,1} \in \mathbf{V}_s} \frac{1}{M^{N-1}} p(\mathbf{R}_{\ell,1}|\mathbf{V}_{\ell,1}) + (1 - C_s/M^{N-1}) p(\mathbf{R}_{\ell,1}|\mathbf{V}_{\ell,1}^{(C_s)}) \right)^{\frac{1}{1+\rho}}. \quad (4.29)$$

We note that $(\sigma/\eta)^{1+\rho}$ is used in (4.20). Since the function $f(x) = x^{1/(1+\rho)}$ is concave for $0 \leq \rho \leq 1$, the inequalities

$$\frac{\sigma_s}{\eta_s} \leq \frac{\sigma}{\eta} \leq \frac{\sigma_u}{\eta_u} \quad (4.30)$$

hold. Hence, $E_0(\rho, N, D = 1)$ in (4.20) is, respectively, upper and lower bounded when using the subset \mathbf{V}_s and replacing the exact ratio σ/η with σ_s/η_s and σ_u/η_u , respectively.

4.3.2 $D > 1$ and Spatially Uncorrelated Fading

In the previous section, we dealt with the special case of $D = 1$, where we reverted to the traditional block fading channel with the fading coherence interval equal to the MSDD observation interval. We outline a procedure to obtain lower bounds on the random coding exponent and Gallager's function $E_0(\rho, N, D)$ for any general block fading channel, i.e. $D > 1$ as well as spatially correlated fading. Although in this case the pdf $p(\mathbf{R}_{\ell,d}|\mathbf{V}_{\ell,d}, \mathbf{H}_\ell)$ is used for Monte-Carlo integrating (3.36), and thus $r(C_s)$ does not appear in the expression for Gallager's function, we extend the application of the bound (4.26) to determine the set \mathbf{V}_s to these cases. The reasons for this are that

(a) $r(C_s)$ (or its bound (4.26)) is still a good indicator whether the terms dominating the argument in (3.36) for given \mathbf{R}_ℓ and \mathbf{H}_ℓ are taken into account and (b) sorting $\mathbf{V}_{\ell,d}$ according to a mixed criterion of $\exp(\lambda(\mathbf{R}_{\ell,d}|\mathbf{V}_{\ell,d}))^{-\rho/(1+\rho)}$ and $p(\mathbf{R}_{\ell,d}|\mathbf{V}_{\ell,d})$ or $p(\mathbf{R}_{\ell,d}|\mathbf{V}_{\ell,d}, \mathbf{H}_\ell)$ would depend on ρ .

We note that in this case, using only the set \mathbf{V}_s does not result in bounds for the Gallager function, but rather *approximates* $E_0(\rho, N, D)$.

4.3.3 Upper Bound to the Random Coding Exponent

We further explore the possibility of upper bounding the random coding exponent, thereby lower bounding the word error rate, by modifying the metric in 4.30. In order to achieve this, we assume that the tree-search based sphere decoder iterates through the ML probabilities in order, and assuming the C_s th iteration is in progress, we approximate all the remaining $M^{N_T R_m} - C_s$ probabilities with the most recent output of the sphere decoder, i.e. $p_N(\mathbf{R}_{\ell,d}|\mathbf{V}_{\ell,d})^{(C_s)}$. We therefore approximate the contributions of the remaining candidate vectors to the most pessimistic value, since our sphere decoder outputs the channel probabilities in ML order. We see that this approximation is actually a very accurate upper bound to the random coding exponent, and using the lower and upper bounds previously derived, the word error rate for any particular block fading channel can be accurately bounded through our analysis and with a computational efficiency many orders of magnitude better than the ML solution.

4.4 Performance Analysis and Results

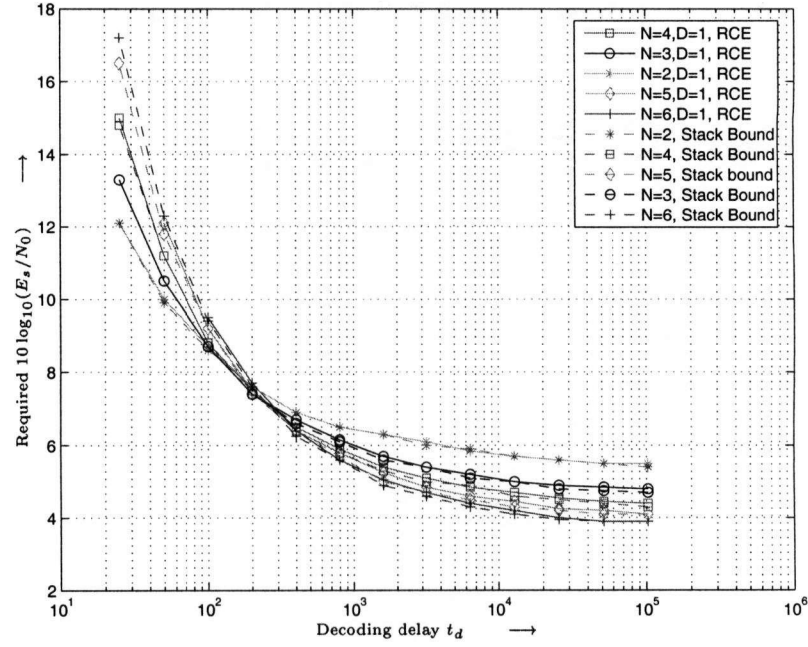
We illustrate the various sphere decoding decoding algorithms explained in this chapter through various illustrative examples. We use these decoding algorithms to upper and lower bound the random coding exponent and correspondingly the probability of

error, and show how our approach results in a significant reduction in computational complexity for the random coding exponent.

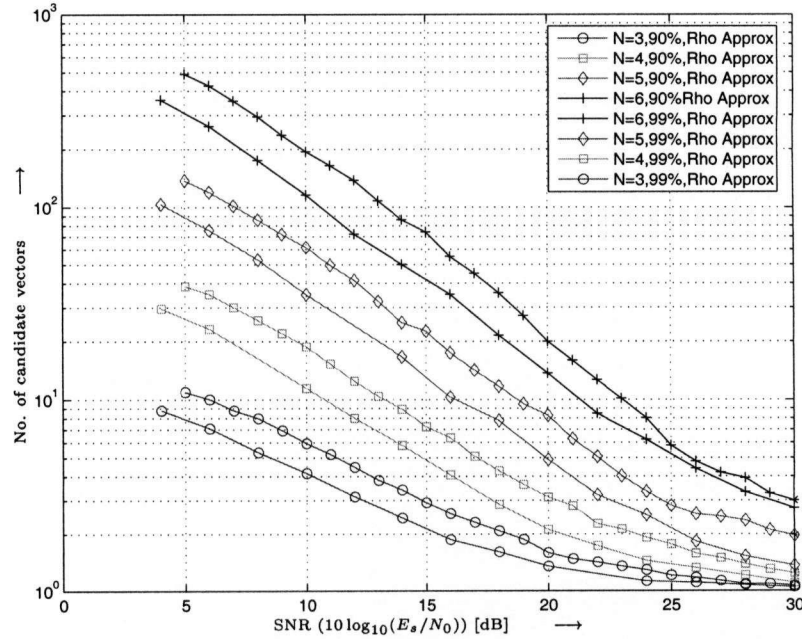
4.4.1 Single Antenna Systems

We illustrate the tree search algorithms for single antenna systems namely the stack algorithm, and show how we can obtain accurate bounds for the random coding exponent using a combination of the tree-search algorithms and a *stopping criteria* ratio. We illustrate how modifying the ratio $r(C_s)$, which acts as an accuracy parameter for our approximation, affects the lower bound to the random coding exponent, and we also show the complexity advantage that we gain from using the modified tree-search decoding algorithm to obtain recursively the best noncoherent metrics. In fig. 4.1(a), we show the Gallager random coding exponent as a function of the decoding delay t_d , and plot it against the SNR ($10 \log_{10}(E_s/N_0)$) required to achieve a $\text{WER} \leq 10e^{-3}$. We show the results for CDD ($N = 2$) as well as MSDD ($N = 3, 4, 5, 6$) for single antenna D4PSK as before. In order to provide a good approximation for the exponent we bound it using the stack algorithm (see section 4.1), and fix the accuracy parameter to 90%.

We can then fix the accuracy for the stack approximations by controlling the parameter C_s . In fig. 4.1 we fix this parameter so that the ratio in 4.26 does not exceed 0.9, thus giving us an accuracy of 90%. We can see that using this criteria to approximate the random coding exponent does not involve a great deal of loss in the approximation, the maximum loss never exceeds 0.1 dB. However, the complexity advantage gained by using efficient stack algorithms is significant, and justifies the negligible loss in accuracy. For example, in fig. 4.1(b) we compute the complexity for the random coding exponent bound in fig. 4.1(a). We can clearly see that for the scenario in fig. 4.1(a), the ML solution would involve $M^{D(N-1)} = 4^5 = 1024$ test vectors and their corresponding



(a) *Random Coding Exponent for D4PSK: Lower bound using Stack Algorithm. Overall accuracy 99%, according to 4.26*



(b) *Complexity analysis for Stack algorithm: D4PSK with $N = 2, 3, 4, 5, 6$, 90%, 99% accuracy of Stack algorithm*

Figure 4.1: *Stack algorithm bounds for D4PSK: MSDD with 99% accuracy*

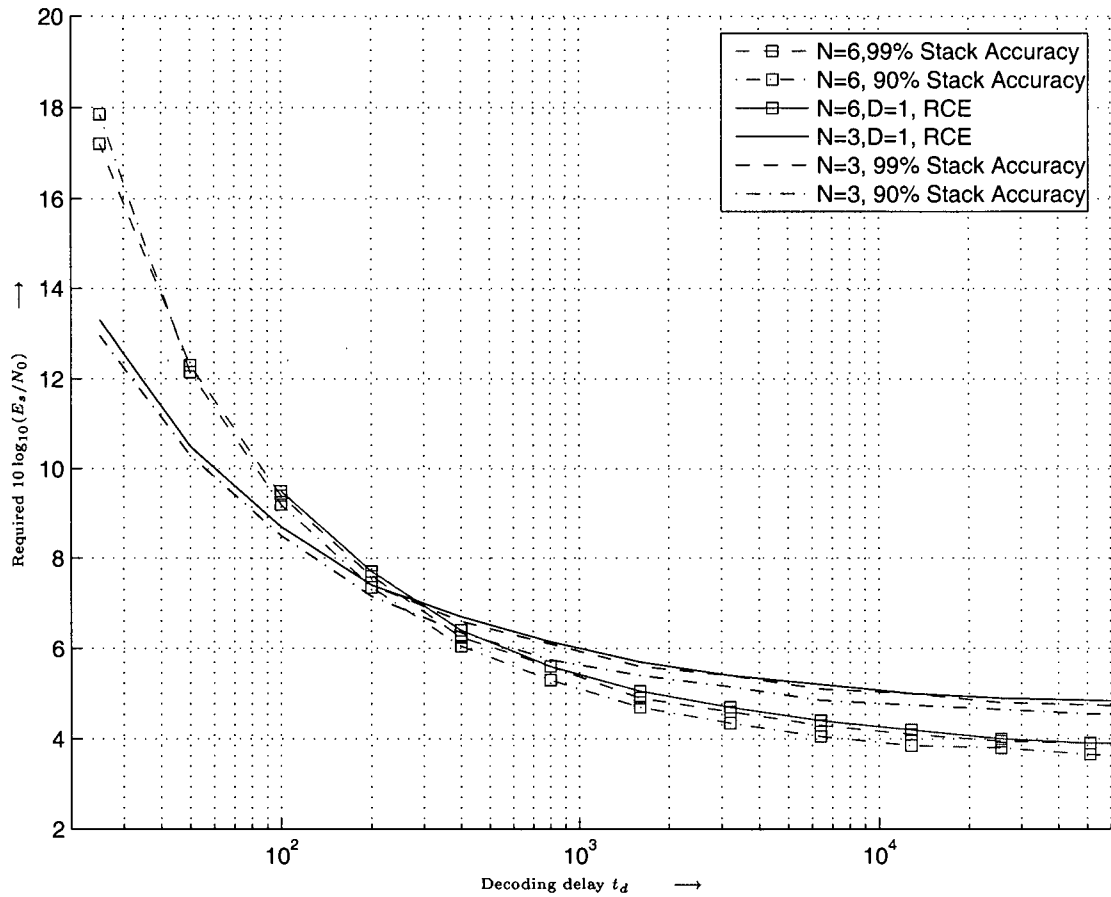
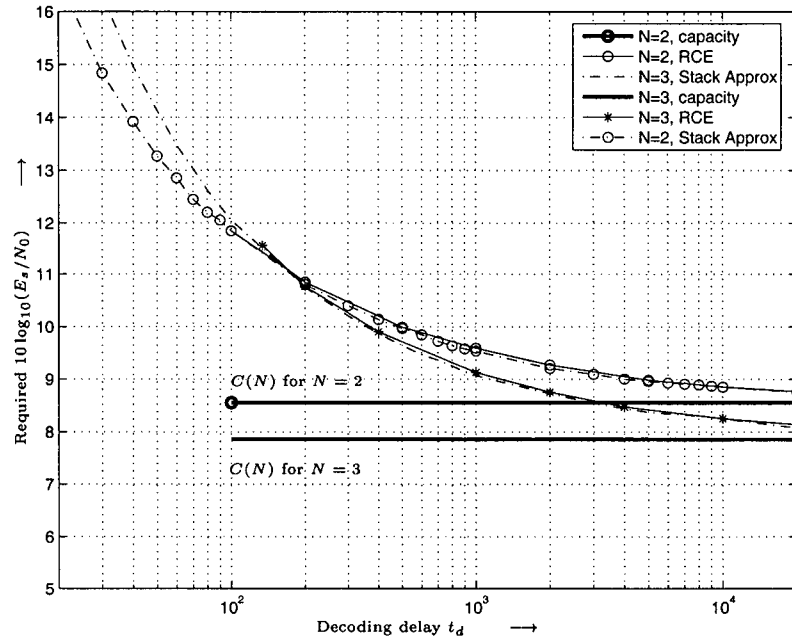
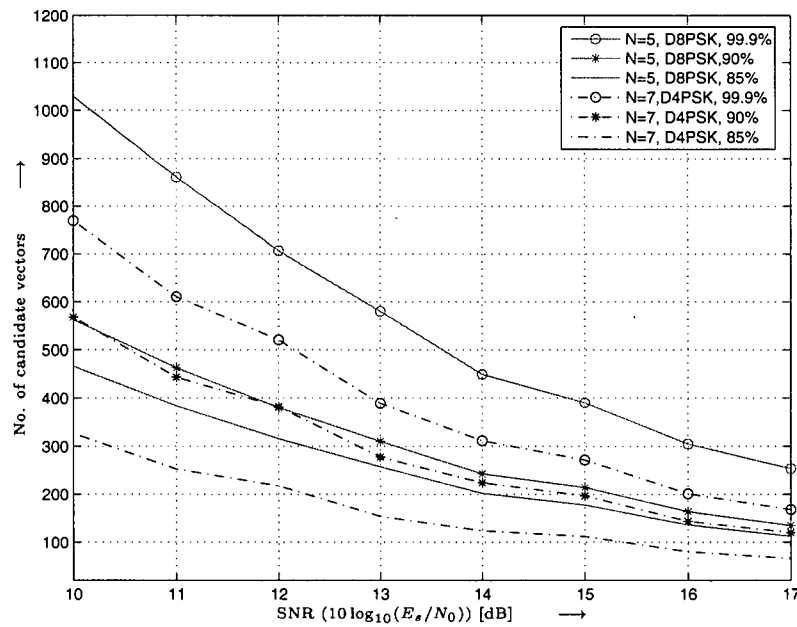


Figure 4.2: *Random Coding Exponent for D4PSK: Lower bounds using Stack Algorithm with different accuracies. CDD ($N = 2$) and MSDD ($N = 6$) with 90% and 99% accuracy*

metrics to be calculated. However, for SNR's of interest to us such as an overall E_s/N_0 of 10 dB, the number of test vector needed to provide a 99% approximation to the true random coding exponent is almost 10 orders of magnitude less (approx. 10^2). Similar results are obtained for smaller values of N , however the cases of interest where the overall signal space would be too large begin from about $N \geq 4$. As a comparison, we also show the reduction in complexity that could be achieved if the accuracy constraint was further reduced to 90%, however the savings in complexity do not justify the negligible improvement to the random coding approximation. This can



(a) Random Coding Exponent for D8PSK: Lower bound using Stack Algorithm. Overall accuracy 99%, according to 4.26



(b) Complexity analysis: D4PSK and D8PSK with $N = 5, 7$, different accuracies of Stack algorithm with same ML search space

Figure 4.3: Complexity and random coding analysis of Stack Algorithm, D4PSK and D8PSK with 99, 90 and 85% accuracy

be seen in fig. 4.2, where we compare the lower bound achieved with different accuracies of the stack algorithm. We see that reducing the complexity from 99% to 90% shows a marked difference in accuracy to the true RCE, however the reduction in complexity is not significant, as we see from fig. 4.1(b). In fig. 4.3, we show the effectiveness of the stack algorithm for D8PSK, and compare the complexity advantage gained with the corresponding D4PSK signal space. In each case, the ML estimate would have to take into account the same number of signal points, however the stack algorithm for each accuracy performs significantly better for D4PSK than D8PSK. The advantage gained is again in orders of magnitude, and is thus an invaluable tool to accurately approximate the random coding exponent for higher dimensions where a ML decoder would be prohibitive. In fig. 4.3(a), we demonstrate the accuracy for smaller values of N i.e CDD with $N = 2$ and MSDD with $N = 3$ and can see that even with extremely limited vector search spaces, the Stack algorithm does an excellent job of approximating the exponent, as the accuracy of the decoder increases with the size of the search space.

4.4.2 Multiple Antenna Systems

Bounds using Sphere Decoding for Spatially Correlated Antennas

Similar to the case of $D > 1$ and spatially uncorrelated fading, we analyze the transmission scheme where spatial correlation results in a loss of diversity. In previous sections, we look at the case of a general block fading channel considering spatially uncorrelated fading, i.e the channel between each transmitter and each receiver is modeled as an independent fading channel, with the fading coherence interval equal to DN . Here we also present and analyze some results for the case of spatially correlated channels, which also provides us with interesting insights into the behavior of such channels. We begin with the case of spatial correlation, where the channel between each transmitter and receiver does not experience independent fading, but is correlated with a param-

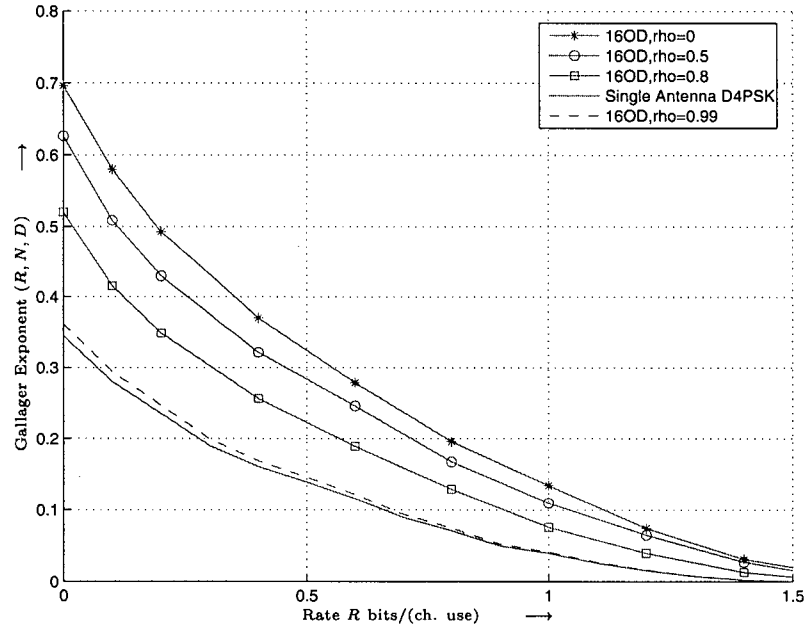
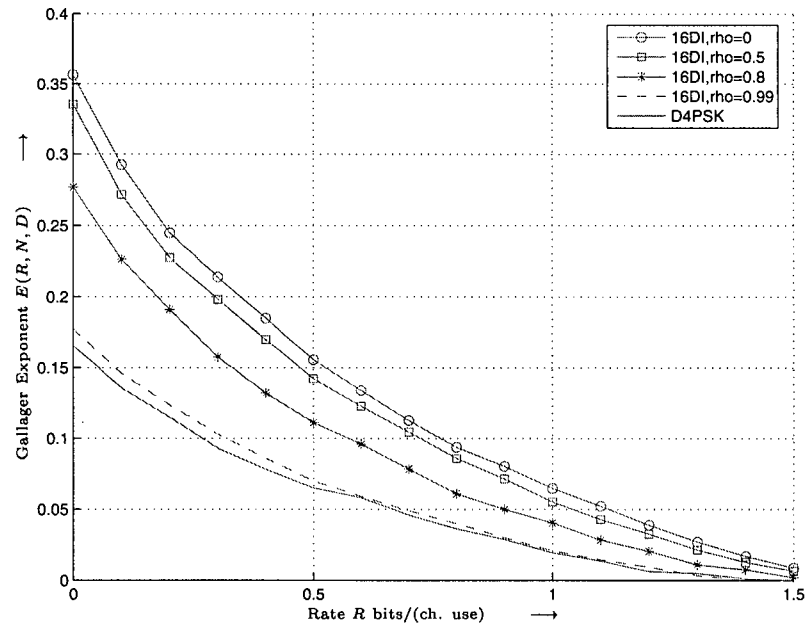
(a) Orthogonal Codes: Coding diversity 10 i.e $n_c/L = 20$ (b) Diagonal constellations: Coding diversity 5 i.e $n_c/L = 40$

Figure 4.4: $E(R, N, D)$ vs R for DSTM, different ST codes with $R_t = 1$ (bit/ch. use), spatially correlated antennas.

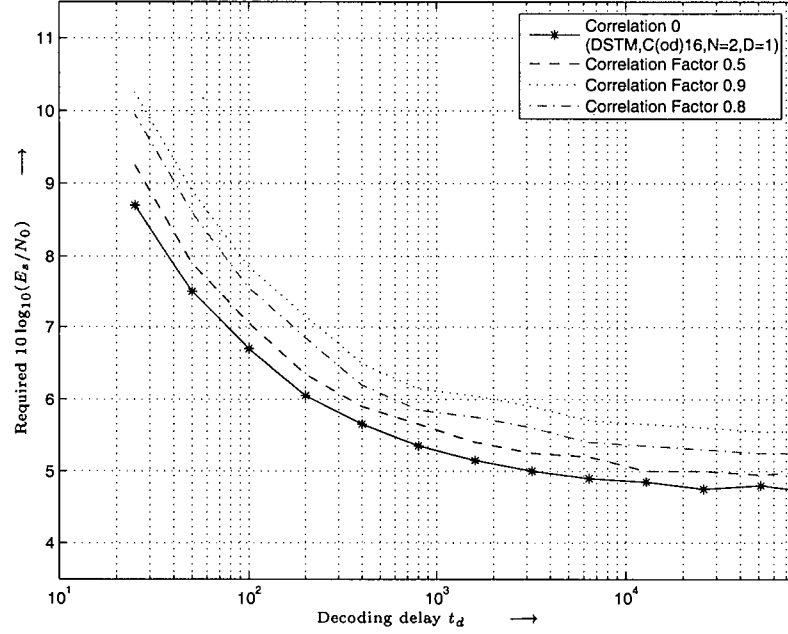


Figure 4.5: Required E_s/N_0 vs decoding delay t_d for DSTM with CDD, $C_{OD}(16)$ with $R_t = 1$ (bit/ch. use), spatially correlated antennas

eter we denote as ρ . We model this spatial correlation, for example with $N_T = 2$ and $N_R = 1$ by generating Gaussian random variables g_1, g_2 such that

$$g_2 = \rho \cdot g_1 + \sqrt{1 - \rho^2} g_1 \quad (4.31)$$

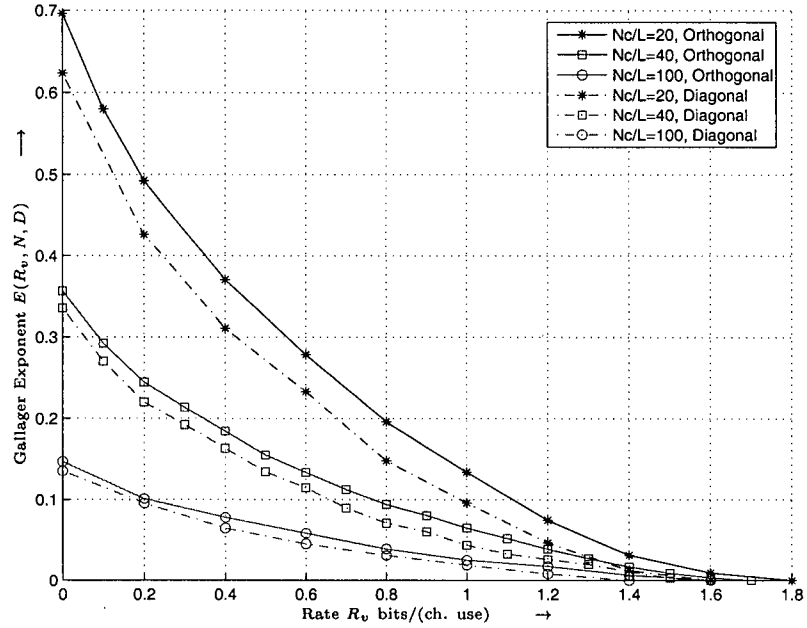
where the greater the value of ρ , greater is the correlation. For $\rho = 0$, we revert to the uncorrelated channel scenario. In fig. 4.4(a), we plot the random coding exponent $E(R, N, D)$ against the rate R for orthogonal constellations, with a target rate $R_t = 1$ bit/(channel use). We plot this exponent for different values of the parameter ρ indicating different levels of correlation between antennas. We find that the performance of the system w.r.t the random coding exponent deteriorated significantly as ρ approaches the fully correlated case i.e $\rho = 1$, and at that value of $\rho = 1$, the performance of the overall space-time model approaches that of an equivalent single antenna system. This is because as the antennas become more correlated, the channel between the $N_T = 2$ transmit antennas and the $N_R = 1$ receive antennas becomes the same, and the overall

effect is the same as if there was a single channel between the transmitter and the receiver. ($\rho = 1$ implies that the channels are the same). Fig. 4.4(b) illustrates the same effect for the case of diagonal constellations with the normalized coherence interval $n_c/L = 40$ implying a coded diversity of 5.

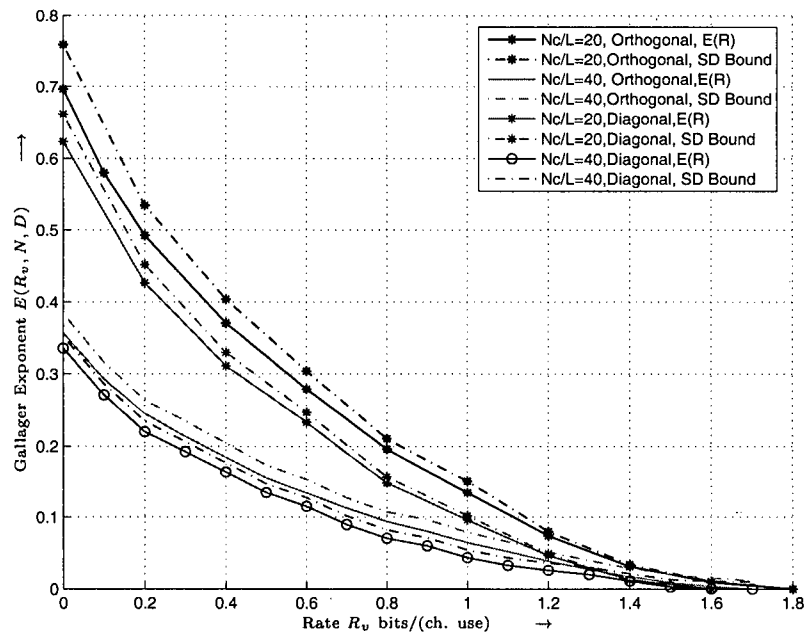
In order to clearly visualize the effect of spatial diversity on the random coding exponent, we plot the required E_s/N_0 vs the decoding delay t_d in PSK samples in fig. 4.5. We can see that corresponding to fig. 4.4, as the spatial correlation between antennas increases, a significant drop in the RCE performance results, and this bears out our $E(R, N, D)$ analysis showing that the performance of DSTM with spatially correlated antennas approaches that of the single antenna transmission system as the correlation factor ρ increases towards 1.

Diagonal vs Orthogonal Constellations

We present some simulation results to justify our claims from the previous section, where we devised tree-search based sphere decoders for orthogonal and diagonal constellations, in order to accurately represent the random coding exponent for differential space time modulation with high dimensions and correspondingly large matrix search spaces. We also look into different space-time constellations as in the capacity analysis in chapter 3 from a random coding exponent point of view. We present simulation results related to the bounds for $E_0(\rho, N, D)$, and to this end, we analyze the sphere decoding algorithms for orthogonal and diagonal constellations first, and demonstrate the effectiveness of these algorithms by comparing the bounds obtained using such efficient decoding algorithms with the true calculation of the random coding exponent. In fig. 4.6(a), we compare the random coding exponent $E(R, N, D)$ for two commonly used space-time constellations namely orthogonal constellations and diagonal constellations. We see that for a modulation rate $R_m = 2$ bits/(channel use), the orthogonal constellations consistently outperforms the diagonal constellation. This is also demon-



(a) $E(R, N, D)$ for DSTM: Orthogonal $C_{OD}(16)$ and diagonal $C_{DI}(16)$ codes, $R_{target} = 1$ bit/(channel use)



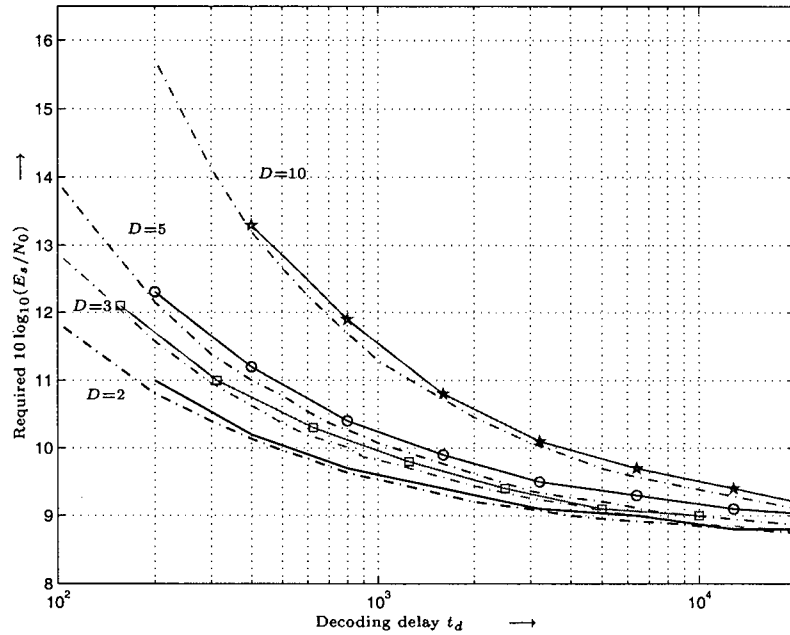
(b) RCE for DSTM: (b) Stack algorithm for orthogonal and diagonal codes, $R_{target} = 1$ bit/(ch. use)

Figure 4.6: Random coding exponent for DSTM: $C_{OD}(16)$ and $C_{DI}(16)$ codes

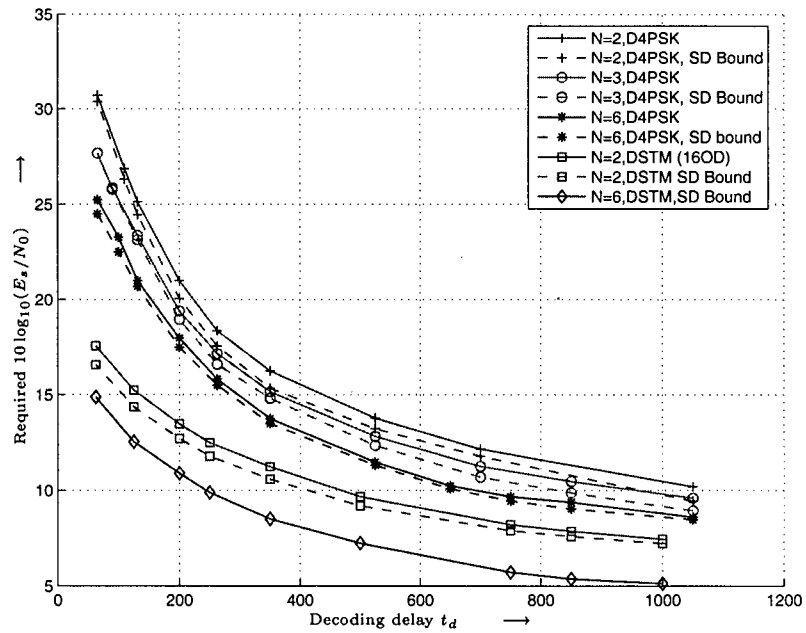
strated by the capacity results from chapter 3. Combining this with the modified block fading channel, we demonstrate also the effect of coding diversity on the random coding exponent, where the normalized channel coherence length n_c/L is changed to reflect different coding diversities. We see that for both diagonal as well as orthogonal constellations, an increase in diversity results in an increase in the random coding exponent performance, reflecting the effect of seeing more independent fading realizations per codeword for the same delay. We also present in fig. 4.6(b) the lower bounds to the exponents in fig. 4.6(a) obtained using the sphere decoding algorithms from section 4.3. We fix the accuracy of the sphere decoding algorithms to 90%, and observe that we can very closely approximate the random coding exponent to within ± 0.1 dB.

Lower Bounds to the Random Coding Exponent using Sphere Decoding

As mentioned, we have described the method to use efficient tree-search based algorithms to bound the random coding exponent for single as well as multiple transmit antenna transmission systems. We have shown illustrative examples in 4.4.2 to show the effect of the modified block fading channel on the random coding exponent with $D > 1$. Here we present some examples analyzing the E_s/N_0 performance of our approximations and bounds to the random coding exponent, and observe the effectiveness of our proposed scheme to accurately lower bound the exponent for different cases of single antenna transmission and DSTM, with the fading coherence interval independent of N . In fig. 4.7(a), we plot the required E_s/N_0 in dB against the decoding delay t_d for single antenna D8PSK, where the MSDD observation interval $N = 2$, and D varies to reflect different fading coherence intervals. We also lower bound the random coding exponent using the stack algorithm with 99% accuracy. We observe that though the overall performance of the system deteriorates with increasing D , because of the reduced diversity observed due to fewer independent fading realizations seen by each codeword, the stack approximation nevertheless gives us an accurate bound to the ran-



(a) SNR-delay analysis, D8PSK with CDD ($N = 2$) and varying $D = 2, 3, 5, 10$ with lower bound using stack approximation: effect of D



(b) Random coding exponent for DSTM vs DPSK, Sphere decoder bounds.

Figure 4.7:

dom coding exponent even for higher values of $D \gg N$. This proves our claim that by a novel combination of sphere decoding/stack algorithms and a performance metric designed to optimize the search space for the random coding exponent calculation, we obtain very accurate bounds to the exponent and at the same time increase the efficiency of the analysis by many orders of magnitude over current analysis techniques. Fig. 4.7(b) shows the application of the sphere decoding algorithm from section 4.2 to lower bound the random coding exponent for DSTM, where the constellation chosen is the orthogonal constellation $\mathbf{C}_{OD}(16)$, and we can compare the performance of single antenna DPSK with DSTM. We see that for the same transmitted vector length DN for both single and multiple antenna systems, the performance of MSDD consistently improves with N for both DPSK and DSTM. This is different from our earlier analysis when we kept the overall coding diversity constant, and hence as D increased, the performance degraded due to reduced independent fading intervals seen by the system. We also note that the stack algorithm for DPSK and the sphere decoding algorithms for DSTM provide an excellent bound to the random coding exponents. Our proposal has the added advantage of being computationally effective as compared to current techniques to analyze the random coding exponent. As an excellent example, we provide the approximation (bound) for the random coding exponent for $N = 6$ with DSTM in fig. 4.7(b), where the computation of the exponent would involve a best metric search through 1,048,576 or 1 million candidate vectors, and would limit the channels which could be analyzed purely due to computational complexity.

Upper Bounds using Sphere Decoding

In this section, we complete our analysis of the random coding exponent for single antenna and multiple antenna systems by attempting to upper bound the exponent with an innovative approximation to the metric in section 4.3. In this case, in fig. 4.8(a) we begin with the single antenna case for the Gallager exponent $E(R, N, D)$ for D4PSK

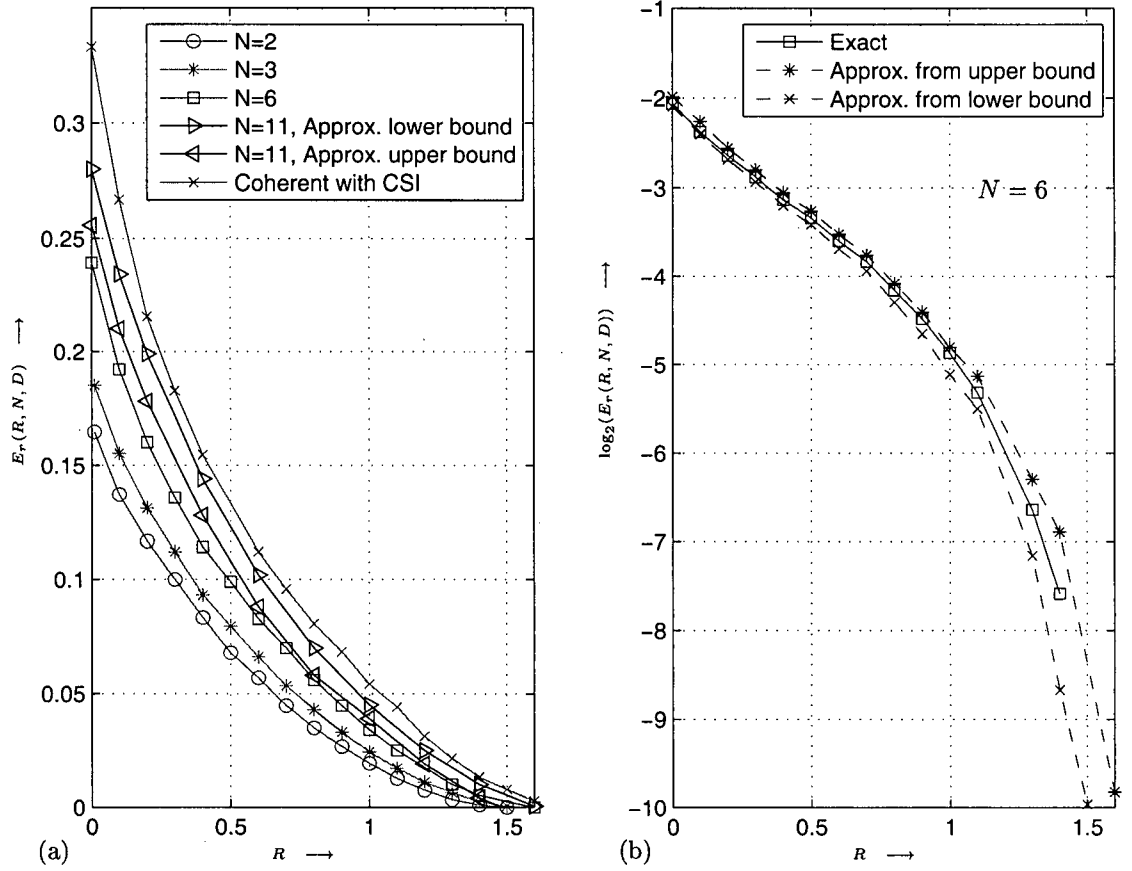


Figure 4.8: $E(R, N, D)$ vs. R : D4PSK and $10 \log_{10}(E_s/N_0) = 10$ dB. (a): MSDD with different N . (b): $N = 6$. Exact expression and approximation for $E_r(R, N, D)$. $n_c/L = 40$.

and $N_T = 1$. We plot the random coding exponent as a function of rate R in fig. 4.8(a) to serve as a reference, and can see that the exponent approaches the performance of coherent detection with CSI as the MSDD observation window N increases. Since the complexity associated with $N \geq 10$ becomes prohibitive to simulate, we present the upper and lower bound approximations using 4.30 to bound the true random coding exponent, and hence the overall WER. Fig. 4.8(b) gives us a much clearer insight into this approximation where we plot the bounds on a logarithmic scale along with the true RCE for $N = 6$. Extending this approximation to the multiple antenna case, in fig. 4.9,

we refer back to fig. 4.7(b) to compare the performance of single antenna DPSK and DSTM, but we upper bound the random coding exponent along with the true RCE. We can see that our analysis does indeed give us an excellent upper bound to the random coding exponent. To this end, we show the SNR $10 \log_{10}(E_b/N_0)$ required to achieve $P_e \leq 10^{-3}$ according to (3.33) as function of code length n_c for transmission with rate $R = 1$ bit per scalar channel use. The normalized coherence length is $n_c/L = 100$, i.e., degrees of freedom L and thus coded diversity increase linearly with code length n_c . Two things are interesting to note. First, increasing N consistently improves power efficiency regardless of n_c . This is decidedly different from [24], where coded diversity was decreasing with N . Second, the gain due to transmit diversity increases with decreasing n_c . This result reiterates the benefits of spatial transmit diversity if coded diversity is limited, e.g. due to delay constraints.

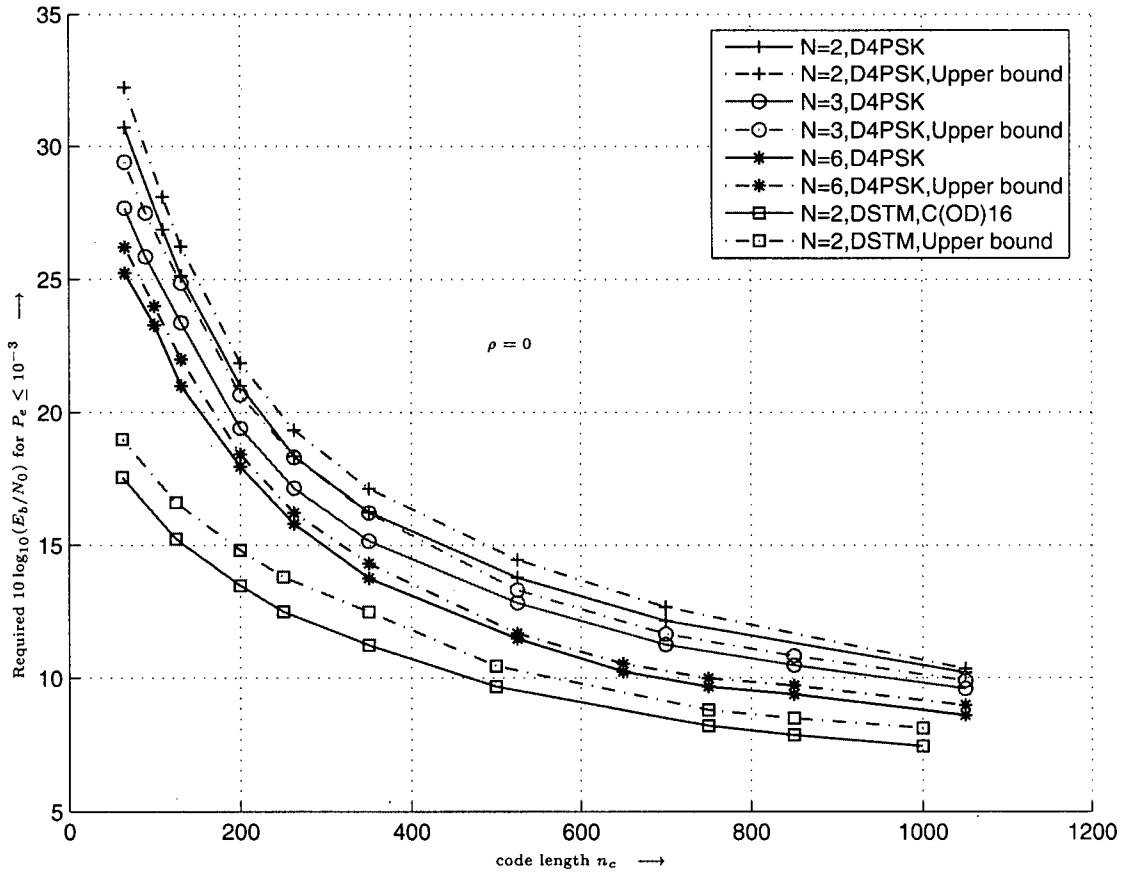


Figure 4.9: Required $10 \log_{10}(E_b/N_0)$ for $P_e \leq 10^{-3}$ vs. n_c . 4DPSK ($N_T = 1$) and 16OD ($N_T = 2$). $n_c/L = 100$, $R = 1$ bit/(scalar channel use): upper bound approximation

Chapter 5

Conclusions and Future Work

In this thesis, we consider coded space-time transmission with multiple antennas over a block fading channel with an arbitrary fading coherence interval and a finite code length, with differential encoding applied at the transmitter. We focus our attention on error exponent analysis of such channels, namely the calculation of the random coding exponent proposed by Gallager. Our work differed from previous work in this field, which limited error exponent analysis to block fading channels with the fading interval equal to the MSDD observation window N .

- To this end, we derived an expression for the random coding exponent for differential space-time modulation and MSDD over a block fading space-time channel where the fading coherence interval is arbitrary. This extension to previous work is particularly important as it allows for a fair comparison of MSDD with different window sizes but fixed *coded diversity*.
- Furthermore, we devised upper and lower bounds and approximations for the random coding exponent using tree-search based sphere decoding algorithms for orthogonal and diagonal space-time codes. These sphere decoding algorithms en-

able efficient Monte-Carlo integration, and reduce the complexity cost by several orders of magnitude at SNR's of interest.

- To achieve this, we proposed a novel performance metric to control the accuracy of the approximation, and correspondingly the computational complexity. We illustrated the potential advantage of our analysis with several numerical results to demonstrate the usefulness of our approximations, the improvements with increasing observation interval N and the benefits of spatial transmit diversity even in highly correlated channels for coded DSTM transmission with short or moderate code lengths.

Through our investigations, we believe that we have formulated a novel and innovative approach to calculate the random coding exponent for a wide range of channels, and have provided researchers with an invaluable technique to calculate such exponents with a performance efficiency increase of several orders. Through our analysis and subsequent simulations, we can summarize our findings as follows:

- The performance of the random coding exponent for DSTM with MSDD approaches that of coherent detection with perfect channel state information.
- The analysis allows for accurate comparisons of systems with fixed coded diversity but different MSDD observation windows. The random coding exponent performance degrades with an increase in the fading coherence interval for the same code length, effectively reducing the overall diversity due to a reduced number of independent fading realizations.
- If the transmit diversity increases, the random coding exponent performance improves correspondingly, reiterating the benefit of our analysis where we enable accurate analysis of block fading channels with arbitrary fading intervals. Previous work limited the analysis of the random coding exponent to block fading

channels without the advantage of analyzing the effect of coding diversity, as the channel coherence interval was equal to N .

- We also devised bounds and approximations for the random coding exponent using efficient decoding techniques such as tree-search based sphere decoding algorithms. We found that the novel performance metric we devised to control the accuracy of the approximation enabled an excellent approximation to the random coding exponent, at the same time reducing the computational complexity of the analysis by several orders of magnitude. Another advantage to our approach was the flexibility to balance the degree of accuracy of the approximation with the complexity cost. We concluded that reducing the accuracy from 99% of the full-search approach (taking all possible candidate vectors into account) to 90% gave us only a marginal improvement in computational complexity, as opposed to the improvement of several orders of magnitude in computational complexity when reducing the accuracy from the full-search to 99%.
- Different from previous work, we also took into consideration the effect of spatial correlation between antennas, and through suitable examples demonstrated the random coding exponent for such systems with varying degrees of correlation. We concluded that as the spatial correlation approached full correlation (the parameter ρ in our analysis approached 1), the performance of DSTM degraded and became equivalent to a single antenna link between the transmitter and the receiver.
- We also successfully outlined a procedure to provide an accurate upper bound to the random coding exponent, and along with the lower bound obtained from the sphere decoding algorithms coupled with a suitable performance metric, we demonstrated a technique to bound the error exponent performance of coded transmission over a multiple antenna link with MSDD.

In our opinion, there are a number of interesting and advantageous research topics for future work in this field. The work done in this thesis can further be used to analyze channels with frequency selective fading or time-varying channels. We can also combine the analysis from this thesis with different coding strategies such as bit-interleaved coded modulation (BICM) or hybrid coded modulation (HCM). We could also consider the application of the analysis from this thesis to the design of noncoherent coded modulation schemes for multiuser communications and for extremely high bandwidth-efficient MIMO systems.

Appendix A

Derivation of Gallager's Function with Coherent metric

In this appendix, we present the derivation of Gallager's function with reduced complexity metric given by 3.36 in detail. The channel is the block fading channel considered throughout this thesis channel, seeing L independent fading realizations, where the channel remains constant for DN symbols (PSK symbols or matrix symbols corresponding to single antenna or DSTM respectively). For simplicity, we denote the N dimensional vectors $\mathbf{R}_{\ell,d}[k], \mathbf{V}_{\ell,d}[k]$ as $\mathbf{R}_i, \mathbf{V}_i$ respectively, we begin with the expression for Gallager's function $E_0(\rho, N, D)$

$$E_0(\rho, N, D) = \frac{-1}{L \cdot D} \log_2 \left[\int_{\mathbf{R} \in \mathcal{C}^{n_w}} \left(\sum_{\mathbf{V} \in \mathcal{V}^{L \cdot D(N-1)}} \prod_{i=1}^{LD} Pr\{\mathbf{V}_i\} \lambda(\mathbf{R}_i | \mathbf{V}_i)^{-\rho/(1+\rho)} \cdot p_{n_w}(\mathbf{R} | \mathbf{V}) \right) \left(\prod_{i=1}^{L \cdot D} \sum_{\mathbf{V}_i \in \mathcal{V}^{N-1}} Pr\{\mathbf{V}_i\} \lambda(\mathbf{R}_i | \mathbf{V}_i)^{1/(1+\rho)} \right)^{\rho} d\mathbf{R} \right] \quad (\text{A.1})$$

where \mathbf{R} is the vector of overlapping symbols of length $L \cdot D \cdot N = \hat{n}_w$, $n_w = LDN$, $n_V = L \cdot D$. In our analysis, without any loss of generality we can consider non-overlapping

blocks, as such a system would have the same performance as a system with overlapping when MSDD is applied and noncoherent transmission is considered. Thus

$$E_0(\rho, N, D) = \frac{-1}{L.D} \log_2 \left[\int_{\mathbf{R} \in \mathbf{C}^{L.D.N}} \left(\sum_{\mathbf{V} \in \mathbf{V}^{L.D.(N-1)}} \prod_{i=1}^{L.D} Pr\{\mathbf{V}_i\} \lambda(\mathbf{R}_i|\mathbf{V}_i)^{-\rho/(1+\rho)} \cdot p(\mathbf{R}|\mathbf{V}) \right) \left(\prod_{i=1}^{L.D} \sum_{\mathbf{V}_i \in \mathbf{V}^{N-1}} Pr\{\mathbf{V}_i\} \lambda(\mathbf{R}_i|\mathbf{V}_i)^{1/(1+\rho)} \right)^\rho d\mathbf{R} \right] \quad (\text{A.2})$$

Since

$$p(\mathbf{R}|\mathbf{V}) = \prod_{\ell=1}^L p(\mathbf{R}_\ell|\mathbf{V}_\ell) \quad (\text{A.3})$$

and

$$\begin{aligned} p(\mathbf{R}_\ell|\mathbf{V}_\ell) &= \int_{\mathbf{H}_\ell} p(\mathbf{H}_\ell) p(\mathbf{R}_\ell|\mathbf{V}_\ell, \mathbf{H}_\ell) d\mathbf{H}_\ell \\ &= \sum_{s_0} \frac{1}{M} \int_{\mathbf{H}_\ell} p(\mathbf{H}_\ell) p(\mathbf{R}_\ell|\mathbf{S}_\ell, \mathbf{H}_\ell) d\mathbf{H}_\ell \end{aligned} \quad (\text{A.4})$$

where s_0 is the reference symbol, we get

$$\begin{aligned} E_0(\rho, N, D) &= \frac{-1}{L.D} \log_2 \left[\int_{\mathbf{R}_\ell \in \mathbf{C}^{L.D.N}} \left(\sum_{\mathbf{V}_\ell \in \mathbf{V}^{L.D.(N-1)}} \prod_{i=1}^{L.D} Pr\{\mathbf{V}_i\} \lambda(\mathbf{R}_i|\mathbf{V}_i)^{-\rho/(1+\rho)} \right. \right. \\ &\quad \left. \prod_{\ell=1}^L \frac{1}{M} \sum_{s_0} \int_{\mathbf{H}_\ell} p(\mathbf{H}_\ell) p(\mathbf{R}_\ell|\mathbf{S}_\ell, \mathbf{H}_\ell) d\mathbf{H}_\ell \right) \\ &\quad \left. \left(\prod_{i=1}^{L.D} \sum_{\mathbf{V}_i \in \mathbf{V}^{N-1}} Pr\{\mathbf{V}_i\} \lambda(\mathbf{R}_i|\mathbf{V}_i)^{1/(1+\rho)} \right)^\rho d\mathbf{R}_\ell \right] \end{aligned} \quad (\text{A.5})$$

Since each block of L samples experiences independent fading, we separate the product

over $L.D$ and obtain

$$E_0(\rho, N, D) = \frac{-1}{L.D} \log_2 \left[\prod_{\ell=1}^L \int_{\mathbf{R}_\ell \in \mathbf{C}^{D.N}} \left(\sum_{\mathbf{V}_\ell \in \mathbf{V}^{D(N-1)}} \prod_{i=1}^D \Pr\{\mathbf{V}_i\} \lambda(\mathbf{R}_i|\mathbf{V}_i)^{-\rho/(1+\rho)} \right. \right. \\ \left. \left. \frac{1}{M} \sum_{s_0} \int_{\mathbf{H}_\ell} p(\mathbf{H}_\ell) p(\mathbf{R}_\ell|\mathbf{S}_\ell, \mathbf{H}_\ell) d\mathbf{H}_\ell \right) \right. \\ \left. \left(\prod_{i=1}^D \sum_{\mathbf{V}_i \in \mathbf{V}^{N-1}} \Pr\{\mathbf{V}_i\} \lambda(\mathbf{R}_i|\mathbf{V}_i)^{1/(1+\rho)} \right)^\rho d\mathbf{R}_\ell \right] \quad (\text{A.6})$$

$$= \frac{-1}{LD} \log_2 \left[\frac{1}{M} \sum_{s_0} \int_{\mathbf{R}_\ell \in \mathbf{C}^{D.N}} \int_{\mathbf{H}_\ell} p(\mathbf{H}_\ell) \left(\sum_{\mathbf{V} \in \mathbf{V}^{D(N-1)}} \prod_{i=1}^D \Pr\{\mathbf{V}_i\} \lambda(\mathbf{R}_i|\mathbf{V}_i)^{-\rho/(1+\rho)} p(\mathbf{R}_\ell|\mathbf{S}_\ell, \mathbf{H}_\ell) \right) \right. \\ \left. \left(\prod_{i=1}^D \sum_{\mathbf{V}_i \in \mathbf{V}^{N-1}} \Pr\{\mathbf{V}_i\} \lambda(\mathbf{R}_i|\mathbf{V}_i)^{1/(1+\rho)} \right)^\rho d\mathbf{H}_\ell d\mathbf{R}_\ell \right] \quad (\text{A.7})$$

Since $p(\mathbf{R}_\ell|\mathbf{S}_\ell, \mathbf{H}_\ell) = \prod_{i=1}^D p(\mathbf{R}_i|\mathbf{S}_i, \mathbf{H}_i)$,

$$E_0(\rho, N, D) = \frac{-1}{D} \log_2 \left[\frac{1}{M} \sum_{s_0} \int_{\mathbf{R} \in \mathbf{C}^{D.N}} \int_{\mathbf{h}} p(\mathbf{H}_\ell) \cdot \prod_{i=1}^D \left\{ \left(\sum_{\mathbf{V}_\ell \in \mathbf{V}^{D(N-1)}} \Pr\{\mathbf{V}_i\} \lambda(\mathbf{R}_i|\mathbf{V}_i)^{-\rho/(1+\rho)} \cdot p(\mathbf{R}_i|\mathbf{S}_i, \mathbf{H}_i) \right) \right. \right. \\ \left. \left. \left(\sum_{\mathbf{V}_i \in \mathbf{V}^{N-1}} \Pr\{\mathbf{V}_i\} \lambda(\mathbf{R}_i|\mathbf{V}_i)^{1/(1+\rho)} \right)^\rho \right\} d\mathbf{H}_\ell d\mathbf{R}_\ell \right] \quad (\text{A.8})$$

Rewriting the above equation using the relation

$$p(\mathbf{R}_\ell, \mathbf{H}_\ell, s_0) = p(\mathbf{R}_\ell, \mathbf{H}_\ell|s_0) \cdot p(s_0) = \frac{1}{M} p(\mathbf{R}_\ell, \mathbf{H}_\ell|s_0) \quad (\text{A.9})$$

we get

$$\begin{aligned}
 E_0(\rho, N, D) = & \frac{-1}{D} \log_2 \left[\sum_{s_0} \int_{\mathbf{R}_\ell \in \mathcal{C}^{D,N}} \int_{\mathbf{H}_\ell} p(\mathbf{R}_\ell, \mathbf{H}_\ell, s_0) \cdot \prod_{i=1}^D \right. \\
 & \left\{ \left(\sum_{\mathbf{V}_i \in \mathcal{V}^{D(N-1)}} P_T\{\mathbf{V}_i\} \lambda(\mathbf{R}_i | \mathbf{V}_i)^{-\rho/(1+\rho)} \cdot \frac{p(\mathbf{R}_i | \mathbf{S}_i, \mathbf{H}_i)}{p(\mathbf{R}_i | \mathbf{H}_i, s_0)} \right) \right. \\
 & \left. \left(\sum_{\mathbf{V}_i \in \mathcal{V}^{N-1}} P_T\{\mathbf{V}_i\} \lambda(\mathbf{R}_i | \mathbf{V}_i)^{1/(1+\rho)} \right)^\rho \right\} d\mathbf{H}_\ell d\mathbf{R}_\ell \left. \right] \quad (\text{A.10})
 \end{aligned}$$

Finally, in order to efficiently use Monte Carlo techniques we write Gallager's function in terms of expected values as

$$\begin{aligned}
 E_0(\rho, N, D) = & -\frac{1}{D} \log_2 \left[\mathcal{E}_{\mathbf{R}_\ell, \mathbf{H}_\ell} \left\{ \prod_{d=1}^D \right. \right. \\
 & \mathcal{E}_{\mathbf{V}_i} \left\{ \exp(\lambda(\mathbf{R}_i | \mathbf{V}_i))^{-\frac{\rho}{1+\rho}} p(\mathbf{R}_i | \mathbf{V}_i, \mathbf{H}_\ell) \right\} \\
 & \left(\mathcal{E}_{\mathbf{V}_i} \left\{ \exp(\lambda(\mathbf{R}_i | \mathbf{V}_i))^{\frac{1}{1+\rho}} \right\} \right)^\rho \\
 & \left. \left. (\mathcal{E}_{\mathbf{V}_i} \{p(\mathbf{R}_i | \mathbf{V}_i, \mathbf{H}_\ell)\})^{-1} \right\} \right]. \quad (\text{A.11})
 \end{aligned}$$

Appendix B

List of Symbols and Acronyms

In this appendix, we summarize the important symbols and list the acronyms used throughout this thesis.

Operators

$\operatorname{argmax}\{.\}$	argument maximizing the expression in brackets
$\operatorname{diag}\{.\}$	diagonal matrix with diagonal entries of vector argument
$\mathcal{E}_X\{.\}$	expectation with respect to X
$\operatorname{Pr}\{.\}$	probability
$\operatorname{Re}\{.\}, \operatorname{Im}\{.\}$	real and imaginary parts of a complex number
$(.)^H$	Hermitian transposition
$(.)^T$	transposition
$(.)^*$	complex conjugation
$. $	magnitude of a complex number

Constants

j	imaginary unit: $j^2 = -1$
π	the number pi: $\pi = 3.14159265358979\dots$
e	Euler number: $e = 2.718281828\dots$
$\mathbf{0}_X$	$X \times X$ all-zero matrix
\mathbf{I}_X	$X \times X$ identity matrix

Sets

\mathcal{C}	signal constellation
\mathbb{C}	matrix (space-time) signal constellation
\mathbb{N}	natural numbers
\mathbb{R}	real numbers
\mathcal{S}	alphabet of scalar transmit symbols
\mathbf{S}	alphabet of matrix (space-time) transmit symbols
\mathcal{V}	alphabet of scalar data symbols
\mathbf{V}	alphabet of matrix (space-time) data symbols
\mathbb{Z}	integer numbers

Sequences

$C[i], C[n]$	binary code symbol
$H[k]$	Ricean fading process
$N[k]$	discrete time zero-mean white Gaussian noise process

$R[k]$	received (matrix) signal sample
$S[k]$	transmit (matrix) symbol
$V[k]$	data (matrix) symbol

Miscellaneous Functions

$\cos(\cdot)$	cosine function
$\det\{\cdot\}$	determinant of matrix argument
$\delta_{i,j}$	Kronecker delta function
$\exp(\cdot)$	exponential function
$\log(\cdot)$	logarithm to base e
$\log_2(\cdot)$	logarithm to base 2
$\log_{10}(\cdot)$	logarithm to base 10
$p(\cdot)$	probability density function
$p(\cdot \cdot)$	conditional probability density function
$\sin(\cdot)$	sine function
$\text{tr}\{\cdot\}$	trace of a matrix

Variables

Scalars

$C(\cdot)$	constellation constrained channel capacity
C_{CSI}	capacity of coherent transmission with perfect CSI

D	number of blocks of length N transmitted over one fading interval
$E(\cdot, \cdot)$	random coding exponent
$E_0(\cdot, \cdot)$	Gallager's function
E_b	average received energy per information bit
E_s	average received energy per symbol
$I(\cdot; \cdot)$	average mutual information
k	scalar and matrix (space-time) symbol discrete-time index, respectively
$\lambda()$	log-likelihood metrics based on conditional PDFs
L	number of independent fading realizations per codeword
ℓ	index of block within one fading realization
M	size of PSK symbol alphabet
n_V	number of vector symbols per code word
n_D	number of vector symbols per code word for arbitrary block fading channel
N	observation window size for noncoherent detection
N_0	single sided power spectral density of passband noise process
N_T	number of transmit antennas
N_R	number of receive antennas
P_W	word error rate
ρ	parameter of random coding exponent
R	data rate in bits per channel use
R_m	modulation rate in bits per channel use
R_V	data rate in bits per vector symbol
σ^2	variance
σ_n^2	noise variance
t_d	decoding delay
T	modulation interval

Matrices

$\mathbf{H}[k]$	matrix of fading samples
$\mathbf{N}[k]$	matrix of noise samples
$\mathbf{R}[k]$	matrix of received signal samples
$\mathbf{S}[k]$	matrix of space-time transmitted signal samples
$\mathbf{S}_D[k]$	diagonal matrix of transmit symbols
$\mathbf{V}[k]$	matrix of space-time data symbols
$\mathbf{V}_D[k]$	diagonal matrix of space-time data symbols
Φ	autocorrelation matrix

Acronyms

AWGN	additive white Gaussian noise
BER	bit error rate
CDD	conventional differential detection
CSI	channel state information
DPSK	differential phase shift keying
dB	decibel
DSTM	differential space-time modulation
MIMO	multiple input multiple output
ML	maximum likelihood
MLD	maximum likelihood decoding
MMSE	minimum mean squared error
MSDD	multiple symbol differential decoding
MSD-SD	multiple symbol differential sphere decoding
pdf	probability density function

PEP	pairwise error probability
PSK	phase shift keying
RCE	random coding exponent
SA	stack algorithm
SD	sphere decoding
SISO	single input single output
SNR	signal-to-noise ratio
ST	space-time
u.i.i.d.	uniformly, independently and identically distributed
WER	word error rate

Bibliography

- [1] S.M. Alamouti. A Simple Transmitter Diversity Scheme for Wireless Communications. *IEEE J. Select. Areas Commun.*, 16(7):1451–1458, October 1998.
- [2] J.A.C. Bingham. *ADSL, VDSL, and Multicarrier Modulation*. John Wiley and Sons, New York, 2000.
- [3] J.K. Cavers. An Analysis of Pilot Symbol Assisted Modulation for Rayleigh Fading Channels. *IEEE Trans. Veh. Technol.*, 40(4):686–693, November 1991.
- [4] N. Chiurtu and T.L. Marzetta. Random Coding Exponents For Unitary Space-Time Modulation. In *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, page 301, Lausanne, June/July 2002.
- [5] K.L. Clarkson, W. Sweldens, and A. Zheng. Fast multiple antenna differential decoding. *IEEE Trans. Commun.*, 49:253–261, February 2001.
- [6] T. Cui and C. Tellambura. Multiple-Symbol Differential Unitary Space-Time Demodulation with Reduced Complexity. *Proc. of IEEE International Conference on Communications (ICC)*, Seoul, May 2005.
- [7] M.O. Damen, A. Chkeif, and J.C. Belfiore. Lattice code decoder for Space-Time codes. *IEEE Commun. Lett.*, pages 161–163, May 2000.

- [8] D. Divsalar and M.K. Simon. Multiple-Symbol Differential Detection of MPSK. *IEEE Trans. Commun.*, 38(3):300–308, March 1990.
- [9] D. Divsalar, M.K. Simon, and M. Shahshahani. The Performance of Trellis-Coded MDPSK with Multiple Symbol Detection. *IEEE Trans. Commun.*, 38(9):1391–1403, September 1990.
- [10] R.G. Gallager. *Information Theory and Reliable Communication*. John Wiley and Sons, 1968.
- [11] C. Gao, A. Haimovich, and D. Lao. Multiple-Symbol Differential Detection for Space-Time Block Codes. *Conference on Information Sciences and Systems, Princeton University*, March 2002.
- [12] B. Hassibi and B.M. Hochwald. Cayley Differential Unitary Space-Time Codes. *IEEE Trans. Inform. Theory*, 48(6):1485–1503, June 2002.
- [13] P. Ho and D. Fung. Error Performance of Multiple-Symbol Differential Detection of PSK Symbols Transmitted over Correlated Rayleigh Fading Channels. *IEEE Trans. Commun.*, 40:25–29, October 1992.
- [14] B.M. Hochwald and T.L. Marzetta. Unitary Space-Time Modulation for Multiple-Antenna Communications in Rayleigh Flat Fading. *IEEE Trans. Inform. Theory*, 46:543–564, March 2000.
- [15] B.M. Hochwald and W. Sweldens. Differential Unitary Space-Time Modulation. *IEEE Trans. Commun.*, 48(12):2041–2052, December 2000.
- [16] B.L. Hughes. Differential Space-Time Modulation. *IEEE Trans. Inform. Theory*, 46(7):2567–2578, November 2000.
- [17] B.L. Hughes. Optimal Space-Time Constellations from Groups. *IEEE Trans. Inform. Theory*, 49(2):401–410, February 2003.

- [18] J. Jaldén, M. Skoglund, and B. Ottersten. On the Random Coding Exponent of Multiple Antenna Systems Using Space-Time Block Codes. In *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, page 188, Chicago, June 2004.
- [19] G. Kaplan and S. Shamai (Shitz). Error Probabilities for the Block-Fading Gaussian Channel. *AEÜ, Int. Journal of Electronics*, 49(4):192–205, 1995.
- [20] R. Knopp. *Coding and Multiple-Access Over Fading Channels*. PhD Thesis, Ecole Polytechnique Fédérale de Lausanne, 1997.
- [21] R. Knopp and H. Leib. M -ary Phase Coding for the Noncoherent AWGN Channel. *IEEE Trans. Inform. Theory*, 40(6):1968–1984, November 1994.
- [22] L. Lampe, R. Schober, V. Pauli and C. Windpassinger. Multiple-Symbol Differential Sphere Decoding. *IEEE Trans. Commun.*, 53(12):1981–1985, December 2005.
- [23] L. Lampe. *Noncoherent Coded Modulation*. PhD Thesis, Universität Erlangen-Nürnberg, 2002.
- [24] L. Lampe and R. Fischer. Random Coding Exponent Based Design of Coded Modulation for Multiple-Symbol Differential Detection. In *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, page 163, Washington, June 2001.
- [25] H. Leib and S. Pasupathy. Optimal Noncoherent Block Demodulation of Differential Phase Shift Keying (DPSK). *Archiv für Elektronik und Übertragungstechnik (International Journal of Electronics)*, 45(5):299–305, 1991.
- [26] C. Ling, W. Mow, K. Li, and A. Kot. Multiple Antenna Differential Lattice Decoding. *IEEE J. Select. Areas Commun.*, 23:1821–1829, September 2005.
- [27] E. Malkamäki and H. Leib. Coded Diversity on Block-Fading Channels. *IEEE Trans. Inform. Theory*, 45(2):771–782, March 1999.

- [28] T.L. Marzetta and B.M. Hochwald. Capacity of a Mobile Multiple-Antenna Communication Link in Rayleigh Flat Fading. *IEEE Trans. Inform. Theory*, 48:139–157, January 1999.
- [29] R.J. McEliece and W.E. Stark. Channels with Block Interference. *IEEE Trans. Inform. Theory*, 30(1):44–53, January 1984.
- [30] J.P. McGeehan and A.J. Bateman. Phase Locked Transparent Tone-In-Band (TTIB). *IEEE Trans. Commun.*, 32:81–87, January 1984.
- [31] U. Mengali and A.N.D'Andrea. *Synchronization Techniques for Digital Receivers*. Plenum Press, New York, 1997.
- [32] K. Miller. *Complex Stochastic Processes*. Addison Wesley, New York, 1974.
- [33] M.L. Moher and J.H. Lodge. TCMP - A Modulation and Coding Strategy for Rician Fading channels. *IEEE J. Select. Areas Commun.*, 7(9):1347–1355, December 1989.
- [34] V. Pauli and L. Lampe. Tree Search Multiple Symbol Differential Decoding for Unitary Space Time Modulation. *Submitted to IEEE Trans. Commun.*, preprint available from <http://www.ece.ubc.ca/~lampe/publicat.html>, 2005., 2005.
- [35] D. Raphaeli. Improvement of Noncoherent Orthogonal Coding by Time Overlapping. *IEEE Proceedings-Communications*, 141(5):309–311, October 1994.
- [36] A. Shokrollahi, B. Hassibi, B.M. Hochwald, and W. Sweldens. Representation Theory for High Rate Multiple-Antenna Code Design. *IEEE Trans. Inform. Theory*, 47(6):2335–2367, September 2001.
- [37] S. Srinivasan, L. Lampe, and V. Pauli. On the Random Coding Exponent for Differential Modulation and Detection. *Submitted to ISIT 2006, Seattle*, January 2006.

- [38] P. Tarasak and V. Bhargava. Reduced Complexity Multiple Symbol Differential Detection of Space-Time Block Code. *Proc. of IEEE Wireless Communications and Networking Conference (WCNC)*, Orlando, Florida, March 2002.
- [39] V. Tarokh and H. Jafarkhani. A Differential Detection Scheme for Transmit Diversity. *IEEE J. Select. Areas Commun.*, 18(7):1168–1174, July 2000.
- [40] V. Tarokh, H. Jafarkhani, and A.R. Calderbank. Space-Time Block Codes from Orthogonal Designs. *IEEE Trans. Inform. Theory*, 45(5):1456–1467, July 1999.
- [41] J. Ventura-Traveset, G. Caire, E. Biglieri, and G. Taricco. Impact of Diversity Reception on Fading Channels with Coded Modulation-Part II: Differential Block Detection. *IEEE Trans. Commun.*, 45(6):676–686, June 1997.
- [42] S.G. Wilson, J. Freebersyser, and C. Marshall. Multi-Symbol Detection of M -DPSK. *Proc. of IEEE Global Telecommunications Conference (Globecom)*, pages 47.3.1–47.3.6, November 1989.
- [43] J. Wolfowitz. Memory Increases Capacity. *Information and Control*, 11:423–428, 1967.