Lattice-Based Precoding And Decoding in MIMO Fading Systems

by

Mahmoud Taherzadeh

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2008

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

In this thesis, different aspects of lattice-based precoding and decoding for the transmission of digital and analog data over MIMO fading channels are investigated:

1) Lattice-based precoding in MIMO broadcast systems: A new viewpoint for adopting the lattice reduction in communication over MIMO broadcast channels is introduced. Lattice basis reduction helps us to reduce the average transmitted energy by modifying the region which includes the constellation points. The new viewpoint helps us to generalize the idea of lattice-reduction-aided precoding for the case of unequal-rate transmission, and obtain analytic results for the asymptotic behavior of the symbol-error-rate for the lattice-reduction-aided precoding and the perturbation technique. Also, the outage probability for both cases of fixed-rate users and fixed sum-rate is analyzed. It is shown that the lattice-reduction-aided method, using LLL algorithm, achieves the optimum asymptotic slope of symbol-error-rate (called the precoding diversity).

2) Lattice-based decoding in MIMO multiaccess systems and MIMO point-to-point systems: Diversity order and diversity-multiplexing tradeoff are two important measures for the performance of communication systems over MIMO fading channels. For the case of MIMO multiaccess systems (with single-antenna transmitters) or MIMO point-to-point systems with V-BLAST transmission scheme, it is proved that lattice-reduction-aided decoding achieves the maximum receive diversity (which is equal to the number of receive antennas). Also, it is proved that the naive lattice decoding (which discards the out-of-region decoded points) achieves the maximum diversity in

V-BLAST systems. On the other hand, the inherent drawbacks of the naive lattice decoding for general MIMO fading systems is investigated. It is shown that using the naive lattice decoding for MIMO systems has considerable deficiencies in terms of the diversity-multiplexing tradeoff. Unlike the case of maximum-likelihood decoding, in this case, even the perfect lattice space-time codes which have the non-vanishing determinant property can not achieve the optimal diversity-multiplexing tradeoff.

3) Lattice-based analog transmission over MIMO fading channels: The problem of finding a delay-limited schemes for sending an analog source over MIMO fading channels is investigated in this part. First, the problem of robust joint source-channel coding over an additive white Gaussian noise channel is investigated. A new scheme is proposed which achieves the optimal slope for the signal-to-distortion-ratio (SDR) curve (unlike the previous known coding schemes). Then, this idea is extended to MIMO channels to construct lattice-based codes for joint source-channel coding over MIMO channels. Also, similar to the diversity-multiplexing tradeoff, the asymptotic performance of MIMO joint source-channel coding schemes is characterized, and a concept called *diversity-fidelity tradeoff* is introduced in this thesis.

Acknowledgements

This work would not exist without help and support from many individuals who deserve sincere recognition.

First and foremost, I am deeply grateful to Professor Amir K. Khandani for providing guidance and inspiration throughout the course of my graduate studies. I would like to thank Amir for his support which was not limited to research work.

I also would like to thank the members of my dissertation committee, Professors Frank Kschischang, Oussama Damen, Murat Uysal, and Wentang Kuo for taking the time to carefully read my thesis and providing me with insightful questions and suggestions.

I would like to thank all members of coding and signal transmission lab, who helped create a very pleasant and friendly learning environment. During the course of my study, I have enjoyed many insightful discussions with them. Specially, I would like to thank Amin Mobasher, Kamyar Moshksar, Mohammad Maddah-Ali, Alireza Bayesteh, and Seyed Abolfazl Motahari.

Contents

List of Figures

List of Notations and Abbreviations

MIMO	Multiple-Input Multiple-Output
MISO	Multiple-Input Single-Output
SIMO	Single-Input Multiple-Output
SISO	Single-Input Single-Output
SNR	Signal to Noise Ratio
SDR	Signal to Distortion Ratio
SINR	Signal to Interference plus Noise Ratio
CDF	Cumulative Distribution Function
PDF	Probability Density Function
AWGN	Additive White Gaussian Noise
Boldface Upper-Case Letters	Matrices
Boldface Lower-Case Letters	Vectors
$(\mathbf{A})^{\mathtt{H}}$	Hermitian transpose (conjugate transpose) of $\mathbf{A}$
$\det(\mathbf{H})$	Determinant of $\mathbf{H}$
$\mathrm{Tr}(\mathbf{H})$	Trace of $\mathbf{H}$
$\mathbf{H} \succeq 0$	Matrix $\mathbf{H}$ is positive semi-definite
$\mathbf{1}_n$	n-dimensional vector with all entries equal to one
$\mathbf{I}_{n\times n}$	n-dimensional identity matrix
$\mathbf{H}\perp\mathbf{G}$	Columns of $\mathbf{H}$ are orthogonal to columns of $\mathbf{G}$
$\Omega(\mathbf{H})$	The sub-space spanned by columns of $\mathbf{H}$
$\mathrm{N}(\mathbf{H})$	The null space of the matrix $\mathbf{H}$

$\mathbf{h}^{(i)}$	The i^{th} column of the matrix $\mathbf{H}$
$[a_{p,q}]_{(p,q)}^{m\times n}$	$m \times n$ matrix with $a_{p,q}$ as entry (p,q)
$\mathbb{C}$	The set of complex numbers
$\mathbb{R}$	The set of real numbers
$\mathbf{x}(i:j)$	A vector including the entries i to j of the vector $\mathbf{x}$
$\lvert\mathcal{S}\rvert$	Cardinality of the set $\mathcal{S}$
N_t	Number of transmit antennas
N_r	Number of receive antennas
$E(.)$	Expectation
$\Pr(.)$	Probability
$\lfloor z \rfloor$	Floor of z
$\lceil z \rceil$	Ceiling of z
$\log$	Logarithm in basis 2
$\Re(x)$	Real part of the complex number x
$\Im(x)$	Imaginary part of the complex number x

Chapter 1

Introduction

In the last decade, there has been a substantial increase in multimedia services in wireless environments and the demand for ubiquitous high-rate connectivity. To meet these goals, designing spectrally efficient signaling schemes is an essential element. An important approach to increase spectral efficiency is to exploit the spatial resources in the system, i.e., to use multiple antennas at the transmitter and the receiver. Communications over multiple-antenna systems is also called multiple-input multiple-output (MIMO) communications. When all the transmit antennas are used by one transmitter, and all the receive antennas are used by one receiver, the system is called *point-to-point MIMO*. When there are several transmitters, but the receive antennas are used by a single user, the system is called *MIMO multiaccess*. Similarly, when there is a single transmitter and several receivers, the systems is called *MIMO broadcast*.

In most wireless environments, the channel between a transmit and a receive antenna is assumed to be a Rayleigh fading channel, i.e. the baseband signal (which can be modeled as a complex number) is assumed to be multiplied by a complex channel gain, with a complex Gaussian distribution, and then added by an additive Gaussian noise [85]. The error rate of uncoded transmission over this channel scales as $\frac{1}{SNR}$, where SNR denotes the signal-to-noise-ratio. It was known that when the environment is highly dispersive, we can reduce the error rate and improve its scaling by exploiting multiple antennas. In the literature, this improvement in the asymptotic scaling of the error rate is known as *diversity*. When the error rate is asymptotically proportional to $\frac{1}{SNR^d}$, for a positive number d, then d is called the *diversity order*. It is known that for a MIMO point-to-point system, diversity order is at most equal to N_tN_r, where N_t and N_r are the numbers of receive and transmit antennas [81].

The other important measure in the communications over wireless fading channels is the average transmission rate. For a single antenna system, the average achievable data rate is asymptotically close to $\log SNR$. In [82], it was shown that when multiple antennas are used in a fading system with independent channel coefficients, the average capacity asymptotically grows as $M \log SNR$, where M is the minimum of the number of transmit and receive antennas. This increase in the rate is called *spatial multiplexing*. In [25], a practical system, called V-BLAST, was proposed for providing spatial multiplexing in MIMO point-to-point systems. This system was based on sending independent symbols over transmit antennas and using sequential minimum mean-squared error (MMSE) detection at the receiver.

Since the seminal works by Telatar [82] and Foschini *et al.* [25], MIMO communications has attracted intensive research. In [3, 80, 81], some coding schemes are proposed to obtain the full diversity, along with a modest rate. Also, for noncoherent communications over time-varying channels, differential space-time modulation schemes (based on unitary matrices) are introduced in [37, 39, 79]. These works gave birth to the field of *Space-Time Coding*, which has attracted many researchers in recent years [17, 31, 32, 40, 41, 46, 67, 70].

To have a unified benchmark for evaluating different space-time coding schemes in terms of both reliability and spectral efficiency, the concept of *Diversity-Multiplexing Tradeoff* was introduced in [96]. For a family of space-time codes $\mathcal{C}_i$, corresponding to the signal-to-noise ratio SNR_i, the normalized rate r is defined as $r = \lim_{SNR_i \to \infty} \frac{R_i}{\log SNR_i}$, where R_i is the rate of the code $\mathcal{C}_i$. Also, the diversity $d(r)$ is defined as $d(r) = \lim_{SNR_i \to \infty} \frac{-\log \Pr(error)}{\log SNR_i}$. For any coding scheme, there is a trade-off between the normalized rate r (which can be any real number between 0 and the multiplexing gain, $\min\{N_t, N_r\}$) and the diversity $d(r)$ (which can be any number between 0 and the maximum diversity order N_tN_r). In [96], the authors have shown that for integer values of r, $0 \leq r \leq \min\{N_t, N_r\}$, the optimum diversity $d^*(r)$ is $d^*(r) = (N_t - r)(N_r - r)$.

Several articles have dealt with achieving the optimum tradeoff with practical schemes. For the encoding part, recently, several lattice codes are introduced which have the non-vanishing determinant property and achieve the optimal trade-off, conditioned on using exact maximum-likelihood decoding [18, 38, 53]. The lattice structure of these codes facilitates the encoding. For the decoding part, various lattice decoders, including the sphere decoder and the lattice-reduction-aided decoder are presented in the literature [14, 91].

Besides point-to-point systems, new information theoretic results [7, 87, 88, 95] have shown that using multiple-antennas can be also useful in multiuser MIMO systems. It has been shown that in a MIMO broadcast system, the sum-capacity grows linearly with the minimum number of the transmit and receive antennas [87, 88, 95].

Another important issue in MIMO systems is the problem of joint source-channel coding, for the transmission of analog sources over MIMO channels. This problem can be very important in the high-rate broadcast of multimedia in wireless environments. While there has been some work on this issue [6, 30], the main questions (about both the theoretical limitations and the design of practical schemes) are not answered yet.

The organization of this thesis is as follows. Sections 1.1 and 1.2 briefly introduce the basic concepts of lattices and *lattice-basis reduction*. The rest of this chapter is devoted to a brief summary of the materials presented in the following chapters. Chapter 2 deals with the precoding for MIMO broadcast channels. In chapters 3 and 4, lattice decoding in MIMO multiaccess and point-to-point systems is considered. Chapter 5 deals with joint source-channel coding and chapter 6 investigates this problem in MIMO systems.

1.1 Lattices

Lattice structures have been frequently used in different communication applications such as quantization or decoding of MIMO systems. A real (or complex) lattice Λ is a discrete set[1] of N-dimensional vectors in the real Euclidean space $\mathbb{R}^N$ (or the complex

[1] A set $\mathcal{S}$ is called discrete if for any point in it, there is a neighborhood that does not contain any other points of the set.

Euclidean space $\mathbb{C}^N$) that forms a group under ordinary vector addition [10, 23, 29]. Every lattice Λ is generated by the integer linear combinations of some set of linearly independent vectors $\mathbf{b}_1, \cdots, \mathbf{b}_M \in \Lambda$, where the integer M, $M \leq N$, is called the dimension of the lattice Λ. The set of vectors $\{\mathbf{b}_1, \cdots, \mathbf{b}_M\}$ is called a basis of Λ, and the matrix $\mathbf{B} = [\mathbf{b}_1, \cdots, \mathbf{b}_M]$, which has the basis vectors as its columns, is called the basis matrix (or the *generator matrix*) of Λ.

For a real lattice Λ with basis vectors $\mathbf{b}_1, \cdots, \mathbf{b}_M \in \Lambda$, the parallelotope consisting of the points

$$a_1\mathbf{b}_1 + \cdots + a_M\mathbf{b}_M \quad (0 \leq a_i < 1)$$

is called the *fundamental parallelotope* or the *fundamental region*[2]. The basis for representing a lattice is not unique, hence, there are various choices for the fundamental region of the lattice. However, the volume of a fundamental region of the lattice is uniquely determined by the lattice and is independent of the choice of the basis. This volume is equal to

$$\mathrm{Vol}(\Lambda) = \left(\det \mathbf{B}\mathbf{B}^{\mathtt{H}}\right)^{\frac{1}{2}}$$

and its square is called the *determinant* or the *discriminant* of the lattice. For any lattice point $\mathbf{v} \in \Lambda$, the set of points whose closest lattice point is $\mathbf{v}$ is called the *Voronoi region* (or the *Voronoi cell*) corresponding to $\mathbf{v}$. The volume of the Voronoi

[2]Similarly, for an M-dimensional complex lattice, we can define the $2M$-dimensional fundamental parallelotope as the region consisting of the points

$$a_1\mathbf{b}_1 + \cdots + a_M\mathbf{b}_M \quad (0 \leq \Re(a_i), \Im(a_i) < 1)$$

regions is equal to the volume of the lattice, and unlike the fundamental region, the the Voronoi regions are independent of the choice of the lattice basis [10].

In Communication applications, lattices are used in constructing signal sets for transmitting digital data, or for quantizing analog data. To construct transmit signals, usually lattice points (or a shifted version of them) inside a certain region $\mathcal{A}$ (which is called *constellation region*) are used as transmit signal constellation. When the constellation region $\mathcal{A}$ is large (compared to the fundamental region of the lattice), the number of signal points can be approximated by $\frac{\text{Vol}(\mathcal{A})}{\text{Vol}(\Lambda)}$. Also, the average transmit energy per dimension can be approximated by the *dimensionless normalized second moment* of the region $\mathcal{A}$ which is defined as

$$G(\mathcal{A}) = \frac{1}{N} \left(\text{Vol}(\mathcal{A})\right)^{-\frac{N+2}{2N}} \int_{\mathcal{A}} \|\mathbf{x}\|^2 \, d\mathbf{x}. \tag{1.1}$$

This type of high-rate approximation is commonly referred to as *continuous approximation* in the literature.

1.2 Lattice-Basis Reduction

Usually a basis consisting of relatively short and nearly orthogonal vectors is desirable. The procedure of finding such a basis for a lattice is called *Lattice Basis Reduction*. A popular criterion for lattice-basis reduction is to find a basis such that $\|\mathbf{b}_1\| \cdot \ldots \cdot \|\mathbf{b}_M\|$ is minimized. Because the volume of the lattice does not change with the change of basis, this problem is equivalent to minimizing the orthogonality defect which is defined as

$$\delta \triangleq \frac{(\|\mathbf{b}_1\|^2 \|\mathbf{b}_2\|^2 \ldots \|\mathbf{b}_M\|^2)}{\det \mathbf{B}^{\mathtt{H}} \mathbf{B}}. \tag{1.2}$$

The problem of finding such a basis is NP-hard [28]. Several distinct sub-optimal reductions have been studied in the literature, including those associated to the names Minkowski, Korkin-Zolotarev, and more recently Lenstra-Lenstra and Lovasz (LLL) [29].

An ordered basis $(\mathbf{b}_1, \cdots, \mathbf{b}_M)$ is a *Minkowski-Reduced Basis* [36] if

- $\mathbf{b}_1$ is the shortest nonzero vector in the lattice Λ, and
- For each $k = 2, ..., M$, $\mathbf{b}_k$ is the shortest nonzero vector in Λ such that $(\mathbf{b}_1, \cdots, \mathbf{b}_k)$ may be extended to a basis of Λ.

Minkowski reduction can be seen as a greedy solution for the lattice-basis reduction problem. However, finding Minkowski reduced basis is equivalent to finding the shortest vector in the lattice and this problem by itself is NP-hard. Thus, there is no polynomial time algorithm for this reduction method.

In [45], a reduction algorithm (called *LLL algorithm*) is introduced which uses the Gram-Schmidt orthogonalization and has a polynomial complexity and guarantees a bounded orthogonality defect. For any ordered basis of Λ, say $(\mathbf{b}_1, \cdots, \mathbf{b}_M)$, one can compute an ordered set of Gram-Schmidt vectors, $\left(\hat{\mathbf{b}}_1, \cdots, \hat{\mathbf{b}}_M\right)$, which are mutually orthogonal, using the following recursion:

$$\begin{aligned} \hat{\mathbf{b}}_i = \mathbf{b}_i - \textstyle\sum_{j=1}^{i-1} \mu_{ij} \hat{\mathbf{b}}_j, \text{ with} \\ \mu_{ij} = \frac{< \mathbf{b}_i, \hat{\mathbf{b}}_j >}{\|\hat{\mathbf{b}}_j\|^2}. \end{aligned} \tag{1.3}$$

where $< \cdot, \cdot >$ is the inner product. An ordered basis $(\mathbf{b}_1, \cdots, \mathbf{b}_M)$ is an *LLL Reduced Basis* [45] if,

- $|\mu_{ij}| \leq \frac{1}{2}$ for $1 \leq i < j \leq M$, and
- $p \cdot \|\hat{\mathbf{b}}_i\|^2 \leq \|\hat{\mathbf{b}}_{i+1} + \mu_{i+1,i}\hat{\mathbf{b}}_i\|^2$

where $\frac{1}{4} < p < 1$, and $\left(\hat{\mathbf{b}}_1, \cdots, \hat{\mathbf{b}}_M\right)$ is the Gram-Schmidt orthogonalization of the ordered basis and $\mathbf{b}_i = \sum_{j=1}^{i} \mu_{ij}\hat{\mathbf{b}}_j$ for $i = 1, ..., M$.

It is shown that the LLL basis-reduction algorithm produces relatively short basis vectors with a polynomial-time computational complexity [45]. The LLL basis reduction has found extended applications in several contexts due to its polynomial-time complexity. In [52], the LLL algorithm is generalized for Euclidean rings (including the ring of complex integers). In this chapter, we will use the following important property of the complex LLL reduction (for $p = \frac{3}{4}$):

Theorem 1.1 (see [52])**.** *Let* Λ *be an* M*-dimensional complex lattice and* $\mathbf{B} = [\mathbf{b}_1 ... \mathbf{b}_M]$ *be the LLL reduced basis of* Λ*. If* δ *is the orthogonality defect of* $\mathbf{B}$*, then,*

$$\sqrt{\delta} \leq 2^{M(M-1)}. \tag{1.4}$$

1.3 Summary of the Dissertation

Chapter 2 presents a lattice-reduction-aided technique for precoding in MIMO broadcast systems. Although a similar technique has been introduced in the past [92], the new viewpoint helps us in generalizing it for unequal-rate transmission, as well as analyzing the asymptotic performance of these lattice-reduction-aided methods. This chapter introduces a concept called precoding diversity for the asymptotic performance of the precoding techniques. It analyzes the asymptotic performance of lattice-based precoding techniques (including the exact lattice precoding and its lattice-reduction-aided approximation), and shows that they achieve the optimum precoding diversity.[3]

Chapter 3 investigates the performance of lattice-reduction-aided decoding in MIMO multiaccess systems and MIMO point-to-point systems with V-BLAST transmission. In this chapter, it is shown that lattice-reduction-aided decoding achieves the optimum receive diversity in these systems. It is also shown that the naive lattice decoding (i.e. lattice decoding without considering the boundary of the constellation) is sufficient for achieving the receive diversity.[4]

Chapter 4 evaluates the diversity-multiplexing tradeoff of lattice-decoded codes in MIMO point-to-point systems. This chapter presents an upper bound on the achievable tradeoff of lattice-decoded codes which are based on full-rate lattices. It shows that when we fix the lattice structure of the space-time code, if we use naive lattice decoding, we cannot achieve the optimum diversity-multiplexing tradeoff. In recent years, different elegant lattice codes have been introduced which achieve the optimal trade-off [18] [53] [38], but they need ML decoding to achieve optimality. On the other hand, there can exist a family of lattice codes (based on different lattice structures for different rates and SNR values) which achieves the optimum tradeoff using the naive lattice decoding [51]. The result of this chapter shows that the problem of achieving the optimum diversity-multiplexing tradeoff by a practical encoding and decoding scheme

[3]The content of Chapter 2 was presented in part in [77, 78], and published in [75].

[4]The content of Chapter 3 was presented in part in [78], and published in [76].

is still open. [5]

Chapter 5 deals with constructing robust joint source-channel codes for a Gaussian channel and an analog source. It introduces a new coding scheme which achieves the optimum scaling of signal-to-distortion-ratio (SDR) for the whole range of SNR. It also presents some bounds on the performance of other alternatives (including general hybrid digital analog codes). These robust joint source-channel codes can be used for sending analog sources over single-input single-output (SISO) fading channels. In fading channels with time-varying channel gains, or when the objective is to transmit common analog data to different users (with different channel gains), it is desirable to have a coding scheme which is robust to the variations of the channel gain. Several analog and hybrid digital-analog schemes have been introduced in the past, but none of them achieve the optimum SDR scaling.[6]

Chapter 6 uses the idea presented in Chapter 5, to construct robust joint source-channel codes for MIMO fading channels. It also introduces a concept called diversity-fidelity tradeoff (which can be regarded as the analog version of the diversity-multiplexing tradeoff). It shows that the proposed coding scheme achieves the optimum diversity-fidelity tradeoff.

Finally, Chapter 7 presents a summary of the thesis contributions and discusses several future research directions.

---

[5] The content of Chapter 4 is also available in [71].

[6] The content of Chapter 5 was presented in part in [73, 74], and is also available in [72].

Chapter 2

Lattice-based precoding in MIMO broadcast systems

In the recent years, communications over multiple-antenna fading channels has attracted the attention of many researchers. Initially, the main interest has been on the point-to-point Multiple-Input Multiple-Output (MIMO) communications [3, 25, 80–82]. In [82] and [25], the authors have shown that the capacity of a MIMO point-to-point channel increases linearly with the minimum number of the transmit and the receive antennas.

More recently, new information theoretic results [7], [95], [88], [87] have shown that in multiuser MIMO systems, one can exploit most of the advantages of multiple-antenna systems. It has been shown that in a MIMO broadcast system, the sum-capacity grows linearly with the minimum number of the transmit and receive antennas [95], [88], [87]. To achieve the sum capacity, some information theoretic schemes, based on dirty-paper coding, are introduced. Dirty-paper coding was originally proposed for the Gaussian interference channel when the interfering signal is known at the transmitter [12]. Some methods, such as using nested lattices, are introduced as practical techniques to achieve the sum-capacity promised by the dirty-paper coding [19]. However, these methods are not easy to implement.

As a simple precoding scheme for MIMO broadcast systems, the channel inversion technique (or zero-forcing beamforming [7]) can be used at the transmitter to separate the data for different users. To improve the performance of the channel inversion technique, a zero-forcing approximation of the dirty paper coding (based on QR decomposition) is introduced in [7] (which can be seen as a scalar approximation of [19]). However, both of these methods are vulnerable to the poor channel conditions, due to the occasional near-singularity of the channel matrix (when the channel matrix has at least one small eigenvalue). This drawback results in a poor performance in terms of the symbol-error-rate for the mentioned methods [55].

In [55], the authors have introduced a *vector perturbation technique* which has a good performance in terms of symbol error rate. Nonetheless, this technique requires a lattice decoder which is an NP-hard problem. To reduce the complexity of the lattice decoder, in [22, 90–92], the authors have used lattice-basis reduction to approximate the closest lattice point (using Babai approximation).

In this chapter, we present a transmission technique for the MIMO broadcast channel based on the lattice-basis reduction. Instead of approximating the closest lattice point in the perturbation problem, we use the lattice-basis reduction to reduce the average transmitted energy by reducing the second moment of the fundamental region generated by the lattice basis. This viewpoint helps us to: (i) achieve a better performance as compared to [90], (ii) expand the idea for the case of unequal-rate transmission, and (iii) obtain some analytic results for the asymptotic behavior (SNR$\longrightarrow \infty$) of the symbol-error-rate for both the proposed technique and the perturbation technique of [55].

The rest of the chapter is organized as the following: Section 2.1 briefly describe the system model. In section 2.2, the proposed method is described and in section 2.3, the proposed approach is extended for the case of unequal-rate transmission. In section 2.4, we consider the asymptotic performance of the proposed method for high SNR values, in terms of the probability of error. We define the precoding diversity and the outage probability for the case of fixed-rate users. It is shown that by using lattice basis reduction, we can achieve the maximum precoding diversity. For the proof, we use a bound on the orthogonal deficiency of an LLL-reduced basis. Also, an upper bound is given for the probability that the length of the shortest vector of a lattice (generated by complex Gaussian vectors) is smaller than a given value. Using this result, we also show that the perturbation technique achieves the maximum precoding diversity. In section 2.5, some simulation results are presented. These results show that the proposed method offers almost the same performance as [55] with a much smaller complexity. As compared to [90], the proposed method offers almost the same performance. However, by sending a very small amount of side information (a few bits for one fading block), the modified proposed method offers a better performance with a similar complexity.

Finally, in section 2.6, some concluding remarks are presented.

2.1 System Model and Problem Formulation

We consider a multiple-antenna broadcast system with N_t transmit antennas and N_r single-antenna users ($N_t \geq N_r$). Consider $\mathbf{y} = [y_1, ..., y_{N_r}]^T$, $\mathbf{x} = [x_1, ..., x_{N_t}]^T$, $\mathbf{w} = [w_1, ..., w_{N_r}]^T$, and the $N_r \times N_t$ matrix $\mathbf{H}$, respectively, as the received signal, the transmitted signal, the noise vector, and the channel matrix. The transmission over the channel can be formulated as,

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}. \tag{2.1}$$

The channel is assumed to be Raleigh, i.e. the elements of $\mathbf{H}$ are i.i.d. with the zero-mean unit-variance complex Gaussian distribution and the noise is i.i.d. additive Gaussian. Moreover, we have the energy constraint on the transmitted signal, $\mathrm{E}(\|\mathbf{x}\|^2) = 1$. The energy of the additive noise is σ^2 per antenna, i.e. $\mathrm{E}(\|\mathbf{w}\|^2) = N_r\sigma^2$. The Signal-to-Noise Ratio (SNR) is defined as $\rho = \frac{1}{\sigma^2}$.

In a broadcast system, the receivers do not cooperate with each other (they should decode their respective data, independently). The main strategy in dealing with this restriction is to apply an appropriate precoding scheme at the transmitter. The simplest method in this category is using the channel inversion technique at the transmitter to separate the data for different users:

$$\mathbf{s} = \mathbf{H}^{+}\mathbf{u}, \tag{2.2}$$

where $\mathbf{H}^{+} = \mathbf{H}^{\mathtt{H}}(\mathbf{H}\mathbf{H}^{\mathtt{H}})^{-1}$, and $\mathbf{H}^{\mathtt{H}}$ is the Hermitian transpose (conjugate transpose) of $\mathbf{H}$. Moreover, $\mathbf{s}$ is the transmitted signal before the normalization ($\mathbf{x} = \dfrac{\mathbf{s}}{\sqrt{\mathrm{E}(\|\mathbf{s}\|^2)}}$ is the normalized transmitted signal), and $\mathbf{u}$ is the data vector, i.e. u_i is the data for the i'th user. For $N_t = N_r$ (the number of transmit antennas and the number of users are equal), the transmitted signal is

$$\mathbf{s} = \mathbf{H}^{-1}\mathbf{u}. \tag{2.3}$$

The problem arises when $\mathbf{H}$ is poorly conditioned and $\|\mathbf{s}\|$ becomes very large, resulting in a high power consumption. This situation occurs when at least one of the singular values of $\mathbf{H}$ is very small which results in vectors with large norms as the columns of $\mathbf{H}^{+}$. Fortunately, most of the time (especially for high SNRs), we can combat the effect of a small singular value by changing the supporting region of the constellation which is the main motivation of this chapter of thesis.

When the data of different users are selected from $\mathbb{Z}[i]$, the overall constellation can be seen as a set of lattice points. In this case, lattice algorithms can be used to modify the constellation. Especially, lattice-basis reduction is a natural solution for modifying the supporting region of the constellation.

2.2 Proposed Approach

Assume that the data for different users, u_i, is selected from the points of the integer lattice (or from the half-integer grid [24]). The data vector $\mathbf{u}$ is a point in the Cartesian product of these sub-constellations. As a result, the overall receive constellation consists of the points from $\mathbb{Z}^{2N_r}$, bounded within a $2N_r$-dimensional hypercube. At the transmitter side, when we use the channel inversion technique, the transmitted signal is a point inside a parallelotope whose edges are parallel to vectors, defined by the columns of $\mathbf{H}^+$. If the data is a point from the integer lattice $\mathbb{Z}^{2N_r}$, the transmitted signal is a point in the lattice generated by $\mathbf{H}^+$. When the squared norm of at least one of the columns of $\mathbf{H}^+$ is too large, some of the constellation points require high energy for the transmission. We try to reduce the average transmitted energy, by replacing these points with some other points with smaller square norms. However, the lack of cooperation among the users imposes the restriction that the received signals should belong to the integer lattice $\mathbb{Z}^{2N_r}$ (to avoid the interference among the users). The core of the idea in this chapter is based on using an appropriate supporting region for the transmitted signal set to minimize the average energy, without changing the underlying lattice. This is achieved through the lattice-basis reduction.

When we use the continuous approximation (which is appropriate for large constellations), the average energy of the transmitted signal is approximated by the second moment of the transmitted region [24]. When we assume equal rates for the users, e.g. R bits per user ($\frac{R}{2}$ bits per dimension), the signal points (at the receiver) are inside a hypercube with an edge of length a where

$$a = 2^{R/2}. \tag{2.4}$$

Therefore, the supporting region of the transmitted signal is the scaled version of the fundamental region of the lattice generated by $\mathbf{H}^+$ (corresponding to its basis) with the scaling factor a. Note that by changing the basis for this lattice, we can change the corresponding fundamental region (a parallelotope generated by the basis of the lattice and centered at the origin). The second moment of the resulting region is proportional to the sum of the squared norms of the basis vectors (see Appendix A). Therefore, we should try to find a basis reduction method which minimizes the sum of the squared norms of the basis vectors. Figure 2.1 shows the application of the lattice basis reduction in reducing the average energy by replacing the old basis with a new basis which has shorter vectors. In this figure, by changing the basis $\mathbf{a}_1, \mathbf{a}_2$ (columns of $\mathbf{H}^+$) to $\mathbf{b}_1, \mathbf{b}_2$ (the reduced basis), the fundamental region $\mathcal{F}$, generated by the original basis, is replaced by $\mathcal{F}'$, generated by the reduced basis.

Among the known reduction algorithms, the Minkowski reduction can be considered as an appropriate greedy algorithm for our problem. Indeed, the Minkowski algorithm is the successively optimum solution because in each step, it finds the shortest vector. However, the complexity of the Minkowski reduction is equal to the complexity of the shortest-lattice-vector problem which is known to be NP-hard [2]. Therefore, we use the LLL reduction algorithm which is a suboptimum solution with a polynomial complexity.

Assume that $\mathbf{B} = \mathbf{H}^+\mathbf{U}$ is the LLL-reduced basis for the lattice obtained by $\mathbf{H}^+$, where $\mathbf{U}$ is an $N_r \times N_r$ unimodular matrix (both $\mathbf{U}$ and $\mathbf{U}^{-1}$ have integer entries). We use $\mathbf{x} = \mathbf{B}\mathbf{u}' = \mathbf{H}^+\mathbf{U}\mathbf{u}'$ as the transmitted signal where

$$\mathbf{u}' = \mathbf{U}^{-1}\mathbf{u} \mod a \tag{2.5}$$

is the precoded data vector, $\mathbf{u}$ is the original data vector, and a is the length of the edges of the hypercube, defined by (2.4). At the receiver side, we use modulo operation

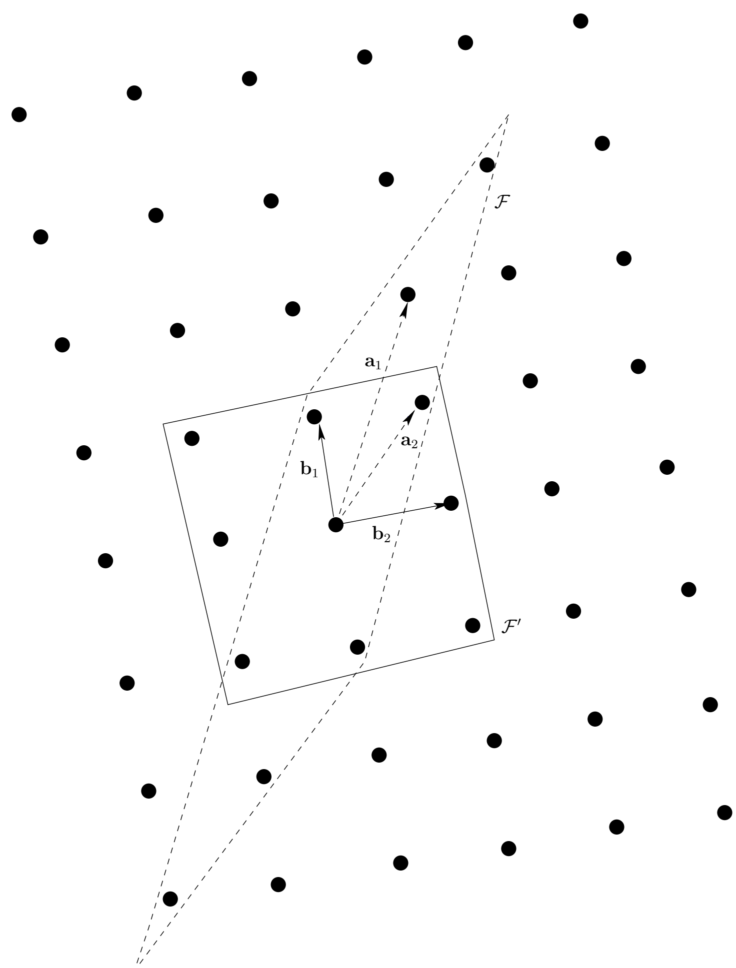

Figure 2.1: Using lattice-basis reduction for reducing the average energy

to find the original data:

$$\mathbf{y} = \mathbf{Hx} + \mathbf{n} = \mathbf{HH}^{+}\mathbf{U}(\mathbf{U}^{-1}\mathbf{u} \mod a) + \mathbf{n} = \mathbf{U}(\mathbf{U}^{-1}\mathbf{u} \mod a) + \mathbf{n} \tag{2.6}$$

$$= \mathbf{UU}^{-1}\mathbf{u} \mod a + \mathbf{n} = \mathbf{u} \mod a + \mathbf{n}. \tag{2.7}$$

In obtaining (2.7) from (2.6), we use the fact that $\mathbf{U}$ and $\mathbf{U}^{-1}$ have integer entries.

In this method, at the beginning of each fading block, we reduce the lattice obtained by $\mathbf{H}^{+}$ and during this block the transmitted signal is computed using (2.5). Neglecting the preprocessing at the beginning of the block (for lattice reduction), the complexity of the precoding is in the order of a matrix multiplication and a modulo operation. Therefore, the complexity of the proposed precoding method is comparable to the complexity of the channel inversion method. However, as we will show by the simulation results, the performance of this method is significantly better, and indeed, is near the performance of the perturbation method, presented in [55].

In the perturbation technique [55], the idea of changing the support region of the constellation has been implemented using a different approach. In [55], $\mathbf{u}' = \mathbf{u} + a\mathbf{l}$ is used as the precoded data, where the integer vector $\mathbf{l}$ is chosen to minimize $\|\mathbf{H}^{+}(\mathbf{u}+a\mathbf{l})\|$. This problem is equivalent to the closest-lattice-point problem for the lattice generated by $a\mathbf{H}^{+}$ (i.e. finding the lattice point which is closer to $-\mathbf{H}^{+}\mathbf{u}$). Therefore, in the perturbation technique, the support region of the constellation is a scaled version of the Voronoi region [11] of the lattice. In the proposed method, we use a parallelotope (generated by the reduced basis of the lattice), instead of the Voronoi region. Although this approximation results in a larger second moment (i.e. higher energy consumption), it enables us to use a simple precoding technique, instead of solving the closest-lattice-point problem.

For the lattice constellations, using a parallelotope instead of the Voronoi region (presented in this chapter) is equivalent with using the Babai approximation instead of the exact lattice decoding (previously presented in [90]). The only practical difference between our lattice-reduction-aided scheme and the scheme presented in [90] is that we reduce $\mathbf{H}^+$, while in [90], $\mathbf{H}^{\mathtt{H}}$ is reduced. As can be observed in the simulation results (figure 2.2), this difference has no significant effect on the performance. However, the new viewpoint helps us in extending the proposed method for the case of variable-rate transmission and obtaining some analytical results for the asymptotic performance.

The performance of the proposed lattice-reduction aided scheme can be improved by combining it with other schemes, such as regularization [54] or the V-BLAST precoding [62], or by sending a very small amount of side information. In the rest of this section, we present two of these modifications.

2.2.1 Regularized lattice-reduction-aided precoding

In [54], the authors have proposed a regularization scheme to reduce the transmitted power, by avoiding the near-singularity of $\mathbf{H}$. In this method, instead of using $\mathbf{H}^+ = \mathbf{H}^{\mathtt{H}}(\mathbf{H}\mathbf{H}^{\mathtt{H}})^{-1}$, the transmitted vector is constructed as

$$\mathbf{x} = \mathbf{H}^{\mathtt{H}}\left(\mathbf{H}\mathbf{H}^{\mathtt{H}} + \alpha\mathbf{I}\right)^{-1}\mathbf{u} \tag{2.8}$$

where α is a positive number. To combine the regularized scheme with our lattice-reduction-aided scheme, we consider $\mathbf{B}_r$ as the matrix corresponding to the reduced basis of the lattice generated by $\mathbf{H}^{\mathrm{H}}\left(\mathbf{H}\mathbf{H}^{\mathrm{H}}+\alpha\mathbf{I}\right)^{-1}$. When we use the regularization, the received signals of different users are not orthogonal anymore and the interference acts like extra noise. Parameter α should be optimized such that the ratio between the power of received signal and the power of the effective noise is minimized [54].

2.2.2 Modified lattice-reduction-aided precoding with small side information

In practical systems, we are interested in using a subset of points with odd coordinates from the integer lattice. In these cases, we can improve the performance of the proposed method by sending a very small amount of side information. When the data vector $\mathbf{u}$ consists of odd integers, using the lattice-basis reduction may result in points with some even coordinates (i.e. $\mathbf{U}^{-1}\mathbf{u}$ has some even elements), instead of points with all-odd coordinates in the new basis. For this case, in (2.5), the set of precoded data $\mathbf{u}'$ is not centered at the origin, hence the transmitted constellation (which includes all the valid points $\mathbf{B}\mathbf{u}'$) is not centered at the origin. Therefore, we can reduce the transmitted energy and improve the performance by shifting the center of the constellation to the origin. It can be shown that the translation vector is equal to $(\mathbf{U}^{-1}[1+i, 1+i, \cdots, 1+i]^{\mathrm{T}} + [1+i, 1+i, \cdots, 1+i]^{\mathrm{T}}) \mod 2$ where $i=\sqrt{-1}$. When we use this shifted version of the constellation, we must send the translation vector to the users (by sending 2 bits per user) at the beginning of the block. However, compared to the size of the block of data, the overhead of these two bits is negligible.

The above idea of using a shift vector can be also used to improve the perturbation technique (if we only use the odd points of the lattice). After reducing the inverse of the channel matrix and obtaining the bits (corresponding to the shift vector) at the beginning of each fading block, the closest point to the signal computed in equation (2.5) can be found by using the sphere decoder. Then, the transmitted signal is obtained by

$$\mathbf{x} = \mathbf{B}\left(\mathbf{u}' + a\mathbf{l} + \mathbf{u}_{par}\right), \tag{2.9}$$

where $\mathbf{u}_{par}$ is the zero-one shift vector, which is computed for users at the beginning of the fading block, and the perturbation vector $\mathbf{l}$ is an even integer vector such that the vector $\mathbf{x}$ has the minimum energy. This method which can be considered as *modified perturbation method* outperforms the perturbation method in [55]. When we are not restricted to the odd lattice points, using (2.9) instead of $\mathbf{H}^{+}\left(\mathbf{u}' + a\mathbf{l}\right)$ does not change the performance of the perturbation method. It only reduces the complexity of the lattice decoder [1].

2.3 Unequal-Rate Transmission

In the previous section, we had considered the case that the transmission rates for different users are equal. In some applications, we are interested in assigning different rates to different users. Consider $R_1, ..., R_{N_r}$ as the transmission rates for the users (we consider them as even integer numbers). Equation (2.5) should be modified as

$$\mathbf{u}' = \mathbf{U}^{-1}\mathbf{u} \mod \mathbf{a}, \tag{2.10}$$

where the entries of $\mathbf{a} = [a_1, ..., a_{N_r}]^T$ are equal to

$$a_i = 2^{R_i/2}. \tag{2.11}$$

Also, at the receiver side, instead of (2.6) and (2.7), we have

$$\mathbf{y} = \mathbf{Hx} + \mathbf{n} = \mathbf{HH}^{+}\mathbf{U}(\mathbf{U}^{-1}\mathbf{u} \mod \mathbf{a}) + \mathbf{n} = \mathbf{U}(\mathbf{U}^{-1}\mathbf{u} \mod \mathbf{a}) + \mathbf{n} \tag{2.12}$$

$$= \mathbf{UU}^{-1}\mathbf{u} \mod \mathbf{a} + \mathbf{n} = \mathbf{u} \mod \mathbf{a} + \mathbf{n}. \tag{2.13}$$

If we are interested in sum-rate, instead of individual rates, we can improve the performance of the proposed method by assigning variable rates to different users. We assume that the sum-rate (rather than the individual rates) is fixed and we want to reduce the average transmitted energy. To simplify the analysis, we use the continuous approximation which has a good accuracy for high rates.

Considering continuous approximation, the sum-rate is proportional to the logarithm of the volume of the lattice with basis $\mathbf{B}$ and the average energy is proportional to the second moment of the corresponding parallelotope, which is proportional to $\sum_{i=1}^{N_r} \|\mathbf{b}_i\|^2 = \text{tr}\mathbf{BB}^{\mathtt{H}}$ (see Appendix A). The goal is to minimize the average energy while the sum-rate is fixed. We can use another lattice generated by $\mathbf{B}'$ with the same volume, where its basis vectors are scaled versions of the vectors of the basis $\mathbf{B}$, according to different rates for different users. Therefore, we can use $\mathbf{B}' = \mathbf{BD}$ instead of $\mathbf{B}$ (where $\mathbf{D}$ is a unit determinant $N_r \times N_r$ diagonal matrix which does not change the volume of the lattice). For a given reduced basis $\mathbf{B}$, the product of the squared norms of the new basis vectors is constant:

$$\begin{aligned}\|\mathbf{b}'_1\|^2\|\mathbf{b}'_2\|^2...\|\mathbf{b}'_{N_r}\|^2 &= (\|\mathbf{b}_1\|^2\|\mathbf{b}_2\|^2...\|\mathbf{b}_{N_r}\|^2)\det\mathbf{D} \\ &= \|\mathbf{b}_1\|^2\|\mathbf{b}_2\|^2...\|\mathbf{b}_{N_r}\|^2 = \text{const.}\end{aligned} \tag{2.14}$$

The average energy corresponding to the new lattice basis should be minimized. When we use the modified basis $\mathbf{B}'$ instead of $\mathbf{B}$, the average energy is proportional to $\sum_{i=1}^{N_r}\|\mathbf{b}'_i\|^2 = \text{tr}\mathbf{B}'\mathbf{B}'^{\text{H}}$ (see Appendix A). According to the arithmetic-geometric mean inequality, $\sum_{i=1}^{N_r}\|\mathbf{b}'_i\|^2 = \text{tr}\mathbf{B}'\mathbf{B}'^{\text{H}}$ is minimized iff

$$\|\mathbf{b}'_1\| = \|\mathbf{b}'_2\| = ... = \|\mathbf{b}'_{N_r}\|. \tag{2.15}$$

Therefore,

$$\min\text{tr}\mathbf{B}'\mathbf{B}'^{\text{H}} = N_r\left(\|\mathbf{b}'_1\|^2\|\mathbf{b}'_2\|^2...\|\mathbf{b}'_{N_r}\|^2\right)^{\frac{1}{N_r}} = N_r\left(\|\mathbf{b}_1\|^2\|\mathbf{b}_2\|^2...\|\mathbf{b}_{N_r}\|^2\right)^{\frac{1}{N_r}} \tag{2.16}$$

Having the matrix $\mathbf{B}$, the columns of matrix $\mathbf{B}'$ can be found using the equation (2.15) and $\text{tr}\mathbf{B}'\mathbf{B}'^{\text{H}}$ can be obtained by (2.16). Now, for the selection of the reduced basis $\mathbf{B}$, we should find $\mathbf{B}$ such that $\|\mathbf{b}_1\|^2\|\mathbf{b}_2\|^2...\|\mathbf{b}_{N_r}\|^2$ is minimized. Because $\det\mathbf{B}^{\text{H}}\mathbf{B} = \det\left(\mathbf{H}^+\right)^{\text{H}}\mathbf{H}^+$ is given, the best basis reduction is the reduction which maximizes $\frac{\|\mathbf{b}_1\|^2\|\mathbf{b}_2\|^2\cdot...\cdot\|\mathbf{b}_{N_r}\|^2}{|\det\mathbf{B}^{\text{H}}\mathbf{B}|}$, or in other words, minimizes the orthogonality defect.

In practice, we use discrete values for the rate, and sometimes, we should assign the rate zero to some users (when their channel is very bad). In this case, for the rate assignment for other users, we use the lattice reduction on the corresponding sublattice. It should be noted that the average transmit power is fixed per channel realization and no long-run averaging is considered, and no long-run power allocation is used.

2.4 Diversity and Outage Probability

In this section, we consider the asymptotic behavior ($\rho \longrightarrow \infty$) of the symbol error rate (SER) for the proposed method and the perturbation technique. We show that for both of these methods, the asymptotic slope of the SER curve is equal to the number of transmit antennas. By considering the outage probability of a fixed-rate MIMO broadcast system, we will show that for the SER curve in high SNR, the slope obtained by the proposed method has the largest achievable value. Also, we analyze the asymptotic behavior of the outage probability for the case of fixed sum-rate. We show that in this case, the slope of the corresponding curve is equal to the product of the number of transmit antennas and the number of single-antenna users.

2.4.1 Fixed-rate users

When we have the Channel-State Information (CSI) at the transmitter, without any assumption on the transmission rates, the outage probability is not meaningful. However, when we consider given rates $R_1, ..., R_{N_r}$ for different users, we can define the outage probability P_{out} as the probability that the point $(R_1, ..., R_{N_r})$ is outside the capacity region.

Theorem 2.1. *For a MIMO broadcast system with N_t transmit antennas, N_r single-antenna receivers ($N_t \geq N_r$), and given rates $R_1, ..., R_{N_r}$,*

$$\lim_{\rho \to \infty} \frac{-\log P_{out}}{\log \rho} \leq N_t. \tag{2.17}$$

Proof. Define P_{out1} as the probability that the capacity of the point-to-point system corresponding to the first user (consisting of N_t transmit antennas and one receive antenna with independent channel coefficients and CSI at the transmitter) is less than R_1:

$$P_{out1} = \Pr\{\log\left(1 + \rho\|\mathbf{h}_1\|^2\right) \leq R_1\} \tag{2.18}$$

where $\mathbf{h}_1$ is the vector defined by the first row of $\mathbf{H}$. Note that the entries of $\mathbf{h}_1$ have iid complex Gaussian distribution with unit variance. Thus, its square norm has a chi square distribution. We have,

$$\Pr\left\{\log\left(2\rho\|\mathbf{h}_1\|^2\right) \leq R_1\right\} \tag{2.19}$$

$$= \Pr\left\{\|\mathbf{h}_1\|^2 \leq \frac{2^{R_1}}{2\rho}\right\} \tag{2.20}$$

$$= \int_0^{\frac{2^{R_1}}{2\rho}} f_{\|\mathbf{h}_1\|^2}(x)\, dx \tag{2.21}$$

$$= \int_0^{\frac{2^{R_1}}{2\rho}} \frac{1}{(N_t-1)!} x^{N_t-1} e^{-x}\, dx \tag{2.22}$$

We are interested in the large values of ρ. For $\rho > 2^{R_1-1}$,

$$\int_0^{\frac{2^{R_1}}{2\rho}} \frac{1}{(N_t-1)!} x^{N_t-1} e^{-x}\, dx \geq \int_0^{\frac{2^{R_1}}{2\rho}} \frac{1}{(N_t-1)!} x^{N_t-1} e^{-1}\, dx \tag{2.23}$$

$$= \frac{e^{-1}}{(N_t-1)!} \int_0^{\frac{2^{R_1}}{2\rho}} x^{N_t-1}\, dx \tag{2.24}$$

$$= \frac{2^{N_t R_1} c}{\rho^{N_t}} \tag{2.25}$$

where $c = \frac{e^{-1}}{2^{N_t} N_t!}$ is a constant number. Now,

$$\log\left(1 + \rho\|\mathbf{h}_1\|^2\right) \leq \log\left(2\rho\|\mathbf{h}_1\|^2\right) \quad \text{for } \rho > \frac{1}{\|\mathbf{h}_1\|^2} \tag{2.26}$$

$$\Longrightarrow \lim_{\rho\to\infty} \frac{-\log \Pr\{\log(1 + \rho\|\mathbf{h}_1\|^2) \leq R_1\}}{\log \rho} \tag{2.27}$$

$$\leq \lim_{\rho\to\infty} \frac{-\log \Pr\{\log(2\rho\|\mathbf{h}_1\|^2) \leq R_1\}}{\log \rho} \tag{2.28}$$

$$\leq \lim_{\rho\to\infty} \frac{-\log \frac{2^{N_t R_1} c}{\rho^{N_t}}}{\log \rho} = N_t \tag{2.29}$$

$$\Longrightarrow \lim_{\rho\to\infty} \frac{-\log P_{out1}}{\log \rho} \leq N_t. \tag{2.30}$$

According to the definition of P_{out1}, $P_{out} \geq P_{out1}$. Therefore,

$$\lim_{\rho\to\infty} \frac{-\log P_{out}}{\log \rho} \leq \lim_{\rho\to\infty} \frac{-\log P_{out1}}{\log \rho} \leq N_t. \tag{2.31}$$

□

We can define the diversity gain of a MIMO broadcast constellation or its *precoding diversity* as $\lim_{\rho\to\infty} \dfrac{-\log P_e}{\log \rho}$ where P_e is the probability of error. Similar to [96, lemma 5], we can bound the precoding diversity by $\lim_{\rho\to\infty} \dfrac{-\log P_{out}}{\log \rho}$. Thus, based on theorem 2.1, the maximum achievable diversity is N_t.

We show that the proposed method (based on lattice-basis reduction) achieves the maximum precoding diversity. To prove this, in lemma 2.1 and lemma 2.2, we relate the length of the largest vector of the reduced basis $\mathbf{B}$ to $d_{\mathbf{H}^{\mathtt{H}}}$ (the minimum distance of the lattice generated by $\mathbf{H}^{\mathtt{H}}$). In lemma 2.3, we bound the probability that $d_{\mathbf{H}^{\mathtt{H}}}$ is too small. Finally, in theorem 2.2, we prove the main result by relating the minimum distance of the receive constellation to the length of the largest vector of the reduced basis $\mathbf{B}$, and combining the bounds on the probability that $d_{\mathbf{H}^{\mathtt{H}}}$ is too small, and the probability that the noise vector is too large.

Lemma 2.1. *Consider* $\mathbf{B} = [\mathbf{b}_1 ... \mathbf{b}_M]$ *as an* $N \times M$ *matrix, with the orthogonality defect* δ*, and* $\mathbf{B}^{-\mathtt{H}} = [\mathbf{a}_1 ... \mathbf{a}_M]$ *as the inverse of its Hermitian (or its pseudo-inverse if* $M < N$*). Then,*

$$\max\{\|\mathbf{b}_1\|, ..., \|\mathbf{b}_M\|\} \leq \frac{\sqrt{\delta}}{\min\{\|\mathbf{a}_1\|, ..., \|\mathbf{a}_M\|\}} \tag{2.32}$$

and

$$\max\{\|\mathbf{a}_1\|, ..., \|\mathbf{a}_M\|\} \leq \frac{\sqrt{\delta}}{\min\{\|\mathbf{b}_1\|, ..., \|\mathbf{b}_M\|\}}. \tag{2.33}$$

Proof. Consider $\mathbf{b}_i$ as an arbitrary column of $\mathbf{B}$. The vector $\mathbf{b}_i$ can be written as $\mathbf{b}'_i + \sum_{i \neq j} c_{i,j} \mathbf{b}_j$, where $\mathbf{b}'_i$ is orthogonal to $\mathbf{b}_j$ for $i \neq j$. Now, $[\mathbf{b}_1 ... \mathbf{b}_{i-1} \mathbf{b}'_i \mathbf{b}_{i+1} ... \mathbf{b}_M]$ can be written as $\mathbf{BP}$ where $\mathbf{P}$ is a unit-determinant $M \times M$ matrix (a column operation matrix):

$$\|\mathbf{b}_1\|^2 ... \|\mathbf{b}_{i-1}\|^2 . \|\mathbf{b}_i\|^2 . \|\mathbf{b}_{i+1}\|^2 ... \|\mathbf{b}_M\|^2 \tag{2.34}$$

$$= \delta \det \mathbf{B}^{\mathtt{H}} \mathbf{B} = \delta \det \mathbf{P}^{\mathtt{H}} \mathbf{B}^{\mathtt{H}} \mathbf{B} \mathbf{P} \tag{2.35}$$

$$= \delta \det \left([\mathbf{b}_1 ... \mathbf{b}_{i-1} \mathbf{b}'_i \mathbf{b}_{i+1} ... \mathbf{b}_M]^{\mathtt{H}} [\mathbf{b}_1 ... \mathbf{b}_{i-1} \mathbf{b}'_i \mathbf{b}_{i+1} ... \mathbf{b}_M] \right). \tag{2.36}$$

According to the Hadamard theorem:

$$\det\left([\mathbf{b}_1...\mathbf{b}_{i-1}\mathbf{b}'_i\mathbf{b}_{i+1}...\mathbf{b}_M]^{\mathtt{H}}[\mathbf{b}_1...\mathbf{b}_{i-1}\mathbf{b}'_i\mathbf{b}_{i+1}...\mathbf{b}_M]\right) \leq \tag{2.37}$$

$$\|\mathbf{b}_1\|^2...\|\mathbf{b}_{i-1}\|^2.\|\mathbf{b}'_i\|^2.\|\mathbf{b}_{i+1}\|^2...\|\mathbf{b}_M\|^2. \tag{2.38}$$

Therefore,

$$\|\mathbf{b}_1\|^2...\|\mathbf{b}_{i-1}\|^2.\|\mathbf{b}_i\|^2.\|\mathbf{b}_{i+1}\|^2...\|\mathbf{b}_M\|^2 \leq \delta\|\mathbf{b}_1\|^2...\|\mathbf{b}_{i-1}\|^2.\|\mathbf{b}'_i\|^2.\|\mathbf{b}_{i+1}\|^2...\|\mathbf{b}_M\|^2 \tag{2.39}$$

$$\Longrightarrow \|\mathbf{b}_i\| \leq \sqrt{\delta}\|\mathbf{b}'_i\|. \tag{2.40}$$

Also, $\mathbf{B}^+\mathbf{B} = \mathbf{I}$ results in $<\mathbf{a}_i, \mathbf{b}_i> = 1$ and $<\mathbf{a}_i, \mathbf{b}_j> = 0$ for $i \neq j$. Therefore,

$$1 \; = \; <\mathbf{a}_i, \mathbf{b}_i> \; = \; <\mathbf{a}_i, (\mathbf{b}'_i + \sum_{i \neq j} c_{i,j}\mathbf{b}_j)> \; = \; <\mathbf{a}_i, \mathbf{b}'_i> \tag{2.41}$$

Now, $\mathbf{a}_i$ and $\mathbf{b}'_i$, both are orthogonal to the $(M-1)$-dimensional subspace generated by the vectors $\mathbf{b}_j$ $(j \neq i)$. Thus,

$$1 = \; <\mathbf{a}_i, \mathbf{b}'_i> \; = \; \|\mathbf{a}_i\|.\|\mathbf{b}'_i\| \geq \|\mathbf{a}_i\|.\frac{\|\mathbf{b}_i\|}{\sqrt{\delta}} \tag{2.42}$$

$$\Longrightarrow 1 \geq \|\mathbf{b}_i\|.\frac{\|\mathbf{a}_i\|}{\sqrt{\delta}} \tag{2.43}$$

$$\Longrightarrow \|\mathbf{b}_i\| \leq \frac{\sqrt{\delta}}{\|\mathbf{a}_i\|} \tag{2.44}$$

The above relation is valid for every i, $1 \leq i \leq M$. Without loss of generality, we can assume that $\max\{\|\mathbf{b}_1\|, ..., \|\mathbf{b}_M\|\} = \|\mathbf{b}_k\|$:

$$\max\{\|\mathbf{b}_1\|, ..., \|\mathbf{b}_M\|\} = \|\mathbf{b}_k\| \leq \frac{\sqrt{\delta}}{\|\mathbf{a}_k\|} \tag{2.45}$$

$$\leq \frac{\sqrt{\delta}}{\min\{\|\mathbf{a}_1\|, ..., \|\mathbf{a}_M\|\}}. \tag{2.46}$$

Similarly, by using (2.44), we can also obtain the following inequality:

$$\max\{\|\mathbf{a}_1\|, ..., \|\mathbf{a}_M\|\} \leq \frac{\sqrt{\delta}}{\min\{\|\mathbf{b}_1\|, ..., \|\mathbf{b}_M\|\}}. \tag{2.47}$$

□

Lemma 2.2. *Consider* $\mathbf{B} = [\mathbf{b}_1 ... \mathbf{b}_M]$ *as an LLL-reduced basis for the lattice generated by* $\mathbf{H}^+$ *and* $d_{\mathbf{H}^{\mathtt{H}}}$ *as the minimum distance of the lattice generated by* $\mathbf{H}^{\mathtt{H}}$*. Then, there is a constant* α_M *(independent of* $\mathbf{H}$*) such that*

$$\max\{\|\mathbf{b}_1\|, ..., \|\mathbf{b}_M\|\} \leq \frac{\alpha_M}{d_{\mathbf{H}^{\mathtt{H}}}}. \tag{2.48}$$

Proof. According to the theorem 1.1,

$$\sqrt{\delta} \leq 2^{M(M-1)}. \tag{2.49}$$

Consider $\mathbf{B}^{-\mathtt{H}} = [\mathbf{a}_1, ..., \mathbf{a}_M]$. By using lemma 2.1 and (2.49),

$$\max\{\|\mathbf{b}_1\|, ..., \|\mathbf{b}_M\|\} \leq \frac{\sqrt{\delta}}{\min\{\|\mathbf{a}_1\|, ..., \|\mathbf{a}_M\|\}} \leq \frac{2^{M(M-1)}}{\min\{\|\mathbf{a}_1\|, ..., \|\mathbf{a}_M\|\}} \tag{2.50}$$

The basis $\mathbf{B}$ can be written as $\mathbf{B} = \mathbf{H}^+\mathbf{U}$ for some unimodular matrix $\mathbf{U}$:

$$\mathbf{B}^{-\mathtt{H}} = ((\mathbf{H}^+\mathbf{U})^{\mathtt{H}})^+ = (\mathbf{U}^{\mathtt{H}}\mathbf{H}^{-\mathtt{H}})^+ = \mathbf{H}^{\mathtt{H}}\mathbf{U}^{-\mathtt{H}}. \tag{2.51}$$

Noting that $\mathbf{U}^{-\mathtt{H}}$ is unimodular, $\mathbf{B}^{-\mathtt{H}} = [\mathbf{a}_1, ..., \mathbf{a}_M]$ is another basis for the lattice generated by $\mathbf{H}^{\mathtt{H}}$. Therefore, the vectors $\mathbf{a}_1, ..., \mathbf{a}_M$ are vectors from the lattice generated by $\mathbf{H}^{\mathtt{H}}$, and therefore, the length of each of them is at least $d_{\mathbf{H}^{\mathtt{H}}}$:

$$\|\mathbf{a}_i\| \geq d_{\mathbf{H}^{\mathtt{H}}} \quad \text{for} \ \ 1 \leq i \leq M \tag{2.52}$$

$$\Longrightarrow \min\{\|\mathbf{a}_1\|, ..., \|\mathbf{a}_M\|\} \geq d_{\mathbf{H}^{\text{H}}} \tag{2.53}$$

$$(2.50) \text{ and } (2.53) \Longrightarrow \max\{\|\mathbf{b}_1\|, ..., \|\mathbf{b}_M\|\} \leq \frac{2^{M(M-1)}}{d_{\mathbf{H}^{\text{H}}}}. \tag{2.54}$$

□

Lemma 2.3. *Assume that the entries of the $N \times M$ matrix $\mathbf{H}$ has independent complex Gaussian distribution with zero mean and unit variance and consider $d_{\mathbf{H}}$ as the minimum distance of the lattice generated by $\mathbf{H}$. Then, there is a constant $\beta_{N,M}$ such that*

$$\Pr\{d_{\mathbf{H}} \leq \varepsilon\} \leq \begin{cases} \beta_{N,M}\varepsilon^{2N} & \textit{for } M < N \\ \beta_{N,N}\varepsilon^{2N}.\max\left\{-(\ln\varepsilon)^{N+1}, 1\right\} & \textit{for } M = N \end{cases}. \tag{2.55}$$

Proof: See Appendix B.

Theorem 2.2. *For a MIMO broadcast system with N_t transmit antennas and N_r single-antenna receivers ($N_t \geq N_r$) and fixed rates $R_1, ..., R_{N_r}$, using the lattice-basis-reduction method,*

$$\lim_{\rho\to\infty} \frac{-\log P_e}{\log\rho} = N_t. \tag{2.56}$$

Proof. Consider $\mathbf{B} = [\mathbf{b}_1...\mathbf{b}_{N_r}]$ as the LLL-reduced basis for the lattice generated by $\mathbf{H}^+$. Each transmitted vector $\mathbf{s}$ is inside the parallelotope, generated by $r_1\mathbf{b}_1, ..., r_{N_r}\mathbf{b}_{N_r}$ (where $r_1, ..., r_{N_r}$ are constant values determined by the rates of the users). Thus, every transmitted vector $\mathbf{s}$ can be written as

$$\mathbf{s} = t_1\mathbf{b}_1 + ... + t_{N_r}\mathbf{b}_{N_r}, \quad \frac{-r_i}{2} \leq t_i \leq \frac{r_i}{2}. \tag{2.57}$$

For each of the transmitted vectors, the energy is

$$P = \|\mathbf{s}\|^2 = \|t_1\mathbf{b}_1 + ... + t_{N_r}\mathbf{b}_{N_r}\|^2 \tag{2.58}$$

$$\Longrightarrow P \leq (\|t_1\mathbf{b}_1\| + ... + \|t_{N_r}\mathbf{b}_{N_r}\|)^2 \tag{2.59}$$

$$\Longrightarrow P \leq \left(\frac{r_1}{2}\|\mathbf{b}_1\| + ... + \frac{r_{N_r}}{2}\|\mathbf{b}_{N_r}\|\right)^2. \tag{2.60}$$

Thus, the average transmitted energy is

$$P_{av} = \mathrm{E}(P) \leq N_r^2 \left(\max\left\{\frac{r_1}{2}\|\mathbf{b}_1\|, ..., \frac{r_{N_r}}{2}\|\mathbf{b}_{N_r}\|\right\}\right)^2 \leq c_1.(\max\{\|\mathbf{b}_1\|^2, ..., \|\mathbf{b}_{N_r}\|^2\}) \tag{2.61}$$

where $c_1 = \frac{N_r^2}{4}\max\left\{r_1^2, ..., r_{N_r}^2\right\}$. The received signals (without the effect of noise) are points from the $\mathbb{Z}^{2N_r}$ lattice. If we consider the normalized system (by scaling the signals such that the average transmitted energy becomes equal to one),

$$d^2 = \frac{1}{P_{av}} \geq \frac{1}{c_1.(\max\{\|\mathbf{b}_1\|^2, ..., \|\mathbf{b}_{N_r}\|^2\})} \tag{2.62}$$

is the squared distance between the received signal points.

For the normalized system, $\frac{1}{\rho}$ is the energy of the noise at each receiver and $\frac{1}{2\rho}$ is the energy of the noise per each real dimension. Using (2.62), for any positive number γ,

$$\begin{aligned} & \Pr\left\{d^2 \leq \frac{\gamma}{\rho}\right\} \\ \leq & \Pr\left\{\frac{1}{c_1\max\{\|\mathbf{b}_1\|^2, ..., \|\mathbf{b}_{N_r}\|^2\}} \leq \frac{\gamma}{\rho}\right\} \end{aligned} \tag{2.63}$$

Using lemma 2.2,

$$\max\{\|\mathbf{b}_1\|, ..., \|\mathbf{b}_{N_r}\|\} \leq \frac{\alpha_{N_r}}{d_{\mathbf{H}^{\mathrm{H}}}} \tag{2.64}$$

$$(2.63), (2.64) \Longrightarrow \Pr\left\{d^2 \le \frac{\gamma}{\rho}\right\} \le \Pr\left\{\frac{d^2_{\mathbf{H}^{\mathtt{H}}}}{c_1\alpha^2_{N_r}} \le \frac{\gamma}{\rho}\right\} = \Pr\left\{d^2_{\mathbf{H}^{\mathtt{H}}} \le \frac{\gamma c_1 \alpha^2_{N_r}}{\rho}\right\} \quad (2.65)$$

The $N_t \times N_r$ matrix $\mathbf{H}^{\mathtt{H}}$ has independent complex Gaussian distribution with zero mean and unit variance. Therefore, by using lemma 2.3 (considering $M = N_r$ and $N = N_t$), we can bound the probability that $d_{\mathbf{H}^{\mathtt{H}}}$ is too small.

Case 1, $N_r = N_t$:

$$\Pr\left\{d^2 \le \frac{\gamma}{\rho}\right\} \le \Pr\left\{d^2_{\mathbf{H}^{\mathtt{H}}} \le \frac{\gamma c_1 \alpha^2_{N_t}}{\rho}\right\} \quad (2.66)$$

$$\le \beta_{N_t,N_t}\left(\frac{\gamma c_1 \alpha^2_{N_t}}{\rho}\right)^{N_t} \max\left\{\left(-\frac{1}{2}\ln\frac{\gamma c_1 \alpha^2_{N_t}}{\rho}\right)^{N_t+1}, 1\right\} \quad (2.67)$$

$$\le \beta_{N_t,N_t}\left(\frac{\gamma c_1 \alpha^2_{N_t}}{\rho}\right)^{N_t} \max\left\{(\ln\rho)^{N_t+1}, 1\right\} \quad \text{for } \gamma > 1 \text{ and } \rho > \frac{1}{c_1\alpha^2_{N_t}} \quad (2.68)$$

$$\le \frac{c_2\gamma^{N_t}}{\rho^{N_t}}(\ln\rho)^{N_t+1} \quad \text{for } \gamma > 1 \text{ and } \rho > \max\left\{\frac{1}{c_1\alpha^2_{N_t}}, e\right\} \quad (2.69)$$

where c_2 is a constant number and e is the Euler number.

If the magnitude of the noise component in each real dimension is less than $\frac{1}{2}d$, the transmitted data will be decoded correctly. Thus, we can bound the probability of error by the probability that $|w_i|^2$ is greater than $\frac{1}{4}d^2$ for at least one i, $1 \le i \le 2N_t$. Therefore, using the union bound,

$$P_e \leq 2N_t \left(\Pr \left\{ |w_1|^2 \geq \frac{1}{4} d^2 \right\} \right) \tag{2.70}$$

$$\begin{aligned} &= 2N_t \left(\Pr \left\{ d^2 \leq \frac{4}{\rho} \right\} . \Pr \left\{ |w_1|^2 \geq \frac{1}{4} d^2 \middle| d^2 \leq \frac{4}{\rho} \right\} \right. \\ &+ \Pr \left\{ \frac{4}{\rho} \leq d^2 \leq \frac{8}{\rho} \right\} . \Pr \left\{ |w_1|^2 \geq \frac{1}{4} d^2 \middle| \frac{4}{\rho} \leq d^2 \leq \frac{8}{\rho} \right\} \\ &\left. + \Pr \left\{ \frac{8}{\rho} \leq d^2 \leq \frac{16}{\rho} \right\} . \Pr \left\{ |w_1|^2 \geq \frac{1}{4} d^2 \middle| \frac{8}{\rho} \leq d^2 \leq \frac{16}{\rho} \right\} + ... \right) \end{aligned} \tag{2.71}$$

$$\begin{aligned} &\leq 2N_t \left(\Pr \left\{ d^2 \leq \frac{4}{\rho} \right\} + \Pr \left\{ \frac{4}{\rho} \leq d^2 \leq \frac{8}{\rho} \right\} . \Pr \left\{ |w_1|^2 \geq \frac{1}{4} . \frac{4}{\rho} \right\} \right. \\ &\left. + \Pr \left\{ \frac{8}{\rho} \leq d^2 \leq \frac{16}{\rho} \right\} . \Pr \left\{ |w_1|^2 \geq \frac{1}{4} . \frac{8}{\rho} \right\} + ... \right) \end{aligned} \tag{2.72}$$

$$\begin{aligned} &\leq 2N_t \left(\Pr \left\{ d^2 \leq \frac{4}{\rho} \right\} + \Pr \left\{ d^2 \leq \frac{8}{\rho} \right\} . \Pr \left\{ |w_1|^2 \geq \frac{1}{\rho} \right\} \right. \\ &\left. + \Pr \left\{ d^2 \leq \frac{16}{\rho} \right\} . \Pr \left\{ |w_1|^2 \geq \frac{2}{\rho} \right\} + ... \right) \end{aligned} \tag{2.73}$$

For the product terms in (2.73), we can bound the first part by (2.69). To bound the second part, we note that w_1 has real Gaussian distribution with variance $\frac{1}{2\rho}$. Therefore,

$$\Pr \left\{ |w_1|^2 \geq \frac{\theta}{\rho} \right\} = 2Q(\sqrt{2\theta}) \leq e^{-\theta} \tag{2.74}$$

Now, for $\rho > \max \left\{ \frac{1}{c_1 \alpha_{N_t}^2}, e \right\}$,

$$(2.69),\ (2.73) \text{ and } (2.74) \Longrightarrow P_e \leq 2N_t \left(\Pr \left\{ |w_1|^2 \geq \frac{1}{4} d^2 \right\} \right) \tag{2.75}$$

$$\leq 2N_t \left(\frac{4^{N_t} c_2}{\rho^{N_t}} (\ln \rho)^{N_t+1} + \sum_{i=0}^{\infty} \frac{2^{N_t(i+3)} c_2}{\rho^{N_t}} (\ln \rho)^{N_t+1} e^{-2^i} \right) \tag{2.76}$$

$$\leq \frac{(\ln \rho)^{N_t+1}}{\rho^{N_t}} . c_2 . 2N_t \left(4^{N_t} + \sum_{i=0}^{\infty} 2^{N_t(i+3)} e^{-2^i} \right) \tag{2.77}$$

$$\leq \frac{c_3 (\ln \rho)^{N_t+1}}{\rho^{N_t}} \tag{2.78}$$

where c_3 is a constant number which only depends on N_t. Thus,

$$\lim_{\rho \to \infty} \frac{-\log P_e}{\log \rho} \geq \lim_{\rho \to \infty} \frac{N_t \log \rho - \log(\ln \rho)^{N_t+1} - \log c_3}{\log \rho} = N_t. \tag{2.79}$$

According to Theorem 2.1, this limit cannot be greater than N_t. Therefore,

$$\lim_{\rho \to \infty} \frac{-\log P_e}{\log \rho} = N_t. \tag{2.80}$$

Case 2, $N_r < N_t$:

For the $N_t \times N_r$ matrix $\mathbf{H}^{\mathtt{H}}$, we use the first inequality in lemma 2.3 (by considering $M = N_r$ and $N = N_t$) to bound the probability that $d_{\mathbf{H}^{\mathtt{H}}}$ is too small:

$$\Pr\left\{ d^2 \leq \frac{\gamma}{\rho} \right\} \leq \Pr\left\{ d_{\mathbf{H}^{\mathtt{H}}}^2 \leq \frac{\gamma c_1 \alpha_{N_r}^2}{\rho} \right\} \tag{2.81}$$

$$\leq \beta_{N_t, N_r} \left(\frac{\gamma c_1 \alpha_{N_r}^2}{\rho} \right)^{N_t} \tag{2.82}$$

$$\leq \beta_{N_t,N_r} \left(\frac{\gamma c_1 \alpha_{N_r}^2}{\rho} \right)^{N_t} \quad \text{for } \gamma > 1 \text{ and } \rho > \frac{1}{c_1 \alpha_{N_r}^2} \tag{2.83}$$

$$\leq \frac{c_2 \gamma^{N_t}}{\rho^{N_t}} \quad \text{for } \gamma > 1 \text{ and } \rho > \frac{1}{c_1 \alpha_{N_r}^2} \tag{2.84}$$

.

The rest of proof is similar to the case 1.

□

Corollary 2.1. *Perturbation technique achieves the maximum precoding diversity in fixed-rate MIMO broadcast systems.*

Proof. In the perturbation technique, for the transmission of each data vector $\mathbf{u}$, among the set $\left\{\mathbf{H}^{+}(\mathbf{u} + a\mathbf{l}) \| \mathbf{l} \in \mathbb{Z}^{2N_r}\right\}$, the nearest point to the origin is chosen. The transmitted vector in the lattice-reduction-based method belongs to that set. Therefore, the energy of the transmitted signal in the lattice-reduction-based method cannot be less than the transmitted energy in the perturbation technique. Thus, the average transmitted energy for the perturbation method is at most equal to the average transmitted energy of the lattice-reduction-based method. The rest of the proof is the same as the proof of theorem 2.2. □

2.4.2 Fixed sum-rate

When the sum-rate R_{sum} is given, similar to the previous part, we can define the outage probability as the probability that the sum-capacity of the broadcast system is less than R_{sum}.

Theorem 2.3. *For a MIMO broadcast system with N_t transmit antennas, N_r single-antenna receivers, and a given sum-rate R_{sum},*

$$\lim_{\rho\to\infty} \frac{-\log P_{out}}{\log \rho} \leq N_t N_r. \tag{2.85}$$

Proof. For any channel matrix $\mathbf{H}$, we have [87]

$$C_{sum} = \sup_{\mathbf{D}} \log |\mathbf{I}_{N_r} + \rho \mathbf{H}^{\mathrm{H}} \mathbf{D} \mathbf{H}| \tag{2.86}$$

where $\mathbf{D}$ is a diagonal matrix with non-negative elements and unit trace. Also, [48]

$$|2\rho \mathbf{H}^{\mathrm{H}} \mathbf{D} \mathbf{H}| \leq \frac{(2\rho \mathrm{tr} \mathbf{H}^{\mathrm{H}} \mathbf{D} \mathbf{H})^{N_r}}{N_r^{N_r}} = \frac{(2\rho \mathrm{tr} \mathbf{H}^{\mathrm{H}} \mathbf{H})^{N_r}}{N_r^{N_r}}. \tag{2.87}$$

The entries of $\mathbf{H}$ have iid complex Gaussian distribution with unit variance. Thus $\mathrm{tr}\rho \mathbf{H}^{\mathrm{H}} \mathbf{H}$ is equal to the square norm of an $N_t N_r$-dimensional complex Gaussian vector and has a chi square distribution with $2N_t N_r$ degrees of freedom. Thus, we have (similar to the equations 2.19-2.25, in the proof of theorem 2.1),

$$\Pr\left\{ \log \frac{(2\rho \mathrm{tr} \mathbf{H}^{\mathrm{H}} \mathbf{H})^{N_r}}{N_r^{N_r}} \leq R_{sum} \right\} \tag{2.88}$$

$$= \Pr\left\{ \mathrm{tr} \mathbf{H}^{\mathrm{H}} \mathbf{H} \leq \frac{2^{\frac{R_{sum}}{N_r}} N_r}{2\rho} \right\} \tag{2.89}$$

$$\geq \frac{2^{N_t R_{sum}} N_r^{N_t N_r} c}{\rho^{N_t N_r}} \qquad (\text{for } \rho > \frac{2^{\frac{R_{sum}}{N_r}} N_r}{2}) \tag{2.90}$$

where c is a constant number. Now,

$$\lim_{\rho\to\infty} \frac{-\log P_{out}}{\log \rho} = \tag{2.91}$$

$$\lim_{\rho\to\infty} \frac{-\log \Pr\left\{\sup_{\mathbf{D}} \log |\mathbf{I}_{N_r} + \rho \mathbf{H}^{\mathrm{H}}\mathbf{D}\mathbf{H}| \leq R_{sum}\right\}}{\log \rho} \tag{2.92}$$

$$\leq \lim_{\rho\to\infty} \frac{-\log \Pr\{\sup_{\mathbf{D}} \log |2\rho \mathbf{H}^{\mathrm{H}}\mathbf{D}\mathbf{H}| \leq R_{sum}\}}{\log \rho}. \tag{2.93}$$

By using (2.87), (2.90), and (2.93):

$$\lim_{\rho\to\infty} \frac{-\log P_{out}}{\log \rho} \leq \lim_{\rho\to\infty} \frac{-\log \frac{2^{N_t R_{sum}} N_r^{N_t N_r} c}{\rho^{N_t N_r}}}{\log \rho} = N_t N_r. \tag{2.94}$$

□

The slope $N_t N_r$ for the SER curve can be easily achieved by sending to only the best user. Similar to the proof of theorem 2.1, the slope of the symbol-error rate curve is asymptotically determined by the slope of the probability that $|h_{max}|$ is smaller than a constant number, where h_{max} is the entry of $\mathbf{H}$ with maximum norm. Due to the iid complex Gaussian distribution of the entries of $\mathbf{H}$, this probability decays with the same rate as $\rho^{-N_t N_r}$, for large ρ. However, although sending to only the best user achieves the optimum slope for the SER curve, it is not an efficient transmission technique because it reduces the capacity to the order of $\log \rho$ (instead of $N_r \log \rho$).

2.5 Simulation Results

Figure 2.2 presents the simulation results for the performance of the proposed schemes, the perturbation scheme [55], and the naive channel inversion approach. The number of the transmit antennas is $N_t = 4$ and there are $N_r = 4$ single-antenna users in the system. The overall transmission rate is 8 bits per channel use, where 2 bits are assigned to each user, i.e. a QPSK constellation is assigned to each user.

By considering the slope of the curves in figure 2.2, we see that by using the proposed reduction-based schemes, we can achieve the maximum precoding diversity, with a low complexity. Also, as compared to the perturbation scheme, we have a negligible loss in the performance (about 0.2 dB). Moreover, compared to the approximated perturbation method [90], we have about 1.5 dB improvement by sending the bits, corresponding to the shift vector, at the beginning of the transmission. Without sending the shift vector, the performance of the proposed method is the same as that of the approximated perturbation method [90]. The modified perturbation method (with sending two shift bits for each user) has around 0.3 dB improvement compared to the perturbation method.

Figure 2.3 compares the regularized proposed scheme with V-BLAST modifications of Zero-Forcing and Babai approximation for the same setting. As shown in the simulation results (and also in the simulation results in [55] and [90]), the modulo-MMSE-VBLAST scheme does not achieve a precoding diversity better than zero forcing (though it has a good performance in the low SNR region). However, combining the lattice-reduction-aided (LRA) scheme with MMSE-VBLAST precoding or other schemes such as regularization improves its performance by a finite coding gain (without changing the slope of the curve of symbol-error-rate). Combining both the regularization and the shift vector can result in better performance compared to other alternatives.

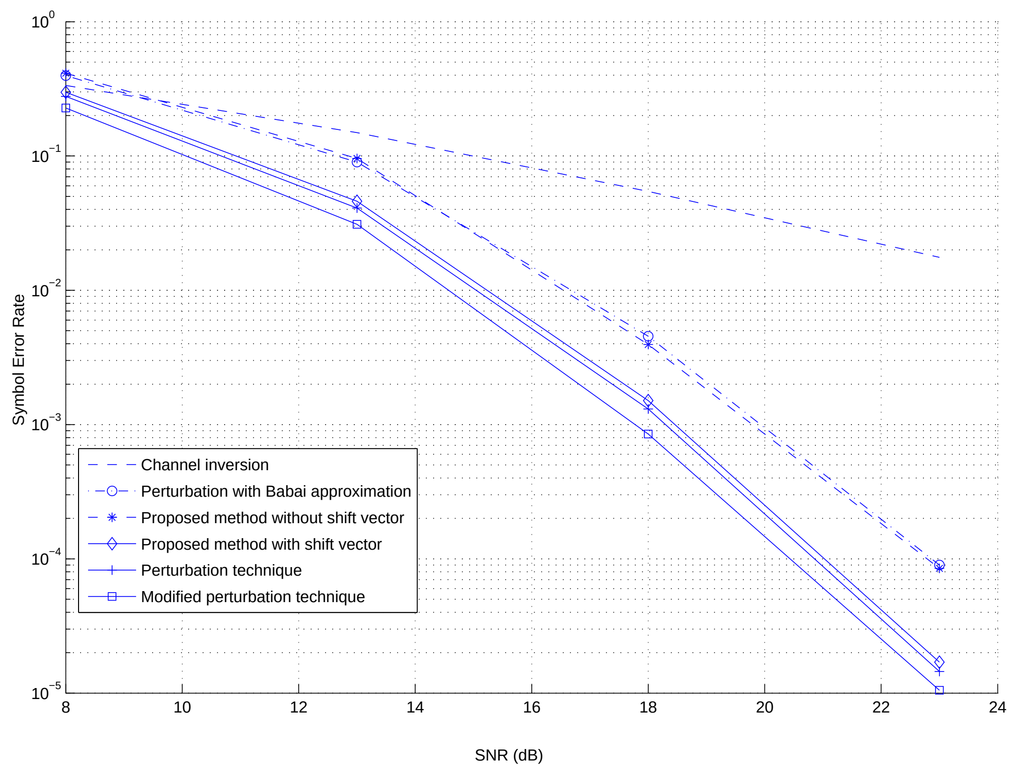

Figure 2.2: Symbol Error Rate of the proposed schemes, the perturbation scheme [55], and the naive channel inversion approach for $N_t = 4$ transmit antennas and $N_r = 4$ single-antenna receivers with the rate $R = 2$ bits per channel use per user.

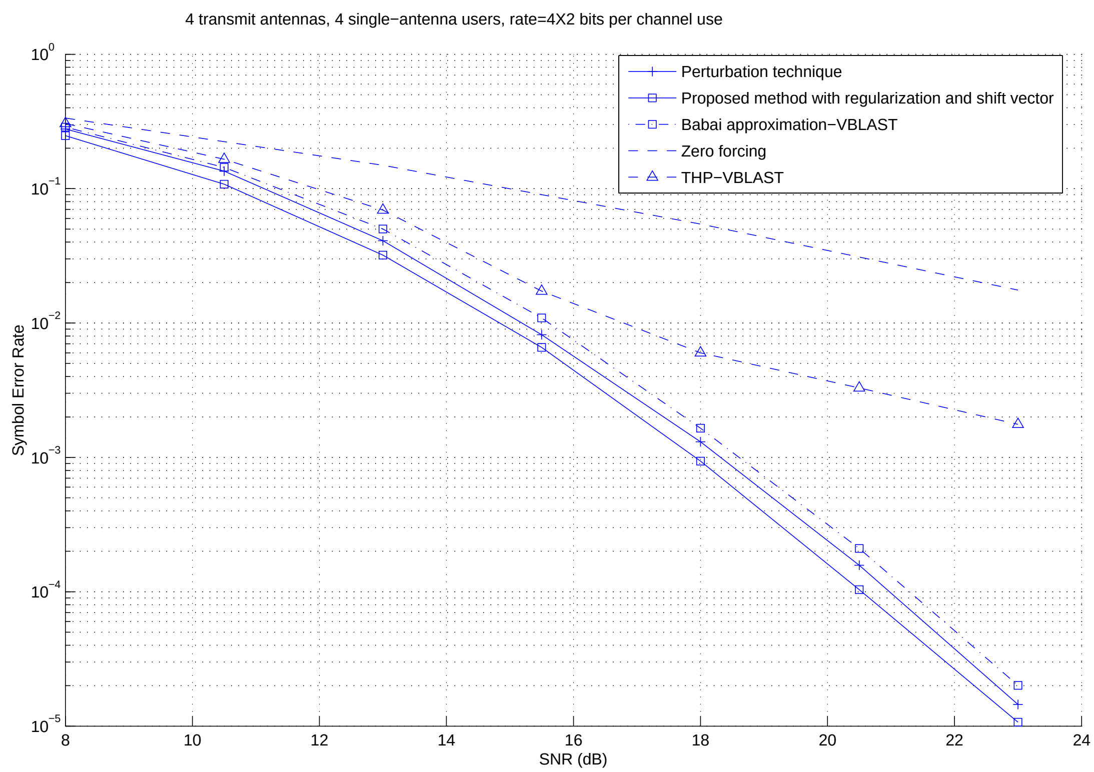

Figure 2.3: Comparison of the regularized proposed scheme with V-BLAST modifications of Zero-Forcing and Babai approximation (for $N_t = 4$ transmit antennas and $N_r = 4$ single-antenna receivers with the rate $R = 2$ bits per channel use per user.).

Figure 2.4 compares the performances of the fixed-rate and the variable-rate transmission using lattice-basis reduction for $N_t = 2$ transmit antennas and $N_r = 2$ users. In both cases, the sum-rate is 8 bits per channel use (in the case of fixed individual rates, a 16QAM constellation is assigned to each user). We see that by eliminating the equal-rate constraint, we can considerably improve the performance (especially, for high rates). In fact, the diversity gains for the equal-rate and the unequal-rate methods are, respectively, N_r and $N_t N_r$.

2.6 Conclusion

A simple scheme for communications in MIMO broadcast channels is introduced which is based on the lattice reduction technique and improves the performance of the channel inversion method. Lattice basis reduction helps us to reduce the average transmitted energy by modifying the region which includes the constellation points. Simulation results show that the performance of the proposed scheme is very close to the performance of the perturbation method. Also, it is shown that by using lattice-basis reduction, we achieve the maximum precoding diversity with a polynomial-time precoding complexity.

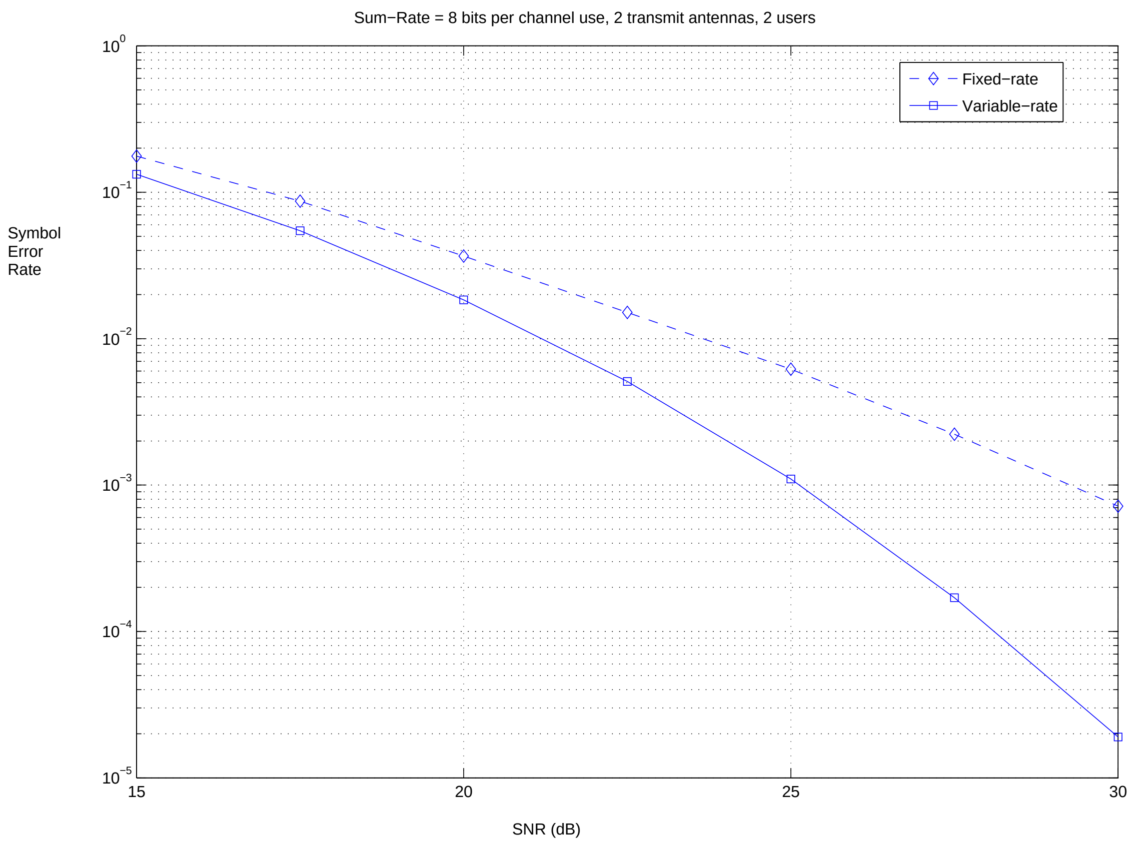

Figure 2.4: Performance comparison between the fixed-rate and the variable-rate transmission for $N_t = 2$ transmit antennas and $N_r = 2$ single-antenna receivers with sum-rate 8 bits per channel use.

Chapter 3

Lattice-reduction-based decoding in MIMO multiaccess systems

In this chapter, we consider the performance of Lattice-reduction-based decoding in MIMO multiaccess systems (or equivalently, in MIMO point-to-point systems with V-BLAST transmission). Most of the sub-optimum decoding methods for BLAST (such as nulling and cancelling, zero forcing and GDFE-type methods) cannot achieve the maximum receive diversity which is equal to the number of receive antennas. In [14], a lattice decoder is proposed for the decoding of BLAST which (according to the simulation results) achieves the maximum diversity. However, its complexity is exponential in terms of the number of antennas. In [91], [94], and [50], an approximation of lattice decoding, using the LLL lattice-basis reduction [45], is introduced which has a polynomial complexity and the simulation results show that it achieves the receive diversity. In this chapter, we give a mathematical proof for achieving the receive diversity by the LLL-aided zero-forcing decoder, which is one of the simplest forms of the lattice-reduction-aided decoders. Also, a similar proof shows that the naive lattice decoding (which discards the out-of-region decoded points) achieves the receive diversity.

In section 3.4, we complement this result by showing that for the special case of equal number of transmit and receive antennas, although the naive lattice decoding (and its LLL-aided approximation) still achieve the maximum receive diversity, their gap with the optimal ML decoding grows unboundedly with SNR.

3.1 System Model

We consider a multiple-antenna system with N_t transmit antennas and N_r receive antennas, where $N_t \leq N_r$. In a multiple-access system, we consider different transmit antennas as different users. We consider vectors $\mathbf{y} = [y_1, ..., y_{N_r}]^T$, $\mathbf{x} = [x_1, ..., x_{N_t}]^T$, $\mathbf{w} = [w_1, ..., w_{N_r}]^T$ and the $N_r \times M$ matrix $\mathbf{H}$, as the received signal, the transmitted signal, the noise vector and the channel matrix, respectively. The following matrix equation describes the channel model:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}. \tag{3.1}$$

The channel is assumed to be Rayleigh and the noise is Gaussian, i.e. the elements of $\mathbf{H}$ are i.i.d with the zero-mean unit-variance complex Gaussian distribution. Also, we have the power constraint on the transmitted signal, $\mathrm{E}|x_i|^2 = P$, where P depends on the size of the constellation. The power of the additive noise is σ^2 per antenna, i.e. $\mathrm{E}\|\mathbf{w}\|^2 = N_r\sigma^2$. Therefore, the signal to noise ratio (SNR) is defined as $\rho = \frac{N_t P}{\sigma^2}$.

In a MIMO multiple-access system or a MIMO point-to-point system with V-BLAST transmission, we send the transmitted vector $\mathbf{x}$ with independent entries from $\mathbb{Z}[i]$, the set of complex Gaussian integers. At the receiver, as the maximum-likelihood (ML) estimate of $\mathbf{x}$, a vector $\widehat{\mathbf{x}}$ should be found among the possible transmitted vectors, such that $\|\mathbf{y}-\mathbf{H}\widehat{\mathbf{x}}\|$ is minimized. For large constellations, the exact ML decoding can be very complex and practically infeasible. Therefore, we need to approximate it by a low-complexity scheme.

As a simple approximation of ML decoding, zero-forcing can be used, which selects $\widehat{\mathbf{x}}$ as the closest integer point to $\mathbf{H}^{-1}\mathbf{y}$. Although zero forcing is very simple to implement, it has a poor performance. Indeed, in zero forcing, $\mathbf{H}^{-1}\mathbf{w}$ is the effective noise, and when $\mathbf{H}$ has a small singular value, $\mathbf{H}^{-1}$ can have very large row vectors, which result in magnifying the effective noise power. To overcome this shortcoming of the zero-forcing decoder, lattice-basis reduction is used in [91], [94], and [50] to enhance the performance of zero forcing and reduce its effective noise.

We can perform two slightly different types of LLL-aided decoding:

Type I) We find $\tilde{\mathbf{x}}$ as the closest integer point to $\mathbf{B}^{\mathtt{H}}\mathbf{y}$ where the $N\times M$ matrix $\mathbf{B}$ is the reduced version of $\mathbf{H}^{-\mathtt{H}}$, i.e. $\mathbf{B}=\mathbf{H}^{-\mathtt{H}}\mathbf{U}$, where $\mathbf{U}$ is an $M\times M$ unimodular matrix (when $M<N$, we use the pseudo-inverse instead of the inverse). The transmitted vector is decoded as,

$$\widehat{\mathbf{x}}=\mathbf{U}^{-\mathtt{H}}\tilde{\mathbf{x}}.$$

In the absence of noise (when $\mathbf{w}=\mathbf{0}$),

$$\widehat{\mathbf{x}}=\mathbf{U}^{-\mathtt{H}}\tilde{\mathbf{x}}=\mathbf{U}^{-\mathtt{H}}\mathbf{B}^{\mathtt{H}}\mathbf{y}=\mathbf{U}^{-\mathtt{H}}\left(\mathbf{H}^{-\mathtt{H}}\mathbf{U}\right)^{\mathtt{H}}\mathbf{y}=\mathbf{U}^{-\mathtt{H}}\mathbf{U}^{\mathtt{H}}\mathbf{H}^{-1}\mathbf{H}\mathbf{x}=\mathbf{x}.$$

In the presence of the noise, $\mathbf{B}^{\mathrm{H}}\mathbf{w}$ can be seen as the effective noise (instead of $\mathbf{H}^{-1}\mathbf{w}$ in the traditional zero forcing).

Type II) We find $\tilde{\mathbf{x}}$ as the closest integer point to ${\mathbf{H}_{red}}^{-1}\mathbf{y}$ where $\mathbf{H}_{red}$ is the reduced version of $\mathbf{H}$ i.e. $\mathbf{H}_{red} = \mathbf{HU}$. The transmitted vector is decoded as,

$$\widehat{\mathbf{x}} = \mathbf{U}\tilde{\mathbf{x}}.$$

In the absence of noise (when $\mathbf{w} = \mathbf{0}$),

$$\widehat{\mathbf{x}} = \mathbf{U}\tilde{\mathbf{x}} = \mathbf{U}{\mathbf{H}_{red}}^{-1}\mathbf{y} = \mathbf{U}\mathbf{U}^{-1}\mathbf{H}^{-1}\mathbf{H}\mathbf{x} = \mathbf{x}$$

In the presence of the noise, ${\mathbf{H}_{red}}^{-1}\mathbf{w}$ is the effective noise.

In the previous works [91] [94] [50], the LLL-aided decoding type II has been used. We show that the type I method is more appropriate to reduce the effective noise, and indeed, has a better performance. In the next section, we present the details of the proof of our main result for the first method and show that a similar proof is valid for the second method.

3.2 Diversity of LLL-aided decoding

For MIMO systems, diversity is defined as $\lim_{\rho\to\infty} \dfrac{-\log P_e}{\log \rho}$. When there is no joint processing among the transmit antennas, the maximum achievable diversity is equal to N, the number of receive antennas [81]. To prove that LLL-aided decoding achieves a diversity order of N, we use Theorem 2.1 to bound δ, the orthogonality defect of the LLL reduction.

In the rest of this section, in the lemmas 3.1 and 3.2, we bound the error probability by the probability of an inequality on $d_{\mathbf{H}}$ (the minimum distance among the points of the lattice generated by $\mathbf{H}$) and the length of the noise vector being valid. In theorem 3.1, we prove the main result by combining the bounds on the probability that $d_{\mathbf{H}}$ is too small (given in lemma 2.3), and the probability that the noise vector is too large.

Lemma 3.1. *Consider* $\mathbf{B} = [\mathbf{b}_1 ... \mathbf{b}_{N_t}]$ *as a reduced basis (LLL) [45] for the lattice generated by* $\mathbf{H}^{-\mathtt{H}}$, $\mathbf{B}^{-\mathtt{H}} = [\mathbf{a}_1 ... \mathbf{a}_{N_t}]$, *and* δ *as the orthogonality defect of the reduction. Then, if the magnitude of the noise vector is less than* $\frac{\min\{\|\mathbf{a}_1\|, ..., \|\mathbf{a}_{N_t}\|\}}{2\sqrt{M\delta}}$, *the LLL-aided decoding method correctly decodes the transmitted signal.*

Proof. When we use the LLL-aided decoding method, we find the nearest integer point to $\mathbf{B}^{\mathtt{H}}\mathbf{y}$. We should show that this point is the same as the transmitted vector; or in other words, all the elements of $\mathbf{B}^{\mathtt{H}}\mathbf{w}$ are in the interval $(-\frac{1}{2}, \frac{1}{2})$. To prove this, we show that $\|\mathbf{B}^{\mathtt{H}}\mathbf{w}\| < \frac{1}{2}$. It is easy to show that,

$$\|\mathbf{B}^{\mathtt{H}}\mathbf{w}\| \leq \sqrt{N_t} \cdot \max\{\|\mathbf{b}_1\|, ..., \|\mathbf{b}_{N_t}\|\} \cdot \|\mathbf{w}\| \tag{3.2}$$

Now, according to (2.32),

$$\max\{\|\mathbf{b}_1\|, ..., \|\mathbf{b}_{N_t}\|\} \leq \frac{\sqrt{\delta}}{\min\{\|\mathbf{a}_1\|, ..., \|\mathbf{a}_{N_t}\|\}} \tag{3.3}$$

Therefore,

$$\|\mathbf{B}^{\mathtt{H}}\mathbf{w}\| \leq \frac{\sqrt{M\delta}.\|\mathbf{w}\|}{\min\{\|\mathbf{a}_1\|, ..., \|\mathbf{a}_{N_t}\|\}} \tag{3.4}$$

By using the assumption of the lemma,

$$\|\mathbf{B}^{\mathtt{H}}\mathbf{w}\| < \frac{\sqrt{M\delta}.\dfrac{\min\{\|\mathbf{a}_1\|, ..., \|\mathbf{a}_{N_t}\|\}}{2\sqrt{M\delta}}}{\min\{\|\mathbf{a}_1\|, ..., \|\mathbf{a}_{N_t}\|\}} \tag{3.5}$$

$$\Longrightarrow \|\mathbf{B}^{\mathtt{H}}\mathbf{w}\| < \frac{1}{2}. \tag{3.6}$$

□

Lemma 3.2. *Consider* $\mathbf{B} = [\mathbf{b}_1 ... \mathbf{b}_{N_t}]$ *as a reduced basis (LLL) [45] and* $d_{\mathbf{H}}$ *as the minimum distance of the lattice generated by* $\mathbf{H}$*, respectively. Then, there is a constant number* c_{N_t} *(independent of* $\mathbf{H}$*) such that the LLL-aided decoding method correctly decodes the transmitted signal, if the magnitude of the noise vector is less than* $c_{N_t} d_{\mathbf{H}}$*.*

Proof. For an LLL reduction,

$$\sqrt{\delta} \leq 2^{M(M-1)}. \tag{3.7}$$

Therefore, if we consider $c_{N_t} = \dfrac{2^{-1-M(M-1)}}{\sqrt{N_t}}$,

$$\|\mathbf{w}\| \leq c_{N_t} d_{\mathbf{H}} \Longrightarrow \|\mathbf{w}\| \leq \frac{1}{2\sqrt{M\delta}} d_{\mathbf{H}} \tag{3.8}$$

The basis $\mathbf{B}$ can be written as $\mathbf{B} = \mathbf{H}^{-\mathtt{H}}\mathbf{U}$ for some unimodular matrix $\mathbf{U}$:

$$\mathbf{B}^{-\mathtt{H}} = (\mathbf{H}^{-\mathtt{H}}\mathbf{U})^{-\mathtt{H}} = \mathbf{H}\mathbf{U}^{-\mathtt{H}} \tag{3.9}$$

Thus, $\mathbf{B}^{-\mathtt{H}} = [\mathbf{a}_1, ..., \mathbf{a}_{N_t}]$ is another basis for the lattice generated by $\mathbf{H}$. Therefore, $\mathbf{a}_1, ..., \mathbf{a}_{N_t}$ are vectors from the lattice generated by $\mathbf{H}$, and therefore, the length of each of them is at least $d_{\mathbf{H}}$. Therefore,

$$\|\mathbf{w}\| \leq \frac{1}{2\sqrt{M\delta}} d_{\mathbf{H}} \leq \frac{1}{2\sqrt{M\delta}} \min\{\|\mathbf{a}_1\|, ..., \|\mathbf{a}_{N_t}\|\}. \tag{3.10}$$

Thus, according to lemma 3.1, LLL-aided decoding method correctly decodes the transmitted signal. □

Theorem 3.1. *For a MIMO multi-access system (or a point-to-point MIMO system with the V-BLAST transmission) with N_t transmit antennas and N_r receive antennas, when we use the LLL lattice-aided-decoding,*

$$\lim_{\rho \to \infty} \frac{-\log P_e}{\log \rho} = N_r. \tag{3.11}$$

Proof. When $\|\mathbf{w}\| \leq c_{N_t} d_{\mathbf{H}}$, according to lemma 3, we have no decoding error. Thus,

$$P_e \leq \Pr\left\{\|\mathbf{w}\| > c_{N_t} d_{\mathbf{H}}\right\} \tag{3.12}$$

$$= \Pr\{c_{N_t}^2 d_{\mathbf{H}}^2 \leq \frac{1}{\rho}\} \cdot \Pr\left\{\|\mathbf{w}\| > c_{N_t} d_{\mathbf{H}} \middle| c_{N_t}^2 d_{\mathbf{H}}^2 \leq \frac{1}{\rho}\right\} +$$

$$\sum_{i=0}^{\infty} \Pr\{\frac{2^i}{\rho} < c_{N_t}^2 d_{\mathbf{H}}^2 \leq \frac{2^{i+1}}{\rho}\} \cdot \Pr\left\{\|\mathbf{w}\| > c_{N_t} d_{\mathbf{H}} \middle| \frac{2^i}{\rho} < c_{N_t}^2 d_{\mathbf{H}}^2 \leq \frac{2^{i+1}}{\rho}\right\} \tag{3.13}$$

$$\leq \Pr\{c_{N_t}^2 d_{\mathbf{H}}^2 \leq \frac{1}{\rho}\} +$$

$$\sum_{i=0}^{\infty} \Pr\{c_{N_t}^2 d_{\mathbf{H}}^2 \leq \frac{2^{i+1}}{\rho}\} \cdot \Pr\left\{\|\mathbf{w}\|^2 \geq \frac{2^i}{\rho}\right\} \tag{3.14}$$

The noise vector has complex Gaussian distribution with variance $\frac{N_t P}{2\rho}$ per each real dimension. Thus, by using the union bound, we can bound the second part of each product term as,

$$\Pr\left\{\|\mathbf{w}\|^2 \geq \frac{\gamma}{\rho}\right\} \leq \sum_{i=1}^{2N_r} \Pr\left\{|w_i|^2 \geq \frac{\gamma}{2N_r\rho}\right\} \leq 4N_r Q\left(\sqrt{\frac{N_t\gamma}{N_r P}}\right) \leq 4N_r e^{-\frac{N_t\gamma}{2N_r P}} \tag{3.15}$$

Also, for the first part of the product terms, we have,

$$\Pr\left\{c_{N_t}^2 d_{\mathbf{H}}^2 \leq \frac{\theta}{\rho}\right\} = \Pr\left\{d_{\mathbf{H}} \leq \sqrt{\frac{\theta}{c_{N_t}^2 \rho}}\right\} \tag{3.16}$$

Bu using lemma 2.3, we can bound (3.16) and by combining it with (3.15), we can bound (3.14).

Case 1, $N_t < N_r$:

$$(3.14) \leq \beta_{N_r,N_t}\left(\frac{1}{c_{N_t}^2\rho}\right)^{N_r} + \sum_{i=0}^{\infty}\beta_{N_r,N_t}\left(\frac{2^{i+1}}{c_{N_t}^2\rho}\right)^{N_r} \cdot 4N_r \cdot e^{-\frac{2^i N_t}{2N_r P}} \tag{3.17}$$

$$= \frac{\beta_{N_r,N_t}}{\rho^{N_r}}\left(\left(\frac{1}{c_{N_t}^2}\right)^{N_r} + \sum_{i=0}^{\infty}\left(\frac{2^{i+1}}{c_{N_t}^2}\right)^{N_r} \cdot 4N_r \cdot e^{-\frac{2^i N_t}{2N_r P}}\right) \tag{3.18}$$

$$\Longrightarrow P_e \leq \frac{c}{\rho^{N_r}} \tag{3.19}$$

where c is a constant[1]. Therefore,

$$\lim_{\rho\to\infty}\frac{-\log P_e}{\log\rho} \geq N_r. \tag{3.20}$$

Case 2, $N_t = N_r = M$:

$$(3.14) \leq \beta_{M,M}\left(\frac{1}{{c_M}^2\rho}\right)^{M}\max\left\{\left(\frac{1}{2}\ln c_M^2\rho\right)^{M+1}, 1\right\} +$$

$$\sum_{i=0}^{\infty}\beta_{M,M}\left(\frac{2^{i+1}}{c_M^2\rho}\right)^{M}\max\left\{\left(\frac{1}{2}\ln\frac{c_M^2\rho}{2^{i+1}}\right)^{M+1}, 1\right\} \cdot 4M \cdot e^{-\frac{2^i M}{2MP}} \tag{3.21}$$

[1]The terms of this series have double exponential parts which ensure its convergence (according to the ratio test).

We are interested in the large values of ρ. For $\rho > c_M^2$ and $\ln\rho > 1$,

$$(3.14) \leq \beta_{M,M}\left(\frac{1}{c_M^2\rho}\right)^M (\ln\rho)^{M+1} + \sum_{i=0}^{\infty}\beta_{M,M}\left(\frac{2^{i+1}}{c_M^2\rho}\right)^M (\ln\rho)^{M+1}\cdot 4M\cdot e^{-\frac{2^i}{2P}} \quad (3.22)$$

$$= \frac{\beta_{M,M}(\ln\rho)^{M+1}}{\rho^M}\left(\left(\frac{1}{c_M^2}\right)^M + \sum_{i=0}^{\infty}\left(\frac{2^{i+1}}{c_M^2}\right)^M \cdot 4M\cdot e^{-\frac{2^i}{2P}}\right) \quad (3.23)$$

$$\Longrightarrow P_e \leq \frac{c'\,(\ln\rho)^{M+1}}{\rho^M} \quad (3.24)$$

where c' is a constant. Therefore,

$$\lim_{\rho\to\infty}\frac{-\log P_e}{\log\rho} \geq \lim_{\rho\to\infty}\frac{\log\rho^M - (M+1)\log(\ln\rho) - \log c'}{\log\rho} = M. \quad (3.25)$$

□

In the above proof, we have considered the LLL-aided decoding type I. In this case, the effective noise vector is equal to $\mathbf{w}' = \mathbf{B}^{\mathtt{H}}\mathbf{w}$, compared to $\mathbf{w}' = \mathbf{H}^{-1}\mathbf{w}$ in zero-forcing. In the previous works [91] [94] [50], the LLL-aided decoding type II has been used. For the type II method, the effective noise vector is equal to $\mathbf{w}' = {\mathbf{H}_{red}}^{-1}\mathbf{w}$ and the average energy of its ith component is proportional to the square norm of the ith column of ${\mathbf{H}_{red}}^{-\mathtt{H}}$. By using inequality (2.33) from lemma 1 (to bound the square norm of the columns of ${\mathbf{H}_{red}}^{-\mathtt{H}}$) and using a similar proof as lemma 3.1, we can show that the results of lemma 3.1 and theorem 3.1 are still valid. Therefore, both of these LLL-aided decoding methods achieve the receive diversity in V-BLAST MIMO systems (or multiple access MIMO systems). However, it is worth noting that the first method is a more natural approach to reduce the power of the entries of the effective noise

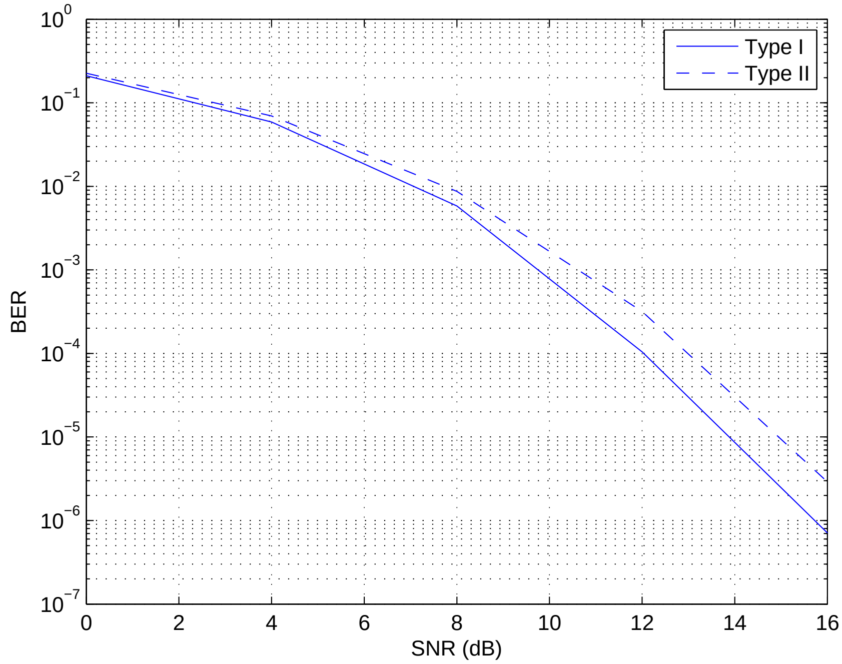

Figure 3.1: Bit-Error-Rate (versus SNR-per-bit) of the two LLL-aided decoding methods for $M = 6$ transmit antennas and $N = 6$ receive antennas with the rate $R = 12$ bits per channel use.

vector, and has a better performance (see figure 3.1). For the case of real lattices, a latice-reduction-aided approach similar to type I is recently studied in [47] and based on the concept of proximity factor, another justification for its superior performance over type II is presented.

3.3 Relation with the naive lattice-decoding

When we have a finite constellation, for each pair of constellation points, the pair-wise error probability can be bounded by Chernoff bound (similar to [81]). By using the union bound, we can show that the exact ML decoding achieves the diversity order of N_r, the number of receive antennas. However, when we use lattice decoding for a finite constellation and consider the out-of-region decoded lattice points as errors, achieving the maximum diversity by lattice decoding is not trivial anymore. Nonetheless, by using lemma 4, we can show that this suboptimum method (called the naive lattice decoding [27]) still achieves the maximum diversity.

Theorem 3.2. *For a MIMO multi-access system (or a point-to-point MIMO system with the V-BLAST transmission method) with N_t transmit antennas and N_r receive antennas, when we use the naive lattice decoding,*

$$\lim_{\rho \to \infty} \frac{-\log P_e}{\log \rho} = N_r. \tag{3.26}$$

Proof. When $\|\mathbf{w}\| \leq \frac{1}{2} d_{\mathbf{H}}$, we have no decoding error. Thus, by using $\frac{1}{2}$ instead of c_{N_t} in the proof of theorem 3.1, we can bound P_e by bounding $\Pr\left\{\|\mathbf{w}\| > \frac{1}{2} d_{\mathbf{H}}\right\}$. Therefore, we can obtain the same result as theorem 3.1. □

In [27], it is shown that for the naive lattice decoding, we can find a family of lattices (generating a family of space-time codes) which achieves diversity order of N_t ($N_t \leq N_r$ is the number of transmit antennas). The current result shows that even if we use the codes generated by the integer lattice, the naive lattice decoding achieves the maximum receive diversity of N_r (number of receive antennas).

3.4 Asymptotic performance of the naive lattice decoding for $N_t = N_r$

We have shown that for $N_r \geq N_t$, the naive lattice decoding achieves the receive diversity in V-BLAST systems (indeed, even its simple latice-reduction-aided approximation still achieves the optimum receive diversity of order N). However, there is a difference between two cases of $N_t < N_r$ and $N_t = N_r$. While for $N_t < N_r$, compared to ML decoding, the performance loss of the naive lattice decoding is bounded in terms of SNR, here we show this is not valid for the case of $N_t = N_r = M$. This dichotomy is related to the bounds on the probability of having a short lattice vector in a lattice generated by a random Gaussian matrix.

In lemma 2.3, an upper bound on the probability of having a short lattice vector is given. For the case of $M \times M$ Gaussian matrices,

$$Prob\{d_{\mathbf{H}} \leq \varepsilon\} \leq C\varepsilon^{2M} \ln\left(\frac{1}{\varepsilon}\right)^{M-1}.$$

The term $\ln\left(\frac{1}{\varepsilon}\right)$ suggests an unboundedly increasing gap between the performance of ML decoding and the naive lattice decoding (though both of them have the same slope M).

In this section, we present a lower bound on the error probability of the naive lattice decoding and show that this unboundedly increasing gap does exist.

Lemma 3.3. *For $M \geq 2$ and $\varepsilon < 1$, for the lattice generated by an $M \times M$ random complex Gaussian matrix $\mathbf{H}$ with zero mean and unit variance, there is a constant C' such that,*

$$Prob\{d_{\mathbf{H}} \leq \varepsilon\} \geq C' \varepsilon^{2M} \ln\left(\frac{1}{\varepsilon}\right). \tag{3.27}$$

Proof: Consider $L_{(\mathbf{v}_1,...,\mathbf{v}_{N_t})}$ as the lattice generated by $\mathbf{v}_1$,$\mathbf{v}_2$,...,$\mathbf{v}_{N_t}$. Each point of $L_{(\mathbf{v}_1,...,\mathbf{v}_{N_t})}$ can be represented by $\mathbf{v}_{(z_1,...,z_{N_t})} = z_1\mathbf{v}_1 + z_2\mathbf{v}_2 + ... + z_{N_t}\mathbf{v}_{N_t}$, where $z_1, ..., z_{N_t}$ are complex integer numbers.

The vectors $\mathbf{v}_1$,$\mathbf{v}_2$,...,$\mathbf{v}_{N_t}$ are independent and jointly Gaussian. Therefore, for every complex vector $\mathbf{b} = (b_1, ..., b_{N_t})$, the vector $\mathbf{v_b} = b_1\mathbf{v}_1 + b_2\mathbf{v}_2 + ... + b_{N_t}\mathbf{v}_{N_t}$ has complex circular Gaussian distribution with the variance

$$\varrho_{\mathbf{b}}^2 = \|\mathbf{b}\|^2 = |b_1|^2 + ... + |b_{N_t}|^2. \tag{3.28}$$

Now, considering the pdf of $\mathbf{v_b}$, we can bound $\Pr\{\|\mathbf{v_b}\| \leq \varepsilon\} = \int_{\|\mathbf{v}\|\leq\varepsilon} f_{\mathbf{v_b}}(\mathbf{v})\, d\mathbf{v}$ by using the fact that $e^{-\frac{\varepsilon^2}{\varrho_{\mathbf{b}}^2}} \leq e^{-\frac{\|\mathbf{v}\|^2}{\varrho_{\mathbf{b}}^2}} \leq 1$ for $\|\mathbf{v}\| \leq \varepsilon$:

$$\int_{\|\mathbf{v}\|\leq\varepsilon} \frac{1}{\pi^{N_t}\varrho_{\mathbf{b}}^{2M}} e^{-\frac{\varepsilon^2}{\varrho_{\mathbf{b}}^2}}\, d\mathbf{v} \leq \int_{\|\mathbf{v}\|\leq\varepsilon} f_{\mathbf{v}}(\mathbf{v})\, d\mathbf{v} \leq \int_{\|\mathbf{v}\|\leq\varepsilon} \frac{1}{\pi^{N_t}\varrho_{\mathbf{b}}^{2M}}\, d\mathbf{v}. \tag{3.29}$$

Thus, because the volume of region of the integral (which is a $2M$-dimensional sphere with radius ε) is proportional to ε^{2M},

$$c_6 \frac{\varepsilon^{2M}}{\|\mathbf{b}\|^{2M}} e^{-\frac{\varepsilon^2}{\|\mathbf{b}\|^2}} \leq \Pr\{\|\mathbf{v_b}\| \leq \varepsilon\} \leq c_7 \frac{\varepsilon^{2M}}{\|\mathbf{b}\|^{2M}}. \tag{3.30}$$

We can represent any M-dimensional complex integer vector as a $2M$-dimensional real integer vector. In our proof, we consider only integer vectors in the set $\mathcal{B}$ which consists of integer vectors $\mathbf{z}$ such that their real entries do not have a nontrivial common divisor and $\|\mathbf{z}\|_\infty \leq \varepsilon^{-\frac{1}{2M}}$ where $\|\cdot\|_\infty$ represents the norm of the largest real entry. First, we show that the number of such integer vectors $\mathbf{z}$ in the region $2^{(k-1)} < \|\mathbf{z}\|_\infty \leq 2^k$ is at least 2^{2Mk}. The total number of integer points in the region $2^{(k-1)} < \|\mathbf{z}\|_\infty \leq 2^k$ is[2] $\left(2^{k+1}+1\right)^{2M} - \left(2^k+1\right)^{2M}$. The number of those points whose entries have a common divisor i is at most equal to the number of integer points in the region $\|\mathbf{z}\|_\infty \leq \frac{2^k}{i}$. Therefore, n_k, the number of integer vectors $\mathbf{z}$ whose entries does not have nontrivial common divisors, can be lower bounded by

$$\begin{aligned}
n_k &\geq \left(\left(2^{k+1}+1\right)^{2M} - \left(2^k+1\right)^{2M}\right) - \sum_{i=2}^{2^k}\left(2\frac{2^k}{i}+1\right)^{2M} \\
&> \left(\left(2^{k+1}+1\right)^{2M} - \left(3\cdot 2^{k-1}\right)^{2M}\right) - \sum_{i=2}^{2^k}\left(3\frac{2^k}{i}\right)^{2M} \\
&> 2^{2kM+2M}\left(1-\left(\frac{3}{4}\right)^{2M} - \left(\frac{3}{2}\right)^{2M}\sum_{i=2}^{\infty}\frac{1}{i^{2M}}\right) \\
&> 2^{2kM+2M}\left(1-\left(\frac{3}{4}\right)^{2M} - \left(\frac{3}{2}\right)^{2M}\cdot\left(\frac{1}{2^{2M}}+\frac{1}{3^{2M}}+\int_3^\infty \frac{1}{x^{2M}}\,dx\right)\right) \\
&= 2^{2kM+2M}\left(1-\left(\frac{3}{4}\right)^{2M} - \left(\frac{3}{4}\right)^{2M} - \left(\frac{1}{2}\right)^{2M} - \left(\frac{3}{2}\right)^{2M}\cdot\frac{1}{3^{2M-1}(2M-1)}\right) \\
&> 2^{2kM+2M}\left(1-2\left(\frac{3}{4}\right)^{2M} - \frac{1}{2^{2M}}.\left(1+\frac{2}{2M-1}\right)\right) \\
&\geq 2^{2kM+2M}\left(1-2\left(\frac{3}{4}\right)^{4} - \frac{1}{2^{4}}.\left(1+\frac{2}{3}\right)\right) > 2^{2kM+2M}\cdot 2^{-4} \geq 2^{2kM} \quad \text{for } M \geq 2.
\end{aligned} \tag{3.31}$$

[2]The number of points in the cube $\|\mathbf{z}\|_\infty \leq 2^k$ is $\left(2^{k+1}+1\right)^{2M}$ and the number of points in the cube $\|\mathbf{z}\|_\infty \leq 2^{(k-1)}$ is $\left(2^k+1\right)^{2M}$.

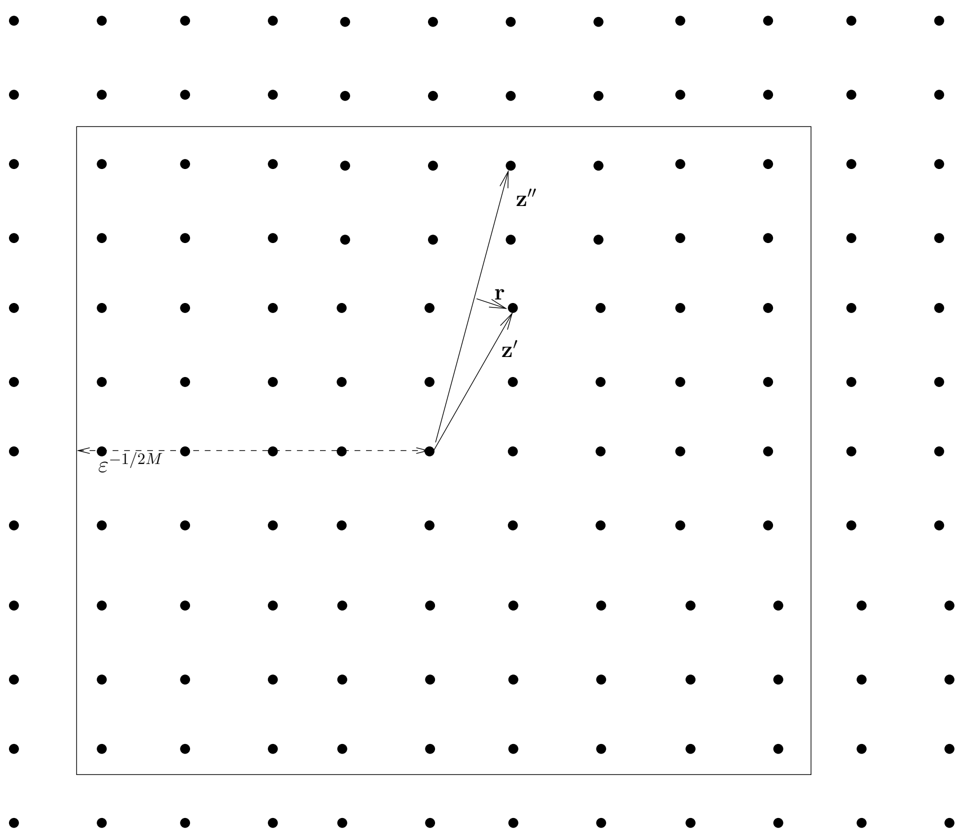

Figure 3.2: integer points in the region $\|\mathbf{z}\|_\infty \leq \varepsilon^{-\frac{1}{2M}}$.

Now, we find an upper bound on $\Pr\{\|\mathbf{v}_{\mathbf{z}'}\| \leq \varepsilon, \|\mathbf{v}_{\mathbf{z}''}\| \leq \varepsilon\}$ for two different complex integer vectors $\mathbf{z}'$ and $\mathbf{z}''$ which belong to $\mathcal{B}$. We can write $\mathbf{z}'$ as $a\mathbf{z}'' + \mathbf{r}$ where a is a complex number and $\mathbf{r}$ is a complex vector, orthogonal to $\mathbf{z}''$. We show that $\|\mathbf{r}\| \geq \frac{1}{\sqrt{2M}}\varepsilon^{\frac{1}{2M}}$. The area of the triangle which has vertexes $\mathbf{0}$, $\mathbf{z}'$, and $\mathbf{z}''$, is equal to $S = \frac{1}{2}\|\mathbf{r}\| \cdot \|\mathbf{z}''\|$. On the other hand, because $\mathbf{0}$, $\mathbf{z}'$, and $\mathbf{z}''$ are integer points, $2S$ should be integer. Also, because the entries of $\mathbf{z}'$ do not have any nontrivial common divisor, $\mathbf{z}'$ cannot be a multiplier of $\mathbf{z}''$ (and vice versa). Because $\mathbf{z}'$ and $\mathbf{z}''$ are not multipliers of each other, S is nonzero. Thus, $S \geq \frac{1}{2}$, hence,

$$\frac{1}{2}\|\mathbf{r}\| \cdot \|\mathbf{z}''\| \geq \frac{1}{2} \tag{3.32}$$

$$\Longrightarrow \|\mathbf{r}\| \geq \frac{1}{\|\mathbf{z}''\|} \geq \frac{1}{\sqrt{2M}\|\mathbf{z}''\|_\infty} \geq \frac{1}{\sqrt{2M}\varepsilon^{-\frac{1}{2M}}} = \frac{1}{\sqrt{2M}}\varepsilon^{\frac{1}{2M}}. \tag{3.33}$$

Now we bound $\Pr\{\|\mathbf{v}_{\mathbf{z}'}\| \leq \varepsilon, \|\mathbf{v}_{\mathbf{z}''}\| \leq \varepsilon\}$. Because $\mathbf{r} \perp \mathbf{z}''$, we can see that $\mathbf{v}_{\mathbf{r}} \perp \mathbf{v}_{\mathbf{z}''}$. Thus, when $\|\mathbf{v}_{\mathbf{r}}\| > \varepsilon$, using the fact that $\mathbf{v}_{\mathbf{a}+\mathbf{b}} = \mathbf{v}_{\mathbf{a}} + \mathbf{v}_{\mathbf{b}}$,

$$\|\mathbf{v}_{\mathbf{z}'}\| = \|\mathbf{v}_{a\mathbf{z}''+\mathbf{r}}\| = \|\mathbf{v}_{a\mathbf{z}''} + \mathbf{v}_{\mathbf{r}}\| \geq \|\mathbf{v}_{\mathbf{r}}\| > \varepsilon$$

Therefore,

$$\begin{aligned} &\Pr\{\|\mathbf{v}_{\mathbf{z}'}\| \leq \varepsilon, \|\mathbf{v}_{\mathbf{z}''}\| \leq \varepsilon\} \\ &\leq \Pr\{\|\mathbf{v}_{\mathbf{r}}\| \leq \varepsilon, \|\mathbf{v}_{\mathbf{z}''}\| \leq \varepsilon\} \end{aligned} \tag{3.34}$$

Based on the orthogonality of $\mathbf{r}$ and $\mathbf{z}''$, $\mathbf{v}_{\mathbf{r}}$ and $\mathbf{v}_{\mathbf{z}''}$ are independent. Thus, using (3.30), (3.33), and noting that $\|\mathbf{z}''\| \geq 1$ (because $\mathbf{z}''$ is a nonzero integer vector):

$$\Pr\left\{\|\mathbf{v}_{\mathbf{z}'}\| \leq \varepsilon, \|\mathbf{v}_{\mathbf{z}''}\| \leq \varepsilon\right\} \leq \Pr\left\{\|\mathbf{v}_{\mathbf{z}''}\| \leq \varepsilon\right\} \cdot \Pr\left\{\|\mathbf{v}_{\mathbf{r}}\| \leq \varepsilon\right\}$$

$$\leq \left(c_7 \frac{\varepsilon^{2M}}{\|\mathbf{z}''\|^{2M}}\right) \cdot \left(c_7 \frac{\varepsilon^{2M}}{\|\mathbf{r}\|^{2M}}\right)$$

$$\leq c_7^2 \varepsilon^{2M} \cdot \frac{\varepsilon^{2M} \left(2M\right)^{N_t}}{\left(\varepsilon^{\frac{1}{2M}}\right)^{2M}}$$

$$= c_8 \varepsilon^{4M-1} \tag{3.35}$$

Now, we use the Bonferroni inequality [26],

$$\Pr\left\{d_{\mathbf{H}} \leq \varepsilon\right\} = \Pr\left\{\exists\, \mathbf{z} \neq 0 :\ \|\mathbf{v}_{\mathbf{z}}\| \leq \varepsilon\right\} \geq \Pr\left\{\exists\, \mathbf{z} :\ \mathbf{z} \in \mathcal{B}, \|\mathbf{v}_{\mathbf{z}}\| \leq \varepsilon\right\}$$

$$\geq \sum_{\mathbf{z} \in \mathcal{B}} \Pr\left\{\|\mathbf{v}_{\mathbf{z}}\| \leq \varepsilon\right\}$$

$$- \sum_{\mathbf{z}', \mathbf{z}'' \in \mathcal{B}} \Pr\left\{\|\mathbf{v}_{\mathbf{z}'}\| \leq \varepsilon, \|\mathbf{v}_{\mathbf{z}''}\| \leq \varepsilon\right\} \tag{3.36}$$

For the first term of (3.36),

$$\sum_{\mathbf{z} \in \mathcal{B}} \Pr\left\{\|\mathbf{v}_{\mathbf{z}}\| \leq \varepsilon\right\} \tag{3.37}$$

$$\geq \sum_{k=0}^{\left\lfloor \log\left(\varepsilon^{-\frac{1}{2M}}\right)\right\rfloor} \sum_{\mathbf{z} \in \mathcal{B}, 2^{k-1} < \|\mathbf{z}\|_\infty \leq 2^k} \Pr\left\{\|\mathbf{v}_{\mathbf{z}}\| \leq \varepsilon\right\} \tag{3.38}$$

By using (3.31), (3.30), and noting that $\|\mathbf{z}\| \leq \sqrt{2M}\|\mathbf{z}\|_\infty$ and $e^{-\frac{\varepsilon^2}{\|\mathbf{z}\|^2}} \geq e^{-1}$ (because $\varepsilon < 1$ and $\|\mathbf{z}\| \geq 1$),

$$(3.38) \geq \sum_{k=0}^{\left\lfloor \log\left(\varepsilon^{-\frac{1}{2M}}\right)\right\rfloor} 2^{2kM} \cdot \frac{c_6 \varepsilon^{2M}}{\left(2^k\right)^{2M} \cdot (2M)^{N_t}} \cdot e^{-1} \tag{3.39}$$

$$\geq \sum_{k=0}^{\left\lfloor \log\left(\varepsilon^{-\frac{1}{2M}}\right)\right\rfloor} c_9 \varepsilon^{2M} \tag{3.40}$$

$$= \left(\left\lfloor \log\left(\varepsilon^{-\frac{1}{2M}}\right)\right\rfloor + 1\right) \cdot c_9 \varepsilon^{2M} \geq c_{10} \varepsilon^{2M} \cdot \ln\left(\frac{1}{\varepsilon}\right). \tag{3.41}$$

For the second term of of (3.36), because the number of complex integers in $\mathcal{B}$ (which is at most the number of integer points in the cube $\|\mathbf{z}\|_\infty \leq \varepsilon^{-\frac{1}{2M}}$) is bounded by $c_{11}\left(\varepsilon^{-\frac{1}{2M}}\right)^{2M} = c_{11}\varepsilon^{-1}$, the number of pairs $(\mathbf{z}', \mathbf{z}'')$ is at most $\left(c_{11}\varepsilon^{-1}\right)^2$. Thus, using (3.35):

$$\sum_{\mathbf{z}', \mathbf{z}'' \in \mathcal{B}} \Pr\left\{\|\mathbf{v}_{\mathbf{z}'}\| \leq \varepsilon, \|\mathbf{v}_{\mathbf{z}''}\| \leq \varepsilon\right\} \tag{3.42}$$

$$\leq \left(c_{11}\varepsilon^{-1}\right)^2 \cdot c_8 \varepsilon^{4M-1} \tag{3.43}$$

$$\leq c_{12} \varepsilon^{4M-3} \tag{3.44}$$

Now, by using (3.41) and (3.44),

$$(3.36) \geq c_{10}\varepsilon^{2M} \ln\left(\frac{1}{\varepsilon}\right) - c_{12}\varepsilon^{4M-3} \tag{3.45}$$

$$\geq C'\varepsilon^{2M} \ln\left(\frac{1}{\varepsilon}\right), \quad \text{for } M \geq 2. \tag{3.46}$$

■

Theorem 3.3. *Consider a MIMO fading channel with $N_t = N_r = M$ transmit and receive antennas and a V-BLAST transmission system. The naive lattice-decoding has an asymptotically unbounded loss, campared to the exact ML decoding.*

Proof: For ML decoding, by using the Chernoff bound for the pairwise error probability and then applying the union bound for the finite constellation, we have [81]

$$P_{error-ML} \leq c_{13}(SNR)^{-M} \tag{3.47}$$

where c_{13} depends on the size of constellation.

For naive lattice decoding,

$$P_{error-NLD} \geq \Pr\left\{d_{\mathbf{H}} \leq \frac{1}{\sqrt{SNR}}\right\} .Q\left(\frac{1}{\sqrt{N_t}}\right)$$

$$\geq c_{14}(SNR)^{-M}\ln(SNR). \tag{3.48}$$

Therefore, although both of them asymptotically have the same slope and achieve the optimal receive diversity of order M, for large SNRs, the gap between their performances is unbounded (with a logarithmic growth, or in other words, $\log\log SNR$ in dB scale). ■

3.5 Conclusions

We have shown that LLL-aided zero-forcing, which is a polynomial-time algorithm, achieves the maximum receive diversity in MIMO systems. By using LLL reduction before zero-forcing, the complexity of the MIMO decoding is equal to the complexity of the zero-forcing method with just an additional polynomial-time preprocessing for the whole fading block. Also, it is shown that by using the naive lattice decoding, instead of ML decoding, we do not lose the receive diversity order.

Chapter 4

Diversity-multiplexing tradeoff of lattice-decoded codes

The optimal diversity-multiplexing trade-off [96] is considered as an important theoretical benchmark for practical systems. For the encoding part, recently, several lattice codes are introduced which have the non-vanishing determinant property and achieve the optimal trade-off, conditioned on using the exact maximum-likelihood decoding [18] [53] [38]. The lattice structure of these codes facilitates the encoding. For the decoding part, various lattice decoders, including the sphere decoder and the lattice-reduction-aided decoder are presented in the literature [14] [91]. To achieve the exact maximum likelihood performance, we need to find the closest point of the lattice inside the constellation region, which can be much more complex than finding the closest point in an infinite lattice. To avoid this complexity, one can perform the traditional lattice decoding (for the infinite lattice) and then, discard the out-of-region points.

This approach is called Naive Lattice Decoding (NLD).

In chapter 3, it is shown that this sub-optimum decoding (and even its lattice-reduction-aided approximation) still achieve the maximum receive diversity in the fixed-rate MIMO systems. Achieving the optimal receive diversity by a low decoding complexity makes lattice-reduction-aided decoding (using the LLL reduction) an attractive choice for different applications. Nonetheless, this work shows that concerning diversity-multiplexing trade-off, the optimality cannot be achieved by the naive-lattice decoding or its approximations.

In [27], using a probabilistic method, a lower bound on the best achievable trade-off, using the naive lattice decoding, is presented. In this chapter, we present an upper bound on the performance of the naive lattice decoding for codes based on full-rate lattices. We show that NLD cannot achieve the optimum diversity-multiplexing trade-off. Also, for the special case of equal number of transmit and receive antennas, we show that even the best full-rate lattice codes (including perfect space-time codes such as the Golden code [53]) cannot perform better than the simple V-BLAST (if we use the naive lattice decoding at the receiver). It should be noted that in this chapter, we have assumed that the underlying lattice is fixed for different rates and SNR values (e.g. lattice codes introduced in [18] [53] [38]). If we relax this restriction, there can exist a family of lattice codes (based on different lattice structures for different rates and SNR values) which achieves the optimum tradeoff using the naive lattice decoding [51].

4.1 Diversity-multiplexing tradeoff for the naive lattice decoding

Similar to Chapter 3, we consider a MIMO system with N_t transmit antennas and N_r receive antennas. We send space-time codewords $\mathbf{X} = [\mathbf{x}_1, ..., \mathbf{x}_T]$ with complex entries ($\mathbf{x}_i \in \mathbb{C}^{N_t}$) and at the receiver, we find $\tilde{\mathbf{x}}_i$ as $\mathbf{H}^{-1}\tilde{\mathbf{y}}_i$ where $[\tilde{\mathbf{y}}_1, ..., \tilde{\mathbf{y}}_T]$ is the closest N_tT-dimensional lattice point to $[\mathbf{y}_1, ..., \mathbf{y}_T]$.

To derive the upper bound on the diversity-multiplexing trade-off of NLD, we first present a lower bound on the probability that the received lattice (the lattice code after passing through the fading channel) has a short vector.

Lemma 4.1. *Assume that the $N_r \times N_t$ matrix $\mathbf{H}$ has i.i.d. complex Gaussian entries with zero mean and unit variance and let $d(\mathbf{H}_T\mathbf{L})$ be the minimum distance of the lattice generated by $\mathbf{H}_T\mathbf{L}$, where $\mathbf{L}$ is the full-rank $N_tT \times N_tT$ generator of a given complex lattice with unit volume and $\mathbf{H}_T$ is the $N_rT \times N_tT$ block diagonal matrix constructed by repeating $\mathbf{H}$ along the main diagonal. We have,*

$$\lim_{\varepsilon \to 0} \frac{\log \Pr\{d(\mathbf{H}_T\mathbf{L}) \leq \varepsilon\}}{\log \varepsilon} \leq 2N_t(N_r - N_t + 1) \tag{4.1}$$

Proof: Let $\sigma_1 \leq \sigma_2 \leq ... \leq \sigma_{N_t}$ be the nonzero singular values of $\mathbf{H}$. Considering the pdf of the singular values of a Gaussian matrix [15], it can be shown that [96]

$$\lim_{\varepsilon \to 0} \frac{\log \Pr\left\{\sigma_1 \leq \varepsilon^{b_1}, ..., \sigma_{N_t} \leq \varepsilon^{b_{N_t}}\right\}}{\log \varepsilon} = \sum_{i=1}^{N_t} 2(N_r - N_t + 2i - 1)b_i \tag{4.2}$$

Thus for any $b_1 > N_t$ and $b_2, \cdots, b_{N_t} > 0$

$$\lim_{\varepsilon \to 0} \frac{\log \Pr\left\{\sigma_1 \leq \varepsilon^{b_1}, ..., \sigma_{N_t} \leq \varepsilon^{b_{N_t}}\right\}}{\log \varepsilon} \geq \lim_{\varepsilon \to 0} \frac{\log \Pr\left\{\sigma_1 \leq \frac{1}{4\sqrt{N_t}}\varepsilon^{N_t}, \sigma_i \leq \frac{1}{4\sqrt{N_t}} \ for\ i > 1\right\}}{\log \varepsilon} \tag{4.3}$$

and by considering $b'_1 = N_t$ and $b'_2, \cdots, b'_{N_t} = 0$,

$$\lim_{\varepsilon \to 0} \frac{\log \Pr\left\{\sigma_1 \leq \frac{1}{4\sqrt{N_t}}\varepsilon^{N_t}, \sigma_i \leq \frac{1}{4\sqrt{N_t}} \ for\ i > 1\right\}}{\log \varepsilon} \geq 2N_t(N_r - N_t + 1). \tag{4.4}$$

By combining (4.3) and (4.4), we have

$$\lim_{\varepsilon \to 0} \frac{\log \Pr\left\{\sigma_1 \leq \frac{1}{4\sqrt{N_t}}\varepsilon^{N_t}, \sigma_i \leq \frac{1}{4\sqrt{N_t}} \ for\ i > 1\right\}}{\log \varepsilon} = 2N_t(N_r - N_t + 1). \tag{4.5}$$

Let $\mathbf{v}_{min}$ be the singular vector of $\mathbf{H}$, corresponding to σ_1. For each N_tT-dimensional complex vector $\mathbf{v} = [a_1\mathbf{v}_{min}^\mathsf{T} \;\; a_2\mathbf{v}_{min}^\mathsf{T} ... \; a_T\mathbf{v}_{min}^\mathsf{T}]^\mathsf{T}$,

$$\|\mathbf{H}_T\mathbf{v}\|^2 = \sum_{i=1}^{T} a_i^2 \|\mathbf{H}\mathbf{v}_{min}\|^2 = \sum_{i=1}^{T} \sigma_1^2 \|a_i\mathbf{v}_{min}\|^2 = \sigma_1^2 \|\mathbf{v}\|^2. \tag{4.6}$$

Thus, assuming $\sigma_1 \leq \frac{1}{4\sqrt{N_t}}\varepsilon^{N_t}$,

$$\|\mathbf{H}_T\mathbf{v}\| \leq \frac{1}{4\sqrt{N_t}}\varepsilon^{N_t}\|\mathbf{v}\|. \tag{4.7}$$

Let $\mathcal{A}$ be a $2N_tT$-dimensional hypercube with edges of length $\frac{1}{\varepsilon^{N_t}}$ whose $2T$ edges are parallel to the subspace spanned by the vectors $\mathbf{v} = [a_1\mathbf{v}_{min}^\mathsf{T} \;\; a_2\mathbf{v}_{min}^\mathsf{T} ... \; a_T\mathbf{v}_{min}^\mathsf{T}]^\mathsf{T}$ and the other $2T(N_t - 1)$ edges are orthogonal to that subspace. The volume of this cube is $\varepsilon^{-2N_t^2T}$. Because the volume of the lattice is 1, for K, the number of lattice points inside this cube, we have[1] $\lim_{\varepsilon\to 0}\frac{K}{\varepsilon^{-2N_t^2T}}=1$.

Now, assuming $\sigma_1 \leq \frac{1}{4\sqrt{N_t}}\varepsilon^{N_t}$ and $\sigma_{N_t} \leq \frac{1}{4\sqrt{N_t}}$, the region $\mathbf{H}_T\mathcal{A}$ is inside a $2N_tT$-dimensional orthotope (in the subspace spanned by $\mathbf{H}_T$) whose $2T$ edges (which correspond to the smallest singular value σ_1) have length $\frac{1}{4\sqrt{N_t}}$ and the length of the other $2T(N_t-1)$ is at most $\frac{1}{4\sqrt{N_t}\varepsilon^{N_t}}$ (because of the bound on the largest singular value σ_{N_t}). The $2T$ smaller edges can be covered by at most $\lceil 4^{-1}\varepsilon^{-1}\rceil \leq 2^{-1}\varepsilon^{-1}$ segments of length $\frac{\varepsilon}{\sqrt{N_t}}$ and the others can be covered by at most $\lceil 4^{-1}\varepsilon^{-(N_t+1)}\rceil \leq 2^{-1}\varepsilon^{-(N_t+1)}$ segments of length $\frac{\varepsilon}{\sqrt{N_t}}$. Thus, this orthotope can be covered by at most $\left(2^{-1}\varepsilon^{-1}\right)^{2T}\left(2^{-1}\varepsilon^{-(N_t+1)}\right)^{2T(N_t-1)} = 2^{-2N_tT}\varepsilon^{-2N_t^2T}$ hypercubes of edge length $\frac{\varepsilon}{\sqrt{N_t}}$. Because $\lim_{\varepsilon\to 0}\frac{K}{\varepsilon^{-2N_t^2T}}=1$, when $\varepsilon\to 0$, the number of these small hypercubes is smaller than the number of lattice points inside them. Thus, based on Dirichlet's box principle, in one of these hypercubes there are at least 2 points of the new lattice, hence $d\left(\mathbf{H}_T\mathbf{L}\right)$ is smaller than the diameter of the small hyper cubes:

$$d_{\mathbf{H}} \leq \sqrt{N_t}\cdot\frac{\varepsilon}{\sqrt{N_t}}. \tag{4.8}$$

Therefore,

$$\lim_{\varepsilon\to 0}\frac{\log\Pr\{d\left(\mathbf{H}_T\mathbf{L}\right)\leq\varepsilon\}}{\log\varepsilon} \leq \lim_{\varepsilon\to 0}\frac{\log\Pr\left\{\sigma_1\leq\varepsilon^{N_t},\sigma_{N_t}\leq\frac{1}{2N_t}\right\}}{\log\varepsilon} = 2N_t(N_r-N_t+1). \tag{4.9}$$

■

[1]When a region is large, the number of lattice points inside the region can be approximated by the ratio between the volume of the region and the volume of the lattice.

Theorem 4.1. *Consider a MIMO fading channel with N_t transmit and N_r receive antennas ($N_t \leq N_r$) with codebooks from an N_tT-dimensional lattice $\mathbf{L}$, which are sent over T channel uses. For the naive lattice decoding, the diversity-multiplexing trade-off of the system is*

$$d_{NLD}(r) \leq N_t(N_r - N_t + 1) - r\left(N_r - N_t + 1\right), \quad \text{for } 0 \leq r \leq N_t. \tag{4.10}$$

Proof: Consider the code of rate R constructed from the lattice. The number of codewords is equal to 2^R. Without any loss of generality, we can assume that the volume of the lattice is fixed and is equal to 1, and the power constraint P is dependent on the rate. To satisfy the power constraint, at least half of the codewords should have power less than $2P$. The number of codewords with power less than $2P$ is equal to the number of lattice points inside a $2N_t$-dimensional sphere whose volume is proportional to P^{N_t}. Thus, by approximating the number of lattice points with the ratio of the volume of the region and the volume of the lattice:

$$2^R \leq c_1 P^{N_t}. \tag{4.11}$$

where c_1 is a constant, independent of SNR[2]. According to the definition of the multiplexing gain, $r = \lim_{SNR\to\infty} \frac{\log R}{\log SNR}$. Using (4.11),

$$\lim_{SNR\to\infty} \frac{\log P}{\log SNR} \geq \frac{\log \frac{1}{N_t} \log R}{\log SNR} = \frac{r}{N_t}. \tag{4.12}$$

[2]Throughout this chapter $c_1, c_2, ...$ are only dependent on size of dimensions.

For the symbol error probability P_e, considering $SNR = \frac{N_t P}{\sigma^2}$,

$$P_e \geq \Pr\left\{d\left(\mathbf{H}_T\mathbf{L}\right) \leq \frac{\sigma}{\sqrt{N_t}}\right\}.Q\left(\frac{1}{2\sqrt{N_t}}\right) = \Pr\left\{d\left(\mathbf{H}_T\mathbf{L}\right) \leq \frac{\sqrt{P}}{\sqrt{SNR}}\right\}.Q\left(\frac{1}{2\sqrt{N_t}}\right). \tag{4.13}$$

Therefore, using lemma 1 (with $\varepsilon = \frac{\sqrt{P}}{\sqrt{SNR}}$) and (4.12),

$$d_{NLD}(r) = \lim_{SNR\to\infty} \frac{-\log P_e}{\log SNR} \leq \lim_{SNR\to\infty} \frac{-\log \Pr\left\{d_{\mathbf{H}} \leq \frac{\sqrt{P}}{\sqrt{SNR}}\right\}}{\log SNR}$$

$$\leq \lim_{SNR\to\infty} \frac{-2N_t(N_r - N_t + 1)\left(\log \frac{\sqrt{P}}{\sqrt{SNR}}\right)}{\log SNR}$$

$$= \lim_{SNR\to\infty} \frac{-2N_t(N_r - N_t + 1)\left(\frac{1}{2}\log P - \frac{1}{2}\log SNR\right)}{\log SNR}$$

$$\leq -\left(\frac{r}{2N_t} - \frac{1}{2}\right) \cdot 2N_t(N_r - N_t + 1)$$

$$= N_t(N_r - N_t + 1) - r\left(N_r - N_t + 1\right). \tag{4.14}$$

■

Corollary 4.1. *In a MIMO fading channel with $N_t = N_r$ transmit and receive antennas, if we use the naive lattice decoding, the diversity-multiplexing trade-off for full-rate lattice code cannot be better than that of V-BLAST.*

Proof: When $N_t = N_r$, according to Theorem 4.1,

$$d_{NLD}(r) \leq N_t - r \tag{4.15}$$

On the other hand, for the V-BLAST system with lattice decoding [85],

$$d_{V-BLAST}(r) = N_t - r \tag{4.16}$$

■

It is interesting to compare this result with the results on lattice space-time codes which have non-vanishing determinants. Although by ML decoding, these codes (such as the 2×2 Golden code) achieve the optimal diversity-multiplexing trade-off, when we replace ML decoding with the naive lattice decoding (and its approximations), their performance is not much better than the simple V-BLAST scheme (specially when the number of transmit and receive antennas are the same)

To better understand the difference between the naive lattice decoding and the ML decoding, we note that for small constellations, when the generator of the received lattice has a small singular value, the minimum distance of the lattice can be much smaller than the minimum distance of the constellation. Figure 4.2 shows this situation for a small 4-point constellation from a 2-dimensional lattice.

We should note that this upper bound is for full-rate lattices. Lattices with lower rate, can provide higher diversity, but their rate is limited by the dimension of the lattice. For example, The Alamouti code, based on QAM constellations, can achieve the full diversity for fixed rates ($r = 0$), but its rate is limited by one.

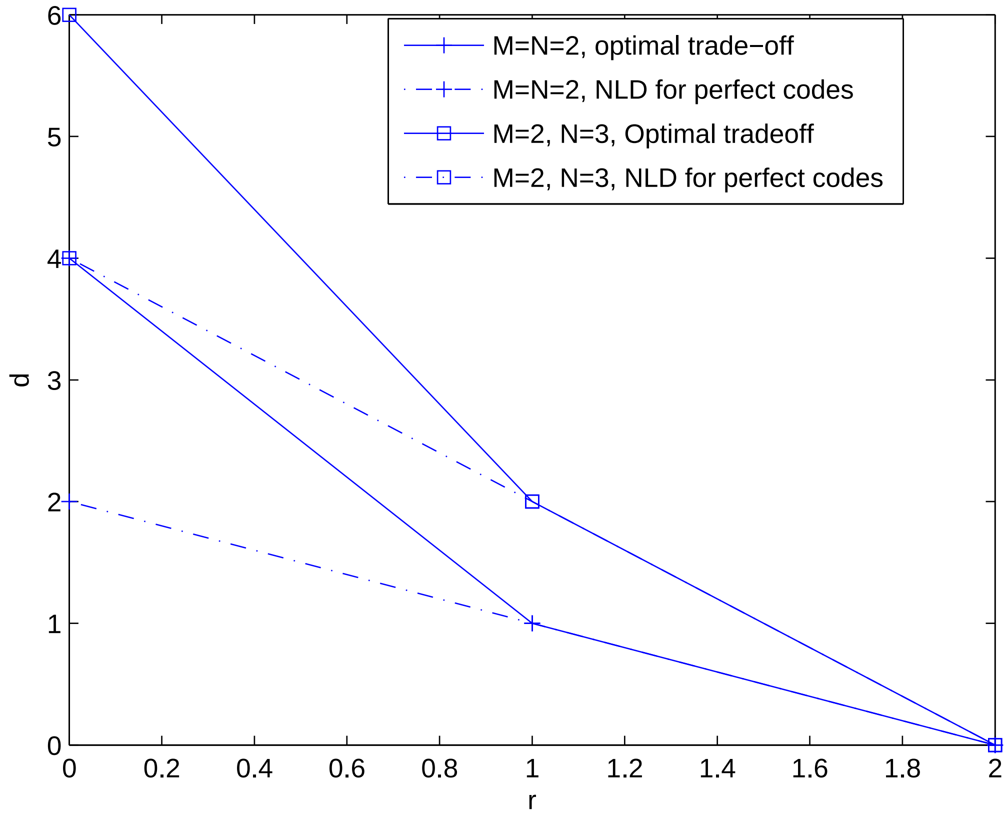

Figure 4.1: Comparison between the optimal diversity-multiplexing tradeoff and the upper bound on the diversity-multiplexing trade-off of full-rate lattice codes (including perfect space-time codes such as the Golden code)

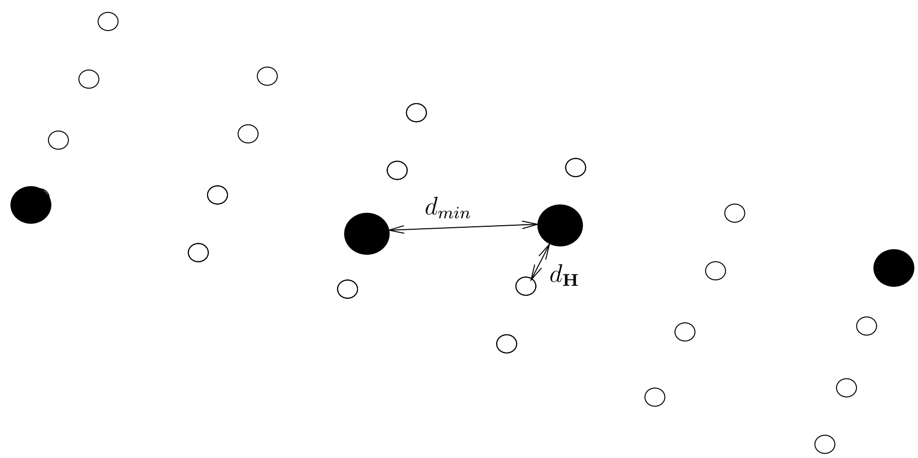

Figure 4.2: Minimum distance of a lattice ($d_{\mathbf{H}} = d\left(\mathbf{H}_T \mathbf{L}\right)$), compared to the minimum distance of a lattice code (d_{min})

4.2 Conclusions

In this chapter, the inherent limitations of the performance of the naive lattice decoding is investigated. The naive lattice decoding and various implementions of it (such as the sphere decoding) and its simple approximated versions (such as the LLL-aided decoding) are very attractive for the practical MIMO systems. Nonetheless, to achieve theoretical benchmarks (such as the diversity-multiplexing trade-off), these techniques are not always sufficient. For the diversity-multiplexing trade-off, although different elegant lattice codes have been introduced which achieve the optimal trade-off [18] [53] [38], they need ML decoding to achieve optimality. On the other hand, there can exist a family of lattice codes (based on different lattice structures for different rates and SNR values) which achieves the optimum tradeoff using the naive lattice decoding [51]. However, the existence proof in [51] does not provide any constructive solution for the encoding of such codes. Therefore, the problem of achieving the optimum diversity-multiplexing tradeoff by a practical encoding and decoding scheme is still open.

Chapter 5

Robust joint source-channel coding in Gaussian channels

In many applications, delay-limited transmission of analog sources over an additive white Gaussian channel is needed. Also, in many cases the exact signal-to-noise ratio is not known at the transmitter, and may vary over a wide range of values. Two examples of this scenario are transmitting an analog source over a quasi-static fading channel and/or multicasting it to different users (with different channel gains).

Without considering the delay limitations, digital codes can theoretically achieve the optimal performance in the Gaussian channel. Indeed, for the ergodic point-to-point channels, Shannon's source-channel coding separation theorem [64] [65] ensures the optimality of separately designing source and channel codes. However, for the case of a limited delay, several articles [4] [21] [20] [56] [89] have shown that joint source-channel codes have a better performance as compared to the separately designed source and channel codes (which are called tandem codes). Also, digital coding is very sensitive to the mismatch in the estimation of the channel signal-to-noise-ratio (SNR).

To avoid the saturation effect of digital coding, various analog and hybrid digital-analog schemes are introduced and investigated in the past [5, 8, 9, 13, 33–35, 44, 57, 66, 68, 69, 83, 84, 86, 93]. Among them, examples of 1-to-2-dimensional analog maps can be found as early as the works of Shannon [66] and Kotelnikov [44] and different variations of *Shannon-Kotelnikov maps* (which are also called *twisted modulations*) and their performance is studied in [33, 84, 93]. Also, in [8] and [86], analog codes based on dynamical systems are proposed. Although these codes can provide asymptotic gains (for high SNR) over the simple repetition code, they suffer from a threshold effect. Indeed, when the SNR becomes less than a certain threshold, the performance of these systems degrades severely. Therefore, the design parameters of these methods should be chosen according to the operating SNR, hence, these methods are still very sensitive to the errors in the estimation of SNR. Also, although the performance of the system is not saturated for the high SNR (unlike digital codes), the scaling of the end-to-end distortion is far from the theoretical bounds. Theoretical bounds on the robustness of joint source channel coding schemes (for the delay-unlimited case) are presented in [49] and [58]. To achieve better SDR scaling, a coding scheme is introduced in [60, 61] which uses B repetitions of a (k,n) binary code to map the digits of the infinite binary expansion of k sample of the source to the digits of a nB-dimensional transmit vector. For this scheme, the banwidth expansion factor is $\eta = \frac{nB}{k}$ and the SDR asymptotically scales as $SDR \propto SNR^B$, while in theory, the optimum scaling is $SDR \propto SNR^\eta$. Thus, this scheme cannot achieve the optimum scaling by using a single mapping.

In this chapter, we address the problem of robust joint source-channel coding, using delay-limited codes. In particular, we show that the optimum slope of the SDR curve can be achieved by a single modulation mapping. The rest of the chapter is organized as the following:

In section 5.1, the system model and the basic concepts are presented. Section 5.2 presents an analysis of some of the previous analog coding schemes, and their limitations. In section 5.3, we introduce a class of joint source-channel codes which have a self-similar structure, and achieve a better asymptotic performance, compared to the analog and hybrid digital - analog coding schemes. The asymptotic performance of these codes, in terms of SDR scaling, is comparable with the scheme presented in [60, 61], but with a simpler structure and a shorter delay. We investigate the limits of the asymptotic performance of self-similar coding schemes and its relation with the Hausdorff dimension of the modulation signal set. In section 5.4, we present a single mapping which achieves the optimum slope of the SDR curve , which is equal to the bandwidth expansion factor. Although this mapping achieves the optimum slope of the SDR curve, its gap with the optimum SDR curve is unbounded (in terms of dB). In section 5.5, we construct a family of robust mappings, which individually achieve the optimum SDR slope, and together, maintain a bounded gap with the optimum SDR curve. We also analyze the limits on the asymptotic performance of delay-limited HDA coding schemes.

5.1 System model and theoretical limits

We consider a memoryless $\{X_i\}_{i=1}^{\infty}$ uniform source with zero mean and variance $\frac{1}{12}$, i.e. $-\frac{1}{2} \leq x_i < \frac{1}{2}$. Also, the samples of the source sequence are assumed independent with identical distributions (i.i.d.). Although the focus of this chapter is on the uniform source, as we discussed in Appendix E, the asymptotic results are valid for all distributions which have a bounded probability density function (including unbounded sources, such as the Gaussian source).

The transmitted signal is sent over an additive white Gaussian noise (AWGN) channel. The problem is to map the one-dimensional signal to the N-dimensional channel space, such that the effect of the noise is minimized. This means that the data x, $-\frac{1}{2} \leq x < \frac{1}{2}$, is mapped to the transmitted vector $\mathbf{s} = (s_1, ..., s_N)$. At the receiver side, the received signal is $\mathbf{y} = \mathbf{s} + \mathbf{z}$ where $\mathbf{z} = (z_1, ..., z_N)$ is the additive white Gaussian noise with variance σ^2.

As an upper bound on the performance of the system, we can consider the case of delay-unlimited. In this case, we can use Shannon's theorem on the separation of source and channel coding. By combining the lower bound on the distortion of the quantized signal (using the rate-distortion formula) and the capacity of N parallel Gaussian channels with the noise variance σ^2, we have [86]

$$D \geq c\sigma^{2N} \tag{5.1}$$

where c is a constant number and D is the average distortion.

5.2 Codes based on dynamical systems and hybrid digital-analog coding

Previously, two related schemes, based on dynamical systems, have been proposed for the scenario of delay-limited analog coding:

1. Shift-map dynamical system [8]

2. Spherical shift-map dynamical system [86]

5.2.1 Shift-map dynamical system

In [8], an analog transmission scheme based on shift-map dynamical systems is presented. In this method, every analog data x is mapped to the modulated vector $(s_1, ..., s_N)$ where

$$s_1 = x \quad \mod 1 \tag{5.2}$$

$$s_{i+1} = b_i s_i \mod 1, \quad \text{for } 1 \leq i \leq N-1 \tag{5.3}$$

where b_i is an integer number, $b_i \geq 2$. The set of modulated signals generated by the shift map consists of $b = b_1 ... b_{N-1}$ parallel segments inside an N-dimensional unit hypercube. In [86], the authors have shown that by appropriately choosing the parameters $\{b_i\}$ for different SNR values, one can achieve the SDR scaling (versus the channel SNR) with the slope $N - \epsilon$, for any positive number ϵ. Indeed, we can have a slightly tighter upper bound on the end-to-end distortion:

Theorem 5.1. *Consider the shift-map analog coding system which maps the source sample to an N-dimensional modulated vector. For any noise variance $\sigma^2 \leq \frac{1}{2}$, we can find a number a such that for the shift-map scheme with the parameters $b_i = a \geq 2$, the distortion of the decoded signal D is bounded as*

$$D \leq c\sigma^{2N}(-\log_2 \sigma)^{N-1} \tag{5.4}$$

where c is some constant number depending only on N.

Proof: See Appendix C. ■

Also, we have the following lower bound on the end-to-end distortion:

Theorem 5.2. *For any shift-map analog coding scheme, the output distortion is lower bounded as*

$$D \geq c'\sigma^{2N}(-\log \sigma)^{N-1} \tag{5.5}$$

where c' is a constant number, independent of σ (depending only on N).

Proof: See Appendix D. ■

5.2.2 Spherical shift-map dynamical system

In [86], a spherical code based on the linear system $\dot{\mathbf{s}}_T = \mathbf{A}\mathbf{s}_T$ is introduced, where $\mathbf{s}_T$ is the $2N$-dimensional modulated signal and $\mathbf{A}$ is a skew-symmetric matrix, i.e. $\mathbf{A}^T = -\mathbf{A}$.

This scheme is very similar to the shift-map scheme. Indeed, with an appropriate change of coordinates, the above modulated signal can be represented as

$$\mathbf{s}_T = \frac{1}{\sqrt{N}} \big(\cos 2\pi x, \cos 2a\pi x, ..., \cos 2a^{N-1}\pi x,$$

$$\sin 2\pi x, \sin 2a\pi x, ..., \sin 2a^{N-1}\pi x\big) \tag{5.6}$$

for some parameter a.

If we consider $\mathbf{s}_{sm}$ as the modulated signal generated by the shift-map scheme with parameters $b_i = a$ in (5.3), then, (5.6) can be written in the vector form as

$$\mathbf{s}_T = \left(Re\left\{e^{\pi i \mathbf{s}_{sm}}\right\}, Im\left\{e^{\pi i \mathbf{s}_{sm}}\right\}\right). \tag{5.7}$$

The relation between the spherical code and the linear shift-map code is very similar to the relation between PSK and PAM modulations. Indeed, the spherical shift-map code and PSK modulation are, respectively, the linear shift-map and PAM modulations which are transformed from the unit interval, $[\frac{-1}{2}, \frac{1}{2})$, to the unit circle.

For the performance of the spherical codes, the same result as Theorem 5.1 is valid. Indeed, for any parameters a and N, the spherical code asymptotically has a saving of $\frac{(2\pi)^2}{12}$ or 5.17 dB in the power. This asymptotic gain results from transforming the unit-interval signal set (with length 1 and power $\frac{1}{12}$) to the unit-circle signal set (with length 2π and power 1) . However, the spherical code uses $2N$ dimensions (compared to N dimensions for the linear shift-map scheme).

For both these methods, for any fixed parameter a, the output SDR asymptotically has linear scaling with the channel SNR. The asymptotic gain (over the simple

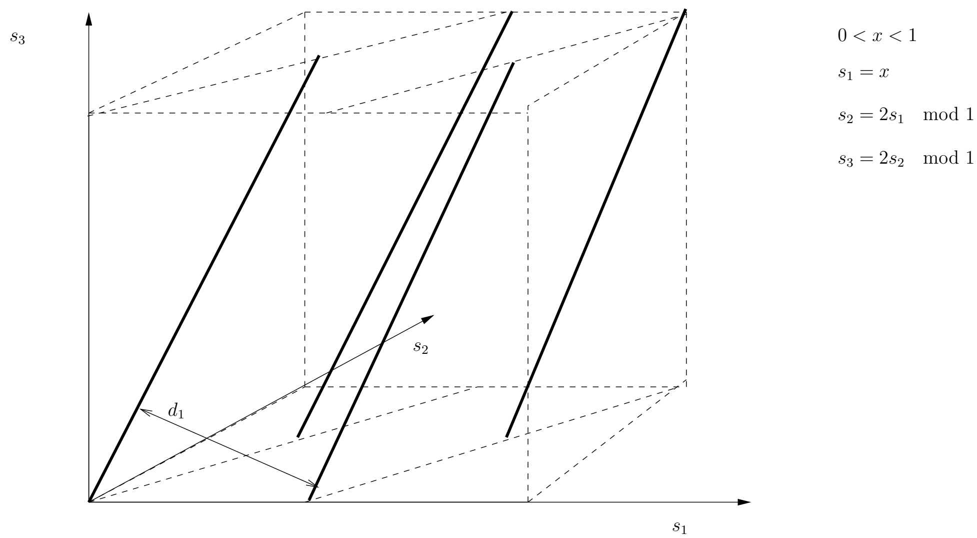

Figure 5.1: The shift-map modulated signal set for $N = 3$ dimensions and $a = 2$.

repetition code) is approximately $a^{2(N-1)}$ (because the modulated signal is stretched approximately a^{N-1} times)[1]. Therefore, a larger scaling parameter a results in a higher asymptotic gain. However, by increasing a, the distance between the parallel segments of the modulated signal set decreases. This distance is approximately $\frac{1}{a}$ and for the low SNRs (when the noise variance is larger than or comparable to $\frac{1}{a}$), jumping from one segment of the modulated signal set to another one becomes the dominant factor in the distortion of the decoded signal which results in a poor performance in this SNR region. Thus, there is a trade-off between the gain in the high-SNR region and the critical noise level which is fatal for the system. By increasing the scaling parameter a, the asymptotic gain increases, but at the same time, a higher SNR threshold is needed

[1]The exact asymptotic gain is equal to the scaling factor of the signal set, i.e. $a^{2(N-1)}\left(1 + \frac{1}{a^2} + ... + \frac{1}{a^{N-1}}\right)$ for the shift map and $\frac{(2\pi)^2}{12}a^{2(N-1)}\left(1 + \frac{1}{a^2} + ... + \frac{1}{a^{N-1}}\right)$ for the spherical shift map.

to achieve that gain. In [59], the authors have combined the dynamical-system schemes with LDPC and iterative decoding to reduce the critical SNR threshold. However, overall behavior of the output distortion is the same for all these methods. Also, in [42] and [43], a scheme is introduced for approaching arbitrarily close to the optimum SDR, for coloured sources. However, it is not delay-limited and it only works for bandwidth expansion 1.

The shift-map analog coding system can be seen as a variation of a Hybrid Digital-Analog (HDA) joint source-channel code. Various types of these hybrid schemes are investigated in [63], [49], [68], [34] and [69]. Indeed, for the shift-map system, we can rotate the modulated signal set such that all the parallel segments of it become aligned in the direction of one of the dimensions. In this case, by changing the support region of the modulated set (which is a rotated N-dimensional cube) to the standard cube, we obtain a new similar modulation which is hybrid digital-analog and has almost the same performance. In the new modulation, the information signal is quantized by a^{N-1} points in an $(N-1)$-dimensional sub-space and the quantization error is sent over the remaining dimension.

Regarding the scaling of the output distorsion, the performance of the shift-map scheme, with appropriate choice of parameters for each SNR, is very close to the theoretical limit. In fact, the output distortion scales as $\sigma^{2N}(-\log \sigma)^{N-1}$, instead of being proportional to σ^{2N}. However, for any fixed set of parameters, the output SNR (versus the input SNR) is saturated by the unit slope (instead of N). This shortcoming is an inherent drawback of schemes like the shift-map code or spherical code (which are based on dynamical systems). Indeed, in [97], it is shown that no single differentiable mapping can achieve an asymptotic slope better than 1. This chapter addresses this shortcoming.

There are some other analog codes in the literature which use different mappings. Analog codes based on the 2-dimensional Shannon map [13] [57] [35] [5], or the tent map [8] are examples of these codes. However, all these codes share the shortcomings of the shift-map code.

5.3 Joint source-channel codes based on fractal sets

In this section, we propose a coding scheme, based on the fractal sets, that can achieve slopes greater than 1 (for the curve of SDR versus SNR).

Scheme I: For the modulating signal x, $-\frac{1}{2} \leq x < \frac{1}{2}$, we consider the binary expansion of $x + \frac{1}{2}$:

$$x + \frac{1}{2} = \left(\overline{0 \cdot b_1 b_2 b_3 ...}\right)_2 . \tag{5.8}$$

Now, we construct s_1, s_2, ..., s_N as

$$s_1 = \left(\overline{0 \cdot b_1 b_{N+1} b_{2N+1} ...}\right)_\alpha \tag{5.9}$$

$$s_2 = \left(\overline{0 \cdot b_2 b_{N+2} b_{2N+2} ...}\right)_\alpha \tag{5.10}$$

$$...$$

$$s_N = \left(\overline{0 \cdot b_N b_{2N} b_{3N} ...}\right)_\alpha \tag{5.11}$$

where $\left(\overline{0 \cdot b_1 b_2 b_3 ...}\right)_\alpha$ is the base-α expansion[2].

Theorem 5.3. *In the proposed scheme, for any $\alpha > 2$, the output distortion D is upper bounded by*

$$D \leq c\sigma^{2\beta}(-\log\sigma)^N \tag{5.12}$$

where c is independent of σ and $\beta = N\frac{\log 2}{\log \alpha}$.

Proof: Consider w_i as the Gaussian noise on the ith dimension:

$$\Pr\left\{|w_i| > 2\sqrt{N}\sigma\sqrt{-\log\sigma}\right\} = \tag{5.13}$$

$$2Q\left(2\sqrt{N}\sqrt{-\log\sigma}\right) \leq e^{-\frac{4N(-\log\sigma)}{2}} = e^{-2N(-\log\sigma)} = \sigma^{2N} \tag{5.14}$$

Now, we bound the distortion, conditioned on $|w_i| \leq 2\sqrt{N}\sigma\sqrt{-\log\sigma}$ for $1 \leq i \leq N$. If the kth digit of s_i and s'_i are different,

$$|s_i - s'_i| \geq \tag{5.15}$$

$$\left(\overline{0 \cdot \underbrace{0...0}_{k-1}1000..}\right)_\alpha - \left(\overline{0 \cdot \underbrace{0...0}_{k-1}0111...}\right)_\alpha \tag{5.16}$$

$$> (\alpha - 2)\alpha^{-(k+1)} \tag{5.17}$$

Therefore, if $|s_i - s'_i| \leq \delta$ for any $\delta > 0$, the first k digits of s_i and s'_i are the same, where $k \geq \left\lfloor -\log_\alpha\left(\frac{\delta}{\alpha-2}\right)\right\rfloor - 1$. Now, by considering $\delta = 4\sqrt{N}\sigma\sqrt{-\log\sigma}$,

[2] In this chapter, we define the base-α expansion, for any real number $\alpha \geq 2$ and any binary sequence $(b_1 b_2 b_3 ...)$, as $\left(\overline{0 \cdot b_1 b_2 b_3 ...}\right)_\alpha \triangleq \sum_{i=1}^{\infty} b_i \alpha^{-i}$.

$$|s_i - s_i'| \le 2|w_i| \le 4\sqrt{N}\sigma\sqrt{-\log\sigma} \tag{5.18}$$

$$\Longrightarrow k \ge \left\lfloor -\log_\alpha\left(\frac{4\sqrt{N}\sigma\sqrt{-\log\sigma}}{\alpha-2}\right)\right\rfloor - 1 \tag{5.19}$$

Therefore, for $1 \le i \le N$, the first

$$\left\lfloor -\log_\alpha\left(\frac{4\sqrt{N}\sigma\sqrt{-\log\sigma}}{\alpha-2}\right)\right\rfloor - 1$$

digits of $s_1, s_2, ..., s_N$ can be decoded without any error, hence, the first

$$N\left(\left\lfloor -\log_\alpha\left(\frac{4\sqrt{N}\sigma\sqrt{-\log\sigma}}{\alpha-2}\right)\right\rfloor - 1\right)$$

bits of the binary expansion of x can be reconstructed perfectly. In this case, the output distortion is bounded by

$$\sqrt{D} \le 2^{-N\left(\left\lfloor -\log_\alpha\left(\frac{4\sqrt{N}\sigma\sqrt{-\log\sigma}}{\alpha-2}\right)\right\rfloor - 1\right)} \tag{5.20}$$

$$\le c_1\sigma^\beta(-\log\sigma)^{\frac{N}{2}} \tag{5.21}$$

where c_1 depends only on α and N.

By combining the upper bounds for the two cases,

$$D \le \Pr\left\{|w_i| > 2\sqrt{N}\sigma\sqrt{-\log\sigma}\right\} + c_1\sigma^\beta(-\log\sigma)^{\frac{N}{2}} \tag{5.22}$$

$$\le c\sigma^{2\beta}(-\log\sigma)^N. \tag{5.23}$$

■

According to the theorem 5.2, for any $\epsilon > 0$, we can construct a modulation scheme that achieves the asymptotic slope of $N - \epsilon$ (for the curve of SDR versus SNR, in terms of dB). As expected (according to the result by Ziv [97]), none of these mappings are differentiable. More generally, Ziv has shown that [97]:

Theorem 5.4. *([97], Theorem 2) For the modulation mapping* $\mathbf{s} = f(x)$*, define*

$$d_f(\Delta) = \mathbb{E}_x \left(\|f(x + \Delta) - f(x)\|^2 \right).$$

If there are positive numbers M, γ, Δ_0 *such that*

$$d_f(\Delta) \leq M\Delta^\gamma \quad \text{for } \Delta \leq \Delta_0 \tag{5.24}$$

then there is constant c *such that*

$$D \geq c\sigma^{\frac{2}{\gamma}}. \tag{5.25}$$

By decreasing α, we can increase the asymptotic slope β. However, it also degrades the low-SNR performance of the system. This phenomenon is observed in figure 5.3.

In scheme I, the signal set is a self-similar fractal [16], where the parameter β, which determines the asymptotic slope of the curve, is the dimension of the fractal. There are different ways to define the fractal dimension. One of them is the Hausdorff dimension. Consider F as a Borel set in a metric space, and $\mathfrak{A}$ as a countable family of sets that covers it. We define $H^s_\varepsilon(F) = \inf \sum_{A \in \mathfrak{A}} (\text{diameter}(A))^s$, where the infimum is over all countable covers that diameter of their sets are not larger than ε. The s-dimensional Hausdorff space is defined as $H^s(F) = \lim_{\varepsilon \to 0} H^s_\varepsilon(F) = \sup_{\varepsilon > 0} H^s_\varepsilon(F)$. It can be shown that there is a critical value s_0 such that for $s < s_0$, this measure is infinite and for $s > s_0$, it is zero [16]. This critical value s_0 is called the Hausdorff dimension of the set F.

Another useful definition is the box-counting dimension. If we partition the space into a grid cubic boxes of size ε, and consider N_ε as the number of boxes which intersect the set F, the box-counting dimension of F is defined as

$$\mathrm{Dim}_b(F) = \lim_{\varepsilon \to 0} \frac{\log N_\varepsilon}{\log \frac{1}{\varepsilon}} \tag{5.26}$$

It can be shown that for regular self-similar fractals, the Hausdorff dimension is equal to the box-counting dimension [16]. Intuitively, theorem 5.3 means that in scheme I among the N available dimensions, only β dimensions are effectively used. Indeed, we can show that for any modulation set with box-counting dimension β, the asymptotic slope of the SDR curve is at most β:

Theorem 5.5. *For a modulation mapping $\mathbf{s} = f(x)$, if the modulation set F has box-counting dimension β, then,*

$$\lim_{\sigma \to 0} \frac{\log D}{\log \sigma} \leq 2\beta. \tag{5.27}$$

Proof: We divide the space to the boxes of size σ. Consider N_σ as the number of cubic boxes that cover F. We divide the source signal set to $4N_\sigma$ segments of length $\frac{1}{4N_\sigma}$. Consider $\mathcal{A}_1, ..., \mathcal{A}_{4N_\sigma}$ as the corresponding N-dimensional optimal decoding region (based on the MMSE criterion), and $\mathcal{B}_1, ..., \mathcal{B}_{4N_\sigma}$ as their intersection with the N_σ cubes (see figure 5.2). Total volume of these $4N_\sigma$ sets is equal to the total volume of the covering boxes, i.e. $N_\sigma \sigma^N$. Thus, at least, half of these sets (i.e. $2N_\sigma$ of them) have volume less than $\frac{1}{2}\sigma^N$. For any of these sets such as $\mathcal{B}_i$ and any box, the volume of intersection of that box with other sets is at least $V_{min} = \sigma^N - \frac{1}{2}\sigma^N = \frac{1}{2}\sigma^N$. For any point in the corresponding segments of the set $\mathcal{B}_i$, the probability of decoding to a wrong segment is lower bounded by the probability of a jump to the neighboring sets in the same box. Because the variance of the additive Gaussian noise is σ^2 per each dimension, and for such a jump the squared norm of the noise at most needs to be $N\sigma^2$ (square of the diameter of the box), the probability of such a jump can be bounded as

$$\Pr(jump) \geq V_{min} \cdot \min_{\|\mathbf{z}\|^2 \leq N\sigma^2} f_{\mathbf{z}}(\mathbf{z}) \tag{5.28}$$

$$\geq \frac{1}{2}\sigma^N \cdot \frac{1}{(2\pi)^{\frac{N}{2}}\sigma^N} e^{-\frac{N\sigma^2}{2\sigma^2}} = \frac{1}{2^{\frac{N}{2}+1}\pi^{\frac{N}{2}}} e^{-\frac{N}{2}}. \tag{5.29}$$

Now, for these segments of the source, consider the subsegments with length $\frac{1}{20N_\sigma}$ at the center of them. For these subsegments whose total length is at least $\frac{1}{20N_\sigma} \cdot 2N_\sigma = \frac{1}{10}$, at least with probability $\Pr(jump)$, we have a squared error which is not less than $\left(\frac{1}{10N_\sigma}\right)^2$. Thus,

$$D \geq \frac{1}{10} \Pr(jump) \cdot \left(\frac{1}{10N_\sigma}\right)^2 = \frac{c}{N_\sigma^2} \tag{5.30}$$

where c only depends on the bandwidth expansion N.

On the other hand, based on the definition of the box-counting dimension,

$$\beta = \lim_{\sigma \to 0} \frac{\log N(\sigma)}{\log \frac{1}{\sigma}}. \tag{5.31}$$

By using (5.30) and (5.31),

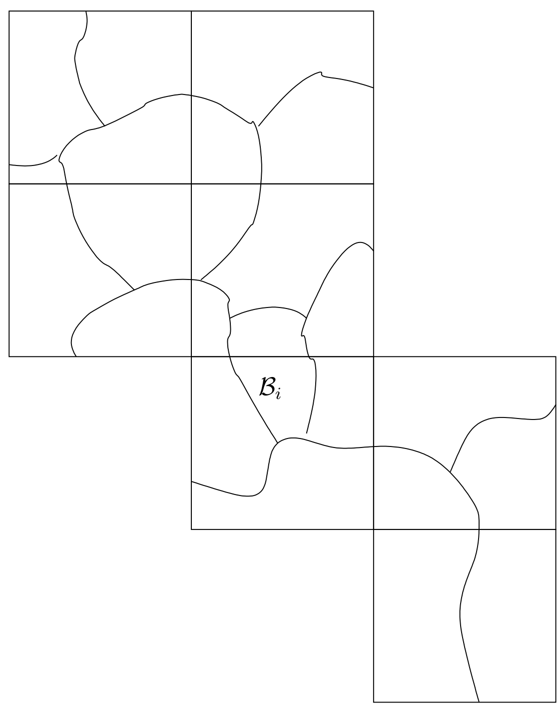

Figure 5.2: Boxes of size σ and their intersections with the decoding regions

$$\lim_{\sigma \to 0} \frac{\log D}{\log \sigma} \leq 2\beta. \tag{5.32}$$

■

It should be noted that theorem 5.5 is valid for all signal sets, not just self-similar signal sets. As a corollary, based on the fact that the box-counting dimension cannot be greater than the dimension of the space [16], theorem 5.5 gives a geometric insight to (5.1).

5.4 Achieving the optimum asymptotic SDR slope using a single mapping

Although scheme I can construct mappings that achieve near-optimum slope for the curve of SDR (versus the channel SNR), none of these mappings can achieve the optimum slope N. To achieve the optimum slope with a single mapping, we slightly modify the scheme I:

For the modulating signal x, consider $x + \frac{1}{2} = \left(\overline{0.b_1 b_2 b_3 ...}\right)_2$. We construct s_1, s_2, ..., s_N as

$$s_1 = \left(\overline{0.b_1 0 b_{\frac{N(N+1)}{2}+1} b_{\frac{N(N+1)}{2}+2} ... b_{\frac{N(N+1)}{2}+N+1} 0 b_{\frac{(2N)(2N+1)}{2}+1} ...}\right)_2 \tag{5.33}$$

$$s_2 = \left(\overline{0.b_2 b_3 0 b_{\frac{(N+1)(N+2)}{2}+1} b_{\frac{(N+1)(N+2)}{2}+2} ... b_{\frac{(N+1)(N+2)}{2}+N+2} 0 ...}\right)_2 \tag{5.34}$$

$$...$$

$$...$$

$$s_N = \left(\overline{0.b_{\frac{N(N-1)}{2}+1} b_{\frac{N(N-1)}{2}+2} ... b_{\frac{N(N+1)}{2}} 0...}\right)_2 \tag{5.35}$$

The difference between this scheme and Scheme I is that instead of assigning the $kN+i$th bit to the signal s_i, the bits of the binary expansion of $x+\frac{1}{2}$ are grouped such that the lth group ($l = kN+i$) consists of l bits and is assigned to the ith dimension.

Theorem 5.6. *Using the mapping constructed by Scheme II, the output distortion D is upper bounded by*

$$D \leq c_1 \sigma^{2N} 2^{c_2\sqrt{-\log\sigma}} \tag{5.36}$$

where c_1 and c_2 are only dependet on N.

Proof: Consider z_i as the Gaussian noise on the ith channel and assume that n is selected such that

$$\sum_{k=1}^{n} kN+i \leq -\log_2\sigma < \sum_{k=1}^{n+1} kN+i \tag{5.37}$$

The probability that $|z_i| \geq 2^{-1-\sum_{k=1}^{n-1} kN+i}$ is negligible. Indeed,

$$\Pr\left\{ |z_i| \geq 2^{-1-\sum_{k=1}^{n-1} kN+i} \middle| -\log_2\sigma \geq \sum_{k=1}^{n} kN+i \right\} \leq \tag{5.38}$$

$$2Q\left(\frac{2^{-1-\sum_{k=1}^{n-1} kN+i}}{2^{-\sum_{k=1}^{n} kN+i}}\right) = 2Q\left(2^{nN-1}\right). \tag{5.39}$$

On the other hand, when $|z_i| < 2^{-1-\sum_{k=1}^{n-1} kN+i}$, the first $\sum_{k=1}^{n-1} kN + i$ bits of s_i can be decoded error-freely. The same is true for all $1 \le i' \le i$, and for $i < i' \le N$, the first $\sum_{k=1}^{n-2} kN + i'$ can be decoded error-freely. Thus, the first $\sum_{j=1}^{(n-1)N+i} j$ bits of x can be decoded error-freely. Now,

$$\sum_{j=1}^{(n-1)N+i} j \ge \tag{5.40}$$

$$N\sum_{k=1}^{n-2} kN + i \ge \tag{5.41}$$

$$N\sum_{k=1}^{n+1} (kN+i) - N\left((n-1)N + i + nN + i + (n+1)N + i\right) \ge \tag{5.42}$$

$$N\sum_{k=1}^{n+1} kN + i - N^2\,(3n+3) \ge \tag{5.43}$$

$$N\sum_{k=1}^{n+1} kN + i - c_3\sqrt{\sum_{k=1}^{n} kN + i} \tag{5.44}$$

where c_3 depends only on N. Therefore, by using the assumption (5.37),

$$\sum_{j=1}^{(n-1)N+i} j \ge \tag{5.45}$$

$$-N\log_2\sigma - c_3\sqrt{-\log_2\sigma} \tag{5.46}$$

Consequently, the output distortion is bounded by

$$D \leq \Pr\left\{|z_i| \geq 2^{-1-\sum_{k=1}^{n-1} kN+i}\right\} + 2^{-2\sum_{j=1}^{(n-1)N+i} j} \tag{5.47}$$

$$\leq 2Q\left(2^{nN-1}\right) + 2^{2N\log_2\sigma + 2c_3\sqrt{-\log_2\sigma}} \tag{5.48}$$

$$\Longrightarrow D \leq c_1\sigma^{2N}2^{c_2\sqrt{-\log\sigma}}. \tag{5.49}$$

■

It should be noted that in the this proof, the assumption of having a uniform distribution is not used, and the above proof is valid for any source whose samples are in the interval $\left[-\frac{1}{2}, \frac{1}{2}\right)$. In Appendix C, we extend the scheme proposed in this section to other sources which are not necessarily bounded.

5.5 Approaching a near-optimum SDR by delay-limited codes

In [49], a family of hybrid digital-analog (HDA) source-channel codes are proposed which together can achieve the optimum SDR curve and each of them only suffers from the mild saturation effect (the asymptotic unit slope for the curve of SDR versus SNR). However, their approach is based on using capacity-approaching digital codes as a component of their scheme. In [58], it is shown that for any joint source-channel code that touches the optimum SDR curve in a certain SNR, the asymptotic slope cannot be better than one.

In this section, we consider the problem of finding a family of delay-limited analog codes which together have a bounded asymptotic loss in the SDR performance (in terms of dB). Results of Section 5.2 show that none the previous analog coding schemes (based on dynamical systems) can construct such a family of codes. In this section, we also show that no HDA source-channel coding scheme can be used to achieve it. In the HDA source-channel coding, in general, to map an M dimensional source to an N dimensional signal set, the source is quantized by κ points which are sent over $N-M$ dimensions and the residual noise is transmitted over the remaining M dimensions. In other words, the region of the source (which is a hypercube for the case of a uniform source) is divided into κ subregions $\mathcal{A}_1, ..., \mathcal{A}_\kappa$. These subregions are mapped to κ parallel subsets of the N dimensional Euclidean space, $\mathcal{A}'_1, ..., \mathcal{A}'_\kappa$, where $\mathcal{A}'_i$ is a scaled version of $\mathcal{A}_i$ with a factor of a.

Theorem 5.7. *Consider a HDA joint source-channel code which maps an M-dimensional uniform source (inside the unit cube) to κ parallel M-dimensional subsets of an N dimensional Euclidean space ($N > M$), with a power constraint of P. If the decoding of digital and analog parts are done separately, for any noise variance σ^2 and any integer κ, the output distortion is bounded by*

$$D \geq c\sigma^{\frac{2N}{M}}(-\log\sigma)^{\frac{N-M}{M}} \tag{5.50}$$

where c is a constant number (independent of SNR).

Proof: See Appendix F. ■

Now, we construct families of delay-limited analog codes which by a proper choice of parameters (according to the channel SNR) have only a bounded asymptotic loss in the SDR performance (in terms of dB).

Type I - Family of piece-wise linear mappings: For any $2^{-k-1} < \sigma \leq 2^{-k}$, for $k > 0$, we construct an analog code as the following:

For $x + \frac{1}{2} = \left(\overline{0 \cdot b_1 b_2 ... b_{Nk-1}}\right)_2 + \frac{\{2^{Nk-1}x\}}{2^{Nk-1}}$, where $\{\cdot\}$ represents the fractional part, we construct $s_1, s_2, ..., s_N$ as

$$s_1 = \sum_{i=1}^{k} (2^{-i} + 2^{-k}(k-i)) b_{(i-1)N+1}$$

$$s_2 = \sum_{i=1}^{k} (2^{-i} + 2^{-k}(k-i)) b_{(i-1)N+2}$$

$$...$$

$$s_{N-1} = \sum_{i=1}^{k} (2^{-i} + 2^{-k}(k-i)) b_{(i-1)N+N-1}$$

$$s_N = \sum_{i=1}^{k-1} (2^{-i} + 2^{-k}(k-i)) b_{(i-1)N+N} + 2^{Nk-k-2} \frac{\{2^{Nk-1}x\}}{2^{Nk-1}} \qquad (5.51)$$

First, we show that the $0 \leq s_i < 2$, for $1 \leq i \leq N$:

$$s_i \leq \sum_{i=1}^{k} (2^{-i} + 2^{-k}(k-i)) + 2^{-k-1} = \sum_{i=1}^{k+1} 2^{-i} + 2^{-k} \sum_{i=1}^{k} (k-i) \qquad (5.52)$$

$$< 1 + 2^{-k} \cdot \frac{k(k-1)}{2} < 2. \qquad (5.53)$$

Therefore, by an appropriate shift (e.g. modifying the transmitted signal set as $\mathbf{s}' = \mathbf{s}$), the transmitted power can be bounded by one. Next, we show that the proposed scheme has a bounded gap (in terms of dB) to the optimum SDR curve:

Theorem 5.8. *In the proposed scheme, the output distortion D is upper bounded by*

$$D \leq c\sigma^{2N} \tag{5.54}$$

where c is a constant, independent of σ.

Proof: The signal set consists of 2^{Nk-1} segments of length 2^{-k-1}, where each of them is a scaled version of a subsegment of the source region (the unit interval), by a factor of 2^{Nk-k-2}.

The probability that the first error occurs in the lth bit ($l = (i-1)N + j$) of x is bounded by $P_l \leq 2Q\left(\frac{k-i}{2}\right) \leq 2Q\left(\frac{k}{2} - \frac{l}{2N}\right)$ and it results in an output squared error of at most $D_l \leq 4^{-1+1} = 4^{-(i-1)N-j+1}$. Therefore, by considering the union-bound over all possible errors, we obtain

$$\begin{aligned} D &\leq \sum_{l=1}^{Nk-1} D_l \cdot P_l + D_{no-bit-error} \\ &\leq \sum_{l=1}^{Nk-1} 4^{-l+1} \cdot 2Q\left(\frac{k}{2} - \frac{l}{2N}\right) + 4^{-(Nk-k-2)}\sigma^2. \end{aligned} \tag{5.55}$$

Now, by using $Q(x) < e^{-\frac{x^2}{2}}$ and $2^{-k+1} < \sigma$ we have

$$D \leq \sum_{l=1}^{kN-1} 2^{-2l+3} e^{-\frac{(k-l/N)^2}{8}} + 4^{-(Nk-k-2)}\sigma^2$$

$$\leq \sum_{l=1}^{kN-N-1} 2^{-2l+3} e^{-\frac{(k-l/N)^2}{8}} + \sum_{l=Nk-N}^{Nk-1} 4^{-l+1} e^{-\frac{(k-l/N)^2}{8}} + 4^{-(Nk-k-2)}\sigma^2$$

$$\leq \sum_{l=1}^{kN-N-1} 2^{-2l+3} e^{-\frac{(k-l/N)^2}{8}} + N \cdot 4^{-(k-1)N} + 4^{-(Nk-k-2)}\sigma^{2N}$$

$$\leq N \sum_{t=1}^{k-1} 4^{-tN+1} \cdot e^{-\frac{(k-t)^2}{8}} + N \cdot 4^{-(k-1)N} + 4^{-(Nk-k-2)}\sigma^{2N}$$

$$\leq N \cdot 4^{-(k-1)N+1} \sum_{t=1}^{\infty} 4^{N(k-t)} \cdot e^{-\frac{(k-t)^2}{8}} + N \cdot 4^{-(k-1)N} + 4^{-(Nk-k-2)}\sigma^{2N}$$

$$\leq c_9 4^{-(k-1)N} + 4^{-(Nk-k-2)}\sigma^2$$

$$\leq c\sigma^{2N}. \tag{5.56}$$

■

It is worth noting that in the proposed family of codes, for each code, the asymptotic slope of the SDR curve is 1 (as we expected from the fact that for each code, the mapping is partially differentiable). We can mix the idea of this scheme with Scheme II of the previous section, to construct a family of mappings where for each of them, the asymptotic slope is N, and together, they maintain a bounded gap with the optimal SDR (in terms of dB):

Type II - Family of robust mappings: For $x + \frac{1}{2} = \left(\overline{0 \cdot b_1 b_2 b_3 ...}\right)_2$, we construct $f_k(x) = (s_1, s_2, ..., s_N)$ as

$$s_1 = \sum_{i=1}^{k} (2^{-i} + 2^{-k}(k-i)) b_{(i-1)N+1} + 2^{-k-1} \left(\overline{0 \cdot b_{kN+1} 0 b_{\frac{kN+N(N+1)}{2}+1} b_{kN+\frac{N(N+1)}{2}+2} \cdots} \right)_2$$

$$s_2 = \sum_{i=1}^{k} (2^{-i} + 2^{-k}(k-i)) b_{(i-1)N+2} + 2^{-k-1} \left(\overline{0 \cdot b_{kN+2} b_3 0 b_{kN+\frac{(N+1)(N+2)}{2}+1} \cdots} \right)_2$$

$$\cdots$$

$$s_N = \sum_{i=1}^{k} (2^{-i} + 2^{-k}(k-i)) b_{(i-1)N+N} + 2^{-k-1} \left(\overline{0 \cdot b_{kN+\frac{N(N-1)}{2}+1} b_{kN+\frac{N(N-1)}{2}+2} \cdots} \right)_2$$

Theorem 5.9. *In the proposed family of mappings (Type II), there are constants c_1, c_2, c_3, independent of σ and k (only dependent on N) such that for every integer $k > 0$, if we use the modulation map $f_k(x)$,*

i) For $2^{-k-1} < \sigma \le 2^{-k}$,

$$D \le c_3 \sigma^{2N}. \tag{5.57}$$

ii) for any $\sigma < 2^{-k-1}$,

$$D \le c_1 \sigma^{2N} 2^{c_2 \sqrt{-\log \sigma}}. \tag{5.58}$$

Proof: i) The probability that the first error occurs in the lth bit ($l = (i-1)N + j < kN$) of x is bounded by $P_l \ge Q(k-i)$ and it results in an output squared error of at most $4^{(i-1)N-j+1}$, and when there is no error in the first Nk bits, the squared error is $D' \le 4^{-Nk}$. Therefore, by considering the union-bound over all possible errors, we have

$$D \le \sum_{l=1}^{Nk-1} D_l \cdot P_l + D_{no-bit-error}$$

$$\leq \sum_{l=1}^{Nk-1} 4^{-l+1} \cdot 2Q\left(\frac{k}{2} - \frac{l}{2N}\right) + 4^{-Nk}$$

Similar to the proof of theorem 8, by using $Q(x) < e^{-\frac{x^2}{2}}$ and $2^{-k+1} < \sigma \leq 2^{-k}$ we have

$$D \leq \sum_{l=1}^{Nk-1} 4^{-l+1} e^{-\frac{(k-l/N)^2}{8}} + \sigma^{2N}$$
$$\leq c_{12} 4^{-kN} + \sigma^{2N}$$
$$\leq c\sigma^{2N}.$$

ii) Consider z_i as the Gaussian noise on the ith channel and assume that n is selected such that

$$k + \sum_{l=1}^{n} lN + i \leq -\log_2 \sigma < k + \sum_{l=1}^{n+1} lN + i \tag{5.59}$$

The probability that $|z_i| \geq 2^{-k+1-\sum_{l=1}^{n-1} lN+i}$ is negligible.

On the other hand, when $|z_i| < 2^{-k+1-\sum_{l=1}^{n-1} lN+i}$, the first $k + \sum_{l=1}^{n-1} lN + i$ bits of s_i can be decoded error-freely. The same is true for all $1 \leq i' \leq i$, and for $i < i' \leq N$, the first $k + \sum_{l=1}^{n-2} lN + i'$ can be decoded error-freely. Thus, the first $kN + \sum_{j=1}^{(n-1)N+i} j$ bits of x can be decoded error-freely. Now, similar to the proof of theorem 6,

$$kN + \sum_{j=1}^{(n-1)N+i} j \geq \tag{5.60}$$

$$N \sum_{k=1}^{n+1} kN + i - c_2 \sqrt{\sum_{k=1}^{n+1} kN + i} \tag{5.61}$$

where c_2 depends only on N. Therefore, by using the assumption (5.59),

$$kN + \sum_{j=1}^{(n-1)N+i} j \geq \tag{5.62}$$

$$-N \log_2 \sigma - c_2 \sqrt{-\log_2 \sigma} \tag{5.63}$$

Therefore, the output distortion is bounded by

$$D \leq 2^{-2 \sum_{j=1}^{(n-1)N+i} j} \tag{5.64}$$

$$\leq 2^{2N \log_2 \sigma + 2c_2 \sqrt{-\log_2 \sigma}} \tag{5.65}$$

$$\Longrightarrow D \leq c_1 \sigma^{2N} 2^{c_2 \sqrt{-\log \sigma}}. \tag{5.66}$$

■

5.6 Simulation results

In figure 5.3, for a bandwidth expansion factor of 4, the performance of Scheme I (with parameters $\alpha = 4$ and 3) is compared with the shift-map scheme with $a = 3$. As we expect, for the shift-map scheme, the signal-to-distortion-ratio (SDR) curve saturates at slope 1, while the new scheme offers asymptotic slopes higher than one. For the new scheme, with parameters $\alpha_1 = 4$ and $\alpha_2 = 3$, the asymptotic slope is respectively $\beta_1 = \frac{4 \log 2}{\log 4} = 2$ and $\beta_2 = \frac{4 \log 2}{\log 3}$ (as expected from Theorem 5.3). Also, we see that the new scheme provides a gracefull degradation in the low SNR region.

Although Schemes I and II deal with the infinite binary expansion of $x + \frac{1}{2}$ (which is not practically feasible), approximated versions of this mapping can be implemented

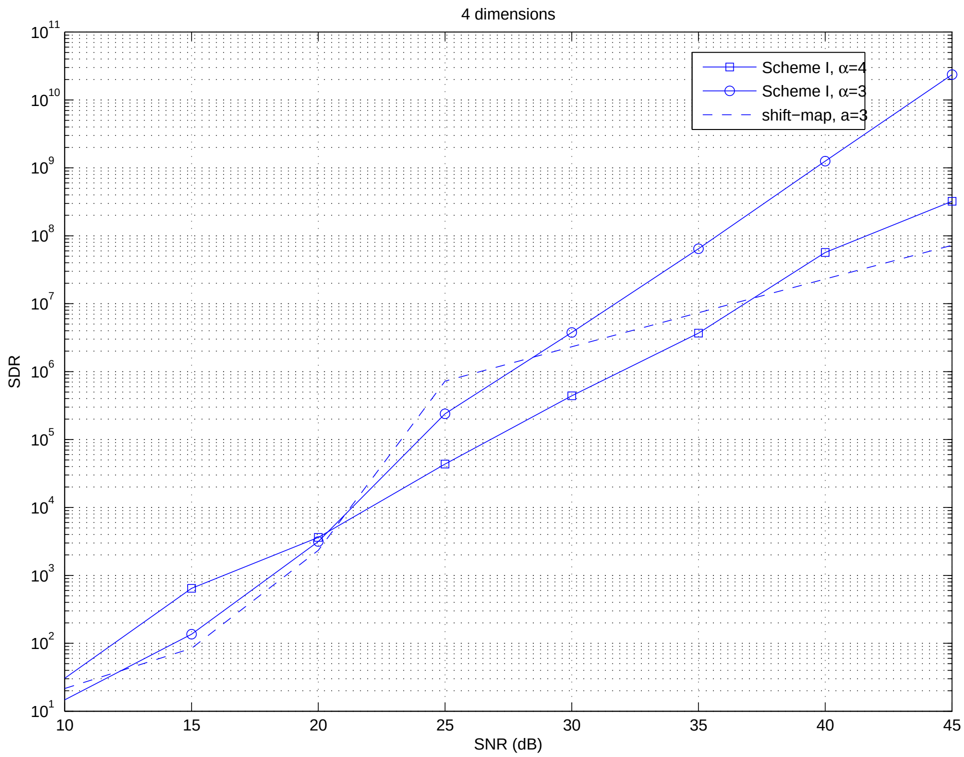

Figure 5.3: The SDR for the first proposed scheme (with $\alpha = 4$ and 3) and the shift-map scheme with $a = 3$. The bandwidth expansion is $N = 4$.

by low-complexity encoders and decoders. A simple approximation of these methods is to split $x + \frac{1}{2}$ as

$$x + \frac{1}{2} = \overline{0.b_1 b_2 ... b_n} + r \tag{5.67}$$

and applying the proposed methods on $\overline{0.b_1 b_2 ... b_n}$ and send r by a linear mapping over one of the dimensions.

5.7 Conclusions

To avoid the mild saturation effect in analog transmission systems and achieving the optimum scaling of the output distortion, we need to use nondifferentiable mappings (more precisely, mappings which are not differentiable on any interval). Two nondifferentiable schemes are introduced in this chapter. Both these schemes outperform the traditional analog schemes, in terms of scaling of the output SDR, and one of them almost achieves the optimum SDR scaling with a simple mapping (it asymptotically achieves slope N for the SDR curve). Also, simulation results show that Piecewise differentiable approximations of these methods offers an acceptable performance (with a low encoding/decoding complexity) for a wide range of channel SNRs.

Chapter 6

Joint source-channel coding over MIMO fading channels

In many applications, such as voice and multimedia transmission in cellular and wireless LAN environments, transmission of analog sources over wireless channels is needed. Results of the research in the past decade shows that using multiple-antenna systems can substantially improve the rate and the reliability of communication in wireless fading environments. However, until now, most of the research has been focused on the transmission of digital data over multiple-input multiple-output (MIMO) channels, and the study of the analog source transmission is still in its early stages.

Although in recent years there has been a colossal amount of research in MIMO communications, the research on MIMO joint source-channel coding is still in its early stages. In [30] and [6] some digital and hybrid digital-analog techniques are examined for joint source-channel coding over MIMO channels, and some bounds on the asymptotic exponents of the average distortion are presented. However, the asymptotic exponents of the average distortion is not a good measure for the performance evaluation of joint source-channel coding schemes in fading environments. The reason is that the average distortion is dominated by rare cases of very bad channels. In these environments, it is more informative to analyze the probability that distortion is greater than a certain level, instead of averaging very large unusual distortions.

Unlike [30] and [6] which are focused on analyzing the asymptotic exponents of the average distortion, here, we analyze the asymptotic exponents of the probability of having a large distortion. This measure, which we call diversity-fidelity tradeoff, can be seen as an analog version of the well-known diversity-multiplexing tradeoff which is proved to be very useful in evaluating various digital space-time coding schemes.

6.1 System Model

We consider a communication system where an analog source of Gaussian independent samples with variance σ_s^2 is to be transmitted over a (N_t, N_r) block fading MIMO channel where N_t and N_r are the number of transmit and receive antennas respectively. Every m samples of the source stacked in a vector $\mathbf{x}_s$ are transmitted over n channel uses. The channel matrix $\mathbf{H}$ is assumed fixed during this period and changes independently for the next n channel uses. We call the ratio $\eta = \frac{n}{m}$ the expansion/contraction factor of the system. In a general setting, the communication strategy consists of source/channel coding and source/channel decoding. As a result of source channel coding, $\mathbf{x}_s$ is mapped into a $N_t \times n$ space-time matrix $\mathbf{X}$ which in turn is received at the receiver side as an $N_r \times n$ matrix $\mathbf{Y}$ where

$$\mathbf{Y} = \sqrt{\frac{\mathrm{SNR}}{N_t}}\mathbf{H}\mathbf{X} + \mathbf{W}$$

in which SNR is the average signal to noise ratio at each receive antenna, and $\mathbf{W}$ is the additive noise matrix at the receiver whose entries are taken to be $\mathcal{CN}(0,1)$. At the receiver side, the source/channel decoder yields an estimation of $\mathbf{x}_s$ out of $\mathbf{Y}$ as $\widehat{\mathbf{x}_s}$. For a specific channel realization $\mathbf{H}$, the distortion measure is

$$D(H) = \mathrm{E}_{\mathbf{x}_s}\{\|\mathbf{x}_s - \widehat{\mathbf{x}_s}\|^2 | \mathbf{H}\}. \tag{6.1}$$

For any specific strategy, we define the $f-$fidelity event as $\mathcal{A}(f) = \{\mathbf{H} : D(\mathbf{H}) > \mathrm{SNR}^{-f}\}$ and we call f the fidelity exponent. For specific values of η, N_t and N_r, we define:

$$d(f) = \sup \lim_{\mathrm{SNR}\to\infty} -\frac{\log \Pr\{\mathcal{A}(f)\}}{\log \mathrm{SNR}} \tag{6.2}$$

where sup is taken over all possible source/channel coding schemes. We call $d(f)$ simply as diversity. In what follows, we offer lower and upper bounds on $d(f)$.

We recall from [96] that if we denote the eigenvalues of $\mathbf{H}\mathbf{H}^{\mathtt{H}}$ by λ_i, setting $\alpha_i = -\frac{\log \lambda_i}{\log \mathrm{SNR}}$, we have[1]:

$$p(\overrightarrow{\alpha}) \doteq \mathrm{SNR}^{-\sum(2i-1+|N_t-N_r|)\alpha_i}. \tag{6.3}$$

[1]In this chapter, we use $a \doteq b$ to denote that a and b are asymptotically equivalent.

On the other hand, in a system of tandem coder, i.e., separate source coder and channel coder, we have:[2]

$$\Pr\{error|\mathbf{H}\} \leq \text{SNR}^{-n[\sum(1-\alpha_i)^+ - r]} \tag{6.4}$$

where $R = r \log \text{SNR}$ is the transmission rate over the channel, i.e., $R = \frac{\log M}{n}$, and M is the number of quantized points in the output of the source coder.

6.2 Upper bound on $d(f)$

To get an upper bound on $d(f)$, we consider the case of delay unlimited where $m, n \to \infty$ and $\frac{n}{m} = \eta$ is a constant. Also, we assume that the transmitter has perfect knowledge of the channel matrix H, and therefore, one may talk about the capacity of this MIMO channel which is given by [96]:

$$R = \sup_{\Sigma:\mathbf{tr}(\Sigma)\leq m} \log\det(\mathbf{I} + \frac{\text{SNR}}{m}\mathbf{H}\Sigma\mathbf{H}^{\mathtt{H}}) \tag{6.5}$$

$$\leq \log\det(\mathbf{I} + \text{SNR}\mathbf{H}\mathbf{H}^{\mathtt{H}}). \tag{6.6}$$

We know that the source rate is $R_s = \eta R$. On the other hand, the Distortion-Rate function of the source is $\mathcal{D}(R_s; \mathbf{H}) = e^{-2R_s}$. Therefore:

$$\mathcal{D}(R_s; \mathbf{H}) \geq \frac{1}{\det(\mathbf{I} + \text{SNR}\mathbf{H}\mathbf{H}^{\mathtt{H}})^{2\eta}}. \tag{6.7}$$

[2] We use x^+ to denote $\max(x, 0)$.

Let us denote the f-fidelity event as $\mathcal{A}_\infty(f)$ here. Thus, we obtain:

$$\Pr\{\mathcal{A}_\infty(f)\} = \Pr\{\mathcal{D}(R_s; H) > f\} \geq \tag{6.8}$$

$$\Pr\{\frac{1}{\det(\mathbf{I} + \text{SNR}\mathbf{H}\mathbf{H}^{\text{H}})^{2\eta}} > \text{SNR}^{-f}\} \tag{6.9}$$

which can be written as:

$$\Pr\{\mathcal{A}_\infty(f)\} \geq \Pr\{\prod_{i=1}^{i=n}(1 + \text{SNR}\lambda_i) < \text{SNR}^{\frac{f}{2\eta}}\}$$

$$\doteq \Pr\{\sum_i (1 - \alpha_i)^+ < \frac{f}{2\eta}\}. \tag{6.10}$$

As a result, based on (6.3) and (6.10), we get:

$$\Pr\{\mathcal{A}_\infty(f)\} \geq \int_{\overrightarrow{\alpha}\in\Delta} \text{SNR}^{-\sum(2i-1+|N_t-N_r|)\alpha_i} d\overrightarrow{\alpha} \tag{6.11}$$

where $\Delta = \{\overrightarrow{\alpha} : \sum_i (1 - \alpha_i)^+ < \frac{f}{2\eta}\}$. Based on [96], we have:

$$\int_{\overrightarrow{\alpha}\in\Delta} \text{SNR}^{-\sum(2i-1+|N_t-N_r|)\alpha_i} d\overrightarrow{\alpha} = \text{SNR}^{-d_{ub}(f)} \tag{6.12}$$

where $d_{ub}(f) = \min_{\overrightarrow{\alpha}\in\Delta} \sum(2i - 1 + |N_t - N_r|)\alpha_i$. According to the results in [96], for integer values of $\frac{f}{2\eta}$, this can be calculated as:

$$d_{ub}(f) = (N_t - \frac{f}{2\eta})(N_r - \frac{f}{2\eta}). \tag{6.13}$$

Clearly, if we let $\Pr\{\mathcal{A}_\infty(f)\} \doteq \text{SNR}^{-d_\infty(f)}$, we have $d_\infty(f) \leq d_{ub}(f)$. Consequently, $d_{ub}(f)$ is an upper bound on $d(f)$.

6.3 Achieving the optimum tradeoff

In this section, we show that for any given bandwidth expansion $\eta = \frac{n}{m}$, the optimum diversity-fidelity tradeoff can be achieved by a family of digital space-time codes $\{\mathcal{C}_k\}$. Consider the $N_t^2 \times N_t^2$ matrix $\mathbf{L}$ as the generator of the underlying lattice of a space-time code with non-vanishing determinant property (e.g. Perfect Codes [18] [53] [38]). We map mN_t^2 samples of the source to an nN_t^2-dimensional vector $\mathbf{s} = [s_1 ... s_{nN_t^2}]^{\mathrm{T}}$, and construct the MIMO codeword by setting $\mathbf{c} = (\mathbf{L} \otimes \mathbf{I}_n)\,\mathbf{s}$ and mapping it to the entries of n consecutive $N_t \times N_t$ space-time matrices.

For the modulating signal $\mathbf{x}_s = (x_1, ..., x_{mN_t^2})$, consider $x_i + \frac{1}{2} = \left(\overline{0.b_{i,1}b_{i,2}b_{i,3}...}\right)_2$. Let $b'_{i+(j-1)mN_t^2} = b_{i,j}$. For the code $\mathcal{C}_k$, we construct s_1, s_2, ..., s_N (where $N = nN_t^2$) as

$$s_1 = \left(\overline{0.b'_1 b'_{N+1} b'_{2N+1} ... b'_{(k-1)N+1}}\right)_2, \tag{6.14}$$

$$s_2 = \left(\overline{0.b'_2 b'_{N+2} b'_{2N+2} ... b'_{(k-1)N+2}}\right)_2, \tag{6.15}$$

...

$$s_N = \left(\overline{0.b'_N b'_{2N} b'_{3N} ... b'_{kN}}\right)_2, \tag{6.16}$$

Lemma 6.1. *[18] Consider $\{\mathcal{C}_i(r)\}$ a sequence of $N_t \times N_t$ lattice space-time codes, designed for rate $R_i = r \cdot \log \mathrm{SNR}_i$ (where $\mathrm{SNR}_i \to \infty$), and assume that its underlying lattice satisfy the nonvanishing determinant property. For $d_{min}(r)$, the minimum distance of the received constellation $\mathcal{C}_{\mathbf{H},i}(r) = \mathbf{H}\mathcal{C}_i(r)$, we have*

$$P_e(r) = \Pr\left\{\frac{d_{min}(r)}{2} \leq \|\mathbf{w}\|\right\} \doteq SNR^{-e(r)} \tag{6.17}$$

where $e(r)$ is the piecewise linear function given by

$$e(r) = (N_t - r)(N_r - r) \tag{6.18}$$

for integer values of r.

Theorem 6.1. *For the proposed family of codes, if we use $\mathcal{C}_k$ for $2^{-(k-1)N_t} > \mathrm{SNR}^{-\frac{f}{2\eta}} \geq 2^{-kN_t}$,*

$$d(f) = (N_t - \frac{f}{2\eta})(N_r - \frac{f}{2\eta})$$

for integer values of $\frac{f}{2\eta}$.

Proof: Because the family of codes $\mathcal{C}_k$ are obtained from a lattice with non-vanishing determinant property, using ML decoding, they achieve the optimum diversity-multiplexing tradeoff. Thus, for this family of codes,

$$P_{error}(r) \doteq \mathrm{SNR}^{-(N_t-r)(N_r-r)} \tag{6.19}$$

for integer values of r, the normalized rate ($r = \frac{R}{\log \mathrm{SNR}}$). In this scheme, code $\mathcal{C}_k$ has rate $R_k = kN_t$ and is used for $2^{-(k-1)N_t} > \mathrm{SNR}^{-\frac{f}{2\eta}} \geq 2^{-kN_t}$. Therefore, $r \doteq \frac{f}{2\eta}$, hence,

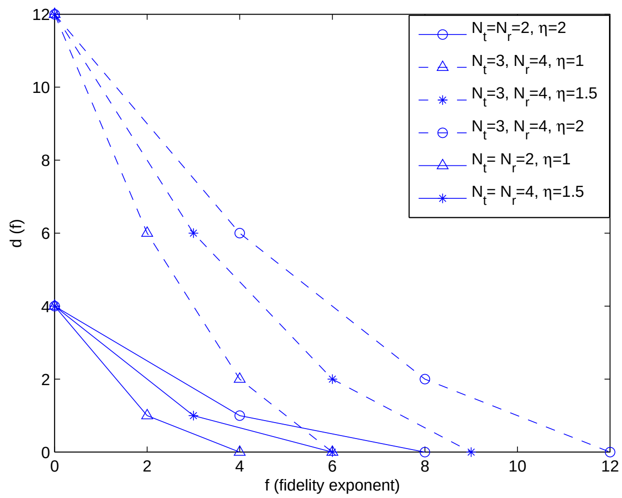

Figure 6.1: Diversity-Fidelity Tradeoff for different numbers of antennas and different bandwidth expansion factors

$$P_{error}(r) \doteq \text{SNR}^{-(N_t - \frac{f}{2\eta})(N_r - \frac{f}{2\eta})}. \tag{6.20}$$

When there is no error in decoding $\mathcal{C}_k$, the $kN = knN_t^2$ bits $b'_1, ..., b'_{kN}$ are decoded correctly, hence the first $\lfloor k\frac{n}{m}N_t^2 \rfloor = \lfloor kN\eta \rfloor$ bits of x_i can be reconstructed without error (for $1 \leq i \leq m$), hence $D \doteq 2^{-2kN\eta} \doteq \text{SNR}^{-f}$. Therefore,

$$\Pr\left\{D > \text{SNR}^{-f}\right\} \leq P_{error}(r) \doteq \text{SNR}^{-(N_t - \frac{f}{2\eta})(N_r - \frac{f}{2\eta})}. \tag{6.21}$$

$$\Longrightarrow d(f) \geq (N_t - \frac{f}{2\eta})(N_r - \frac{f}{2\eta}). \tag{6.22}$$

■

6.4 Conclusions

Similar to the diversity-multiplexing tradeoff which is well known for the high-SNR asymptotic performance of space-time codes, diversity-fidelity tradoff can be used as a benchmark for evaluating various MIMO source-channel codes. It is shown in this chapter that this tradeoff can be achieved by a simple modification of the well-known DMT-achieving codes. However, though the optimum tradeoff can be achieved for different numbers of antennas and arbitrary bandwidth expansion factors, the proposed scheme is not robust (it is dependent on the SNR of the channel). The other problem is complexity of the decoding which is based on ML decoding and is not practical for high-fidelity cases. Therefore, the problem of achieving the optimum (or near-optimum) diversity-fidelity tradeoff by practical and robust codes remains open.

Chapter 7

Conclusion and Future Research

This dissertation focuses on the problem of signaling and fairness for the MIMO multi-user systems.

In Chapter 2, a new viewpoint for adopting the lattice reduction in communication over MIMO broadcast channels is introduced. Lattice basis reduction helps us to reduce the average transmitted energy by modifying the region which includes the constellation points. The new viewpoint helps us to generalize the idea of lattice-reduction-aided precoding for the case of unequal-rate transmission, and obtain analytic results for the asymptotic behavior (for large SNR) of the symbol-error-rate for the lattice-reduction-aided precoding and the perturbation technique. Also, the outage probability for both cases of fixed-rate users and fixed sum-rate is analyzed. It is shown that the lattice-reduction-aided method, using LLL algorithm, achieves the optimum asymptotic slope of symbol-error-rate (called the precoding diversity).

Chapter 3 considers the performance of lattice decoding and its lattice-reduction-aided approximation in MIMO multiaccess systems (or equivalently, MIMO point-to-point systems with V-BLAST transmission). In this chapter, it is proven that in these systems, lattice-reduction-aided decoding achieves the maximum receive diversity (which is equal to the number of receive antennas). Also, it is proven that the naive lattice decoding (which discards the out-of-region decoded points) achieves the maximum diversity.

In Chapter 4, the inherent limitations of the performance of the naive lattice decoding is investigated. The naive lattice decoding and various implementions of it (such as the sphere decoding) and its simple approximated versions (such as the LLL-aided decoding) are very attractive for the practical MIMO systems. Nonetheless, to achieve theoretical benchmarks (such as the rate-diversity trade-off), these techniques are not always sufficient.

In Chapter 5, the problem of sending an analog source over an additive white Gaussian noise channel is considered. The traditional analog coding schemes suffer from the threshold effect. We introduce a robust scheme for analog conding. Unlike the previous methods, the new method asymptotically achieves the optimal scaling of the signal-to-distortion-ratio (SDR) without being affected by the threshold effect. Also, we show that approximated versions of these techniques perform well for the practical applications, with a low complexity in encoding/decoding.

Finally, in Chapter 6, the problem of joint source-channel coding over delay-limited MIMO channels is investigated and similar to the concept of diversity-multiplexing tradeoff (which was defined for the transmission of digital sources over MIMO channels), we have shown that the theoretical limits of transmission of analog sources over MIMO channels can be formulated by a tradeoff, called diversity-fidelity tradeoff. It is shown that the optimum tradeoff can be achieved by a family of lattice joint-source channel codes.

7.1 Future Research Directions

The dissertation can be continued in several directions as briefly explained in what follows.

In chapter five, only the case of delay-limited constellation are studied. Constructing robust joint source-channel codes which approach the optimum performance for a certain SNR value (by having an arbitrarily small gap to the optimum point), is an interesting research direction.

For chapter six, generalizing the concept of diversity-fidelity tradeoff to MIMO multiuser systems and generalizing the ideas of this chapter to construct codes for them is worth trying. Also, the problem of constructing robust MIMO codes, that achieve the optimum diversity-fidelity tradeoff in entire range of SNR, remains open.

Appendix A

Second moment of a parallelotope

In this Appendix, we compute the second moment of a parallelotope whose centroid is the origin and its edges are equal to the basis vectors of the lattice.

Assume that $\mathcal{A}$ is an M-dimensional parallelotope and X is its second moment. The second moment of $\frac{1}{2}\mathcal{A}$ is $\left(\frac{1}{2}\right)^{M+2} X$. The parallelotope $\mathcal{A}$ can be considered as the union of 2^M smaller parallelotopes which are constructed by $\pm\frac{1}{2}\mathbf{b}_1, \pm\frac{1}{2}\mathbf{b}_2, ..., \pm\frac{1}{2}\mathbf{b}_M$, where $\mathbf{b}_i$, $1 \le i \le M$, is a basis vector. These parallelotopes are translated versions of $\frac{1}{2}\mathcal{A}$ with the translation vectors $T_i = \pm\frac{1}{2}\mathbf{b}_1 \pm \frac{1}{2}\mathbf{b}_2 \pm .... \pm \frac{1}{2}\mathbf{b}_M$, $1 \le i \le 2^M$. The second moments of these parallelotopes are equal to $\left(\frac{1}{2}\right)^{M+2} X + \|T_i\|^2 \mathrm{Vol}(\frac{1}{2}\mathcal{A})$, $1 \le i \le 2^M$. By the summation over all these second moments, we can find the second moment of $\mathcal{A}$.

$$X = \sum_{i=1}^{2^M} \left[\left(\frac{1}{2}\right)^{M+2} X + \|T_i\|^2 . \mathrm{Vol}(\frac{1}{2}\mathcal{A}) \right] \tag{A.1}$$

$$= \left(\frac{1}{2}\right)^2 X + 2^{M-2}(\|\mathbf{b}_1\|^2 + ... + \|\mathbf{b}_M\|^2).\mathrm{Vol}(\frac{1}{2}\mathcal{A}) \tag{A.2}$$

$$= \frac{1}{4}X + \frac{1}{4}(\|\mathbf{b}_1\|^2 + ... + \|\mathbf{b}_M\|^2).\mathrm{Vol}(\mathcal{A}) \tag{A.3}$$

$$\Longrightarrow X = \frac{1}{3}(\|\mathbf{b}_1\|^2 + ... + \|\mathbf{b}_M\|^2).\mathrm{Vol}(\mathcal{A}). \tag{A.4}$$

Appendix B

Proof of Lemma 2.3

Lemma 3 states that the probability that a lattice, generated by M independent N-dimensional complex Gaussian vectors, $N \geq M$, with a unit variance per each dimension, has a nonzero point inside a sphere (centered at origin and with the radius ε) is bounded by $\beta_{N,M}\varepsilon^{2N}$ for $N > M$, and $\beta_{N,M}\varepsilon^{2N} \max\left\{(-\ln \varepsilon)^{N+1}, 1\right\}$ for $N = M \geq 2$. We can assume that $\varepsilon < 1$ (for $\varepsilon \geq 1$, lemma 3 is trivial because the probability is bounded).

B.0.1 Case 1: $M = 1$

When $M = 1$, the lattice consists of the integer multiples of the basis vector $\mathbf{v}$. If the norm of one of these vectors is less than ε, then the norm of $\mathbf{v}$ is less than ε. Consider the variance of the components of $\mathbf{v}$ as ϱ^2. The vector $\mathbf{v}$ has an N-dimensional complex

Gaussian distribution, $f_{\mathbf{v}}(\mathbf{v})$. Therefore, the probability of this event is,

$$\Pr\{\|\mathbf{v}\| \leq \varepsilon\} = \int_{\|\mathbf{v}\|\leq\varepsilon} f_{\mathbf{v}}(\mathbf{v})\, d\mathbf{v} \leq \int_{\|\mathbf{v}\|\leq\varepsilon} \frac{1}{\pi^N \varrho^{2N}}\, d\mathbf{v} \leq \beta_{N,1}\frac{\varepsilon^{2N}}{\varrho^{2N}}. \tag{B.1}$$

When the variance of the components of $\mathbf{v}$ is equal to one, we have,

$$\Pr\{\|\mathbf{v}\| \leq \varepsilon\} \leq \beta_{N,1}\varepsilon^{2N}. \tag{B.2}$$

B.0.2 Case 2: $N > M > 1$

Consider $L_{(\mathbf{v}_1,\ldots,\mathbf{v}_M)}$ as the lattice generated by $\mathbf{v}_1$,$\mathbf{v}_2$,...,$\mathbf{v}_M$. Each point of $L_{(\mathbf{v}_1,\ldots,\mathbf{v}_M)}$ can be represented by $\mathbf{v}_{(z_1,\ldots,z_M)} = z_1\mathbf{v}_1 + z_2\mathbf{v}_2 + \ldots + z_M\mathbf{v}_M$, where $z_1, \ldots, z_M$ are complex integer numbers. The vectors $\mathbf{v}_1$,$\mathbf{v}_2$,...,$\mathbf{v}_M$ are independent and jointly Gaussian. Therefore, for every integer vector $\mathbf{z} = (z_1, \ldots, z_M)$, the entries of the vector $\mathbf{v}_{(z_1,\ldots,z_M)}$ have complex Gaussian distributions with the variance

$$\varrho_{\mathbf{z}}^2 = \|\mathbf{z}\|^2\varrho^2 = \left(|z_1|^2 + \ldots + |z_M|^2\right)\varrho^2. \tag{B.3}$$

Therefore, according to the lemma for $M = 1$,

$$\Pr\left\{\|\mathbf{v}_{(z_1,\ldots,z_M)}\| \leq \varepsilon\right\} \leq \beta_{N,1}\frac{\varepsilon^{2N}}{\left(|z_1|^2 + \ldots + |z_M|^2\right)^N}. \tag{B.4}$$

Now, by using the union bound,

$$\Pr\{d_{\mathbf{H}} \leq \varepsilon\} \leq \sum_{\mathbf{z}\neq 0} \Pr\left\{\|\mathbf{v}_{(z_1,\ldots,z_M)}\| \leq \varepsilon\right\} \tag{B.5}$$

$$\leq \sum_{\mathbf{z}\neq 0} \beta_{N,1}\frac{\varepsilon^{2N}}{\left(|z_1|^2 + \ldots + |z_M|^2\right)^N} \tag{B.6}$$

$$= \beta_{N,1} \left(\sum_{1 \le \|\mathbf{z}\| < 2} \frac{\varepsilon^{2N}}{\|\mathbf{z}\|^{2N}} + \sum_{2 \le \|\mathbf{z}\| < 3} \frac{\varepsilon^{2N}}{\|\mathbf{z}\|^{2N}} \right.$$

$$\left. + \sum_{3 \le \|\mathbf{z}\| < 4} \frac{\varepsilon^{2N}}{\|\mathbf{z}\|^{2N}} + \ldots \right). \tag{B.7}$$

The M-dimensional complex integer points $\mathbf{z} = (z_1, ..., z_M)$, such that $k \le \|\mathbf{z}\| < k+1$, can be considered as the centers of disjoint unit-volume cubes. All these cubes are inside the region between the $2M$-dimensional spheres, with radii $k-1$ and $k+2$. Therefore, the number of M-dimensional complex integer points $\mathbf{z} = (z_1, ..., z_M)$, such that $k \le \|\mathbf{z}\| < k+1$, can be bounded by the volume of the region between these two $2M$-dimensional spheres. Thus, this number is bounded by $c_1 k^{2M-1}$ for some constant[1] c_1. Therefore,

$$\sum_{k \le \|\mathbf{z}\| < k+1} \frac{\varepsilon^{2N}}{\|\mathbf{z}\|^{2N}} \le c_1 k^{2M-1} \frac{\varepsilon^{2N}}{k^{2N}} \tag{B.8}$$

$$\text{(B.7)}, \text{(B.8)} \Longrightarrow \Pr\{d_{\mathbf{H}} \le \varepsilon\} \le c_1 \beta_{N,1} \varepsilon^{2N} + 2^{2M-1} c_1 \beta_{N,1} \frac{\varepsilon^{2N}}{2^{2N}} +$$

$$+ 3^{2M-1} c_1 \beta_{N,1} \frac{\varepsilon^{2N}}{3^{2N}} + \ldots \tag{B.9}$$

$$\le c_1 \beta_{N,1} \varepsilon^{2N} \sum_{k=1}^{\infty} \frac{1}{k^{2N-2M+1}}. \tag{B.10}$$

According to the assumption of this case, $N > M$; hence, $2N - 2M + 1 \ge 2$. Therefore, the above summation is convergent:

[1] Throughout this proof, $c_1, c_2, ...$ are some constant numbers.

$$\Pr\{\|\mathbf{v}\| \leq \varepsilon\} \leq \beta_{N,M}\varepsilon^{2N}. \tag{B.11}$$

B.0.3 Case 3: $N = M > 1$

Each point of $L_{(\mathbf{v}_1,\ldots,\mathbf{v}_N)}$ can be represented by $z\mathbf{v}_N - \mathbf{v}$, where $\mathbf{v}$ belongs to the lattice $L_{(\mathbf{v}_1,\ldots,\mathbf{v}_{N-1})}$ and z is a complex integer. Consider $\mathcal{S}_\mathbf{v}$ as the sphere with radius ε and centered at $\mathbf{v}$. Now, $z\mathbf{v}_N - \mathbf{v}$ belongs to $\mathcal{S}_\mathbf{0}$ iff the $z\mathbf{v}_N$ belongs to $\mathcal{S}_\mathbf{v}$. Also, the sphere $\mathcal{S}_\mathbf{v}$ includes a point $z\mathbf{v}_N$ iff $\mathcal{S}_{\mathbf{v},z}$ includes $\mathbf{v}$, where $\mathcal{S}_{\mathbf{v},z} = \frac{\mathcal{S}_\mathbf{v}}{z}$ is the sphere centered at $\mathbf{v}/z$ with radius $\dfrac{\varepsilon}{|z|}$ (see figure 4). Therefore, the probability that a lattice point exists in $\mathcal{S}_\mathbf{v}$ is equal to the probability that $\mathbf{v}_N$ is in at least one of the spheres $\{\mathcal{S}_{\mathbf{v},z}\}$, $z \neq 0$.

If we consider $d_\mathbf{H}$ as the minimum distance of $L_{(\mathbf{v}_1,\ldots,\mathbf{v}_N)}$ and R as an arbitrary number greater than 1:

$$\Pr\{d_\mathbf{H} \leq \varepsilon\} = \Pr\left\{\left(L_{(\mathbf{v}_1,\ldots,\mathbf{v}_N)} - 0\right) \cap \mathcal{S}_\mathbf{0} \neq \varnothing\right\} = \Pr\left\{\mathbf{v}_N \in \bigcup_{\mathbf{v}} \bigcup_{z \neq 0} \mathcal{S}_{\mathbf{v},z}\right\} \tag{B.12}$$

$$\leq \Pr\left\{\mathbf{v}_N \in \bigcup_{\|\frac{\mathbf{v}}{z}\| \leq R} \mathcal{S}_{\mathbf{v},z}\right\} + \Pr\left\{\mathbf{v}_N \in \bigcup_{\|\frac{\mathbf{v}}{z}\| > R} \mathcal{S}_{\mathbf{v},z}\right\} \tag{B.13}$$

In the second term of (B.13), all the spheres have centers with norms greater than R and radii less than 1 (because $|z| \geq 1$). Therefore,

$$\bigcup_{\|\frac{\mathbf{v}}{z}\| > R} \mathcal{S}_{\mathbf{v},z} \subset \{\mathbf{x} \mid \|\mathbf{x}\| > R - 1\} \tag{B.14}$$

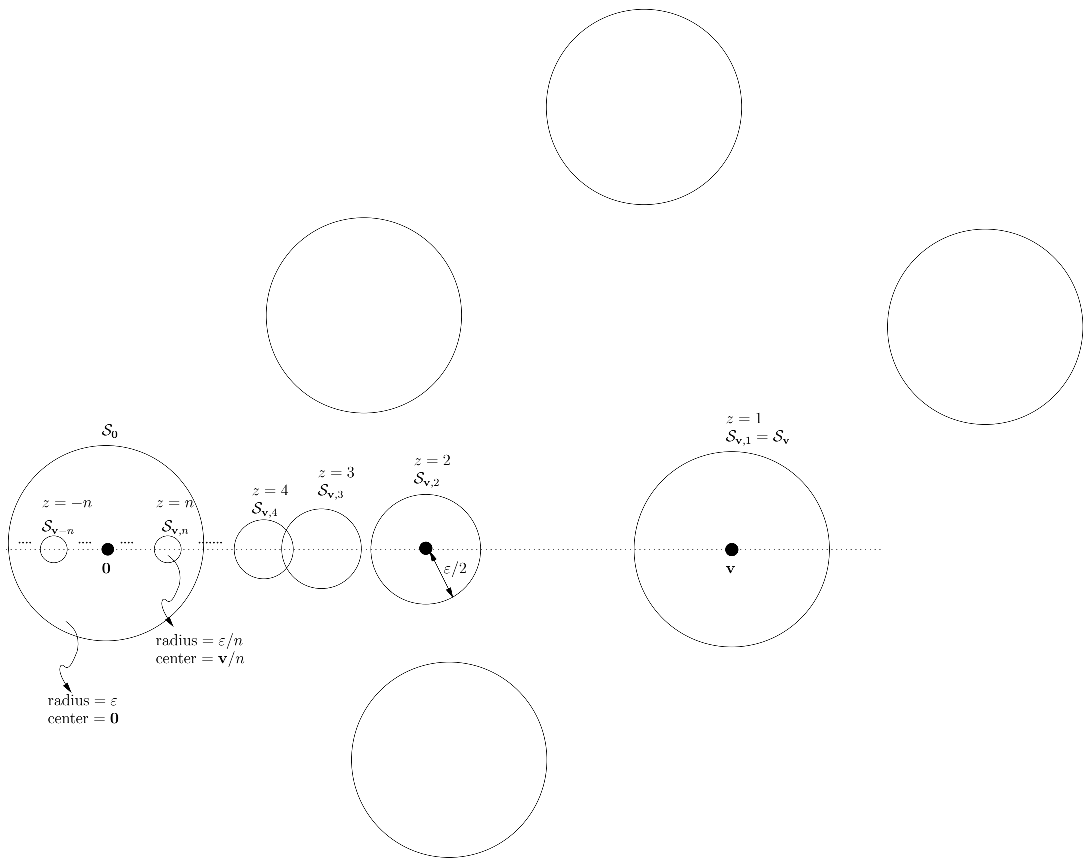

Figure B.1: The family of spheres $\mathcal{S}_{\mathbf{v},z}$

$$\text{(B.14)} \Longrightarrow \text{(B.13)} \leq \Pr\left\{\mathbf{v}_N \in \bigcup_{\|\frac{\mathbf{v}}{z}\| \leq R} \mathcal{S}_{\mathbf{v},z}\right\} + \Pr\left\{\|\mathbf{v}_N\| > R - 1\right\} \tag{B.15}$$

$$\leq \Pr\left\{\mathbf{v}_N \in \bigcup_{\|\frac{\mathbf{v}}{z}\| \leq R, |z| \leq \varepsilon^{1-N}} \mathcal{S}_{\mathbf{v},z}\right\} + \Pr\left\{\mathbf{v}_N \in \bigcup_{\|\frac{\mathbf{v}}{z}\| \leq R, |z| > \varepsilon^{1-N}} \mathcal{S}_{\mathbf{v},z}\right\} + \Pr\left\{\|\mathbf{v}_N\| > R - 1\right\}. \tag{B.16}$$

We bound the first term of (B.16) as the following:

$$\Pr\{\mathbf{v}_N \in \bigcup_{\|\frac{\mathbf{v}}{z}\| \leq R, |z| \leq \varepsilon^{1-N}} \mathcal{S}_{\mathbf{v},z}\} \tag{B.17}$$

$$\leq \Bigg(\sum_{\|\mathbf{v}\| \leq 2R} \sum_{|z| \geq 1} \Pr\{\mathbf{v}_N \in \mathcal{S}_{\mathbf{v},z}\} + \sum_{2R < \|\mathbf{v}\| \leq 3R} \sum_{|z| \geq 2} \Pr\{\mathbf{v}_N \in \mathcal{S}_{\mathbf{v},z}\} +$$

$$\ldots + \sum_{\lfloor \varepsilon^{1-N} \rfloor R < \|\mathbf{v}\| \leq (\lfloor \varepsilon^{1-N} \rfloor + 1) R} \quad \sum_{|z| \geq \lfloor \varepsilon^{1-N} \rfloor} \Pr\{\mathbf{v}_N \in \mathcal{S}_{\mathbf{v},z}\} \Bigg). \tag{B.18}$$

Noting that the pdf of $\mathbf{v}_N$ is less than or equal to $\frac{1}{\pi^N}$,

$$\text{(B.17)} \leq \frac{1}{\pi^N} \Bigg(\sum_{\|\mathbf{v}\| \leq 2R} \sum_{|z| \geq 1} \mathrm{Vol}(\mathcal{S}_{\mathbf{v},z}) + \sum_{2R < \|\mathbf{v}\| \leq 3R} \sum_{|z| \geq 2} \mathrm{Vol}(\mathcal{S}_{\mathbf{v},z}) +$$

$$\ldots + \sum_{\lfloor \varepsilon^{1-N} \rfloor R < \|\mathbf{v}\| \leq (\lfloor \varepsilon^{1-N} \rfloor + 1) R} \quad \sum_{|z| \geq \lfloor \varepsilon^{1-N} \rfloor} \mathrm{Vol}(\mathcal{S}_{\mathbf{v},z}) \Bigg) \tag{B.19}$$

$$\leq \frac{1}{\pi^N} \Bigg(\sum_{\|\mathbf{v}\| \leq 2R} \sum_{|z| \geq 1} \frac{c_2 \varepsilon^{2N}}{|z|^{2N}} + \sum_{2R < \|\mathbf{v}\| \leq 3R} \sum_{|z| \geq 2} \frac{c_2 \varepsilon^{2N}}{|z|^{2N}} +$$

$$... + \sum_{\lfloor \varepsilon^{1-N} \rfloor R < \|\mathbf{v}\| \leq (\lfloor \varepsilon^{1-N} \rfloor + 1)R} \quad \sum_{|z| \geq \lfloor \varepsilon^{1-N} \rfloor} \frac{c_2 \varepsilon^{2N}}{|z|^{2N}} \Bigg). \tag{B.20}$$

By using (B.8), for one-dimensional complex vector $\mathbf{z} = z$,

$$\sum_{|z| \geq i} \frac{c_2 \varepsilon^{2N}}{|z|^{2N}} = \sum_{k=i}^{\infty} c_2 . \sum_{k \leq |z| \leq k+1} \frac{\varepsilon^{2N}}{|z|^{2N}} \leq \sum_{k=i}^{\infty} \frac{c_1 c_2 \varepsilon^{2N}}{k^{2N-1}} \leq \frac{c_3 \varepsilon^{2N}}{i^{2N-2}} \tag{B.21}$$

Now,

$$(\text{B.21}) \Longrightarrow (\text{B.17}) \leq \frac{1}{\pi^N} \Bigg(\sum_{\|\mathbf{v}\| \leq 2R} c_3 \varepsilon^{2N} + \sum_{2R < \|\mathbf{v}\| \leq 3R} \frac{c_3 \varepsilon^{2N}}{2^{2N-2}} +$$

$$... + \sum_{\lfloor \varepsilon^{1-N} \rfloor R < \|\mathbf{v}\| \leq (\lfloor \varepsilon^{1-N} \rfloor + 1)R} \frac{c_3 \varepsilon^{2N}}{\lfloor \varepsilon^{1-N} \rfloor^{2N-2}} \Bigg). \tag{B.22}$$

Assume that the minimum distance of $L_{(\mathbf{v}_1,\ldots,\mathbf{v}_{N-1})}$ is d_{N-1}. The spheres with the radius $d_{N-1}/2$ and centered by the points of $L_{(\mathbf{v}_1,\ldots,\mathbf{v}_{N-1})}$ are disjoint. Therefore, the number of points from the $(N-1)$-dimensional complex lattice $L_{(\mathbf{v}_1,\ldots,\mathbf{v}_{N-1})}$, such that $\|\mathbf{v}\| \leq 2R$, is bounded by $\frac{c_4(2R+d_{N-1}/2)^{2N-2}}{d_{N-1}^{2N-2}}$ (it is bounded by the ratio between the volumes of $(2N-2)$-dimensional spheres with radii $2R + d_{N-1}/2$ and $d_{N-1}/2$). Also, the number of points from $L_{(\mathbf{v}_1,\ldots,\mathbf{v}_{N-1})}$, such that $(k-1)R < \|\mathbf{v}\| \leq kR$, is bounded by $\frac{c_4(kR)^{2N-3}(R+d_{N-1})}{d_{N-1}^{2N-2}}$ (it is bounded by the ratio between the volumes of the region defined by $(k-1)R - d_{N-1}/2 < \|\mathbf{x}\| \leq kR + d_{N-1}/2$ and the sphere with radius $d_{N-1}/2$):

$$(B.22) \leq \frac{c_5(2R + d_{N-1}/2)^{2N-2}}{d_{N-1}^{2N-2}}.\varepsilon^{2N} + \frac{c_5 R^{2N-3}(R + d_{N-1})}{d_{N-1}^{2N-2}}.\varepsilon^{2N} \sum_{k=2}^{\lfloor \varepsilon^{1-N} \rfloor} \frac{1}{k} \tag{B.23}$$

$$(B.22) \leq \frac{c_5(2R + d_{N-1}/2)^{2N-2}}{d_{N-1}^{2N-2}}.\varepsilon^{2N} + \frac{c_5 R^{2N-3}(R + d_{N-1})}{d_{N-1}^{2N-2}}.\varepsilon^{2N}. \ln(\varepsilon^{1-N}) \tag{B.24}$$

$$\leq c_6\varepsilon^{2N}.\max\left(\frac{R^{2N-2}}{d_{N-1}^{2N-2}},1\right).\max\{-\ln\varepsilon,1\}. \tag{B.25}$$

According to the proof of the case 2, we have $\Pr\{d_{N-1}\leq\eta\}\leq\beta_{N,N-1}\eta^{2N}$. Therefore,

$$\mathrm{E}_{d_{N-1}}\left\{\max\left(\frac{R^{2N-2}}{d_{N-1}^{2N-2}},1\right)\right\} \tag{B.26}$$

$$\leq 1.\Pr\{d_{N-1}>R\}+2^{2N-2}.\Pr\left\{\frac{1}{2}R<d_{N-1}\leq R\right\}+3^{2N-2}.\Pr\left\{\frac{1}{3}R<d_{N-1}\leq\frac{1}{2}R\right\}+... \tag{B.27}$$

$$\leq 1+2^{2N-2}.\Pr\{d_{N-1}\leq R\}+3^{2N-2}.\Pr\left\{d_{N-1}\leq\frac{1}{2}R\right\}+... \tag{B.28}$$

$$\leq 1+\sum_{k=1}^{\infty}\frac{(k+1)^{2N-2}}{k^{2N}}.R^{2N}\beta_{N,N-1}\leq c_7R^{2N} \tag{B.29}$$

$$\Longrightarrow \mathrm{E}_{d_{N-1}}\left\{c_6\varepsilon^{2N}.\max\left(\frac{R^{2N-2}}{d_{N-1}^{2N-2}},1\right).\max\{-\ln\varepsilon,1\}\right\}\leq c_8\varepsilon^{2N}.R^{2N}.\max\{-\ln\varepsilon,1\}. \tag{B.30}$$

To bound the second term of (B.16), we note that for $|z|\geq\varepsilon^{1-N}$, the radii of the spheres $\mathcal{S}_{\mathbf{v},z}$ are less or equal to ε^N, and the centers of these spheres lie on the $(N-1)$-dimensional complex subspace containing $L_{\mathbf{v}_1,\ldots,\mathbf{v}_{N-1}}$. Also, the norm of these centers are less than R. Therefore, all of these spheres are inside the region $\mathcal{A}$ which is an orthotope centered at the origin, with $2N$ real dimensions (see figure 5):

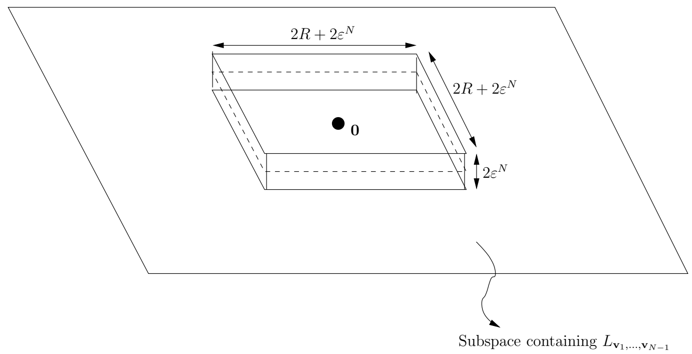

Figure B.2: The orthotope $\mathcal{A}$

$$\bigcup_{\|\frac{\mathbf{v}}{z}\|\leq R, |z|>\varepsilon^{1-N}} \mathcal{S}_{\mathbf{v},z} \subset \mathcal{A} \tag{B.31}$$

$$\Longrightarrow \Pr\{\mathbf{v}_N \in \bigcup_{\|\frac{\mathbf{v}}{z}\|\leq R, |z|>\varepsilon^{1-N}} \mathcal{S}_{\mathbf{v},z}\} \leq \frac{1}{\pi^N}\mathrm{Vol}\left(\bigcup_{\|\frac{\mathbf{v}}{z}\|\leq R, |z|>\varepsilon^{1-N}} \mathcal{S}_{\mathbf{v},z}\right) \leq \frac{1}{\pi^N}\mathrm{Vol}(\mathcal{A}) \tag{B.32}$$

$$\leq \frac{1}{\pi^N}(2\varepsilon^N)^2(2R+\varepsilon^N)^{2N-2} \tag{B.33}$$

Also, according to the Gaussian distribution of the entries of $\mathbf{v}_N$ (which have Variance $\frac{1}{2}$ on each real dimension), we can bound the third term of (B.16) as,

$$\Pr\{\|\mathbf{v}_N\| > R-1\} \leq 2NQ\left(\sqrt{\frac{R-1}{N}}\right) \leq c_9 e^{-\left(\frac{R-1}{\sqrt{2N}}\right)^2}. \tag{B.34}$$

By using (B.30), (B.33), and (B.34),

$$\Pr\{d_{\mathbf{H}} \leq \varepsilon\} \leq c_8\varepsilon^{2N}.R^{2N}.\max\{-\ln\varepsilon, 1\} + \frac{1}{\pi^N}(2\varepsilon^N)^2(2R+\varepsilon^N)^{2N-2} + c_9 e^{-\left(\frac{R-1}{\sqrt{2N}}\right)^2}. \tag{B.35}$$

The above equation is true for every $R > 1$. Therefore, using $R = \sqrt{2}N\sqrt{-\ln(\varepsilon^{2N})}+1$,

$$\Pr\{d_{\mathbf{H}} \leq \varepsilon\} \leq \beta_{N,N}\varepsilon^{2N}.\max\left\{(-\ln\varepsilon)^{N+1}, 1\right\}. \tag{B.36}$$

Appendix C

Proof of Theorem 5.1

The set of modulated signals consists of a^{N-1} parallel segments where the projection of each of them on the ith dimension has the length $a^{-(i-1)}$, hence, each segment has the length $\sqrt{1 + a^{-2} + ... + a^{-2(N-1)}}$. By considering the distance of their intersections with the hyperspace orthogonal to the Nth dimension (which is at least a^{-1}) and the angular factor of these segments, respecting to the s_N-axis, we can bound the distance between two parallel segments of the modulated signal set as

$$d \geq \frac{a^{-1}}{\sqrt{1 + a^{-2} + ... + a^{-2(N-1)}}} \geq \frac{a^{-1}}{\sqrt{1 + 2^{-2} + ... + 2^{-2(N-1)}}} \geq \frac{a^{-1}}{2} \tag{C.1}$$

First, we consider the case of $\sigma \leq \frac{1}{8\sqrt{N}}$. Consider $a = \left\lfloor \frac{1}{8\sqrt{N}\sigma\sqrt{-\log\sigma}} \right\rfloor$. Probability of a jump to a wrong segment (during the decoding) is bounded by

$$\Pr(jump) \leq Q\left(\frac{d}{2\sigma}\right) \leq Q\left(\frac{a^{-1}}{4\sigma}\right) \tag{C.2}$$

$$\leq Q\left(\frac{8\sqrt{N}\sigma\sqrt{-\log\sigma}}{4\sigma}\right). \tag{C.3}$$

By using $Q(x) \leq e^{-\frac{x^2}{2}}$,

$$\Pr(jump) \leq e^{\frac{-(2\sqrt{N}\sqrt{-\log\sigma})^2}{2}} = e^{2N\log\sigma} = \sigma^{2N}. \tag{C.4}$$

On the other hand, each segment of the modulated signal set is a segment of the source signal set, stretched by a factor of $a^{N-1}\sqrt{1+a^{-2}+...+a^{-2(N-1)}}$ (its length is changed from $\frac{1}{a^{N-1}}$ to $\sqrt{1+a^{-2}+...+a^{-2(N-1)}}$). Therefore, assuming the correct segment decoding, the average distortion is the variance of the channel noise divided by $\left(a^{N-1}\sqrt{1+a^{-2}+...+a^{-2(N-1)}}\right)^2$:

$$\mathrm{E}\left\{|\tilde{x}-x|^2|no\ jump\right\} = \tag{C.5}$$

$$\frac{\sigma^2}{\left(a^{N-1}\sqrt{1+a^{-2}+...+a^{-2(N-1)}}\right)^2} \leq \tag{C.6}$$

$$\frac{\sigma^2}{a^{2(N-1)}} = \frac{\sigma^2}{\left\lfloor\frac{1}{8\sqrt{N}\sigma\sqrt{-\log\sigma}}\right\rfloor^{2(N-1)}} \leq c_1\sigma^{2N}\left(-\log\sigma\right)^{N-1} \tag{C.7}$$

where $\tilde{x}$ is the estimate of x and c_1 is independent[1] of a and σ and only depends on N. Now, because $\mathrm{E}\{|\tilde{x}-x|^2|jump\}$ and $\Pr(no\ jump)$ are bounded by 1,

$$D = \Pr(jump)\cdot\mathrm{E}\left\{|\tilde{x}-x|^2|jump\right\} + \Pr(no\ jump)\cdot\mathrm{E}\left\{|\tilde{x}-x|^2|no\ jump\right\} \tag{C.8}$$

[1]Throughout this paper, $c_1, c_2, ...$ are constants, independent of σ (they may depend on N).

$$\Longrightarrow D \leq \Pr(jump) + \mathrm{E}\left\{|\tilde{x} - x|^2 | no\ jump\right\} \tag{C.9}$$

$$\leq c_2 \sigma^{2N} (-\log \sigma)^{N-1}, \quad \text{for } \sigma \leq \frac{1}{8\sqrt{N}}. \tag{C.10}$$

On the other hand, for $\sigma > \frac{1}{8\sqrt{N}}$,

$$D \leq 1 \Longrightarrow D \leq \left(\frac{1}{8\sqrt{N}}\right)^{-2N} \left(-\log\left(\frac{1}{8\sqrt{N}}\right)\right)^{-(N-1)} \sigma^{2N} (-\log \sigma)^{N-1} \tag{C.11}$$

$$= c_3 \sigma^{2N} (-\log \sigma)^{N-1}. \tag{C.12}$$

Therefore, by combining these two bounds together, we obtain:

$$D \leq c \sigma^{2N} (-\log \sigma)^{N-1}. \tag{C.13}$$

Appendix D

Proof of Theorem 5.2

We consider two cases:

Case 1) $a \leq \frac{4}{\sigma\sqrt{-\log\sigma}}$:

Each segment of the modulated signal set is a segment of the source signal set, scaled by a factor of $a^{N-1}\sqrt{1+a^{-2}+...+a^{-2(N-1)}}$, hence,

$$D \geq \mathrm{E}\left\{|\tilde{x}-x|^2|no\ jump\right\} \tag{D.1}$$

$$= \frac{\sigma^2}{\left(a^{N-1}\sqrt{1+a^{-2}+...+a^{-2(N-1)}}\right)^2} \tag{D.2}$$

$$\geq \frac{\sigma^2}{2a^{2(N-1)}} \tag{D.3}$$

$$\geq c_1\sigma^{2N}(-\log\sigma)^{N-1} \tag{D.4}$$

Case 2) $\frac{2^{l+1}}{\sigma\sqrt{-\log\sigma}} < a \leq \frac{2^{l+2}}{\sigma\sqrt{-\log\sigma}}$, for $l \geq 1$:

In this case, we bound the output distortion by the average distortion caused by a large jump to another segment. Consider z_1 as the additive noise in the first dimension.

For any point in the interval $-\frac{1}{2}+(k-1)a^{-1} < x \leq -\frac{1}{2}+ka^{-1}$ (for $1 \leq k \leq a-2^{l+1}$), when $z_1 > 2^{l+1}a^{-1}$, for any point $x' \leq x+2^l a^{-1}$, $f\left(x'+2^l a^{-1}\right)$ is closer to the received point $f(x)+\mathbf{z}$. Therefore, the decoded signal is $\tilde{x} > x+2^l a^{-1}$. Thus, in this case, the squared error is at least $\left(2^l a^{-1}\right)^2$. Therefore, the average distortion is lower bounded by

$$D \geq \Pr\left\{-\frac{1}{2} < x \leq \frac{1}{2} - 2^{l+1}a^{-1}\right\} \cdot \Pr\left\{z_1 > 2^{l+1}a^{-1}\right\} \cdot \left(2^l a^{-1}\right)^2 \qquad \text{(D.5)}$$

$$= \left(1-2^{l+1}a^{-1}\right) \cdot Q\left(\frac{2^{l+1}a^{-1}}{\sigma}\right) \cdot \left(2^l a^{-1}\right)^2 \qquad \text{(D.6)}$$

$$\geq \left(1-\frac{\sigma\sqrt{-\log\sigma}}{2}\right) \cdot Q\left(\frac{\sigma\sqrt{-\log\sigma}}{\sigma}\right) \cdot \left(\frac{\sigma\sqrt{-\log\sigma}}{2^2}\right)^2 \qquad \text{(D.7)}$$

$$= \left(1-\frac{\sigma\sqrt{-\log\sigma}}{2}\right) \cdot Q\left(\sqrt{-\log\sigma}\right) \cdot \frac{\sigma^2\left(-\log\sigma\right)}{2^4} \qquad \text{(D.8)}$$

By using $e^{-x^2} < Q(x) < e^{\frac{-x^2}{2}}$, for $x > 1$,

$$D \geq \left(1-\frac{\sigma\sqrt{-\log\sigma}}{2}\right) \cdot \sigma \cdot \frac{\sigma^2\left(-\log\sigma\right)}{2^4} \qquad \text{(D.9)}$$

$$\Longrightarrow D \geq c_2\sigma^3\left(-\log\sigma\right). \qquad \text{(D.10)}$$

By combining the bounds (for two cases),

$$D \geq \min\left\{c_2\sigma^3\left(-\log\sigma\right), c_1\sigma^{2N}(-\log\sigma)^{N-1}\right\} \tag{D.11}$$

$$D \geq c'\sigma^{2N}(-\log\sigma) \quad \text{for} \;\; N \geq 2. \tag{D.12}$$

Appendix E

Coding for unbounded sources

Consider $\{X_i\}_{i=1}^{\infty}$ as an arbitrary memoryless i.i.d source. We show that the results of Section V be extended for unbounded sources, to construct robust joint-source channel codes with a constraint on the average power. Without any loss of generality, we can assume the variance of the source as 1. For any source sample x, we can write it as $x = x_1 + x_2$ where x_1 is an integer and $-\frac{1}{2} \leq x_2 < \frac{1}{2}$, and $x_2 + \frac{1}{2} = \left(\overline{0 \cdot b_1 b_2 b_3 ...}\right)_2$. Now, we construct the N-dimensional transmission vector as $\mathbf{s}' = (s'_1, s'_2, ..., s'_N) = \left(x_1 + s_1 - \frac{1}{2}, s_2 - \frac{1}{2}, ..., s_N\right)$, where $s_1, ..., s_N$ are constructed using (5.35) in section V. Similar to the proof Theorem 6, we can show that the output distortion D is upper bounded by

$$D \leq c_1 \sigma^{2N} 2^{c_2 \sqrt{-\log \sigma}} \tag{E.1}$$

where c_1 and c_2 are constant. Thus, we only need to show that the average transmitted power is bounded. For $s'_2, ..., s'_N$, the transmitted power is bounded as $|s'_i|^2 \leq \frac{1}{4}$. For

s_1',

$$|s_i'|^2 = \left|x_1 + s_1 - \frac{1}{2}\right|^2 \leq \left(|x_1| + \left|s_1 - \frac{1}{2}\right|\right)^2 \tag{E.2}$$

$$\leq \left(|x| + \frac{1}{2} + \frac{1}{2}\right)^2 = (|x| + 1)^2 \tag{E.3}$$

$$\Longrightarrow \mathrm{E}\left\{|s_i'|^2\right\} \leq \left(\sqrt{\mathrm{E}|x|^2} + 1\right)^2 \tag{E.4}$$

$$\leq (1+1)^2 = 4. \tag{E.5}$$

Appendix F

Proof of Theorem 5.7

Without any loss of generality, we can consider $P = 1$. We consider two cases for a, the scaling factor:

Case 1) $a \leq 2^{N-M+3}\sigma^{-\frac{(N-M)}{M}}(-\log\sigma)^{\frac{-(N-M)}{2M}}$:

Each subset of the modulated signal set is the scaled version of a segment of the source signal set by a factor of a, hence, we can lower bound the distortion by only considering the case the subset is decoded correctly and there is no jump to adjacent subsets,

$$D \geq \mathrm{E}\left\{|\tilde{x} - x|^2 | no\ jump\right\} \tag{F.1}$$

$$= \frac{\sigma^2}{a^2} \tag{F.2}$$

$$\geq c_4 \sigma^{\frac{2N}{M}}(-\log\sigma)^{\frac{N-M}{M}} \tag{F.3}$$

Case 2) $2^{l+2(N-M)} < \frac{a}{\sigma^{-\frac{(N-M)}{M}}(-\log\sigma)^{\frac{-(N-M)}{2M}}} \leq 2^{l+1+2(N-M)}$ for $l \geq 3$:

In this case, we bound the output distortion by the average distortion caused by a jump to another subset, during the decoding. Without loss of generality, we can consider $\sigma < \left(\frac{1}{16}\right)^{\frac{M}{N-M}}$, hence $2^{-l}a > 8$. First, we show that there are two constants c_5 and c_6 (independent of a and σ) such that probability of an squared error of at least $c_5\left(2^{-l}a\right)^{-2}$ is lower bounded by

$$\Pr(jump) \geq c_6 Q\left(\sqrt{-\log\sigma}\right) \geq c_6\sigma \tag{F.4}$$

By considering the power constraint, the maximum distance of each source sample to its quantization point is upper bounded by

$$d_{max} \leq \frac{1}{a}. \tag{F.5}$$

Also, by considering the volumes of quantization regions $\mathcal{A}_1, ..., \mathcal{A}_k$ and their scaled versions, κ is lower bounded by

$$\kappa \geq \left(\frac{1}{2d_{max}}\right)^M \geq \frac{a^M}{2^M}. \tag{F.6}$$

We can partition the M-dimensional uniform source to $n = \left(\left\lfloor\frac{1}{2^l d_{max}}\right\rfloor\right)^M \geq \left(\frac{1}{2^{l+1}d_{max}}\right)^M \geq \left(\frac{a}{2^{l+1}}\right)^M$ cubes of size $s = \frac{1}{\left\lfloor\frac{1}{2^l d_{max}}\right\rfloor} \geq 2^l d_{max}$. We consider $\mathcal{B}_i$ as the union of the quantization regions whose center is in the ith cube $(1 \leq i \leq n)$. Because the decoding of digital and analog parts are done separately, the $(N-M)$-dimensional subspace (dedicated to send the quantization points) can be partitioned to n decoding subsets, corresponding to regions $\mathcal{B}_1, ..., \mathcal{B}_n$. If we consider $\mathcal{C}_1, ..., \mathcal{C}_n$ the intersections of these decoding regions and the $(N-M)$-dimensional cube of size 4, centered at the origin, at least $\frac{n}{2}$

of them have volume less than $2\left(\frac{(4)^{N-M}}{n}\right) \leq \frac{(4)^{N-M}}{\left(2^{-l-1}a\right)^{M}} \leq 2\sigma^{N-M}(-\log\sigma)^{\frac{N-M}{2}}$. This volume is less than the volume of an $(N-M)$-dimensional sphere of radius $\sigma(-\log\sigma)^{\frac{1}{2}}$. Thus, for any point inside $\mathcal{B}_i$ with this property, the probability of being decoded to a wrong subset $\mathcal{B}_j$ is at least $Q\left(\sqrt{-\log\sigma}\right) \geq \sigma$. Now, for the cubes corresponding to these subsets, we consider points inside a smaller cube of size $\frac{s}{4}$.

For these points, at least with probability σ, decoder finds a wrong quantization region where the distance of its center and the original point is at least $\frac{s-\frac{s}{4}}{2} > 2^{l-1}d_{max}$, hence, the final squared error is at least $\left(2^{l-1}d_{max} - d_{max}\right)^2 \geq c_5\left(2^{-l}a\right)^{-2}$.

Because at least half of the n subsets have the mentioned property, the overall probability of having this kind of points as the source is at least $\frac{1}{2}4^{-M}$. Therefore, the unconditioned probability of such an error in the decoding is at least $\frac{1}{2}4^{-M}\sigma$

Now, by considering the lower bound on the probability of this event and the distortion caused by this jump,

$$
\begin{aligned}
D &\geq c_7\sigma\left(2^{-l}a\right)^{-2} \\
&\geq c_8\sigma\cdot\sigma^{\frac{2(N-M)}{M}}(-\log\sigma)^{\frac{-(N-M)}{M}} \\
&= c_8\sigma^{\frac{2N-M}{M}}(-\log\sigma)^{\frac{N-M}{M}}.
\end{aligned} \tag{F.7}
$$

By considering the minimum of (F.3) and (F.7),

$$
D \geq c\sigma^{\frac{2N}{M}}(-\log\sigma)^{\frac{N-M}{M}}. \tag{F.8}
$$

References

[1] E. Agrell, T. Eriksson, A. Vardy, , and K. Zeger. Closest point search in lattices. *IEEE Trans. Info. Theory*, 48:2201–2214, August 2002.

[2] M Ajtai. The shortest vector problem in L2 is NP-hard for randomized reductions. In *Proceedings of the thirtieth annual ACM symposium on Theory of computing*, pages 10–19, 1998.

[3] S. M. Alamouti. A simple transmitter diversity scheme for wireless communications. *IEEE J. Selec. Areas Commun.*, 16:1451–1458, Oct. 1998.

[4] E Ayanoglu and R M Gray. The design of joint source and channel trellis waveform coders. *IEEE Trans. Info. Theory*, pages 855–865, Nov. 1987.

[5] Xiaodong Cai and J W. Modestino. Bandwidth expansion shannon mapping for analog error-control coding. In *Proc. 40th annual Conference on Information Sciences and Systems (CISS)*, March 2006.

[6] G. Caire and K. Narayanan. On the distortion SNR exponent of hybrid digital analog space time coding. *IEEE Trans. Info. Theory*, 53:2867 – 2878, August 2007.

[7] G. Caire and S. Shamai. On the acheivable throughput of a multiple-antenna Gaussian broadcast channel. *IEEE Trans. Info. Theory*, pages 1691–1706, July 2003.

[8] B Chen and G W Wornell. Analog error correcting codes based on chaotic dynamical systems. *IEEE Trans. Info. Theory*, pages 1691–1706, July 1998.

[9] S.-Y. Chung. *On the Construction of some capacity-approaching coding schemes.* PhD thesis, MIT, 2000.

[10] J H Conway and N J A Sloane. *Sphere packing, lattices and groups.* Springer-Verlag, 1999.

[11] J H Conway and N J A Sloane. *Sphere packing, lattices and groups.* Springer-Verlag, 1999.

[12] M. Costa. Writing on dirty paper. *IEEE Trans. Info. Theory*, 29:439–441, May 1983.

[13] H. Coward and T. A. Ramstad. Quantizer optimization in hybrid digitalanalog transmission of analog source signals. In *Proc. Intl. Conf. Acoustics, Speech, Signal Proc*, 2000.

[14] M O Damen, A Chkeif, and J C Belfiore. Lattice code decoder for space-time codes. *IEEE Trans. Info. Theory*, 50:968 – 985, June 2004.

[15] A Edelman. *Eigenvalues and Condition Numbers of Random Matrices*. PhD thesis, MIT, 1989.

[16] Gerald A Edgar. *Measure, Topology and Fractal Geometry*, chapter 6. Springer-Verlag, 1990.

[17] H. ElGamal and M. O. Damen. Universal space-time coding. *IEEE Trans. Info. Theory*, 49:1097–1119, May 2003.

[18] P Elia, K R Kumar, S A Pawar, and P V Kumar H.-F Lu. Explicit spacetime codes achieving the diversitymultiplexing gain tradeoff. *IEEE Trans. Info. Theory*, 52:3869–3884, Sep. 2006.

[19] U. Erez, S. Shamai, and R. Zamir. Capacity and lattice strategies for canceling known interference. *IEEE Trans. Info. Theory*, pages 3820 – 3833, Nov. 2005.

[20] N Farvardin. A study of vector quantization for noisy channels. *IEEE Trans. Info. Theory*, pages 799–809, July 1990.

[21] N Farvardin and V Vaishampayan. Optimal quantizer design for noisy channels: An approach to combined source-channel coding. *IEEE Trans. Info. Theory*, pages 827–838, Nov. 1987.

[22] R.F.H. Fischer. Lattice-reduction-aided equalization and generalized partial response signaling for point-to-point transmission over flat-fading MIMO channels.

In *6th International ITG-Conference on Source and Channel Coding*, Munich, Germany, April 2006.

[23] G D Forney and L F Wei. Coset codes. I. Introduction and geometrical classification. *IEEE Trans. Info. Theory*, 34:1123 – 1151, Sep. 1988.

[24] G D Forney and L F Wei. Multidimensional constellations-Part I: Introduction, figures of merit, and generalized cross constellations. *IEEE J. Select Areas Commun.*, 7:877–892, August 1989.

[25] G. J. Foschini. Layered space-time architechture for wireless communication in a fading environment when using multi-element antennas. *Bell labs. Tech. J.*, 1(2):41–59, 1996.

[26] J Galambos and I Simonelli. *Bonferroni-type inequalities with applications.* Springer-Verlag, 1996.

[27] H E Gamal, G Caire, and M O Damen. Lattice coding and decoding achieve the optimal diversity-multiplexing tradeoff of MIMO channels. *IEEE Trans. Info. Theory*, 50:968 – 985, June 2004.

[28] Martin Grotschel, Laszlo Lovasz, and Alexander Schrijver. *Geometrical algorithms and combinatorial optimization*, chapter 5. Springer-Verlag, 1993.

[29] P. M. Gruber and C. G. Lekkerkerker. *Geometry of numbers.* North-Holland, 1987.

[30] D. Gunduz and E. Erkip. Distortion exponents of mimo fading channels. In *IEEE Information Theory Workshop*, page 694698, July 2006.

[31] G. Han and J. Rosenthal. A numerical approach for designing unitary space-time codes with large diversity product and diversity sum. In *Proceedings of ISIT2003*, page 178.

[32] B. Hassibi and B. Hochwald. Cayley differential unitary space-time codes. *IEEE Trans. Info. Theory*, 48:1485–1502, June 2002.

[33] F Hekland. *On the design and analysis of Shannon-Kotelnikov mappings for joint source-channel coding*. PhD thesis, Norwegian University of Science and Technology, 2007.

[34] F. Hekland, G. E. Oien, and T. A. Ramstad. Design of VQ-based hybrid digital-analog joint source-channel codes for image communication. In *Data Compression Conference*, pages 193 – 202, March 2005.

[35] F. Hekland, G. E. Oien, and T. A. Ramstad. Using 2:1 shannon mapping for joint source-channel coding. In *Proc. Data Compression Conference*, March 2005.

[36] B. Helfrich. Algorithms to construct minkowski reduced and hermit reduced lattice bases. *Theoretical Computer Sci. 41*, pages 125–139, 1985.

[37] B. Hochwald and W. Sweldens. Differential unitary space-time modulation. *IEEE Trans. Commun.*, 48:2041–2052, Dec. 2000.

[38] Lu Hsiao-Feng and P V Kumar. A unified construction of space-time codes with optimal rate-diversity tradeoff. *IEEE Trans. Info. Theory*, 51:1709–1730, May 2005.

[39] B. Hughes. Differential space-time modulation. *IEEE Trans. Info Theory*, 46:2567–2578, Nov. 2000.

[40] Y. Jing and B. Hassibi. Design of fully-diverse multiple-antenna codes based on *Sp(2). submitted to IEEE Trans. Info. Theory*, 2003.

[41] Y Jing and B Hassibi. Three-transmit-antenna space-time codes based on SU(3). *submitted to IEEE Transactions on Information Theory*, 2003.

[42] Y Kochman and R Zamir. Analog matching of colored sources to colored channels. In *IEEE International Symposium on Information Theory*, pages 1539 – 1543, July 2006.

[43] Y Kochman and R Zamir. Approaching R(D)=C in colored joint source/channel broadcasting by prediction. In *IEEE International Symposium on Information Theory*, June 2007.

[44] V A Kotel'nikov. *The Theory of Optimum Noise Immunity.* McGraw-Hill, New York, 1959.

[45] A. K. Lenstra, H. W. Lenstra, and L. Lovasz. Factoring polynomials with rational coefficients. *Mathematische Annalen*, 261:515–534, 1982.

[46] Xue-Bin Liang and Xiang-Gen Xia. Unitary signal constellations for differential space-time modulation with two transmit antennas: parametric codes, optimal designs, and bounds. *IEEE Transactions on Information Theory*, pages 2291–2322, Aug. 2002.

[47] C Ling. Approximate lattice decoding: Primal versus dual basis reduction. In *Proceedings IEEE International Symposium on Information Theory*, pages 1–5, 2006.

[48] Marvin Markus and Henryk Minc. *A survey of matrix theory and matrix inequalities*. Dover, 1964.

[49] Udar Mittal and Nam Phamdo. Hybrid digital-analog (HDA) joint source-channel codes for broadcasting and robust communications. *IEEE Trans. Info. Theory*, pages 1082–1102, May 2002.

[50] W H Mow. Universal lattice decoding: Principle and recent advances. *Wireless Communications and Mobile Computing*, pages 553–569, August 2003.

[51] Y. Nam and H. ElGamal. On the optimality of lattice coding and decoding in multiple access channels. In *IEEE International Symposium on Information Theory*, 2007.

[52] H Napias. A generalization of the LLL algorithm over euclidean rings or orders. *Journal de thorie des nombres de Bordeaux*, 8:387–396, 1996.

[53] F. Oggier, G. Rekaya, J.-C. Belfiore, and E Viterbo. Perfect spacetime block codes. *IEEE Trans. Info. Theory*, 52:3885–3902, Sep. 2006.

[54] C. B. Peel, B. M. Hochwald, and A. L. Swindlehurst. A vector-perturbation technique for near-capacity multiple-antenna multi-user communications-Part I: channel inversion and regularization. *IEEE Trans. Comm.*, pages 195 – 202, Jan. 2005.

[55] C. B. Peel, B. M. Hochwald, and A. L. Swindlehurst. A vector-perturbation technique for near-capacity multiple-antenna multi-user communications-Part II: Perturbation. *IEEE Trans. Comm.*, pages 537 – 544, March 2005.

[56] N Phamdo, N Farvardin, and T Moriya. A unified approach to treestructured and multistage vector quantization for noisy channels. *IEEE Trans. Info. Theory*, pages 835–850, May 1993.

[57] T. A. Ramstad. Shannon mappings for robust communication. *Telektronikk*, 98(1):114128, 2002.

[58] Z Reznic, M Feder, and R Zamir. Distortion bounds for broadcasting with bandwidth expansion. *IEEE Trans. Info. Theory*, pages 3778–3788, Aug. 2006.

[59] I Rosenhouse and A J. Weiss. Combined analog and digital error-correcting codes for analog information sources. *To appear in IEEE Trans. Communications.*

[60] N Santhi and A Vardy. Analog codes on graphs. In *Proceedings of ISIT*, page 13, July 2003.

[61] N Santhi and A Vardy. Analog codes on graphs. *Available online: http://arxiv.org/abs/cs/0608086*, 2006.

[62] D. A. Schmidt, M Joham, and W Utschick. Minimum mean square error vector precoding. In *PIMRC 2005*, pages 107–111, 2005.

[63] S Shamai (Shitz), S Verdu, and R. Zamir. Systematic lossy source/channel coding. *IEEE Trans. Info. Theory*, page 564 579, March 1998.

[64] C Shannon. A mathematical theory of communication. *Bell Syst. Tech. J.*, 1948.

[65] C Shannon. Coding theorems for a discrete source with a fidelity criterion. *IRE Nat. Conv. Rec.*, pages 149–163, Mar. 1959.

[66] C E Shannon. Communication in the presence of noise. *Proceedings of IRE*, 37:10–21, 1949.

[67] A. Shokrollahi. Computing the performance of unitary space-time group codes from their character table. *IEEE Information Theory*, pages 1355 –1371, June 2002.

[68] M Skoglund, N Phamdo, and F F Alajaji. Design and performance of VQ-based hybrid digital-analog joint source-channel codes. *IEEE Trans. Info. Theory*, 48:708 – 720, March 2002.

[69] M Skoglund, N Phamdo, and F F Alajaji. Hybrid digital analog source channel coding for bandwidth compression/expansion. *IEEE Trans. Info. Theory*, 52:3757–3763, Aug. 2006.

[70] M Taherzadeh and A Khandani. Spectrally efficient differential spacetime coding using non-full-diverse constellations. *IEEE Transactions on Communications*, 54:1887 – 1895, Oct. 2006.

[71] M Taherzadeh and A Khandani. On the limitations of the naive lattice decoding. *Submitted to IEEE Transactions on Information Theory, Revised,*, 2007.

[72] M Taherzadeh and A Khandani. Robust joint source-channel coding for delay-limited applications. *submitted to IEEE Transactions on Information Theory*, 2008.

[73] M Taherzadeh and A K Khandani. Analog coding for delay-limited applications. In *Proceedings of 41st Conference on Information Sciences and Systems*, March 2007.

[74] M Taherzadeh and A K Khandani. Robust joint source-channel coding for delay-limited applications. In *ISIT 2007*, June 2007.

[75] M Taherzadeh, A Mobasher, and A Khandani. Communication over MIMO broadcast channels using lattice-basis reduction. *IEEE Transactions on Information Theory*, 53:4567 – 4582, Dec. 2007.

[76] M Taherzadeh, A Mobasher, and A Khandani. LLL reduction achieves the receive diversity in MIMO decoding. *IEEE Transactions on Information Theory*, 53:4801 – 4805, Dec. 2007.

[77] M. Taherzadeh, A. Mobasher, and A. K. Khandani. Communication over MIMO broadcast channels using lattice-basis reduction. In *Proceedings of 42nd Allerton Conference*, Sep. 2004.

[78] M. Taherzadeh, A. Mobasher, and A. K. Khandani. LLL lattice-basis reduction achieves the maximum diversity in mimo systems. In *Proceedings ISIT 2005*, Sep. 2005.

[79] V. Tarokh and H. Jafarkhani. A differential detection space-time block coding for time-selective fading channels. *IEEE J. Selec. Areas Commun.*, 18:1169–1174, July 2000.

[80] V. Tarokh, H. Jafarkhani, and A. R. Calderbank. Space-time block codes from orthogonal designs. *IEEE Trans. Info. Theory*, 45:1456–1467, July 1999.

[81] V. Tarokh, N. Seshadri, and A. R. Calderbank. Space-time codes for high data rate wireless communication: Performance criterion and code construction. *IEEE Trans. Info. Theory*, 44:744–765, Mar. 1998.

[82] I. E. Telatar. Capacity of multi-antenna gaussian channels. *Europ. Trans. Telecommun.*, pages 585–595, Nov. 1999.

[83] C M Thomas, C L May, and G R Welti. Design of signals for analog communications. *IEEE Trans. on Info. Theory*, 16:581–587, Sep. 1970.

[84] U Timor. Design of signals for analog communications. *IEEE Trans. on Info. Theory*, 16:581–587, Sep. 1970.

[85] David Tse and Pramod Viswanath. *Fundamentals of Wireless Communication.* Cambridge University Press, 2005.

[86] V A Vaishampayan and I R Costa. Curves on a sphere, shift-map dynamics, and error control for continuous alphabet sources. *IEEE Trans. Info. Theory*, pages 1691–1706, July 2003.

[87] S. Vishwanath, N. Jindal, and A. Goldsmith. Duality, acheivable rates and sum capacity of Gaussian MIMO broadcast channels. *IEEE Trans. Info. Theory*, pages 2658–2658, August 2003.

[88] P. Viswanath and D. Tse. Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality. *IEEE Trans. Info Theory*, pages 1912–1921, August 2003.

[89] M Wang and T R Fischer. Trellis-coded quantization designed for noisy channels. *IEEE Trans. Info. Theory*, pages 1792–1802, Nov. 1994.

[90] C. Windpassinger, R. F. H. Fischer, and J. B. Huber. Lattice-reduction-aided broadcast precoding. In *5th International ITG Conference on Source and Channel Coding (SCC)*, pages 403–408, Erlangen, Germany, January 2004.

[91] C. Windpassinger and R.F.H. Fischer. Low-complexity near-maximum-likelihood detection and precoding for MIMO systems using lattice reduction. In *Proceedings of Information Theory Workshop*, 2003.

[92] C. Windpassinger, R.F.H. Fischer, and J B Huber. Lattice-reduction-aided broadcast precoding. *IEEE Trans. Communications*, pages 2057–2060, Dec. 2004.

[93] J M Wozencraft and I M Jacobs. *Principles of Communication Engineering*. Wiley, New York, 1965.

[94] H Yao and G W Wornell. Lattice-reduction-aided detectors for MIMO communication systems. In *Proceeding of CISS 2004*, Nov 2002.

[95] W. Yu and J. Cioffi. Sum capacity of a Gaussian vector broadcast channel. In *Proceedings IEEE International Symposium on Information Theory*, page 498, 2002.

[96] L. Zheng and D. Tse. Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels. *IEEE Trans. Info. Theory*, pages 1073–1096, May 2003.

[97] J. Ziv. The behavior of analog communication systems. *IEEE Trans. Info. Theory*, pages 1691–1706, Aug. 1970.