TEL AVIV UNIVERSITY אוניברסיטת תל-אביב
FACULTY OF ENGINEERING הפאקולטה להנדסה

THE IBY AND ALADAR FLEISCHMAN FACULTY OF ENGINEERING
DEPARTMENT OF ELECTRICAL ENGINEERING - SYSTEMS

Precoding for Interference Cancellation at Low SNR using Nested Codes

Thesis submitted for the degree
"Master of Science in Electrical Engineering"

by

Tal Philosof

Submitted to the Senate of Tel-Aviv University

Jan, 2003

TEL AVIV UNIVERSITY אוניברסיטת תל-אביב
FACULTY OF ENGINEERING הפאקולטה להנדסה

THE IBY AND ALADAR FLEISCHMAN FACULTY OF ENGINEERING
DEPARTMENT OF ELECTRICAL ENGINEERING - SYSTEMS

Precoding for Interference Cancellation at Low SNR using Nested Codes

Thesis submitted for the degree
"Master of Science in Electrical Engineering"

by

Tal Philosof

Submitted to the Senate of Tel-Aviv University

This research work was carried out at the
Ddepartment of Electrical Engineering - Systems,
Tel-Aviv University
Under the supervision of
Dr. Ram Zamir

Jan, 2003

Acknowledgments

I wish to thank Dr. Ram Zamir for his supervising and the professional guidance and for helpful discussions and comments in this research, and the special opportunity to know an interesting area beyond the scope of this research.

I would like to acknowledge Uri Erez from the department of Electrical Engineering - system of Tel-Aviv university for his enthusiasm to help and the inspiring discussions and ideas during this work.

I would like to thank Dr. G. D. Forney for his helpful comments for the main part (chapter 4) of this work.

To Ronit for the support and understanding all along.
And to my parents.

Abstract

Transmission with side information at the transmitter was introduced for discrete memoryless channel (DMC) with causal side information of the channel state by Shannon, in 1958. Shannon proved that the capacity is equal to that of derived DMC channel with no side information. Gelfand and Pinsker, in 1980, showed how to achieve the capacity of DMC channel with non-causal side information using random binning. Furthermore, Costa in "writing on dirty paper" showed that for i.i.d Gaussian noise and i.i.d Gaussian interference known non-causally, the capacity is like there is no interference.

Meanwhile, shaping and precoding techniques were commonly used for AWGN channels, and inter-symbol interference (ISI) channels. Tomlinson-Harshima, in 1971, introduced a precoding technique for ISI channels which cancels ISI at the transmitter and achieves coding gain but without shaping gain. A well known improvement to Tomlinson-Harshima precoding is the combined shaping and precoding scheme of Eyouboglu and Forney, trellis precoding. Usually, the Tomlinson-Harshima precoding, and trellis precoding techniques have been used for high SNR in ISI-channels.

Nested codes were considered by several authors. Conway and Sloane used nested lattice codes structure. Forney also considered nested lattice codes and applied these codes to trellis shaping and trellis precoding. Nested codes as a binning scheme were introduced by Shamai, Verdu and Zamir, in 1996, using linear codes, and extended in 1998 by Zamir and Shamai to nested lattice codes. Barron, Chen and Wornell pointed out the application of nested codes to digital Watermarking and to Costa's problem.

Recently, Erez, Shamai and Zamir, developed structured coding schemes for cancelling known interference using nested lattice strategies/codes ("lattice precoding"). ESZ pointed out the connection between shaping and precoding techniques and the information theoretical problem of transmission with side information. The lattice precoding scheme uses a fine code and a coarse code,

where the fine code determines the coding gain while the coarse code the shaping gain. Furthermore, ESZ showed the connection between lattice precoding and multi-terminal network problems.

Most of the existing precoding schemes were design for high SNR. In this work we consider practical aspects of precoding at low SNR. We investigate the interference channel shaping gap bound and present schemes for optimal transmission at low SNR. First, we present a one dimensional (causal) precoding, and we show improvement for TH precoding with feedforward equalizer (FFE) using convolutional code which combines MMSE estimation and dithering at the low SNR regime. Although the scheme is used for ISI-channels, in fact it may be used for any interference channel known at the transmitter.

The main result of our study is for high dimensional (non-causal) precoding. We propose a low rate precoding scheme which combines MMSE estimation, dithering and a variant of nested codes, based on concatenation of a "syndrome dilution" code and a "syndrome-to-coset" modulation code, which approaches the interference channel capacity at low SNR. We show practical implementation based on nested trellis codes which is a variant of trellis precoding for the low SNR regime.

Nomenclature

$\mathbf{x}$	Vector representation of $[x_1, x_2, \ldots, x_n]^T$
$\boldsymbol{I(X;Y)}$	Mutual information between X and Y
$\boldsymbol{h(X)}$	Differential entropy of X
$\mathbb{R}^{\boldsymbol{N}}$	The continuous K dimensional space
$\mathbb{Z}^{\boldsymbol{N}}$	The discrete integer K dimensional space
$\boldsymbol{GF(p^m)}$	Galois field where p is prime number and m is integer
$\boldsymbol{\mathcal{C}_c(n_c, k_C)}$	Linear fine code
$\boldsymbol{\mathcal{C}_s(n_s, k_s)}$	Linear shaping code
$\boldsymbol{G_c}$	Generator matrix of $\mathcal{C}_c$ code
$\boldsymbol{G_s}$	Generator matrix of $\mathcal{C}_s$ code
$\boldsymbol{H_s}$	Parity check matrix of $\mathcal{C}_s$ code
$\boldsymbol{H_s^{-1}}$	Pseudo right inverse of parity check matrix
$\boldsymbol{R}$	Transmission rate bit/dim
$\mathbb{Z}^{\boldsymbol{N}}$	The integer lattice
$\boldsymbol{\Lambda_c}$	Fine lattice
$\boldsymbol{\Lambda_s}$	Coarse lattice
$\boldsymbol{G(\Lambda)}$	Normalized second moment of lattice Λ
$\boldsymbol{\Lambda_1/\Lambda_2/\Lambda_3\ldots}$	Lattices chain
$\boldsymbol{G(Z)}$	The Z-transform of the series g_n
$\boldsymbol{R_{xx}[l]}$	The auto-correlation function of a discrete process x_n
$\boldsymbol{R_{xy}[l]}$	The cross-correlation function of a discrete processes x_n, y_n
$\boldsymbol{S_{xx}(e^{jw})}$	The spectrum of a discrete process x_n
$\boldsymbol{S_{xy}(e^{jw})}$	The cross-spectrum of a discrete processes x_n, y_n
$\boldsymbol{SNR}$	Signal to noise ratio
$\boldsymbol{LCM(x,y)}$	Least common multiplier of x and y

Contents

List of Figures

Chapter 1

Preface

1.1 Introduction

In this study we consider the interference channel:

$$Y = X + S + Z, \tag{1.1}$$

where X is the channel input with power constraint $EX^2 \leq P_X$, Z is i.i.d Gaussian noise $Z \sim N(0, \sigma_z^2)$, and S is an arbitrary interference that is known to the transmitter as shown in Figure 1.1. Let the SNR be P_X/σ_z^2. In this

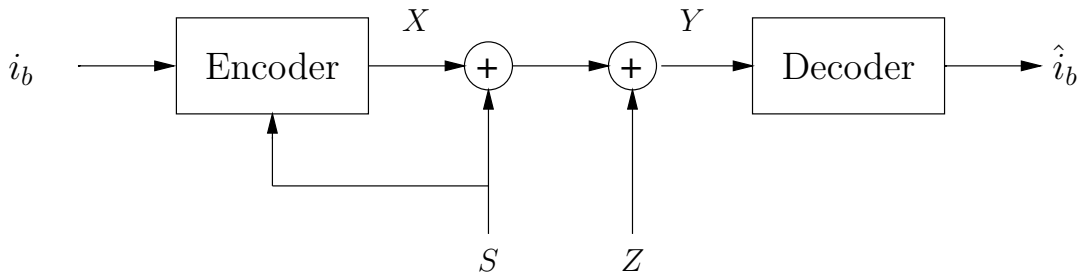

Figure 1.1: Interference channel

model the interference signal may be known to the transmitter causally, meaning only the past and the current interference states are known, or non-causally with finite anticipation or even infinite anticipation. The above model does not explicitly define the interference signal, therefore it can describe several

scenarios such as:

• ISI channel - where the causal part of the ISI, which is caused by previous symbols known to the transmitter, may play the role of S.

• Vector transmission over broadcast channel (MIMO channel) - the transmitter, which is common to all users, has multiple antennas [2, 28, 14]. The received signal of user i is interfered with a signal which has been transmitted for user j. Assuming that the channel is fully known, the interference signal of user i which is caused by user j is the known interference (S) to the i transmitter.

• Analog and digital transmission - digital transmission over the traditional analog communication channels without any effects on the analog transmission is required. For this hybrid transmission the analog signal acts as known interference signal which is known to the digital transmitter. In these channels the digital transmission is needed to operate at a *very* low SNR, since the digital signal has to be significant below the noise level in order not to damage the analog transmission.

• Digital watermarking/information embedding - in this case the side information signal S is considered as a "host" signal which carries the information where the watermark code distorts the host signal as shown in [1].

1.2 Background

1.2.1 Precoding

Tomlinson-Harshima (TH) precoding [25, 15] is used for ISI channel and it can be derived from DFE-equalizer. The DFE-equalizer handles the ISI using

feedforward and feedback equalizers, where the feedback equalizer operates on the former tentative symbols decision (post cursor). In TH precoding the feedback equalizer has been transferred to the transmitter, in which the past symbols are known exactly, furthermore TH precoding uses one dimensional modulo operation at the transmitter and receiver. In this way TH precodoer pre-compensate the post cursor ISI. Since the modulo operation is one dimensional, the TH precoding can introduce only coding gain by using coded transmission such as [26, 27]. Therefore, TH precoding has no shaping gain. A well known improvement to the TH precoding is the combined shaping and precoding scheme of Eyuboglu and Forney (EF), trellis precoding [10]. In this scheme EF combined between TH precoding and trellis shaping [17]. The trellis shaping technique shapes the transmitted symbols to have nearly Gaussian distribution, in order to gain significant part of the entire shaping gain (1.53 dB at high SNR). This is achieved by convolutional code and Viterbi algorithm (VA) for minimum norm search, effectively in multi-dimensional space the transmitted signal is enforced to be in Voronoi like region [16]. For high rate transmission, by using the continues approximation [11], the transmitted signal is uniformly distributed over Voronoi region, therefore having nearly Gaussian distribution on each dimension. The trellis precoding combines the trellis shaping and TH precoding with coded transmission, thus trellis precoding achieves both shaping and coding gains over ISI channel.

1.2.2 Side Information at the Transmitter

For general discrete memoryless channel $p(y|x,s)$, with memoryless states s known causally to the transmitter, Shannon [23] showed that the capacity is

equal to the capacity of a derived DMC channel. The input alphabet of the derived channel is the all possible mappings

$$t : S \to X,$$

meaning $|T| = |x|^{|S|}$. The output y of the derived channel is related to the input t by

$$p(y|t) = \sum_s p(s)p(y|x = t(s), s). \tag{1.2}$$

Therefore, the capacity is

$$C = \max_{p(t)} I(Y;T). \tag{1.3}$$

Gelfand-Pinsker [13] showed using random binning, that the capacity of DMC with random state S known non-causally to the encoder is given by

$$C = \max_{p(u,x|s)} \{I(U;Y) - I(U;S)\}, \tag{1.4}$$

where the maximum is over all the joint distributions of the form $p(s)p(u,x|s)p(y|x,s)$, and U is random variable from a finite set. This result can be extended to continuous memoryless channels. Costa [5] showed that for Gaussian i.i.d interference $\mathbf{S} \sim N(0, QI)$ that is known non-causally at the transmitter, and Gaussian i.i.d noise $\mathbf{Z} \sim N(0, \sigma_z^2 I)$ the capacity is

$$C = \frac{1}{2}\log_2\Big(1 + \frac{P}{\sigma_z^2}\Big). \tag{1.5}$$

This result shows, that the capacity of Gaussian interference channel is equal to non interference case.

1.2.3 Lattice

Definition 1.1. *Lattice Λ is a set of $\mathcal{L} = \{\mathbf{l_i}\}$ where $\mathbf{l_i} \in \mathbb{R}^N$ so that $\mathbf{l_i} + \mathbf{l_j} \in \mathcal{L}$ $\forall i, j$ and $l_0 = \underline{0} \in \mathcal{L}$, therefore lattice is additive group.*

The lattice Λ is spanned by N basis vectors $\mathbf{g_i},\ i = 1, \ldots, N$, which are the columns of the $N \times N$ generator matrix G, therefore we have

$$\Lambda = \{G\mathbf{m},\ \mathbf{m} \in \mathbb{Z}^N\}.$$

The lattice partition $\{\mathcal{V}_i\}$ for each points is a collection of disjoint regions which tessellate $\mathbb{R}^N$, and satisfy $\mathcal{V}_i = \mathbf{l}_i + \mathcal{V}_0$ where $\mathcal{V}_0$ is the basic cell of the lattice Λ. In case that all the points in $\mathcal{V}_i$ region are nearest to the lattice point $\mathbf{l}_i$ than from all other lattice points $\mathbf{l}_j,\ j \neq i$ (in sense of Euclidian distance), then the lattice has Voronoi partition. Now, $\mathcal{V}_i$ is the i-th cell with volume V, which is equal to all cells. The normalized second moment [4] is given by

$$G_N(\Lambda) = \frac{1}{N} \frac{\int_{\mathcal{V}_0} \| \mathbf{x} \|^2 \, d\mathbf{x}}{V^{1+2/N}}, \tag{1.6}$$

for high dimensional lattice, $N \to \infty$, the minimum value of the normalized second moment is $G_N(\Lambda) \approx 1/2\pi e$ [29]. The second moment per dimension of a uniform distribution over $\mathcal{V}_0$ is

$$\sigma^2 = \frac{1}{N} \frac{\int_{P_0} \| \mathbf{x} \|^2 \, d\mathbf{x}}{V} = G_N(\Lambda) V^{2/N}. \tag{1.7}$$

The ultimate shaping gain is the reduction in average power of lattice Λ relative to the N-cube lattice $\mathbb{Z}^N$ with the same volume, for Λ lattice and using continues approximation [11] is given by

$$\gamma_s = \frac{V^{2/N}}{12\sigma^2} = \frac{1}{12 G_N(\Lambda)}. \tag{1.8}$$

For $G_N(\Lambda) \approx 1/2\pi e$, the ultimate shaping gain is $\gamma_s = 1.423 = 1.53\ dB$

1.2.4 Lattice Precoding

Nested block codes are considered in [24, 30] using linear/lattice codes for Wyner-Ziv encoding. The application of nested codes to digital watermarking

and to Costa's problem has been considered by Barron, Chen and Wornell in [1]. In [8, 31] Erez, Shamai, and Zamir showed a coding scheme for cancelling known interference using nested lattice strategies/codes, lattice precoding. Their scheme implies on a connection between precoding schemes and the transmission with side information as handled in [5, 13]. The lattice precoding scheme uses a fine lattice - Λ_c in which a coarse lattice - Λ_s is nested, where the basic cell of the coarse lattice, $\mathcal{V}_s$, defines the region of the code, and the codewords are points of the fine lattice, as shown in Figure 1.2. Therefore, the coarse lattice determines the shaping gain while the fine lattice the coding gain. Although Λ_c and Λ_s belong to the same multi-dimensional space, their *effective* dimensions N_c and N_s may differ. The ESZ result is a gener-

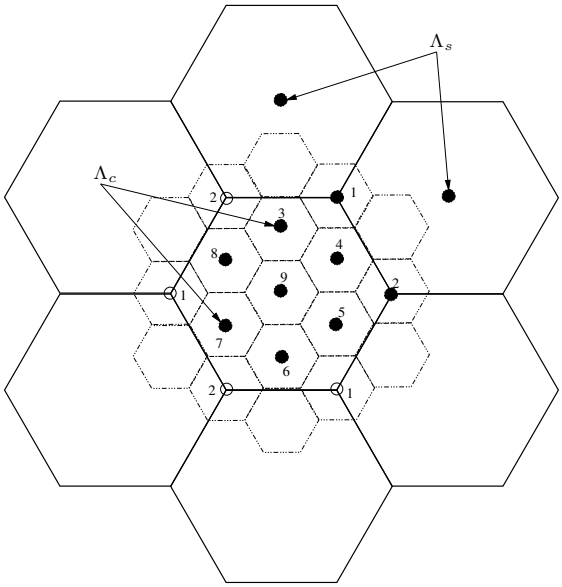

Figure 1.2: Nested lattices

alization of Costa's result in the sense that for any interference the capacity (1.5) is achieved. In order to achieve this result, the lattice strategies incorporate common randomness ("dither"), which is distributed uniformly over $\mathcal{V}_s$,

and minimum mean squared error (MMSE) estimation. The lattice precoding scheme may be regarded as a generalization of Tomlinson-Harashima (TH) precoding [25, 15], where S plays the role of post cursor ISI, and the scalar modulo operation of TH is equivalent to $\mathbb{Z}^{N_s}$ coarse lattice, implying effectively one dimensional ($N_s = 1$) shaping, thus no shaping gain.

1.3 Main Achievements

In this work we investigate the shaping gap for interference channel and some aspects of one dimensional (causal) precoding at low SNR. Furthermore, we present modifications of existing shaping and precoding scheme for transmission at the low SNR regime using lattice strategies.

Initially, we investigate one dimensional precoding scheme which is mainly introduced for ISI-channels and we show an improvement for TH precoding with feedforward equalizer (FFE) which combines MMSE estimation and dithering at low SNR.

The main part of this work deals with high dimensional (non-causal) precoding. We present nested lattices construction and investigate its loss from the capacity. Furthermore, we show the "generalized inflated lattice lemma", and derive a transmission scheme which combines MMSE estimation, dithering and a variant of nested codes based on concatenation, which approaches the interference channel capacity at low SNR. We provide simulation results for several configurations of nested trellis codes, which reduce the gap to capacity.

1.3.1 Organization of the Thesis

Chapter 1 - is a introduction and background in precoding schemes, transmission with side information at the transmitter and nested lattices.

Chapter 2 - summarizes the lattice strategies and focuses on the behavior of the shaping gap for the interference channel.

Chapter 3 - considers one dimensional (casual) precoding for ISI-channels.

Chapter 4 - focuses on high dimensional (non-casual) precoding for the interference channel at low SNR.

Chapter 5 - summary.

Chapter 2

Lattice Strategies

In this chapter we show the precoding shaping gap behavior and investigate some aspects of transmission at low SNR using lattice strategies over the interference channel.

2.1 Lattice Strategies - Background

In order to achieve the interference channel capacity (1.1), ESZ proposed the following lattice strategies transmission.

Encoder: transmit the error vector between $\mathbf{v}$ and $\alpha\mathbf{s} + \mathbf{d}$.

$$\mathbf{x} = \Big[\mathbf{v} - \alpha\mathbf{s} - \mathbf{d}\Big] \, mod \, \Lambda_s, \tag{2.1}$$

where $\mathbf{v}$ is a member of the basic Voronoi region of Λ_s, which carries the information. $\mathbf{s}$ is the interference vector, and $\mathbf{d} \sim U(\mathcal{V}_s)$ is the dither signal.

Decoder: the channel output $\mathbf{y}$ is multiplied by α and then reduced modulo Λ_s, i.e,

$$\mathbf{y}' = \alpha\mathbf{y} \bmod \Lambda_s. \tag{2.2}$$

Using the "inflated lattice Lemma" [9, 8] ESZ showed, that the resulting channel is equivalent interference-free modulo-Λ_s channel, for a later use we show

the proof.

Lemma 2.1. *(**Inflated lattice lemma[8]**) The channel defined by (1.1), (2.1) and (2.2) satisfies*

$$\mathbf{Y}' = \Big[\mathbf{V} + \mathbf{Z}'\Big] \ mod\, \Lambda_s, \tag{2.3}$$

with

$$\mathbf{Z}' = \Big[(1-\alpha)\mathbf{U} + \alpha\mathbf{Z}\Big] \ mod\, \Lambda_s, \tag{2.4}$$

where $\mathbf{U}$ *is distributed uniformity over Voronoi region of* Λ_s *and is statistically independent of* $\mathbf{Z}$, $\mathbf{V}$.

Proof. From (1.1), (2.1) and (2.2)

$$\begin{aligned} Y &= X + S + Z && (2.5)\\ \mathbf{x} &= \Big[\mathbf{v} - \alpha\mathbf{s} - \mathbf{d}\Big] \bmod \Lambda_s && (2.6)\\ \mathbf{y}' &= \Big[\alpha\mathbf{y} + \mathbf{d}\Big] \bmod \Lambda_s && (2.7) \end{aligned}$$

$\mathbf{y}'$ can be written as

$$\begin{aligned}
\mathbf{y}' &= \Big[\alpha(\mathbf{x}+\mathbf{s}+\mathbf{z})+\mathbf{d}\Big] \bmod \Lambda_s && (2.8)\\
&= \Big\{\alpha\Big[(\mathbf{v}-\mathbf{d}-\alpha\mathbf{s})\bmod \Lambda_s+\mathbf{s}+\mathbf{z}\Big]+\mathbf{d}\Big\} \bmod \Lambda_s && (2.9)\\
&= \Big\{\alpha\Big[(\mathbf{v}-\mathbf{d}-\alpha\mathbf{s})-Q_{\Lambda_s}(\mathbf{v}-\mathbf{d}-\alpha\mathbf{s})+\\
&\quad +\mathbf{s}+\mathbf{z}+\mathbf{d}/\alpha\Big]\Big\} \bmod \Lambda_s && (2.10)\\
&= \Big\{\alpha\Big[\mathbf{v}-Q_{\Lambda_s}(\mathbf{v}-\mathbf{d}-\alpha\mathbf{s})+\frac{1-\alpha}{\alpha}\Big[\mathbf{v}-Q_{\Lambda_s}(\mathbf{v}-\mathbf{d}-\alpha\mathbf{s})\Big]-\\
&\quad -\frac{1-\alpha}{\alpha}\Big[\mathbf{v}-Q_{\Lambda_s}(\mathbf{v}-\mathbf{d}-\alpha\mathbf{s})\Big]+(1-\alpha)\mathbf{s}+\\
&\quad +\frac{1-\alpha}{\alpha}\mathbf{d}+\mathbf{z}\Big]\Big\} \bmod \Lambda_s && (2.11)\\
&= \Big\{\alpha\Big[\frac{1}{\alpha}\Big[\mathbf{v}-Q_{\Lambda_s}(\mathbf{v}-\mathbf{d}-\alpha\mathbf{s})\Big]+\\
&\quad +\frac{1-\alpha}{\alpha}\Big[Q_{\Lambda_s}(\mathbf{v}-\mathbf{d}-\alpha\mathbf{s})-\mathbf{v}+\mathbf{d}+\alpha\mathbf{s}\Big]+\mathbf{z}\Big]\Big\} \bmod \Lambda_s && (2.12)\\
&= \Big\{\mathbf{v}-Q_{\Lambda_s}(\mathbf{v}-\mathbf{d}-\alpha\mathbf{s})+\\
&\quad +(1-\alpha)\Big[Q_{\Lambda_s}(\mathbf{v}-\mathbf{d}-\alpha\mathbf{s})-\mathbf{v}+\mathbf{d}+\alpha\mathbf{s}\Big]+\alpha\mathbf{z}\Big\} \bmod \Lambda_s && (2.13)\\
&= \Big\{\mathbf{v}+(1-\alpha)\Big[Q_{\Lambda_s}(\mathbf{v}-\mathbf{d}-\alpha\mathbf{s})-(\mathbf{v}-\mathbf{d}-\alpha\mathbf{s})\Big]+\\
&\quad +\alpha\mathbf{z}\Big\} \bmod \Lambda_s. && (2.14)
\end{aligned}$$

Let $\mathbf{U}\triangleq Q_{\Lambda_s}(\mathbf{v}-\mathbf{d}-\alpha\mathbf{s})-(\mathbf{v}-\mathbf{d}-\alpha\mathbf{s})$ we have that

$$\mathbf{y}'=\Big[\mathbf{v}+(1-\alpha)\mathbf{U}+\alpha\mathbf{z}\Big] \bmod \Lambda_s, \qquad (2.15)$$

since the dither $\mathbf{d}\sim U(\mathcal{V}_s)$, and $\mathbf{U}$ is the quantization error of the quantizer Q_{Λ_s}, by the dithered quantization property $\mathbf{U}\sim U(\mathcal{V}_s)$. The lemma is proved

using the following definition

$$\mathbf{Z}' = \Big[(1-\alpha)\mathbf{U} + \alpha\mathbf{Z}\Big] \bmod \Lambda_s. \tag{2.16}$$

□

The "effective noise", $\mathbf{Z}'$, distribution is a mixture of the Gaussian noise distribution and the "self noise" component $(1-\alpha)\mathbf{U}$, where $\mathbf{U}$ is uniformly distributed over $\mathcal{V}_s$, folded into $\mathcal{V}_s$. For the achievable rate let $\mathbf{V} \sim U(\mathcal{V}_s)$, and $\alpha = \frac{P_X}{P_X+\sigma_z^2}$, therefore

$$E\mathbf{Z}'^2 \leq (1-\alpha)^2 Var(\mathbf{U}) + \alpha^2 Var(\mathbf{Z}) = \frac{\sigma_z^2 P_X}{\sigma_z^2 + P_X} \tag{2.17}$$

the rate which is achieved using lattice strategies is

$$\begin{aligned}
R &= \frac{1}{N} I(\mathbf{V};\mathbf{Y}') && (2.18)\\
&= \frac{1}{N} h(\mathbf{Y}') - \frac{1}{N} h(\mathbf{Y}'|\mathbf{V}) && (2.19)\\
&= \frac{1}{N} h(\mathbf{Y}') - \frac{1}{N} h(\mathbf{Z}') && (2.20)\\
&= \frac{1}{N} h(\mathbf{Y}') - \frac{1}{N} h(\mathbf{Z}') && (2.21)\\
&\geq \frac{1}{N}\log(V_s) - \frac{1}{2}\log\left(2\pi e \frac{\sigma_z^2 P_X}{\sigma_z^2 + P_X}\right) && (2.22)\\
&= \frac{1}{2}\log\left(\frac{P_X}{G_N(\Lambda_s)}\right) - \frac{1}{2}\log\left(2\pi e \frac{\sigma_z^2 P_X}{\sigma_z^2 + P_X}\right) && (2.23)\\
&= \frac{1}{2}\log\left(1 + \frac{P_X}{\sigma_z^2}\right) - \frac{1}{2}\log(2\pi e G_N(\Lambda_s)). && (2.24)
\end{aligned}$$

Where the inequality (2.22) is from that the differential entropy is bounded by Gaussian distribution [6, 12], and since $E\mathbf{Z}'^2 \leq \frac{\sigma_z^2 P_X}{\sigma_z^2+P_X}$. Furthermore, for high dimensional lattice so that the normalized second moment $G_N(\Lambda_s) \approx 1/2\pi e$, and $\mathbf{Z}'$ is close to Gaussian in the divergence sense [29], i.e, $\frac{1}{N}h(\mathbf{Z}') \to \frac{1}{2}\log\left(2\pi e \frac{\sigma_z^2 P_X}{\sigma_z^2+P_X}\right)$. Therefore, the lattice strategies can achieve $R = \frac{1}{2}\log(1 + \frac{P_X}{\sigma_z^2})$.

2.1.1 Nested Lattice Codes

The lattice precoding scheme uses nested lattices structure for transmission scheme. Let Λ_c and Λ_s lattices in $\mathbb{R}^N$, with lattice decoding function, nearest neighbor quantizers, Q_{Λ_c}, Q_{Λ_s}. Basic Voronoi cells $\mathcal{V}_c = \{\mathbf{x} : Q_{\Lambda_c}(\mathbf{x}) = 0\}$ and $\mathcal{V}_s = \{\mathbf{x} : Q_{\Lambda_s}(\mathbf{x}) = 0\}$, and cell volumes V_c, V_s, so that

a) $\Lambda_s \subseteq \Lambda_c$ - Λ_s is nested in Λ_c.

b) The lattices second moment are

$$\begin{aligned} \sigma_s^2 &= P_X \\ \sigma_c^2 &= \frac{P_X \sigma_z^2}{P_X + \sigma_z^2}. \end{aligned} \tag{2.25}$$

c) As in [30], Λ_c, Λ_s are good lattices with high dimensions and normalized second moments

$$\begin{aligned} \lim_{N\to\infty} G_N(\Lambda_s) &= 1/2\pi e \\ \lim_{N\to\infty} G_N(\Lambda_c) &= 1/2\pi e. \end{aligned} \tag{2.26}$$

2.1.2 Transmission Scheme

Message selection: each message selects a unique coset leader of Λ_s in Λ_c, i.e $\mathbf{v} \in \{\Lambda_c \bmod \Lambda_s\}$, or equivalently $\mathbf{v} \in \{\Lambda_c \cap \mathcal{V}_s\}$, the rate is

$$R = \frac{1}{N}\log(\frac{V_s}{V_c}) = \frac{1}{2}\log(\frac{\sigma_s^2}{\sigma_c^2}) = \frac{1}{2}\log(1 + \frac{P_X}{\sigma_z^2}). \tag{2.27}$$

Encoding: transmits the error vector between $\alpha\mathbf{s} + \mathbf{d}$ and the selected coset leader $\mathbf{v}$, i.e,

$$\mathbf{x} = \Big[\mathbf{v} - \alpha\mathbf{s} - \mathbf{d}\Big] \bmod \Lambda_s, \tag{2.28}$$

where $\mathbf{s}$ is the interference vector and $\mathbf{d}$ is the dither. Since $\mathbf{x}$ is distributed uniformly over $\mathcal{V}_s$ independent of $\mathbf{s}$ and $\mathbf{v}$ by the property of dithered quanti-

zation. The power constraint is $\frac{1}{N}E\{\| \mathbf{x} \|^2\} = \sigma_s^2 = P_X$.

Decoding: to reconstruction the transmitted massage the receiver has to estimate the transmitted coset leader, it is done by

$$\hat{\mathbf{V}} = Q_{\Lambda_c}(\alpha\mathbf{Y} + \mathbf{d}) \bmod \Lambda_s \tag{2.29}$$

$$= Q_{\Lambda_c}\left(\left[\alpha\mathbf{Y} + \mathbf{d}\right] \bmod \Lambda_s\right). \tag{2.30}$$

Figure 2.1 shows schematically the encoder and decoder.

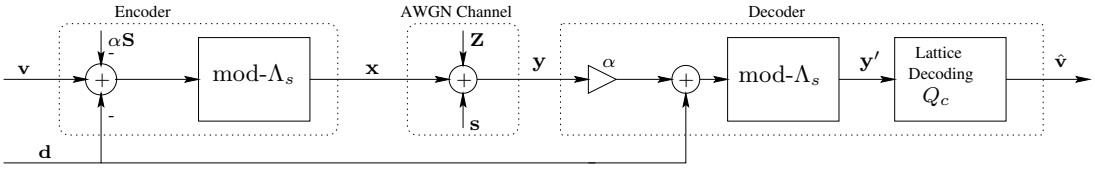

Figure 2.1: Nested lattices scheme

2.2 Precoding "Shaping Gap"

The ultimate precoding "shaping gap" is the power (or SNR) gap between precoding using one dimensional coarse lattice and precoding using high dimensional coarse lattice with $G(\Lambda_s) \approx 1/2\pi e$ (which achieves the capacity), for a fixed mutual information. For shaping gap analysis we use Lemma 2.1 where the modulo additive channel is given by

$$\mathbf{Y}' = \left[\mathbf{V} + \mathbf{Z}'\right] \bmod \Lambda_s, \tag{2.31}$$

the equivalent noise is

$$\mathbf{Z}' = \left[\alpha\mathbf{Z} + (1-\alpha)\mathbf{U}\right] \bmod \Lambda_s, \tag{2.32}$$

where $\mathbf{U} \sim U(\mathcal{V}_s)$. Multiplying $\mathbf{Y}'$ by $1/\sqrt{\alpha}$ does not change the mutual information therefore we can write

$$\mathbf{Y}'/\sqrt{\alpha} = \Big[\mathbf{V}/\sqrt{\alpha} + \mathbf{Z}'/\sqrt{\alpha}\Big] \bmod (\Lambda_s/\sqrt{\alpha}). \tag{2.33}$$

Let

$$\mathbf{Y}'' \triangleq \mathbf{Y}'/\sqrt{\alpha}, \;\; \mathbf{V}' \triangleq \mathbf{V}/\sqrt{\alpha}, \;\; \mathbf{Z}'' \triangleq \mathbf{Z}'\sqrt{\alpha}, \;\; \Lambda_s' \triangleq \Lambda_s/\sqrt{\alpha}, \tag{2.34}$$

thus,

$$\mathbf{Y}'' = \Big[\mathbf{V}' + \mathbf{Z}''\Big] \bmod \Lambda_s', \tag{2.35}$$

where

$$\mathbf{Z}'' = \Big[\frac{\alpha}{\sqrt{\alpha}}\mathbf{Z} + \frac{1-\alpha}{\sqrt{\alpha}}\mathbf{U}\Big] \bmod (\mathbf{\Lambda_s}/\sqrt{\alpha}). \tag{2.36}$$

Let $\tilde{\mathbf{Z}}'' \triangleq \frac{\alpha}{\sqrt{\alpha}}\mathbf{Z} + \frac{1-\alpha}{\sqrt{\alpha}}\mathbf{U}$, for $\alpha = \frac{P_X}{P_X+\sigma_z^2}$,

$$\begin{aligned} E\tilde{\mathbf{Z}}''^2 &= \frac{P_X\sigma_z^2}{P_X+\sigma_z^2}\frac{P_X+\sigma_z^2}{P_X} = \sigma_z^2 \\ E\mathbf{V}'^2 &= \frac{P_X}{\frac{P_X}{P_X+\sigma_z^2}} = P_X + \sigma_z^2 \end{aligned}$$

for $N_s = 1$ we have

$$h(Y''^{(N_s=1)}) = \frac{1}{2}\log\Big(12(P_X^{(N_s=1)} + \sigma_z^2)\Big) \tag{2.37}$$

for $N_s \to \infty$

$$h(Y''^{(N_s\to\infty)}) = \frac{1}{2}\log\Big(2\pi e(P_X^{(N_s\to\infty)} + \sigma_z^2)\Big). \tag{2.38}$$

Since the shaping gap is the transmitted power reduction from $N_s = 1$ to $N_s \to \infty$, for constant mutual information, thus

$$I(Y''^{(N_s=1)}; V'^{(N_s=1)}) = I(Y''^{(N_s\to\infty)}; V'^{(N_s\to\infty)}) \tag{2.39}$$

$$h(Y''^{(N_s=1)}) - h(Z''^{(N_s=1)}) = h(Y''^{(N_s\to\infty)}) - h(Z''^{(N_s\to\infty)}) \tag{2.40}$$

$$\begin{aligned} \frac{1}{2}\log_2\Big(12(P_X^{(N_s=1)} + \sigma_z^2)\Big) &= \frac{1}{2}\log_2\Big(2\pi e(P_X^{(N_s\to\infty)} + \sigma_z^2)\Big) + \\ + \Big[h(Z''^{(N_s=1)}) - h(Z''^{(N_s\to\infty)})\Big]. \end{aligned} \tag{2.41}$$

Let define $\left[h(Z''^{(N_s=1)}) - h(Z''^{(N_s\to\infty)})\right] \stackrel{\Delta}{=} \frac{1}{2}\log_2 \Delta_z$, therefore

$$P_X^{(N_s=1)} + \sigma_z^2 = \frac{\pi e}{6}(P_X^{(N_s\to\infty)} + \sigma_z^2)\Delta_z \tag{2.42}$$

$$\frac{P_X^{(N_s=1)}}{P_X^{(N_s\to\infty)}} = \Delta_z\frac{\pi e}{6} + \frac{\sigma_z^2}{P_X^{(N_s=1)}}\Big(\Delta_z\frac{\pi e}{6} - 1\Big). \tag{2.43}$$

1) For high SNR - $\alpha \to 1$, from Figure 2.2 the entropy difference is

$$\lim_{\alpha\to 1}\left[h(Z''^{(N_s=1)}) - h(Z''^{(N_s\to\infty)})\right] = 0,$$

thus $\Delta_z = 1$ and $\gamma_{ps} = \frac{P_X^{(N_s=1)}}{P_X^{(N_s\to\infty)}} = \frac{\pi e}{6} = 1.53\ dB$

2) For $SNR = 1 = 0\ dB$ - we have $\sigma_z^2 = P_X^{(N_s=1)}$ and

$$\gamma_{ps} = \frac{P_X^{(N_s=1)}}{P_X^{(N_s\to\infty)}} = \frac{\pi e}{6}\cdot\frac{1}{2 - \frac{\pi e}{6}\Delta_z}\cdot\Delta_z$$

from Figure 2.2 we can deduce that

$$\Delta_z = 2^{2\left[h(Z''^{(N_s=1)}) - h(Z''^{(N_s\to\infty)})\right]} = 2^{-2\Delta h|_{SNR=0\ dB}} = 2^{-0.084} = 0.943$$

thus

$$\gamma_{ps} = \frac{P_X^{(N_s=1)}}{P_X^{(N_s\to\infty)}} = \frac{\pi e}{6}\cdot 1.521\cdot 0.943 = 2.04 = 3.1\ dB.$$

3) For $SNR \to 0$ - the shaping gap continue to increase.

Unlike the interference-free AWGN channel, this gap is particularly significant at the low SNR regime for the interference channel. Figure 2.3 shows the mutual information of a one dimensional lattice strategies achieved by (2.3), the mutual information achieved by a uniform input over an interference-free AWGN channel, and the capacity $C = \frac{1}{2}\log(1 + SNR)$. We refer to the SNR gap between the first two and the latter as "shaping gap", at high SNR the

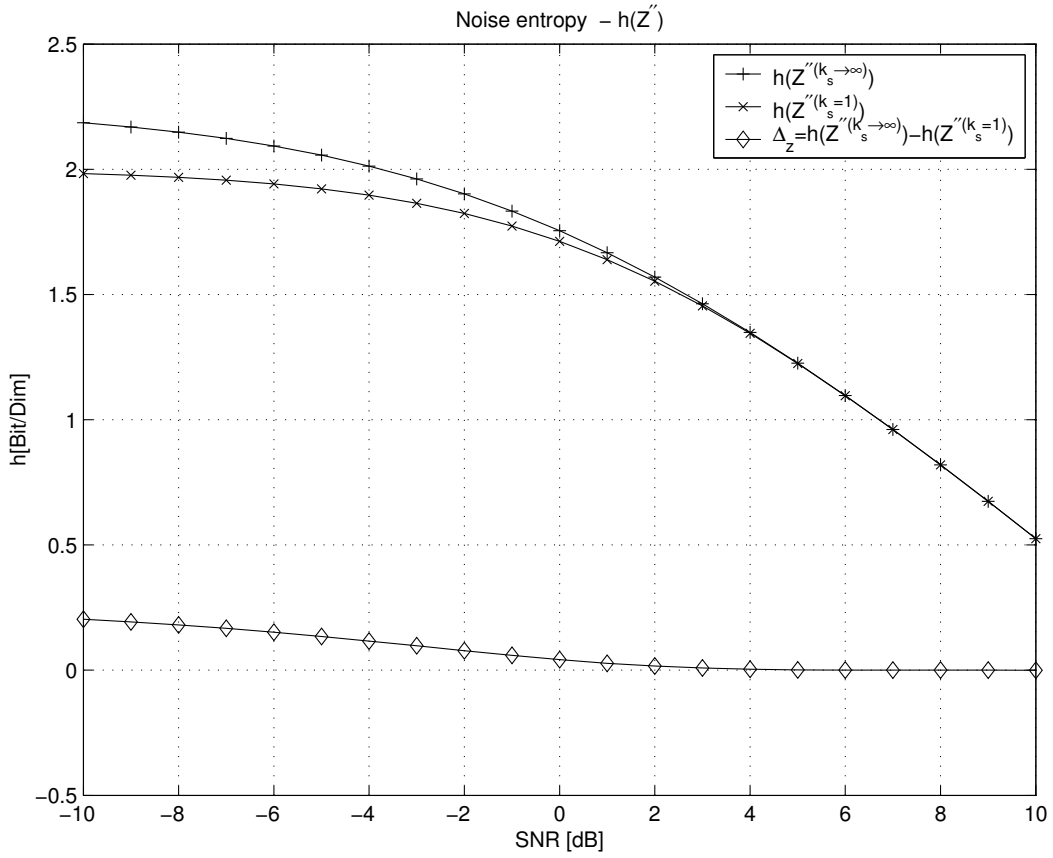

Figure 2.2: Noise entropy

interference-free AWGN and the interference channels have an identical shaping gap of 1.53 dB. At low SNR, the interference-free AWGN channel has no shaping gap, on the other hand the interference channel has over 1.53 dB precoding shaping gap. From Figure 2.3, at $SNR = 0\ dB$ the precoding shaping gap is 3.1 dB, while the shaping gap of interference-free AWGN channel is close to zero.

The rate loss between one dimensional precoding scheme and N_s dimensional precoding scheme is bounded by $\frac{1}{2}\log(2\pi e G_{N_s}(\Lambda_s)) \leq 0.254$ bit/dim, [8], for *any* SNR. Nevertheless, the capacity curve with respect to the SNR

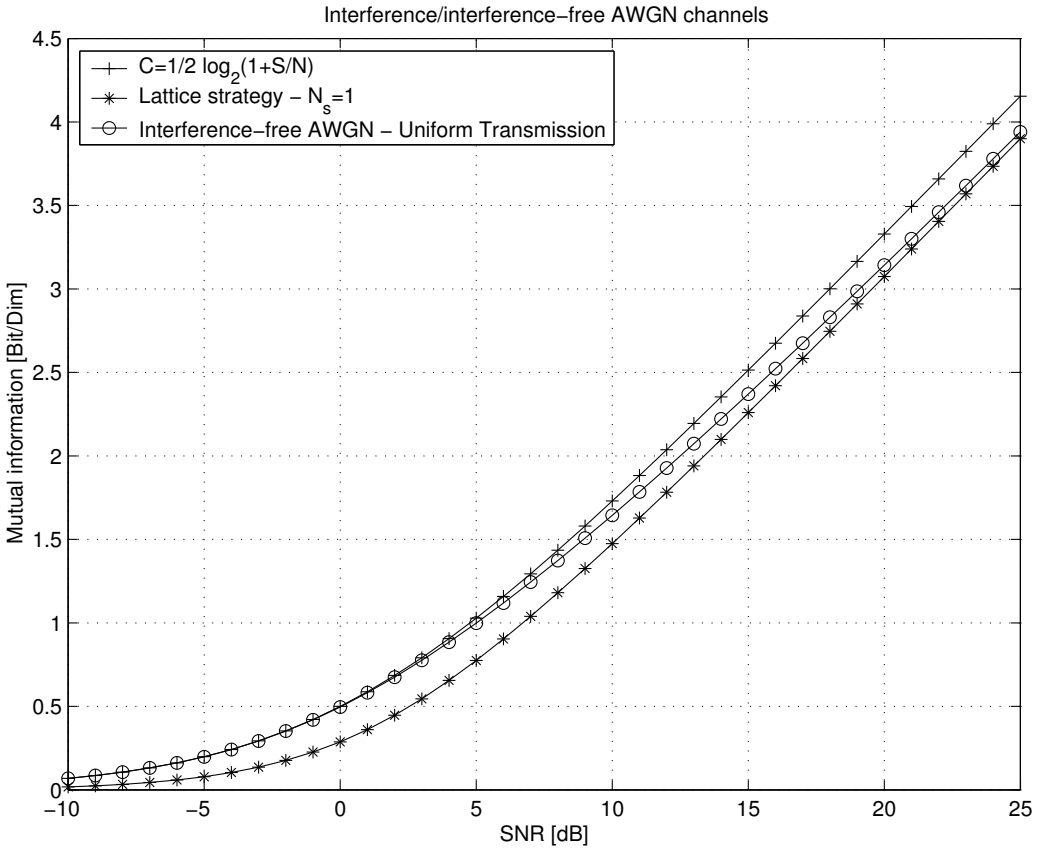

Figure 2.3: Precoding shaping gap versus interference-free AWGN shaping gain

(in logarithmic scale) is more sensitive at low SNR (the "6 dB per bit" is not valid at low SNR), i.e, the loss in dB increases. Figure 2.8 also illustrates this gap clearly, where the rate is drawn with respect to $\frac{SNR}{2R}$ (which is the E_b/N_0 the SNR per bit) as will explain later.

2.3 Optimal MMSE Factor

For high dimensional lattice with $G(\Lambda_s) \approx 1/2\pi e$, ESZ showed that the optimal MMSE factor $\alpha^* \triangleq \frac{P_X}{P_X+\sigma_z^2}$ achieves the capacity. We consider the optimal α that maximizes the mutual information for low dimensional transmission.

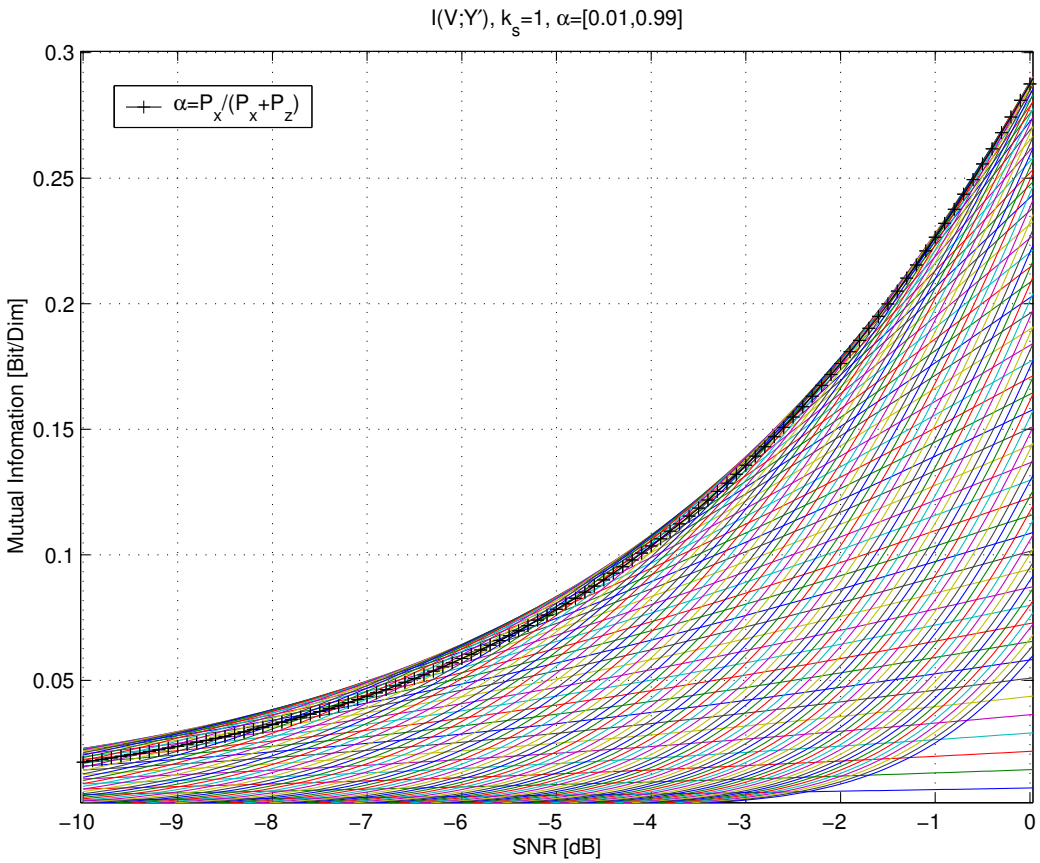

Figure 2.4: Mutual Information envelope

The value of α^* is derived to bring the $E\mathbf{Z}'^2$ to minimum [8], i.e, optimum MMSE estimation. On the other hand for low dimensional transmission we are intersting to bring $I(\mathbf{V};\mathbf{Y}')$ to maximum. For high SNR or high dimensional transmission scheme $\mathbf{Z}'$ is close to Gaussian. In these cases the maximization of the mutual information $I(\mathbf{V};\mathbf{Y}')$ is equivalent to minimize $h(\mathbf{Z}')$, and in Gaussian case it is equivalent to minimize $E\mathbf{Z}'^2$. Therefore, at high SNR or high dimensional precoding scheme α^* is the optimum α to maximize $I(\mathbf{V};\mathbf{Y}')$, while at low SNR and low dimensional precoding scheme the optimum α will not match α^*, as shown in Figure 2.4 and Figure 2.5 which are drawn for one

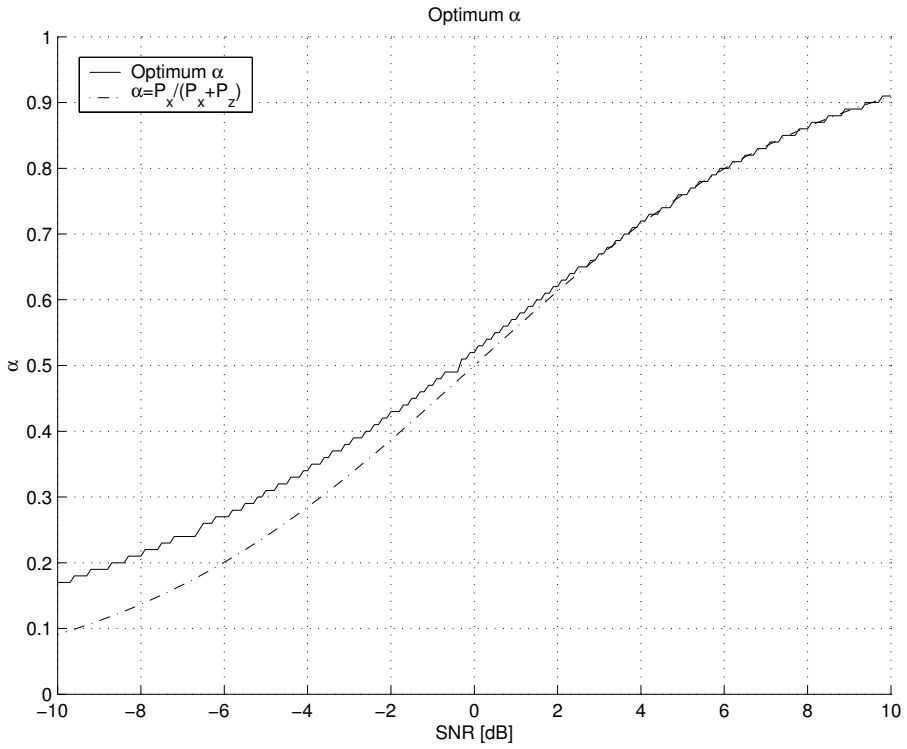

Figure 2.5: Optimum α

dimensional coarse lattice, $N_s = 1$. The gain of using the optimum α with respect to α^* is about 1 dB at $SNR = -10$ dB. Generally, as long as the dimension increases this gain is decrease.

2.4 Time-Sharing Transmission at Low SNR

The rates achieve in one dimensional lattice strategies with optimum MMSE factor as shown in Figure 2.4, can be improved by time sharing. The mutual information of one dimensional transmission at low SNR (for $SNR < 1$), using lattice strategies with optimum MMSE factor, is concave function with respect to SNR (in linear scale), as shown in Figure 2.6.

Therefore, using time sharing between $SNR = 0$ and any point in the interval

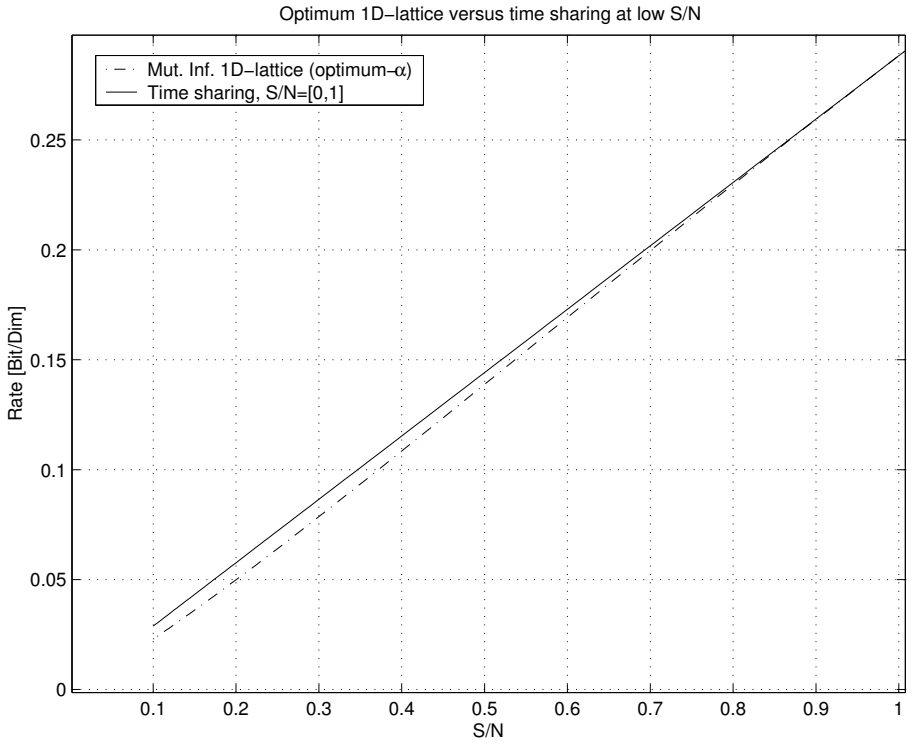

Figure 2.6: Time-sharing linear scale

$SNR \in [0, 1]$ higher rates can be achieved. Numerical result shows that the critical SNR for time sharing in one dimensional transmission is $SNR_c = 1$ $(0\ dB)$ as shown in Figure 2.6. The gain of the time sharing with respect to one dimensional lattice strategies with optimum MMSE factor is about 1 dB at $SNR = -10\ dB$ as shown in Figure 2.7.

Above $SNR = 1$ the mutual information for one dimensional transmission is convex function with respect to SNR, therefore no time sharing is needed. Generally, for high dimensional coarse lattice the critical SNR, SNR_c, will be lower then $SNR = 1$, in the limit of infinite dimensional lattice with $G(\Lambda_s) \approx 1/2\pi e$ where $SNR_c = 0$ is the optimal point for time sharing, meaning no time sharing is needed.

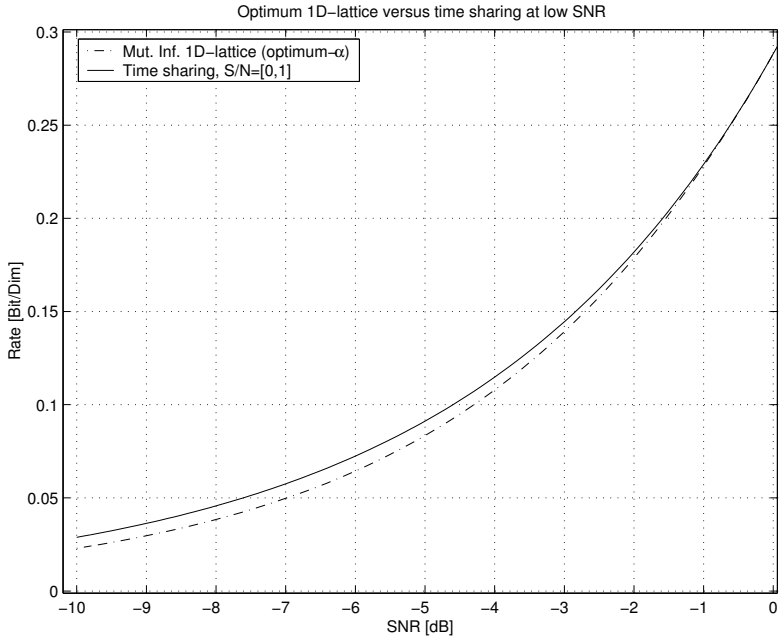

Figure 2.7: Time-sharing gain

Asymptotic precoding shaping gap with respect to time sharing - The ultimate precoding shaping gap with respect to one dimension time sharing (it is not the same as for one dimensional lattice strategies), is obtained for $SNR \to 0$. From Figure 2.6 the rate at $SNR = 1$ is 0.2883 bit/dim, therefore the achievable rates by time sharing for the interval $[0, 1]$ are

$$R^{1D} = 0.2883 \cdot SNR. \tag{2.44}$$

Using the well known approximation, that

$$\log(1 + x) = x \log(e),\ x \ll 1, \tag{2.45}$$

for $C = \frac{1}{2}\log(1 + SNR)$, we have

$$C = \frac{1}{2} SNR \log(e). \tag{2.46}$$

The shaping gap is obtained for $C = R^{1D}$

$$\gamma_{ps} = \frac{1}{0.5766} \log(e) = 3.98\ dB. \tag{2.47}$$

Therefore, the ultimate precoding shaping gap with respect to time sharing is 3.98 dB.

The shaping gap at the low SNR regime is well understood by drawing the mutual information with respect to $\frac{SNR}{2I(\mathbf{V};\mathbf{Y}')}$ (or $\frac{SNR}{2R}$), which is equivalent to E_b/N_0 (SNR per bit) and E_b is the energy per information bit. Since the discrete time channel model (contain the optimum match filter detection as front end) has noise variance $N_0/2$, in order to have the same SNR as the continues waveform channel with bandwidth $W = \frac{1}{2T}$ where $1/T$ is the symbols rate. And $E_b = E_s/R$ where E_s is the symbol energy, therefore

$$\frac{E_b}{N_0} = \frac{E_s/R}{2\sigma_Z^2} = \frac{E_s/\sigma_Z^2}{2R} = \frac{SNR}{2R}. \tag{2.48}$$

Figure 2.8 shows the capacity, the mutual information of one dimensional lattice strategies and one dimensional lattice strategies with time sharing at the interval $SNR \in [0, 1]$ (where $SNR_c = 1$ is critical SNR for optimal transmission with time sharing at the low SNR regime). For $R \to 0$, the Shannon limit is at $E_b/N_0 = -1.59\ dB$, the gap between the lattice strategy with time sharing and the capacity is bounded by 4 dB as shown above an in Figure 2.8.

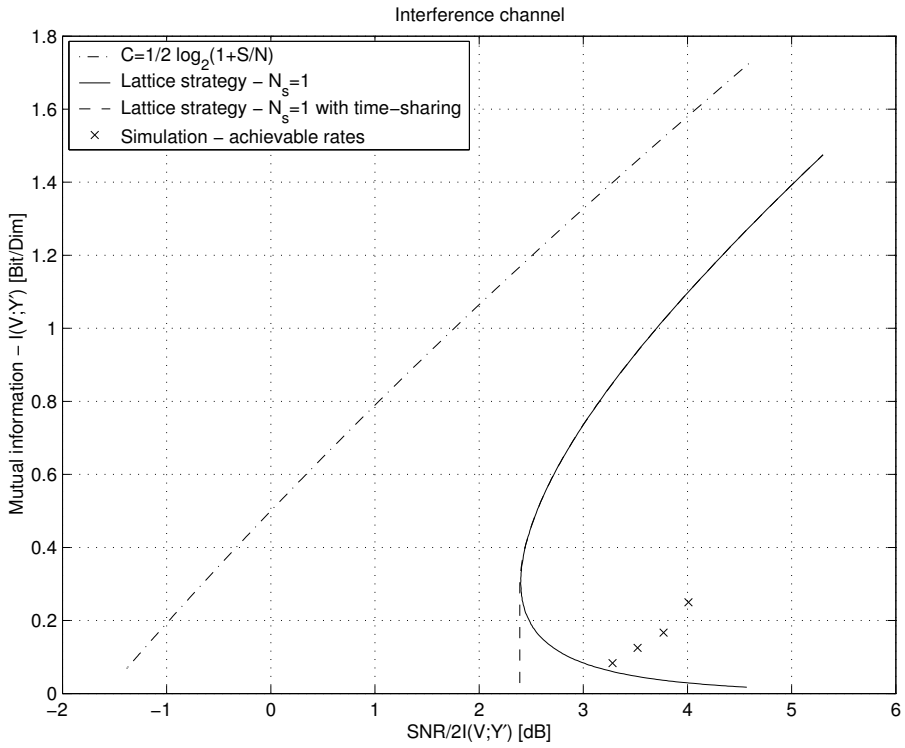

Figure 2.8: Precoding shaping gap vs $\frac{SNR}{2I(\mathbf{V};\mathbf{Y}')}$.

Chapter 3

One Dimensional (Causal) Precoding

In this chapter we introduce the one dimensional precoding scheme, that is one dimensional coarse lattice Λ_s, based on lattice precoding scheme. However, the fine lattice Λ_c is high dimensional lattice. The traditional TH precoding is a degenerate case of lattice precoding scheme where the coarse lattice is one dimensional. In this chapter we focus on ISI channel, therefore we begin with the MMSE decision feedback equalizer (DFE) background as shown in [3].

3.1 MMSE Decision Feedback Equalizer

The discrete model of the ISI channel is given by

$$\tilde{y}_n = x_n \star c_n + \tilde{n}_n, \tag{3.1}$$

where $x_n, \tilde{y}_n$ are the channel input and output, respectively, $\tilde{n}_n$ is Gaussian noise, and c_n is channel response. Using the whitened matched filter (WMF) model [3], which includes the matched filter (MF) and the whitened filter as front end [22], we have

$$y_n = x_n \star h_n + n_n, \tag{3.2}$$

where h_n has causal response and $n_n \sim (0, \sigma_n^2)$ is i.i.d Gaussian noise. The MMSE Decision feedback Equalizer (MMSE-DFE) is a well known technique to minimize the ISI by designing of two filters, feedforward and feedback filters, in order to have minimum mean square error (MMSE) before the decision, as shown in Figure 3.1. The feedforward filter $A(Z)$ is non-casual filter and its

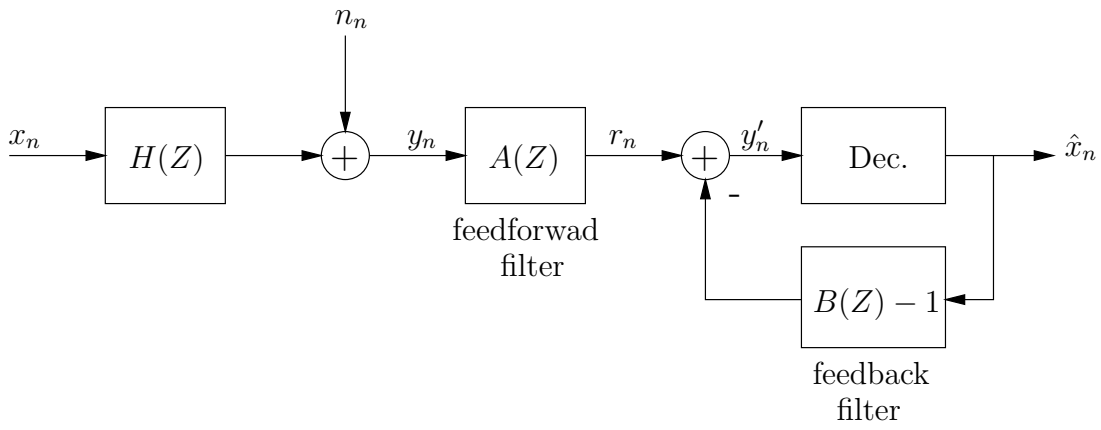

Figure 3.1: DFE structure

output is given by

$$r_n = x_n \star h_n \star a_n + n_n \star a_n. \tag{3.3}$$

Let $z_n \triangleq n_n \star a_n$ and $t_n \triangleq h_n \star a_n$, we have that

$$\begin{aligned} r_n &= x_n \star t_n + z_n & (3.4)\\ &= \sum_{k=-\infty}^{\infty} t_k x_{n-k} + z_n & (3.5)\\ &= t_0 x_n + \sum_{k=-\infty}^{-1} t_k x_{n-k} + \sum_{k=1}^{\infty} t_k x_{n-k} + z_n & (3.6) \end{aligned}$$

where $\sum_{k=-\infty}^{-1} t_k x_{n-k}$ is the pre-cursor ISI which depends on future symbols, and $\sum_{k=1}^{\infty} t_k x_{n-k}$ is the post-cursor ISI which depends on past symbols. Assuming that the decision is correct the feedback filter uses the decided symbols

to eliminate the post cursor ISI. The feedforward and the feedback filters are designed so that the residual error, the noise and the pre-cursor, will be minimized in the sense of mean square error, The error sequence is given by

$$\begin{aligned} E(Z) &= A(Z)Y(Z) - [B(Z) - 1]X(Z) - X(Z) && (3.7) \\ &= A(Z)Y(Z) - B(Z)X(Z). && (3.8) \end{aligned}$$

For optimal estimation the orthogonal condition is $R_{ey}[l] = 0$ or $S_{ey}(Z) = 0$, thus we have that

$$S_{ey}(Z) = A(Z)S_{yy}(Z) - B(Z)S_{xy}(Z) = 0. \tag{3.9}$$

The cross spectrum S_{xy} is given by

$$S_{xy}(Z) = P_X^2 H^*(1/Z^*), \tag{3.10}$$

and S_{yy} is

$$S_{yy}(Z) = P_X^2 H(Z)H^*(1/Z^*) + \sigma_n^2. \tag{3.11}$$

Assuming that $S_{yy}(Z)$ is agree with Paley Wiener condition $\int_{-\pi}^{\pi} |\log(S_{yy}(e^{jw}))| dw < \infty$ and $0 < S_{yy} < \infty$, the spectral factorization of $S_{yy}(Z)$ is

$$S_{yy}(Z) = S_0 G(Z)G^*(1/Z^*), \tag{3.12}$$

where $G(Z)$ is monic, causal, stable, minimum phase and

$$S_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} log|S_{yy}(e^{jw})|^2 dw. \tag{3.13}$$

Therefore we have,

$$S_{yy}(Z) = P_X^2 H(Z)H^*(1/Z^*) + \sigma_n^2 = S_0 G(Z)G^*(1/Z^*). \tag{3.14}$$

Substitute $S_{yy}(Z)$ and $S_{xy}(Z)$ into (3.9), we have that

$$A(Z) = B(Z)\frac{S_{xy}(Z)}{S_{yy}(Z)} = B(Z)\frac{P_X^2 H^*(1/Z^*)}{S_0 G(Z)G^*(1/Z^*)} \tag{3.15}$$

Now, the error sequence is

$$E(Z) = B(Z)\left[\Big(\frac{P_X^2}{S_0}\Big)\frac{Y(Z)H^*(1/Z^*)}{G(Z)G^*(1/Z^*)} - X(Z)\right] = B(Z)E'(Z). \tag{3.16}$$

Therefore, we have

$$\begin{aligned}
S_{e'e'}(Z) &= \Big(\frac{P_X^2}{S_0}\Big)^2 \frac{S_{yy}(Z)H(Z)H^*(1/Z^*)}{[G(Z)G^*(1/Z^*)]^2} - \\
&\quad -2\Big(\frac{P_X^2}{S_0}\Big)\frac{S_{xy}H^*(1/Z^*)}{G(Z)G^*(1/Z^*)} + P_X^2 && (3.17)\\
&= \Big(\frac{P_X^2}{S_0}\Big)\frac{S_0 H(Z)H^*(1/Z^*)}{G(Z)G^*(1/Z^*)} - \\
&\quad -2\Big(\frac{P_X^2}{S_0}\Big)\frac{P_X^2 H(Z)H^*(1/Z^*)}{G(Z)G^*(1/Z^*)} + P_X^2 && (3.18)\\
&= \Big(\frac{P_X^2}{S_0}\Big)\frac{S_0 G(Z)G^*(1/Z^*) - P_X^2 H(Z)H^*(1/Z^*)}{G(Z)G^*(1/Z^*)} && (3.19)\\
&= \Big(\frac{P_X^2}{S_0}\Big)\frac{\sigma_n^2}{G(Z)G^*(1/Z^*)} && (3.20)
\end{aligned}$$

To have white error spectrum we have to choose

$$B(Z) = G(Z), \tag{3.21}$$

thus $B(Z)$ is causal filter. The feedback and the feedforward filters are

$$\begin{aligned}
\tilde{B}(Z) &\triangleq B(Z) - 1 = G(Z) - 1 && (3.22)\\
A(Z) &= \Big(\frac{P_X^2}{S_0}\Big)\frac{H^*(1/Z^*)G(Z)}{G(Z)G^*(1/Z^*)} = \Big(\frac{P_X^2}{S_0}\Big)\frac{H^*(1/Z^*)}{G^*(1/Z^*)}, && (3.23)
\end{aligned}$$

where $A(Z)$ is non-causal filter. The MSE and the SNR are given by

$$\begin{aligned}
MSE_{MMSE-DFE} &= \frac{1}{2\pi}\int_{-\pi}^{\pi} S_{ee}(e^{jw})dw = \Big(\frac{P_X^2}{S_0}\Big)\sigma_n^2 && (3.24)\\
SNR_{MMSE-DFE} &= \frac{S_0}{\sigma_n^2}\,. && (3.25)
\end{aligned}$$

$Y'(Z)$ can be written as

$$\begin{aligned} Y'(Z) &= X(Z)H(Z)A(Z) - X(Z)[B(Z)-1] + N(Z)A(Z) && (3.26)\\ &= \Big[X(Z)H(Z)+N(Z)\Big]\Big(\frac{P_X^2}{S_0}\Big)\frac{H^*(1/Z^*)}{G^*(1/Z^*)} - \\ &\quad -X(Z)G(Z)+X(Z) && (3.27)\\ &= X(Z)+N(Z)\Big(\frac{P_X^2}{S_0}\Big)\frac{H^*(1/Z^*)}{G^*(1/Z^*)} - \Big(\frac{\sigma_n^2}{S_0}\Big)\frac{X(Z)}{G^*(1/Z^*)} && (3.28)\\ &= X(Z)+E(Z) && (3.29) \end{aligned}$$

where (3.28) is from (3.16). The error sequence e_n is white process with variance $Ee_n^2 = P_X^2\frac{\sigma_n^2}{S_0}$. It is a mixture of colored Gaussian noise and uniform precursor interference, which is equivalently a white process. This result gives a biased MMSE receiver, meaning $E\{y'_n - x_n|x_n\} \neq 0$, as can be shown in (3.28). The unbiased MMSE-DFE receiver is accepted by multiplication of y'_n with $\frac{S_0}{S_0-\sigma_n^2}$. The biased MMSE-DFE has minimum MSE, while the unbiased receiver has minimum probability of error. The unbiased filters are given by

$$\tilde{B}_U(Z) = \frac{S_0}{S_0-\sigma_n^2}[G(Z)-1] \quad (3.30)$$

$$A_U(Z) = \Big(\frac{P_X^2}{S_0-\sigma_n^2}\Big)\frac{H^*(1/Z^*)}{G^*(1/Z^*)}\,. \quad (3.31)$$

The unbiased MSE and SNR are given by

$$\begin{aligned} MSE_{MMSE-DFE_u} &= P_X\sigma_n^2/(S_0-\sigma_n^2) && (3.32)\\ SNR_{MMSE-DFE_u} &= SNR_{MMSE-DFE} - 1 = (S_0-\sigma_n^2)/\sigma_n^2\,. && (3.33) \end{aligned}$$

3.1.1 TH Precoding

For TH precoding the feedback filter is moved to the transmitter as shown in Figure 3.2, TH-FFE can use the transmitted past symbols instead of the

decided symbols. TH uses the modulo operation to the interval $[-L/2, L/2)$, i.e, one dimensional modulo operation.

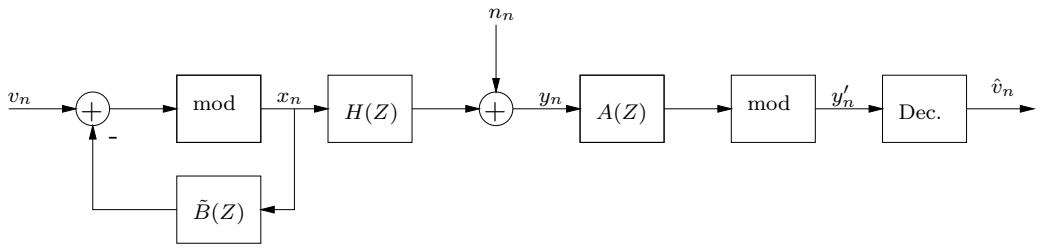

Figure 3.2: TH-FFE structure

3.2 TH-FFE and Lattice Precoding Schemes Equivalent

We apply the knowledge from lattice precoding to improve the tractional TH-precoding. Therefore, we use the unbiased MMSE-DFE feedforward and feedback filters $A_U(Z), \tilde{B}_U(Z)$, respectively, combined with the lattice precoding scheme, as shown in Figure 3.3 . Now, we identify the components in term of

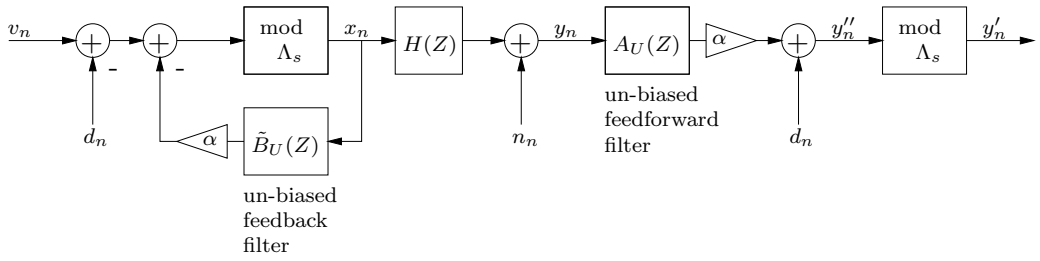

Figure 3.3: Lattice precoding for ISI channel

lattice precoding. The receiver input is given by

$$y_n = x_n \star h_n + n_n \tag{3.34}$$

and we have that

$$y''_n = \Big[\alpha\{x_n \star t_n + z_n + \frac{d_n}{\alpha}\}\Big], \tag{3.35}$$

where $z_n \triangleq n_n \star a_n$, d_n is the dither signal and $t_n \triangleq h_n \star a_n$. Since we use the unbiased MMSE-DFE filters we have that

$$t_0 = (h_n \star a_n)|_{\,n=0} = \sum_{k=0}^{\infty} h_k a_{-k} = 1\ . \tag{3.36}$$

y''_n can be written as

$$\begin{aligned} y''_n &= \alpha\Big\{\sum_{k=-\infty}^{\infty} t_k x_{n-k} + z_n\Big\} + d_n && (3.37)\\ &= \alpha\Big\{x_n + \sum_{k=-\infty}^{-1} t_k x_{n-k} + \sum_{k=1}^{\infty} t_k x_{n-k} + z_n\Big\} + d_n && (3.38)\\ &= \alpha\Big\{x_n + S^U_{pre-cursor} + S^U_{post-cursor} + z_n\Big\} + d_n && (3.39)\\ &= x_n + \alpha S^U_{pre-cursor} - (1-\alpha)x_n + \alpha S^U_{post-cursor} + \alpha z_n + d_n && (3.40) \end{aligned}$$

where

$$S^U_{pre-cursor} \triangleq \sum_{k=-\infty}^{-1} t_k x_{n-k} \tag{3.41}$$

$$S^U_{post-cursor} \triangleq \sum_{k=1}^{\infty} t_k x_{n-k} \tag{3.42}$$

are the pre-cursor and post cursor relative to unbiased filters. The feedback filter is

$$b_n = \begin{cases} t_n\ n \geq 1 \\ 0\ else \end{cases}\ . \tag{3.43}$$

The transmit symbol is given by

$$x_n = \left[v_n - \alpha(x_n \star b_n) - d_n\right] \bmod \Lambda_s \tag{3.44}$$

$$x_n = \left[v_n - \alpha \sum_{k=-\infty}^{k=\infty} b_k x_{n-k} - d_n\right] \bmod \Lambda_s \tag{3.45}$$

$$x_n = \left[v_n - \alpha \sum_{k=1}^{k=\infty} t_k x_{n-k} - d_n\right] \bmod \Lambda_s \tag{3.46}$$

$$x_n = \left[v_n - \alpha S^U_{post-cursor} - d_n\right] \bmod \Lambda_s. \tag{3.47}$$

y'_n can be written as

$$y'_n = y''_n \bmod \Lambda_s \tag{3.48}$$

$$= \Big\{\left[v_n - \alpha S^U_{post-cursor} - d_n\right] \bmod \Lambda_s + \alpha S^U_{pre-cursor} - \\ -(1-\alpha)x_n + \alpha z_n + \alpha S^U_{post-cursor} + d_n\Big\} \bmod \Lambda_s \tag{3.49}$$

$$= \left[v_n + \alpha S^U_{pre-cursor} - (1-\alpha)x_n + \alpha z_n\right] \bmod \Lambda_s \tag{3.50}$$

Therefore, we have the modulo additive channel

$$y'_n = \left[v_n + \tilde{N}\right] \bmod \Lambda_s, \tag{3.51}$$

where the effective noise is

$$\tilde{N} \triangleq \left[\alpha S^U_{pre-cursor} - (1-\alpha)x_n + \alpha z_n\right] \bmod \Lambda_s\ . \tag{3.52}$$

According to (3.28) we can identify the effective noise $\alpha S^U_{pre-cursor} - (1-\alpha)x_n + \alpha z_n$ as the error sequence e_n, which is white process, and it also contain a self noise expression $(1-\alpha)x_n$.

The optimum α can be found by the minimization of $E\tilde{N}^2$, therefore we have

$$\alpha = \frac{P_X}{P_X + E[S^U_{pre-cursor}]^2 + Ez_n^2}, \tag{3.53}$$

where $E[S^U_{pre-cursor}]^2 = P_X \sum_{k=-\infty}^{-1} t_k^2$, and $Ez_n^2 = \sigma_n^2 \sum_{k=-\infty}^{\infty} a_k^2$. The value of α is equal to the biased factor $\frac{S_0-\sigma_n^2}{S_0}$ which used in [3]. Assuming that the equivalent channel (3.51) represents a vector transmission, i.e, v_n is distributed uniformly over a basic cell of high dimensional lattice Λ_s (ignoring that the vector of the post cursor ISI is not fully known to the transmitter), therefore v_n and x_n have nearly one dimensional Gaussian distributions, thus $\tilde{N}$ is also Gaussian with $\tilde{N} \sim (0, P_X \frac{\sigma_n^2}{S_0})$, the capacity is

$$C = h(v_n) - h(\tilde{N}) = \frac{1}{2} \log(\frac{S_0}{\sigma_n^2}), \tag{3.54}$$

which is equal to $\frac{1}{2}\log(1 + SNR_{MMSE-DFE_U})$, that is the channel capacity [3]. Using post cursor ISI as known side information for high dimension lattice transmission is handled using interleaver, see [31].

3.3 A Practical Precoding Scheme for the ISI Channel

We use the above scheme for ISI channel combined with convolutional code. We use the binary lattice chain $\mathbb{Z}^N/\Lambda_c/\Lambda_s = 2^m\mathbb{Z}^N$, effectively Λ_s is one dimensional lattice. The fine lattice, Λ_c is high dimensional lattice spanned by binary convolutional code $\mathcal{C}_c$.

3.3.1 Transmitter

Since the coarse lattice is $2^m\mathbb{Z}^N$ the dither signal is distributed uniformly over $[-L/2, L/2)^N$, where $L = 2^m$. The modulo operation is the one dimensional modulo operation reduced to the interval $[-L/2, L/2)$, as shown in Figure 3.4.

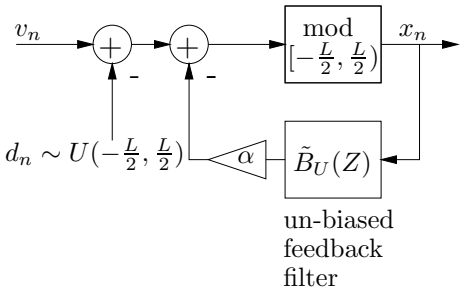

Figure 3.4: Transmitter

3.3.2 Receiver

The receiver is shown in Figure 3.5 uses the un-biased feedforword filter, and α is computed by (3.53). For the equivalent channel

$$y'_n = \left[v_n + \tilde{N}\right] \bmod \Lambda_s, \tag{3.55}$$

where

$$\tilde{N} \triangleq \left[\alpha S^U_{pre-cursor} - (1-\alpha)x_n + \alpha z_n\right] \bmod \Lambda_s \tag{3.56}$$

the optimal receiver has to perform ML decoding. Since we are using convolutional code VA may be used as ML decoder. However, $\tilde{N}$ is not Gaussian, therefore the optimum receiver has to use matched metric to $\tilde{N}$.

The $\tilde{N}$ distribution is a mixture of: αz_n - color Gaussian noise, $\alpha S^U_{pre-cursor} =$

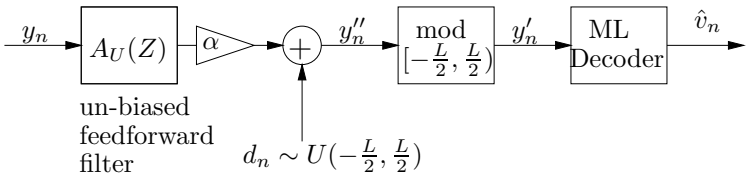

Figure 3.5: Receiver

$\alpha \sum_{k=-\infty}^{-1} t_k x_{n-k}$ - is sum of uniform random variables $u_i \sim [-\alpha t_i \cdot L/2, \alpha t_i \cdot L/2]$ and $(1-\alpha)x_n$ - is uniform random variable. However, $\tilde{N}$ is white process, where

the distribution of $\tilde{N}$ is $f_{\tilde{N}}(x)$. The matched metric to $\tilde{N}$ is given by

$$M(x) = -\log_{10}(f_{\tilde{N}}(x)) \tag{3.57}$$

3.4 Performance Results

I. We compare the traditional TH-FFE to our scheme, while for comparison, we use the metric in (3.57) for the optimal ML decoder, and the Euclidian metric as sub-optimal decoder. The ISI channel is

$$H(Z) = 1 + 0.75Z^{-1} + 0.25Z^{-2} \tag{3.58}$$

where the ISI-channel capacity which is calculated by "water pouring" optimization is shown in Figure 3.6 with the AWGN channel capacity.

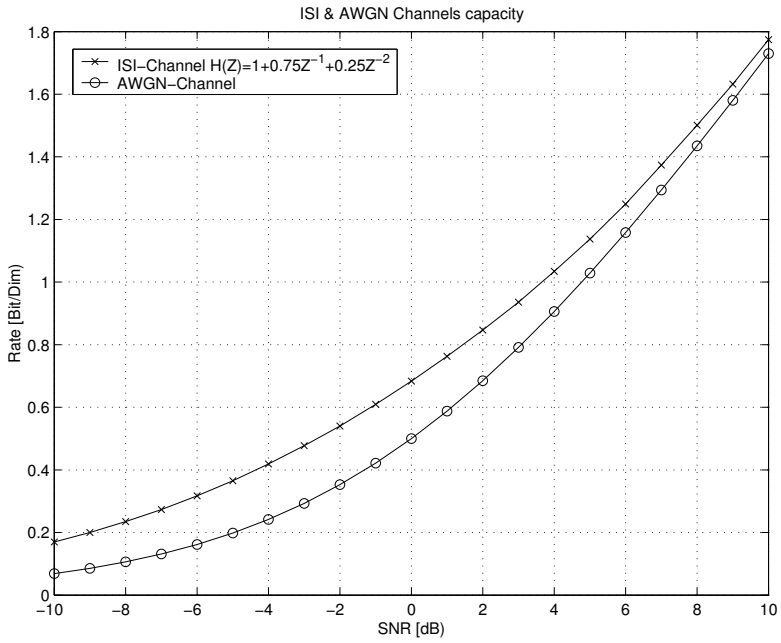

Figure 3.6: ISI-channel and AWGN channel capacity

Figure 3.7 and Figure 3.8 are drawn for rates $R = \frac{1}{8}, \frac{1}{9}$ bit/dim, respectively. There is a gain of about 1 dB respect to the standard TH-FFE precoding

at low SNR. The difference between the optimal ML decoder and the sub optimal decoder is negligible, which implies that the $\tilde{N}$ has nearly Gaussian distribution.

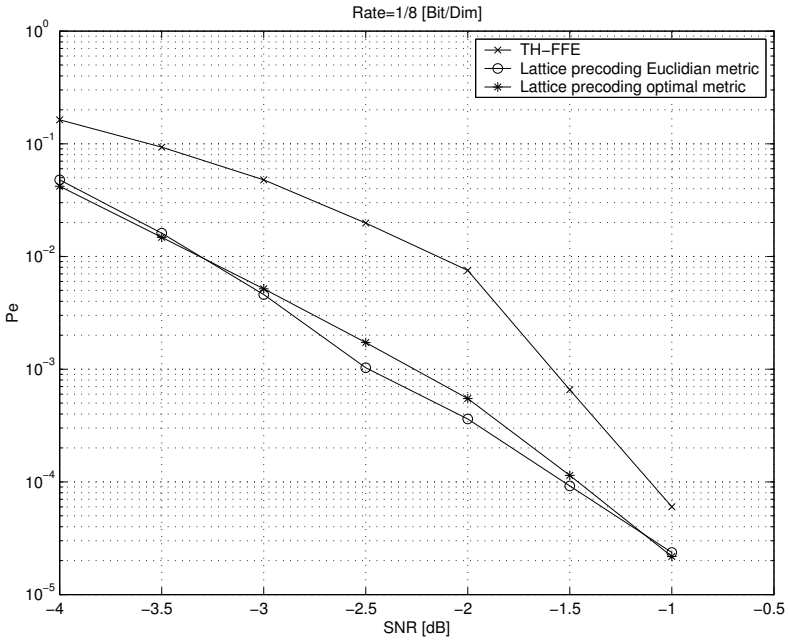

Figure 3.7: TH-FFE versus lattice precoding, $R = \frac{1}{8}$ Bit/Dim

II. A degenerate case for the ISI-channel (3.58), is

$$y_n = x_n + 0.75x_{n-1} + 0.25x_{n-2} + z_n = x_n + s_n + z_n, \tag{3.59}$$

where the post cursor ISI is the known interference $s_n = 0.75x_{n-1} + 0.25x_{n-2}$. The capacity of this interference channel is $C = \frac{1}{2}\log(1 + SNR)$ which is less then the ISI-channel (3.58) capacity as shown in Figure 3.6, since we restrict of using zero decision delay (which is optimal at high SNR) as uses in ZF-DFE.

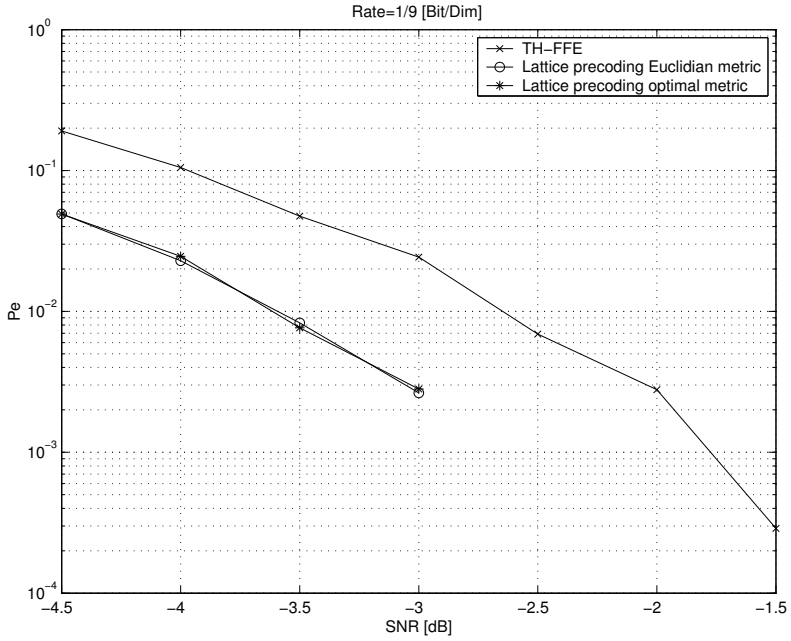

Figure 3.8: TH-FFE versus lattice precoding, $R = \frac{1}{9}$ Bit/Dim

We show the performance results for the following three systems:

System 1: is the usual ZF for TH precoding without MMSE factor.

System 2: includes MMSE, dither and Euclidian metric decoder.

System 3: includes MMSE, dither and optimal metric decoder (according the equivalent noise).

The systems have feedback filters $\tilde{B}_U(Z) = 0.75Z^{-1} + 0.25Z^{-2}$ and feedforward filters $A_U(Z) = 1$.

As it can be shown in Figure 3.9 and Figure 3.10 system 1 could not compete the low SNR condition and has poor performance. The matched decoder for the equivalent noise has better performance, about $0.2 - 0.4\ dB$ (the gap between system 2 and 3).

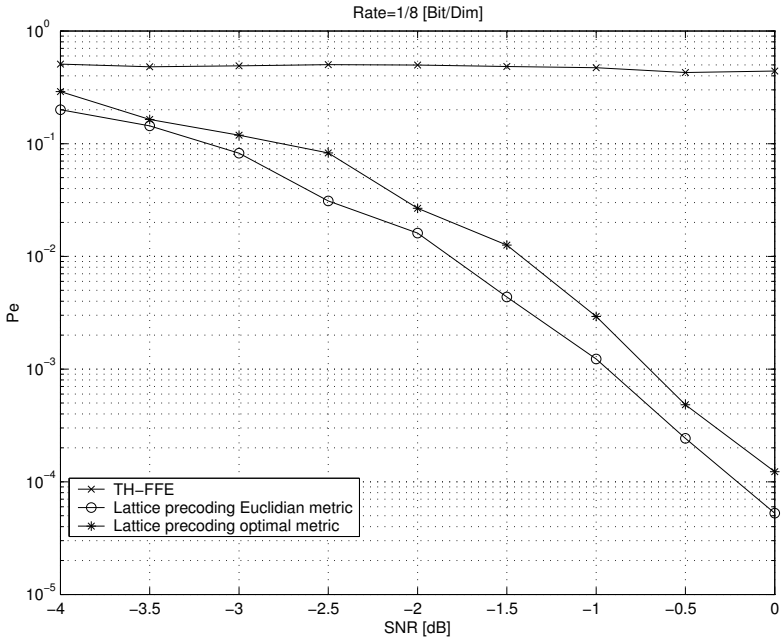

Figure 3.9: TH-ZF versus lattice precoding, $R = \frac{1}{8}$ Bit/Dim

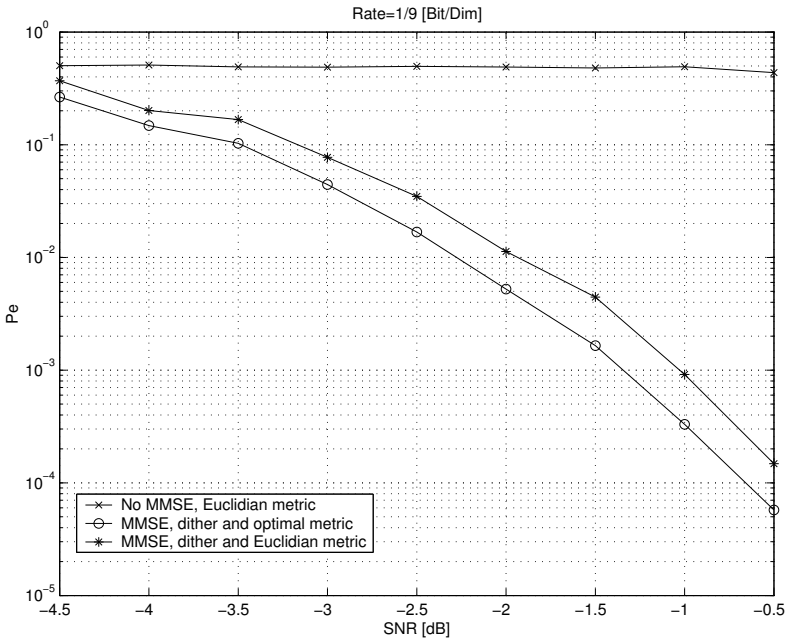

Figure 3.10: TH-ZF versus lattice precoding, $R = \frac{1}{9}$ Bit/Dim

Chapter 4

High Dimensional (Non-Causal) Precoding

In this chapter we present a nested lattices construction. Furthermore, we propose a low rate precoding scheme which combines MMSE estimation, dithering and a variant of nested codes, based on concatenation of a "syndrome dilution" code and a "syndrome-to-coset" modulation code.

4.1 The Nested Lattices Construction

We present the nested lattices construction for Λ_c and Λ_s so that $\mathbb{Z}^N/\Lambda_c/\Lambda_s/p\mathbb{Z}^N$ where p is prime number. The nested lattices construction uses a fundamental method to construct lattice known as "construction A" [4].

Definition 4.1. *Construction A - Let $\mathcal{C}$ be an (n,k) linear code over $GF(p)$. $\mathbf{x} = (x_1, \ldots, x_n)$ is a lattice, Λ, point in $\mathbb{R}^n$ iff $\mathbf{x}$ is congruent (modulo p) to a codeword of $\mathcal{C}$.*

Explicitly, $\mathcal{C}$ is in the region of n-cube $[0, p)^n$, the lattice Λ is obtained by tessellating $\mathbb{R}^n$ with translations of $\mathcal{C}$, i.e., $\Lambda = \mathcal{C} + p\mathbb{Z}^n$, thus Λ is a sub-lattice of $\mathbb{Z}^n$. Furthermore, "construction A" of random ensembles of lattices achieve

the interference-free AWGN channel capacity [18, 7], therefore good lattices can be generated by "construction A". Generally, a sub-lattice Λ_2 of Λ_1 induces a nested lattice partition Λ_1/Λ_2 of Λ_1 into $|\{\Lambda_1 \bmod \Lambda_2\}|$ cosets of Λ_2.

Initially, we construct Λ'_s by "construction A" with linear code $\mathcal{C}_s(n_s, k_s)$ over $\mathbb{Z}_p$, meaning $\Lambda'_s = \mathcal{C}_s + p\mathbb{Z}^{n_s}$, thus Λ'_s is a sub-lattice of $\mathbb{Z}^{n_s}$. The code $\mathcal{C}_s$ contains all the codewords $\{\mathbf{c}_s\}$ so that $\mathbf{c}_s \in \mathbb{Z}_p^{n_s}$ and whose syndrome, $\mathbf{s_y} \triangleq H_s\mathbf{c}_s$, is equal to zero where $\mathbf{s_y} \in \mathbb{Z}_p^{n_s-k_s}$. The coset leader group is the set $\{\mathbb{Z}^{n_s} \bmod \Lambda'_s\}$, i.e., the coset with minimum norm, while the coset representative group includes a unique representative of each coset, but not necessarily the coset leaders. The coset representative and coset leader groups have an equal number of elements and they are equivalent modulo Λ'_s. There are $p^{n_s}/|\mathcal{C}_s| = p^{n_s}/p^{k_s} = p^{n_s-k_s}$ coset representatives, which are equal to number of $\mathcal{C}_s$ syndromes. Each syndrome $\mathbf{s_y}$ corresponds to a unique coset representative. In order to obtain a $\mathcal{C}_s$ coset from a specific syndrome, we can apply the "pseudo" right inverse parity check matrix on the syndrome [10], $\mathbf{t} = H_s^{-1}\mathbf{s_y}$ so that $\mathbf{s_y} = H_s\mathbf{t}$, that is $\mathbf{t}$ represents a coset of $\mathcal{C}_s$. As shown in Lemma 4.1

Lemma 4.1. *For the lattice chain $\mathbb{Z}^{n_s}/\Lambda'_s/p\mathbb{Z}^{n_s}$, assuming that Λ'_s has been constructed by "construction A" with the linear code $\mathcal{C}_s$ over $GF(p)$. Then each syndrome of $\mathcal{C}_s$ has a unique coset representative of Λ'_s in $\mathbb{Z}^{n_s}$ in the region of the basic cell of $p\mathbb{Z}^{n_s}$.*

Proof. $\mathbf{s_y}$ is a syndrome of $\mathcal{C}_s$, let multiply $\mathbf{s_y}$ by H_S^{-1}, the pseudo right inverse of parity check matrix. Thus, $\mathbf{t} = H_s^{-1}\mathbf{s_y}$ represents a coset of the $\mathcal{C}_s$ code. Now applying the "construction A" to obtain $\mathbf{v}'$, we are interested only in $\mathbf{v}' \in \{\text{basic cell of } \mathrm{p}\mathbb{Z}^{\mathrm{n_s}}\}$.

For $\mathbf{s_y} = \mathbf{0}$ we have that $\mathbf{t} \in \mathcal{C}_s$, thus $\mathbf{v}' \in \{\Lambda'_s \bmod p\mathbb{Z}^{n_s}\}$.

For $\mathbf{s_y} \neq \mathbf{0}$, we have that $\mathbf{t} \in \mathcal{C}_s + \mathbf{a}$, where $\mathbf{a} \in \mathbb{Z}_p^{n_s}$ is a coset of $\mathcal{C}_s$, thus $\mathbf{v}' \in \Lambda'_s + \mathbf{a}$ in the region of the basic cell of $p\mathbb{Z}^{n_s}$.

For the mapping uniqueness, we have to show that for fixed H_s^{-1}, each syndrome maps to a different coset. Assuming that $\mathbf{s_{y_1}} \neq \mathbf{s_{y_2}}$ are mapped to the same coset $\{\Lambda'_s + \mathbf{a_1}\} = \{\Lambda'_s + \mathbf{a_2}\}$, then $\mathbf{t_1} = \mathcal{C}_1 + \mathbf{a_1}$ and $\mathbf{t_2} = \mathcal{C}_2 + \mathbf{a_1}$ where $c_1,\ c_2 \in \mathcal{C}_s$ have the same coset in $\mathcal{C}_s$ code. Therefore, $\mathbf{s_{y_1}} = \mathbf{s_{y_2}}$ which is opposed to the assumption, then each syndrome has unique coset and unique coset representative. Generally, H_s^{-1} is not a unique pseudo right inverse parity check matrix. Therefore, we can have several different mappings from syndromes to coset representatives, but each map is one to one mapping. □

Finally, a sub-lattice Λ_c of $\mathbb{Z}^N$ is obtained by dilution of $\mathbb{Z}^N$ codewords, which is accomplished by dilution of $\mathcal{C}_s$ syndromes. The set of $\mathcal{C}_s$ syndromes, $\{S\}$, is additive group over $GF(p)$, furthermore the remaining syndromes should be a sub group of $\{S\}$ in order to span the lattice Λ_c. Specifically, it is achieved by linear code $\mathcal{C}_c(n_c, k_c)$ over $GF(p)$, which produces a subgroup of syndromes. The nested lattices construction is shown schematically in Figure 4.1.

It should be noted, as a result of the proof of Lemma 4.1, that $\mathbf{t} = H_s^{-1}\mathbf{s_y}$ can be regarded as coset representative of Λ_s in Λ_c in the region of the basic cell of $p\mathbb{Z}^N$.

Now, we consider the dimensions of the nested lattices construction. Λ_c is sub-lattice of $\mathbb{Z}^N$, where $\mathbb{Z}^N$ is $\frac{LCM(n_c, n_s-k_s)}{n_s-k_s}$ times Cartesian product of $\mathbb{Z}^{n_s}$ thus $N = n_s\frac{LCM(n_c, n_s-k_s)}{n_s-k_s}$, and Λ_s is $\frac{LCM(n_c, n_s-k_s)}{n_s-k_s}$ times Cartesian product of

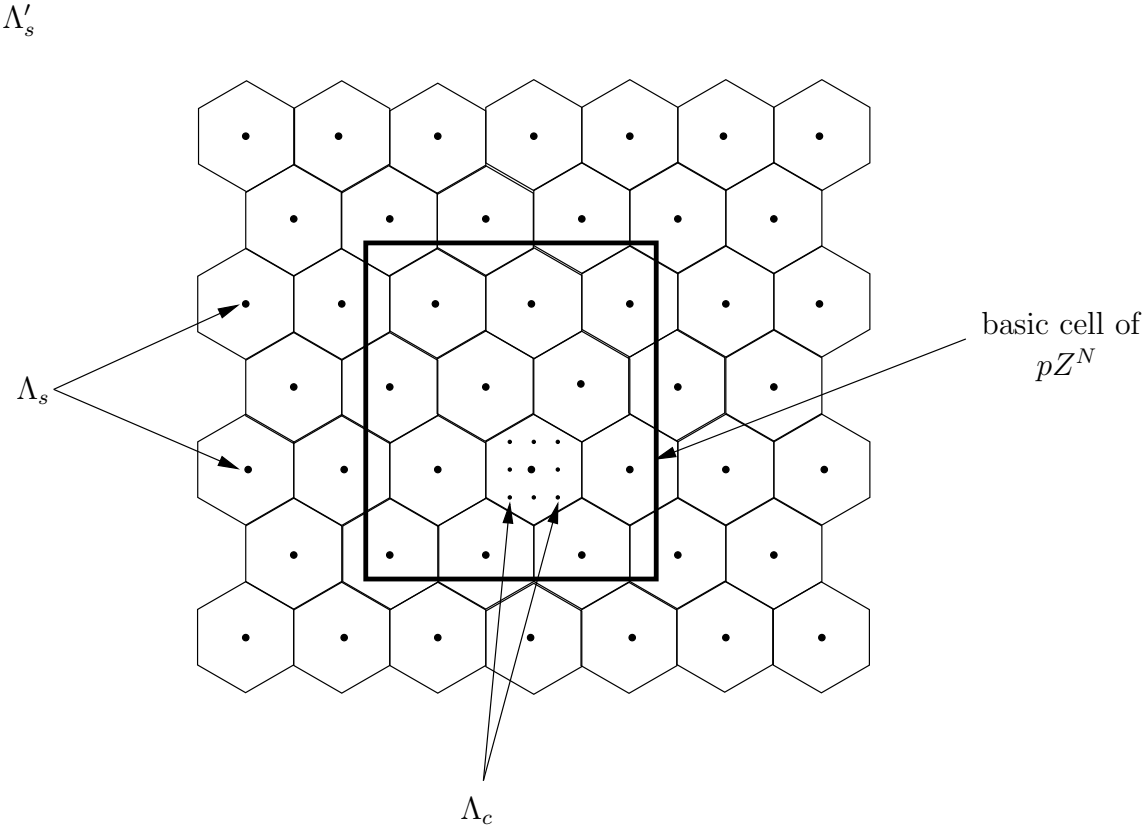

Figure 4.1: Nested lattices structure

The construction implies effective rate of $\frac{1}{N}\log_2(V_s/V_c)$ bit/dim, where V_c and V_s are the cells volume of Λ_c and Λ_s, respectively. The rate is given by

$$\begin{aligned} R &= \frac{1}{N}\log_2(V_s/V_c) && (4.1)\\ &= \frac{1}{N}\log_2(p^{k_c\frac{LCM(n_c,n_s-k_s)}{n_c}}) && (4.2)\\ &= \frac{k_c(n_s-k_s)}{n_c n_s}\log_2 p \text{ bit/dim.} && (4.3) \end{aligned}$$

Assuming that n_c is a multiple of n_s-k_s, then the construction $\mathbb{Z}^{\tilde{N}}/\Lambda_c/\Lambda_s/p\mathbb{Z}^{\tilde{N}}$ has $\tilde{N} = n_s\frac{n_c}{n_s-k_s}$ and the rate is unchanged.

An important property of this construction is shown in the following corollaries.

Corollary 4.0.1. *For the above construction, the tessellation of* $\mathbb{R}^N$ *by the basic cell of* $p\mathbb{Z}^N$*, span the lattices* Λ_c*,* Λ_s *in* $\mathbb{R}^N$*.*

Proof. Let define the coset leader group of $p\mathbb{Z}^N$ in Λ_s by $[\Lambda_s/p\mathbb{Z}^N]_{cl}$, and the coset leader group of Λ_s in Λ_c by $[\Lambda_c/\Lambda_s]_{cl}$, we have that

$$\Lambda_s = p\mathbb{Z}^N + [\Lambda_s/p\mathbb{Z}^N]_{cl}, \tag{4.4}$$

therefore tessellation of $\mathbb{R}^N$ by the basic cell of $p\mathbb{Z}^N$ span Λ_s. Λ_c can be written as

$$\Lambda_c = \Lambda_s + [\Lambda_c/\Lambda_s]_{cl} = p\mathbb{Z}^N + [\Lambda_s/p\mathbb{Z}^N]_{cl} + [\Lambda_c/\Lambda_s]_{cl}, \tag{4.5}$$

therefore tessellation of $\mathbb{R}^N$ by the basic cell of $p\mathbb{Z}^N$ span also Λ_c. □

Corollary 4.0.2. *Each coset of* Λ_s *in* Λ_c *has equal number of elements in the basic cell of* $p\mathbb{Z}^N$*.*

Proof. We have two kinds of Λ_s cells in the basic cell of $p\mathbb{Z}^N$:

1. The intact Λ_s cells - which contribute an equal number of elements from each coset of Λ_s in Λ_c.

2. The fragmented Λ_s cells - since the tessellation of $\mathbb{R}^N$ by the basic cell of $p\mathbb{Z}^N$ span Λ_s all the broken cells of Λ_s must compose an integer number of Λ_s cells, which contribute an equal number of elements from each coset of Λ_s in Λ_c.

□

This corollary can be applied for coset of Λ_s in $\mathbb{R}^N$, i.e, each coset of Λ_s in $\mathbb{R}^N$ over the basic cell of $p\mathbb{Z}^N$ has equal number of elements.

4.2 Transmission Scheme

The transmission scheme is based on the following generalization of the "inflated lattice lemma" [8].

Lemma 4.2. *(**Generalized inflated lattice lemma**) For the nested lattices chain $\Lambda_G/\Lambda_c/\Lambda_s/\Lambda_A$ and the channel defined by (1.1).*

Encoder:

$$\mathbf{x} = [\mathbf{v}' - \alpha\mathbf{s} - \tilde{\mathbf{d}}] \ mod \ \Lambda_s, \tag{4.6}$$

where $\mathbf{v}' \in \{\Lambda_c \cap \mathcal{V}_A\}$, $\tilde{\mathbf{d}} \sim U(\mathcal{V}_A)$ and $\mathcal{V}_A$ is the basic cell Λ_A.

Decoder:

$$\mathbf{y}' = [\alpha\mathbf{y} + \tilde{\mathbf{d}}] \ mod \ \Lambda_A. \tag{4.7}$$

For $\mathbf{v} \in \{\Lambda_c \cap \mathcal{V}_s\}$ the equivalent channel satisfies

$$\mathbf{Y}' = [\mathbf{v} + \mathbf{B} + \mathbf{Z}'] \ mod \ \Lambda_A, \tag{4.8}$$

with

$$\mathbf{Z}' = [(1-\alpha)\mathbf{U} + \alpha\mathbf{Z}] \ mod \ \Lambda_A, \tag{4.9}$$

where $\mathbf{B} = Q_{\Lambda_s}(\mathbf{v}') - Q_{\Lambda_s}(\mathbf{v}' - \alpha\mathbf{s} - \tilde{\mathbf{d}}) \in \Lambda_s$ and $\mathbf{U} \sim U(\mathcal{V}_s)$.

Proof. From (1.1), (4.6) and (4.7)

$$\begin{aligned} Y &= X + S + Z & (4.10)\\ \mathbf{x} &= \left[\mathbf{v}' - \alpha\mathbf{s} - \tilde{\mathbf{d}}\right] \bmod \Lambda_s & (4.11)\\ \mathbf{y}' &= \left[\alpha\mathbf{y} + \tilde{\mathbf{d}}\right] \bmod \Lambda_A & (4.12) \end{aligned}$$

$\mathbf{y}'$ can be written as

$$\begin{aligned}
\mathbf{y}' &= \Big[\alpha(\mathbf{x}+\mathbf{s}+\mathbf{z})+\tilde{\mathbf{d}}\Big] \bmod \Lambda_A && (4.13)\\
&= \Big\{\alpha\Big[(\mathbf{v}'-\tilde{\mathbf{d}}-\alpha\mathbf{s})\bmod \Lambda_s+\mathbf{s}+\mathbf{z}\Big]+\tilde{\mathbf{d}}\Big\} \bmod \Lambda_A && (4.14)\\
&= \Big\{\alpha\Big[(\mathbf{v}'-\tilde{\mathbf{d}}-\alpha\mathbf{s})-Q_{\Lambda_s}(\mathbf{v}'-\tilde{\mathbf{d}}-\alpha\mathbf{s})+\\
&\quad +\mathbf{s}+\mathbf{z}+\tilde{\mathbf{d}}/\alpha\Big]\Big\} \bmod \Lambda_A && (4.15)\\
&= \Big\{\alpha\Big[\mathbf{v}'-Q_{\Lambda_s}(\mathbf{v}'-\tilde{\mathbf{d}}-\alpha\mathbf{s})+\frac{1-\alpha}{\alpha}\Big[\mathbf{v}'-Q_{\Lambda_s}(\mathbf{v}'-\tilde{\mathbf{d}}-\alpha\mathbf{s})\Big]-\\
&\quad -\frac{1-\alpha}{\alpha}\Big[\mathbf{v}'-Q_{\Lambda_s}(\mathbf{v}'-\tilde{\mathbf{d}}-\alpha\mathbf{s})\Big]+(1-\alpha)\mathbf{s}+\\
&\quad +\frac{1-\alpha}{\alpha}\tilde{\mathbf{d}}+\mathbf{z}\Big]\Big\} \bmod \Lambda_A && (4.16)\\
&= \Big\{\alpha\Big[\frac{1}{\alpha}\Big[\mathbf{v}'-Q_{\Lambda_s}(\mathbf{v}'-\tilde{\mathbf{d}}-\alpha\mathbf{s})\Big]+\\
&\quad +\frac{1-\alpha}{\alpha}\Big[Q_{\Lambda_s}(\mathbf{v}'-\tilde{\mathbf{d}}-\alpha\mathbf{s})-\mathbf{v}'+\tilde{\mathbf{d}}+\alpha\mathbf{s}\Big]+\mathbf{z}\Big]\Big\} \bmod \Lambda_A && (4.17)\\
&= \Big\{\mathbf{v}'-Q_{\Lambda_s}(\mathbf{v}'-\tilde{\mathbf{d}}-\alpha\mathbf{s})+\\
&\quad (1-\alpha)\Big[Q_{\Lambda_s}(\mathbf{v}'-\tilde{\mathbf{d}}-\alpha\mathbf{s})-(\mathbf{v}'-\tilde{\mathbf{d}}-\alpha\mathbf{s})\Big]+\alpha\mathbf{z}\Big\} \bmod \Lambda_A && (4.18)
\end{aligned}$$

Let

$$\mathbf{U}\triangleq Q_{\Lambda_s}(\mathbf{v}'-\tilde{\mathbf{d}}-\alpha\mathbf{s})-(\mathbf{v}'-\tilde{\mathbf{d}}-\alpha\mathbf{s}), \qquad (4.19)$$

since $\mathbf{v}=\mathbf{v}' \bmod \Lambda_s = \mathbf{v}'-Q_{\Lambda_S}(\mathbf{v}')$, therefore $\mathbf{v}'=\mathbf{v}+Q_{\Lambda_s}(\mathbf{v}')$ and $\mathbf{y}'$ can be written as

$$\begin{aligned}
\mathbf{y}' &= \Big[\mathbf{v}'-Q_{\Lambda_s}(\mathbf{v}'-\tilde{\mathbf{d}}-\alpha\mathbf{s})+(1-\alpha)\mathbf{U}+\alpha\mathbf{z}\Big] \bmod \Lambda_A && (4.20)\\
&= \Big[\mathbf{v}+Q_{\Lambda_s}(\mathbf{v}')-Q_{\Lambda_s}(\mathbf{v}'-\tilde{\mathbf{d}}-\alpha\mathbf{s})+\\
&\quad +(1-\alpha)\mathbf{U}+\alpha\mathbf{z}\Big] \bmod \Lambda_A && (4.21)
\end{aligned}$$

Let

$$\mathbf{B} \triangleq Q_{\Lambda_s}(\mathbf{v}') - Q_{\Lambda_s}(\mathbf{v}' - \tilde{\mathbf{d}} - \alpha\mathbf{s}), \tag{4.22}$$

thus $\mathbf{B} \in \Lambda_s$, furthermore let

$$\mathbf{Z}' \triangleq \Big[(1-\alpha)\mathbf{U} + \alpha\mathbf{Z}\Big] \bmod \Lambda_A. \tag{4.23}$$

Therefore, $\mathbf{y}'$ is given by

$$\mathbf{y}' = \Big[\mathbf{v} + \mathbf{B} + \mathbf{Z}'\Big] \bmod \Lambda_A. \tag{4.24}$$

Since $\tilde{\mathbf{d}} \bmod \Lambda_s$ is uniformly distributed over $\mathcal{V}_s$ as shown in corollary 4.0.2, let $\mathbf{d} \triangleq \tilde{\mathbf{d}} \bmod \Lambda_s$, where $\mathbf{d} \sim U(\mathcal{V}_s)$. Therefore, $\mathbf{U}$ can be written as

$$\begin{aligned}
\mathbf{U} &= Q_{\Lambda_s}(\mathbf{v}' - \tilde{\mathbf{d}} - \alpha\mathbf{s}) - (\mathbf{v}' - \tilde{\mathbf{d}} - \alpha\mathbf{s}) && (4.25)\\
&= [\tilde{\mathbf{d}} + \alpha\mathbf{s} - \mathbf{v}'] \bmod \Lambda_s && (4.26)\\
&= [\mathbf{d} + Q_{\Lambda_s}(\tilde{\mathbf{d}}) + \alpha\mathbf{s} - (\mathbf{v} + Q_{\Lambda_s}\mathbf{v}')] \bmod \Lambda_s && (4.27)\\
&= [\mathbf{d} + \alpha\mathbf{s} - \mathbf{v}] \bmod \Lambda_s && (4.28)\\
&= Q_{\Lambda_s}(\mathbf{v} - \mathbf{d} - \alpha\mathbf{s}) - (\mathbf{v} - \mathbf{d} - \alpha\mathbf{s}), && (4.29)
\end{aligned}$$

by the dithered quantization property $\mathbf{U} \sim U(\mathcal{V}_s)$ and independent in $\mathbf{v}$, $\mathbf{v}'$. □

The distribution of $\mathbf{x}$ is also uniform over $\mathcal{V}_s$ by the same arguments has been used for $\mathbf{U}$. Therefore, we state the following corollary.

Corollary 4.0.3. *In the "generalized inflated lattice lemma", the transmitted signal* $\mathbf{x}$ *is given by*

$$\mathbf{x} = [\mathbf{v}' - \alpha\mathbf{s} - \tilde{\mathbf{d}}] \ mod\, \Lambda_s = [\mathbf{v} - \alpha\mathbf{s} - \mathbf{d}] \ mod\, \Lambda_s, \tag{4.30}$$

where $\mathbf{d} \sim U(\mathcal{V}_s)$ *and* $\mathbf{x}$ *is uniformly distributed over* $\mathcal{V}_s$ *and independent of* $\mathbf{v}$, $\mathbf{v}'$, $\mathbf{s}$. *The second moment of* $\mathbf{x}$ *is* $P_X = \sigma_s^2$, *where* σ_s^2 *is the lattice* Λ_s *second moment.*

Applying the generalized inflated lattice lemma to the nested lattices construction, where $\Lambda_G = \mathbb{Z}^N$ and $\Lambda_A = p\mathbb{Z}^N$, the equivalent channel till $\mathbf{y}'$ is given by $\mathbf{Y}' = [\mathbf{v} + \mathbf{B} + \mathbf{Z}']$ mod $p\mathbb{Z}^N$ and $\mathbf{B} \in \Lambda_s$. The effect of $\mathbf{B}$ can be cancelled by ML decoder on $\mathbf{y}'$ for the nearest Λ_c codeword in the basic cell of $p\mathbb{Z}^N$. Then, by calculating the syndrome which corresponds to the Λ_c codeword. The information bits are reconstructed directly from the syndrome (the fine code).

4.2.1 Transmitter

The information bits have to be translated into coset representative, this is done by the coset selector which composes from two parts, as shown in Figure 4.2

Shaping code $\mathbb{Z}^{\mathbf{N}}/\mathbf{\Lambda_s}$ - is equivalently the coarse code encoder with rate $R_s = \frac{n_s - k_s}{n_s}$. The input to this block is a syndrome of $\mathcal{C}_s$ code. The shaping code block, equivalently, transforms a syndrome of $\mathcal{C}_s$ to coset representative in the partition $\mathbb{Z}^N/\Lambda_s$ in the region of the basic cell of $p\mathbb{Z}^N$. The mapping from syndrome of $\mathcal{C}_s$ to coset representative of Λ_s in $\mathbb{Z}^N$ is one to one mapping using the pseudo right inverse parity check matrix H_s^{-1}, as shown in Lemma 4.1 where the coset representatives are in different cosets.

Coset dilution - is equivalently the fine code encoder with rate $R_{\mathcal{C}_c} = \frac{k_c}{n_c}$, the coset dilution output can be referred as a syndrome of $\mathcal{C}_s$ code. The fine code dilutes syndromes of $\mathcal{C}_s$, or equivalently dilutes cosets representative of Λ_s in $\mathbb{Z}^N$. The coset dilution output is a codeword of $\mathcal{C}_c$ code, therefore it is a subgroup of syndromes over $GF(p)$. The result of Lemma 4.1 can be applied for any subgroup of syndromes. Therefore, we can not have any two

distinct syndromes which are mapped into two cosets representative with the same coset leader. Since we use only a subgroup of syndromes, each syndrome is map to unique coset representative of Λ_s in Λ_c. Therefore, the modulo Λ_s operation is with out loss of information.

The dither $\tilde{\mathbf{d}}$ and the interference $\alpha\mathbf{s}$ are subtracted from the coset representative $\mathbf{v}'$ and $\mathbf{x}$ is obtained by the modulo Λ_s as shown in Lemma 4.2.

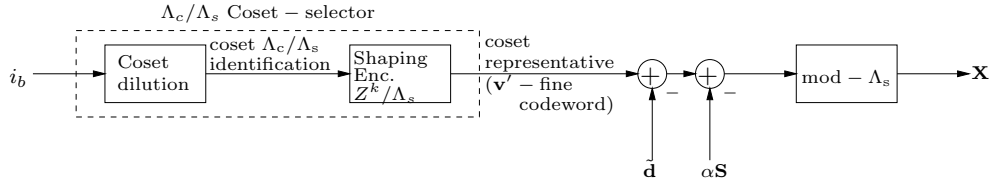

Figure 4.2: Transmitter

A detailed "rate flow diagram" of the encoder is illustrated in Figure 4.3, where the m_s, m_c integers are determined by $LCM(n_c, n_s - k_s) = m_s(n_s - k_s) = m_c n_c$. A K-tuples of p-ary symbols is segmented into m_c vectors of length k_c, each of the m_c vector is encoded by $R_{\mathcal{C}_c} = k_c/n_c$ fine encoder. Therefore, at the fine encoder output there are m_c vectors of length n_c. The $m_c n_c$ p-ary symbols are segmented into m_s vectors of $n_s - k_s$ symbols, each of the m_s vector is encoded by $R_s = \frac{n_s - k_s}{n_s}$ shaping encoder, thus at the shaping encoder output there are m_s vectors of length n_s symbols. The encoder rate is given by

$$R = \frac{K}{N} = \frac{m_c k_c}{m_s n_s} = \frac{k_c(n_s - k_s)}{n_c n_s} \log_2 p \tag{4.31}$$

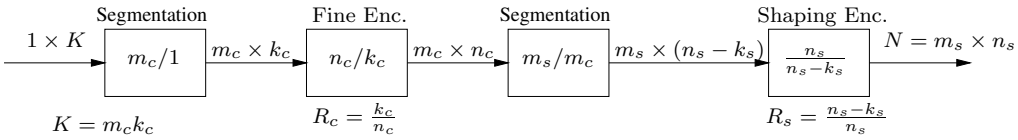

Figure 4.3: Encoder rate flow diagram

4.2.2 Receiver

The decoder of the generalized inflated lattice lemma for $\Lambda_A = \mathbb{Z}^N$ and $\Lambda_G = p\mathbb{Z}^N$ is given by

$$\mathbf{y}' = [\alpha \mathbf{y} + \tilde{\mathbf{d}}] \bmod p\mathbb{Z}^N, \tag{4.32}$$

and the equivalent channel till $\mathbf{y}'$ is

$$\mathbf{Y}' = [\mathbf{v} + \mathbf{Z}' + \mathbf{B}] \bmod p\mathbb{Z}^N, \tag{4.33}$$

where $\mathbf{B} \in \Lambda_s$, the equivalent noise $\mathbf{Z}' = [(1-\alpha)\mathbf{U} + \alpha\mathbf{Z}] \bmod p\mathbb{Z}^N$ and $\mathbf{U} \sim U(\mathcal{V}_s)$.

For high dimensional lattice chain with Voronoi partition and Λ_s with $G_N(\Lambda_s) \approx 1/2\pi e$, the equivalent noise approaches to Gaussian distribution in the equivalent modulo additive channel. In this case, it is equivalent to use quantizer $Q_{\Lambda_c}(\cdot)$ over the basic cell of $p\mathbb{Z}^N$ which estimates the transmitted coset of Λ_c, and alternatively the following ML decoder,

$$\hat{\mathbf{v}} = \underset{\mathbf{v} \in \{\Lambda_c \bmod p\mathbb{Z}^N\}}{\arg\max}\ p(\mathbf{y}'|\mathbf{v}).$$

Since Λ_c is Voronoi partition, then $Q_{\Lambda_c}(\cdot)$ is nearest neighbor decoding and ML decoder for Gaussian noise is also nearest neighbor decoding.

Corollary 4.0.4. *For $N \to \infty$ and $G_N(\Lambda_s) \approx 1/2\pi e$, we have that the following receivers are equivalent:*

1. Calculate $\mathbf{y}' = \left[\alpha\mathbf{y} + \tilde{\mathbf{d}}\right] \ mod\, p\mathbb{Z}^N$.

then, perform ML decoding

$$\hat{\mathbf{v}} = \underset{\mathbf{v}\in\{\Lambda_c \ mod\, p\mathbb{Z}^N\}}{\arg\max}\ p(\mathbf{y}'|\mathbf{v}).$$

2. Calculate $\mathbf{y}' = \left[\alpha\mathbf{y} + \tilde{\mathbf{d}}\right] \ mod\, p\mathbb{Z}^N$.

then,

$$\hat{\mathbf{v}} = Q_{\Lambda_c}(\mathbf{y}').$$

The receiver is shown in Figure 4.4.

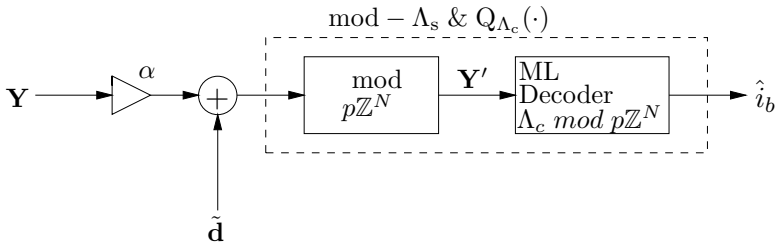

Figure 4.4: Receiver

4.2.3 Nesting by Concatenated Codes and Syndrome Dilution

Although the transmission scheme has been constructed from nested lattice codes, it can be interpreted as structure of concatenated codes. This structure is composed of an inner code - performs "syndrome-to-coset-modulation" for constellation shaping, and an outer code - adds redundancy for error correction (equivalent to "syndromes dilution") to enhance the noise immunity. Furthermore, the outer code is responsible for the coding gain while the inner code for the shaping gain.

4.3 The Construction Optimality

The nested lattices construction uses the lattice partition $\mathbb{Z}^N/\Lambda_s/p\mathbb{Z}^N$ to construct Λ_s, we determine the bound of the shaping gain which is achieved by this construction. The one dimensional distribution $f_\lambda(x)$ with second moment P_X which maximize the differential entropy $h(X)$ constraint that x is in finite interval, I, is the truncated Gaussian distribution [11],

$$f_\lambda(x) = k_\lambda \exp\{-\lambda x^2\},\ x \in I, \tag{4.34}$$

where k_λ is chosen so that $\int_{x\in I} f_\lambda(x)dx = 1$. The construction of the shaping code adds another constraint that each coset of $p\mathbb{Z}$ in $\mathbb{R}$ has an equal probability, now the constant k_λ depends on x and is given by

$$k_\lambda(x) = \frac{1}{\frac{|I|}{p}\sum_{i=0}^{i=p-1} exp\Big\{ -\lambda\Big([x+ip] \bmod I\Big)^2\Big\}},\ x \in I \tag{4.35}$$

and the pdf is

$$f_\lambda(x) = \frac{\exp\{-\lambda x^2\}}{\frac{|I|}{p}\sum_{i=0}^{i=p-1} exp\Big\{ -\lambda\Big([x+ip] \bmod I\Big)^2\Big\}},\ x \in I. \tag{4.36}$$

The entropy $h_\lambda(X)$ and the average power P_X are

$$h_\lambda(X) = -\int_{x\in I} f_\lambda(x)\log_2 f_\lambda(x)dx \tag{4.37}$$

$$P_X = \int_{x\in I} x^2 f_\lambda(x)dx. \tag{4.38}$$

The modulo-Λ_s additive noise channel is

$$\mathbf{Y}' = \Big[\mathbf{V} + \mathbf{Z}'\Big] \bmod \Lambda_s$$

where $\mathbf{Z}'$ is equivalent noise. For one dimensional Λ_s, the mutual information $I_{(N_s=1)}(Y', V)$ with $\alpha = \frac{P_X}{P_X+\sigma_z^2}$ is a function of P_X/σ_z^2. Therefore $I_{(N_s=1)}(Y', V)$

can be written as

$$I_{(N_s=1)}(Y';V) = I_{(N_s=1)}(P_{(N_s=1)}/\sigma_z^2). \tag{4.39}$$

Generally, There is a high dimensional precoding with any $N_s = \mathcal{N}_s$ so that the mutual information can achieves

$$I_{(N_s=\mathcal{N}_s)}(\mathbf{Y}';\mathbf{V}) = \frac{1}{\mathcal{N}_s}h_\lambda(\mathbf{Y}') - \frac{1}{\mathcal{N}_s}h(\mathbf{Z}') \tag{4.40}$$

where $\mathbf{Y}' \in I^{\mathcal{N}_s}$ and $\mathbf{Z}' = (1-\alpha)\mathbf{U} - \alpha\mathbf{Z}$. $\mathbf{U}$ has pdf $f_\lambda(\mathbf{u})$, $\mathbf{u} \in I^{\mathcal{N}_s}$ and $\mathbf{Z} \sim N(0, I\sigma_z^2)$.

The shaping gain is the power reduction from $N_s = 1$ to $N_s = \mathcal{N}_s$ for fixed mutual information, thus

$$I_{(N_s=\mathcal{N}_s)}(\mathbf{Y}';\mathbf{V}) = I_{(N_s=1)}(P_{(N_s=1)}/\sigma_z^2), \tag{4.41}$$

then $P_{(N_s=1)} = \sigma_z^2 \cdot {I_{(N_s=1)}}^{-1}\Big(I_{(N_s=\mathcal{N}_s)}(\mathbf{Y}';\mathbf{V})\Big)$ where ${I_{(N_s=1)}}^{-1}(R)$ is the inverse mutual information of one dimensional lattice strategy. The shaping gap is given by

$$\gamma_{ps} = \frac{P_{(N_s=1)}}{P_{(N_s=\mathcal{N}_s)}} = \frac{\sigma_z^2}{P_{(N_s=\mathcal{N}_s)}} \cdot {I_{(N_s=1)}}^{-1}\Big(I_{(N_s=\mathcal{N}_s)}(\mathbf{Y}';\mathbf{V}))\Big) \tag{4.42}$$

where $I_{(N_s=\mathcal{N}_s)}(\mathbf{Y}';\mathbf{V})$ is given by (4.40).

The lattice Λ_s has been constructed by "construction A" with the linear code $\mathcal{C}_s(n_s, k_s)$, thus we have the lattices chain $\mathbb{Z}^{n_s}/\Lambda_s/p\mathbb{Z}^{n_s}$ where $p^{k_s} = |\{\Lambda_s \bmod p\mathbb{Z}^{n_s}\}|$. Assume a dense lattice Δ^{n_s}, for continues approximation, where $\mathbb{Z}^{n_s}$ is a sub-lattice of Δ^{n_s}, i.e, we have the lattices chain $\Delta^{n_s}/\mathbb{Z}^{n_s}/\Lambda_s/p\mathbb{Z}^{n_s}$. Since

$$|\{\Delta^{n_s} \bmod \Lambda_s\}| \approx 2^{n_s h_\lambda(X)} \tag{4.43}$$

$$|\{\Delta^{n_s} \bmod p\mathbb{Z}^{n_s}\}| \approx 2^{n_s \log_2 p}. \tag{4.44}$$

Thus,

$$\begin{aligned} 2^{k_s \log_2 p} &= |\{\Lambda_s \bmod p\mathbb{Z}^{n_s}\}| && (4.45)\\ &= \frac{|\{\Delta^{n_s} \bmod p\mathbb{Z}^{n_s}\}|}{|\{\Delta^{n_s} \bmod \Lambda_s\}|} && (4.46)\\ &\approx 2^{n_s \log_2 p - n_s h_\lambda(X)} && (4.47) \end{aligned}$$

where $\approx$ is due to the continues approximation. The normalized rate per dimension of the quantizer Λ_s over the basic cell of $p\mathbb{Z}^{n_s}$ is given by

$$R_{\Lambda_s} = \frac{k_s}{n_s}\log_2 p = \log_2 p - h_\lambda(X) \text{ bit/dim}, \tag{4.48}$$

or equivalently the shaping code rate is

$$R_{\mathcal{C}_s} = 1 - \frac{h_\lambda(X)}{\log_2 p}. \tag{4.49}$$

In Figure 4.5, the precoding shaping gain (4.42) of Λ_s is drawn versus the rate R_{Λ_s} (4.48) for the lattices chain $\mathbb{Z}^{n_s}/\Lambda_s/A\mathbb{Z}^{n_s}$ with $A = 2, 3, 4, 5, 8$ at $SNR = 0\ dB$.

For $A = 2$ the maximal shaping gain is 2.91 dB at rate $R_{\Lambda_s} = 0.4$ bit/dim, using shaping code with rate $R_{\mathcal{C}_s} = 1/2$, we can achieve 2.85 dB shaping gain from the entire 3.1 dB at $SNR = 0\ dB$. For $A = 3$ the optimum shaping gain is 3.08 dB at rate $R_{\Lambda_s} = 0.73$ bit/dim which can achieve using shaping code with rate $R_{\mathcal{C}_s} = 0.46$ over $GF(3)$, for shaping code with $R_{\mathcal{C}_s} = 1/2$ over $GF(3)$ the shaping gain is 3.07 dB. For $A = 5$ with shaping code rate of $R_{\mathcal{C}_s} = 1/2$ the entire 3.1 dB can be achieved. For $A = 4$ with rate $R_{\Lambda_s} = 1$ bit/dim the total 3.1 dB is achieved, but the lattice construction should be done over the ring $\mathbb{Z}_4$ (not the field $GF(4)$).

For high SNR, Forney showed in [16] that a binary lattice has a loss of 0.1 dB for shaping code rate $R_{\mathcal{C}_s} = 1/2$, while for $SNR = 0\ dB$ the loss is

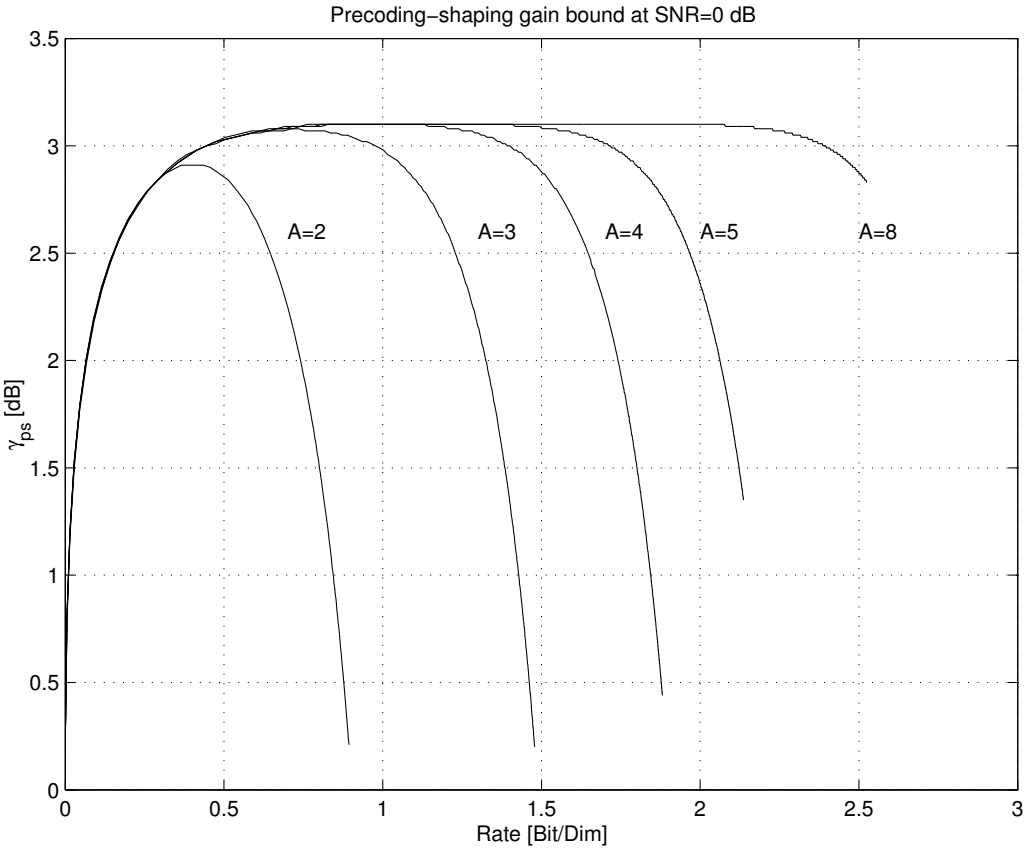

Figure 4.5: Precoding shaping gain bound of lattice construction $\mathbb{Z}^{n_s}/\Lambda_s/A\mathbb{Z}^{n_s}$ at $SNR = 0\ dB$ as function of A

0.25 dB. Generally, as long as the SNR decreases the nested lattices construction has to use a larger p, in order to achieve the entire precoding shaping gain.

4.4 Nested Trellis Codes

A straightforward implementation of the above scheme, for $p = 2$, uses binary convolutional codes $\mathcal{C}_c(n_c, k_c)$, $\mathcal{C}_s(n_s, k_s)$ to construct the nested binary lattices construction $\mathbb{Z}^N/\Lambda_c/\Lambda_s/2\mathbb{Z}^N$. The loss of using $p = 2$ with shaping code rate, $R_s = 1/2$, is 0.25 dB from the entire 3.1 dB shaping gap at SNR = 0 dB, as

shown in the previous section.

Although, the convolutional code can be regarded as infinite length block code. The "construction A" using convolutional code span a lattice (trellis) with a finite effective dimension which depends on the convolutional code power. In our case the two convolutional codes $\mathcal{C}_c$, $\mathcal{C}_s$ span the nested lattice Λ_c, Λ_s with effective dimensions N_c and N_s, respectively. Since Λ_c determines the coding gain, and Λ_s the shaping gain, these values can be dominated by the the convolutional codes power (constrain length) of $\mathcal{C}_c$ and $\mathcal{C}_s$, respectively.

The codes: We use two convolutional binary codes:

Fine code $\mathcal{C}_c(n_c, k_c)$ - with generator matrix $G_c(D)$.

Coarse code $\mathcal{C}_s(n_s, k_s)$ - with generator matrix $G_s(D)$ and parity check matrix $H_s(D)$. As shaping encoder we use the pseudo right inverse matrix of the parity check matrix, $H_s^{-1}(D)$.

Transmitter

The output of the fine code encoder $G_c(D)$ is a syndrome sequence, $S(D)$, of $\mathcal{C}_s$ code. The shaping code encoder output, $T(D)$, is a coset sequence of the code $\mathcal{C}_s$. The sequence $T(D) \in \{0, 1\}$ can be regarded as a series of real value r_n from multidimensional space. Using Lemma 4.2, the dither signal $\tilde{d}_n \sim U([-1, 1))$. The modulo - Λ_s operation can be written as

$$\Big[\mathbf{x}\Big] \bmod \Lambda_s = \mathbf{x} - Q_{\Lambda_s}(\mathbf{x})$$

The lattice quantizer Q_{Λ_s} with Voronoi partition quantize $\mathbf{x}$ to the nearest Λ_s codeword. The quantizer Q_{Λ_s} is equivalent to search the nearest $\mathcal{C}_s$ codeword, in sense of Euclidian distance. practically, it can be preformed using VA for

$\mathcal{C}_s$ code with Euclidian metrics [10], as shown in Figure 4.6. For long delayed

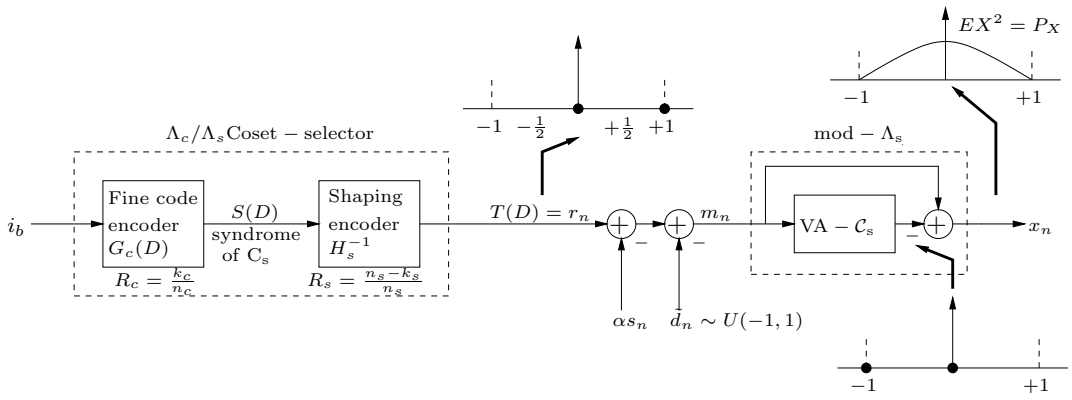

Figure 4.6: Transmitter

VA decision, the VA performs almost nearest neighbor quantization. The transmitted vector is close to uniform distribution over the basic cell of high dimensional Λ_s lattice. The one dimensional output pdf is close to Gaussian distribution.

Receiver

To preform the ML decoding we use VA to search the most likelihood sequence of codeword in Λ_c in the region of the basic cell of $2\mathbb{Z}^N$, as shown in Figure 4.7. The ML decoder is needed to search over all the cosets of $2\mathbb{Z}^N$ in Λ_c. Therefore,

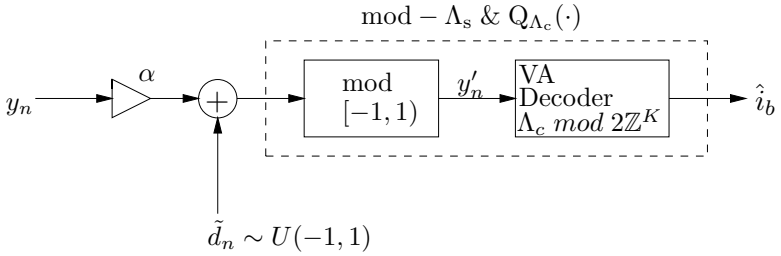

Figure 4.7: Receiver

we need to build the equivalent generator matrix, $\tilde{G}(D)$, which span all the coset leader of $2\mathbb{Z}^N$ in Λ_c.

The generation of $\tilde{G}(D)$ - for the coset representative generation we can use the same structure as used in the transmitter, including the fine encoder and shaping encoder. In order to have all the cosets in Λ_c, it is needed to run over all the possibilities in the same coset of the coset representative. This accomplished by adding all the $\mathcal{C}_s$ codewords to the coset representative path, as shown in Figure 4.8. Since we use $GF(2)$ codes all the additions are over $GF(2)$.

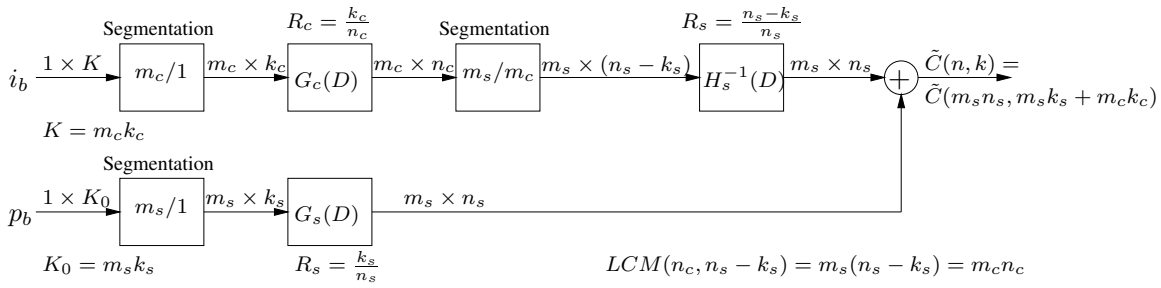

Figure 4.8: $\tilde{G}(D)$ construction

The generator matrix $\tilde{G}(D)$ dimensions are $n = m_s n_s$, and $k = m_c k_c + m_s k_s$.

4.4.1 Performance Results

In this section, we show the performance results of several codes and some scenarios for the nested trellis codes. A detailed example of $\tilde{G}(D)$ generation is shown in the Appendix. The $\tilde{G}(D)$ matrix uses at the decoder for VA decoding. For the other codes the final generator matrix is also given in the Appendix.

Shaping Versus Un-Shaped Precoding Systems

To show the precoding shaping gain that is achieved by high dimensional lattice compared to the case of one dimensional coarse lattice, we use two scenarios: **1)** For the shaped precoding system we use fine code $\mathcal{C}_c$ with $R_{\mathcal{C}_c} = 1/4$ and $\nu_c = 3$, and shaping code with $R_{\mathcal{C}_s} = 1/2$ and $\nu_s = 6$. Another case is shown for shaping code with $R_{\mathcal{C}_s} = 1/2$ and $\nu_s = 8$. In both cases the transmission rate is $\frac{1}{8}$ bit/dim. These codes were compared to the un-shaped precoding system, where the coarse lattice is equivalent to $\mathbb{Z}_s^N$ with the same fine lattice Λ_c and the same rate, as shown in Figure 4.9.

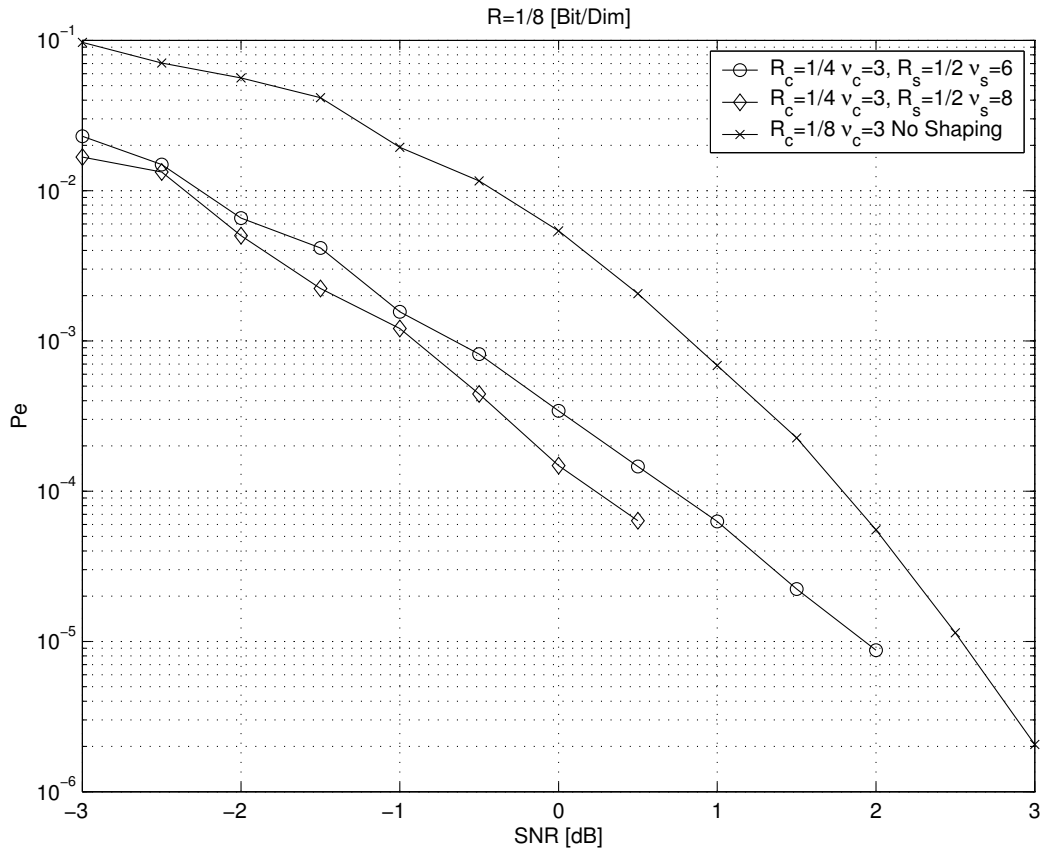

Figure 4.9: Shaping versus un-shaped precoding systems, $R = \frac{1}{8}$ bit/dim

In around $SNR = 0\ dB$ it can be shown a precoding shaping gain of about 1.8 dB between the un-shaped precoding system with shaping code of $\nu_s = 8$. **2)** For the shaped precoding system we use fine code $\mathcal{C}_c$ with $R_{\mathcal{C}_c} = 1/6$ and $\nu_c = 3$, and shaping code with $R_{\mathcal{C}_s} = 1/2$ and $\nu_s = 6$. The rate is $\frac{1}{12}$ bit/dim. The code was compared to the un-shaped precoding system with the same fine lattice Λ_c and the same rate, as shown in Figure 4.10.

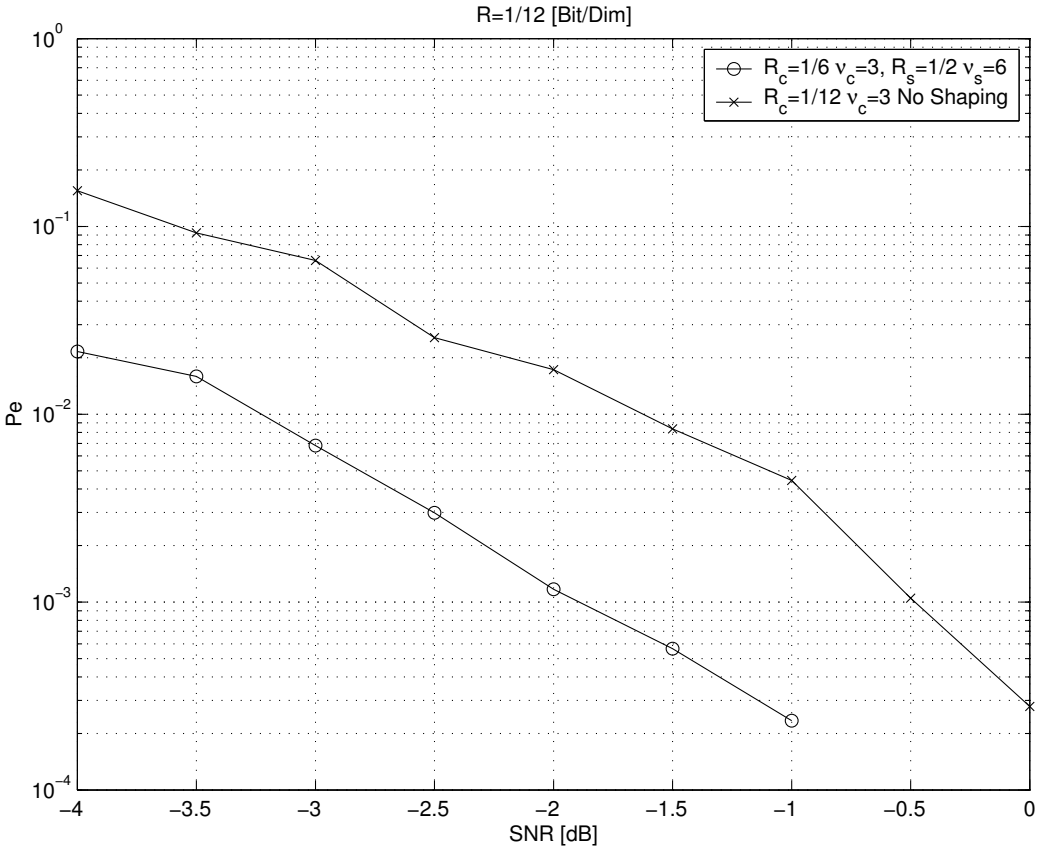

Figure 4.10: Shaping versus un-shaped precoding systems, $R = \frac{1}{12}$ bit/dim

Interference Effects

In our precoding scheme, we expect to have the same performance for any interference. We compare Gaussian and uniform interferences with low inter-

ference power ($-10\ dB$), mid-range power ($0\ dB$), and high power ($+10\ dB$) with respect to the transmitted power.

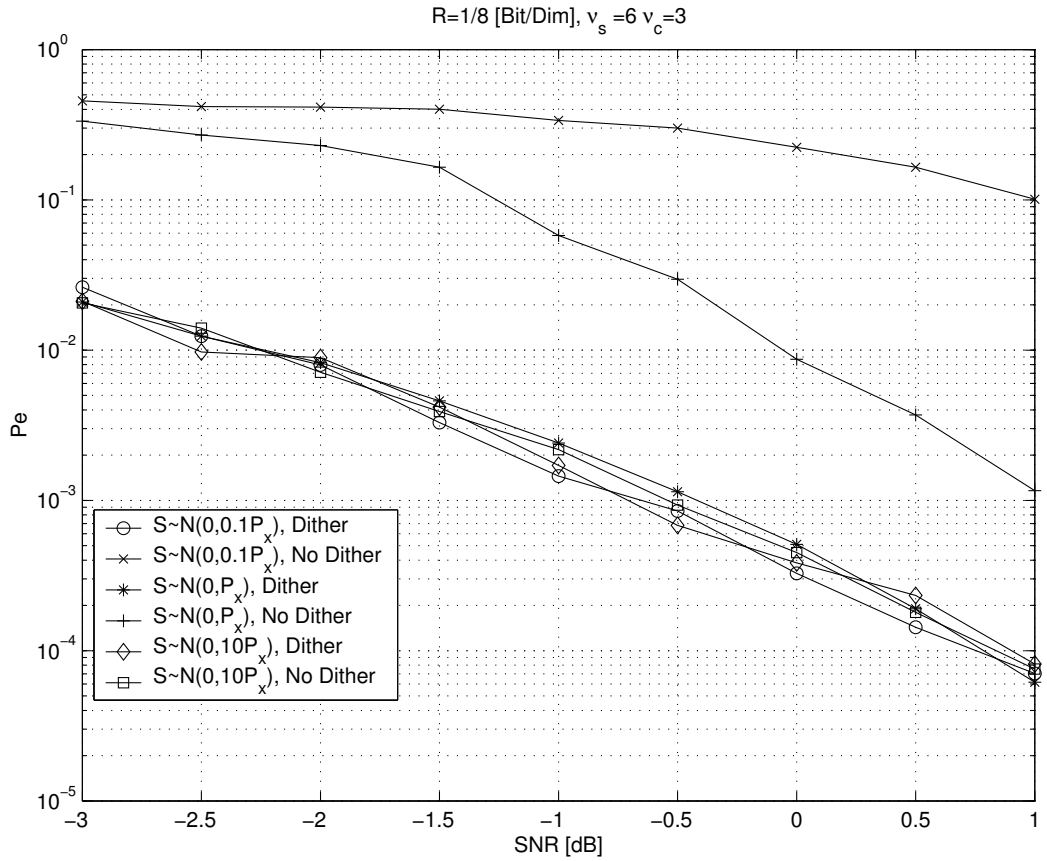

Figure 4.11: Gaussian interference

From Figure 4.11 and Figure 4.12 it can be shown that the interference distribution does not change the performance. Generally, it is correct for any distribution. In both distribution we can notice that the power level of the interference does not change the performance, unless we use a dither. The dither has significant importance to deal with arbitrary interference, even in extreme case where the interference is deterministic [8]. In case we do not use

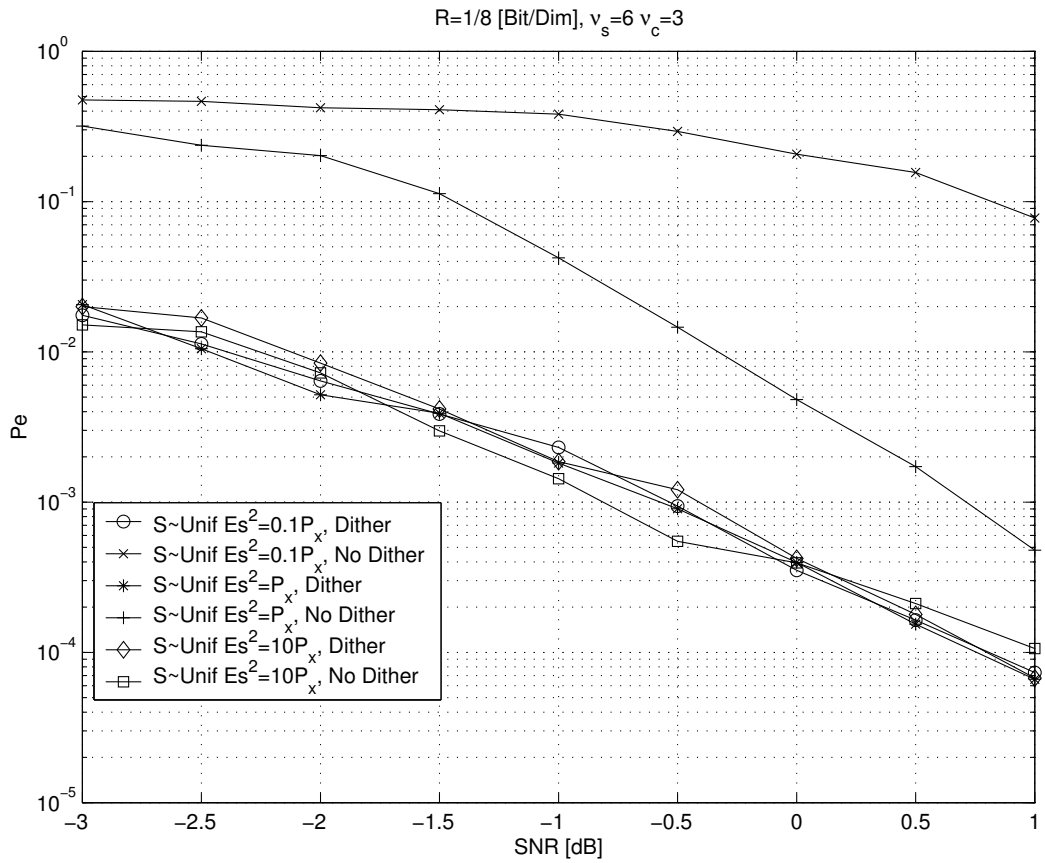

Figure 4.12: Uniform interference

a dither we have poorer performance, especially at mid-range and low interferences power. It is interesting to see that for high power interferences the probability of error with or without dither is the same, since we use continues distributed interferences, then in high variance the interferences have nearly uniform distribution over the interval $[-1, 1)$, thus the interference plays the role of the dither.

MMSE Factor and Dither Effects

In order to analyze the importance of the dither and the MMSE estimation α at low SNR, we explore our scheme with and without a dither and MMSE estimation.

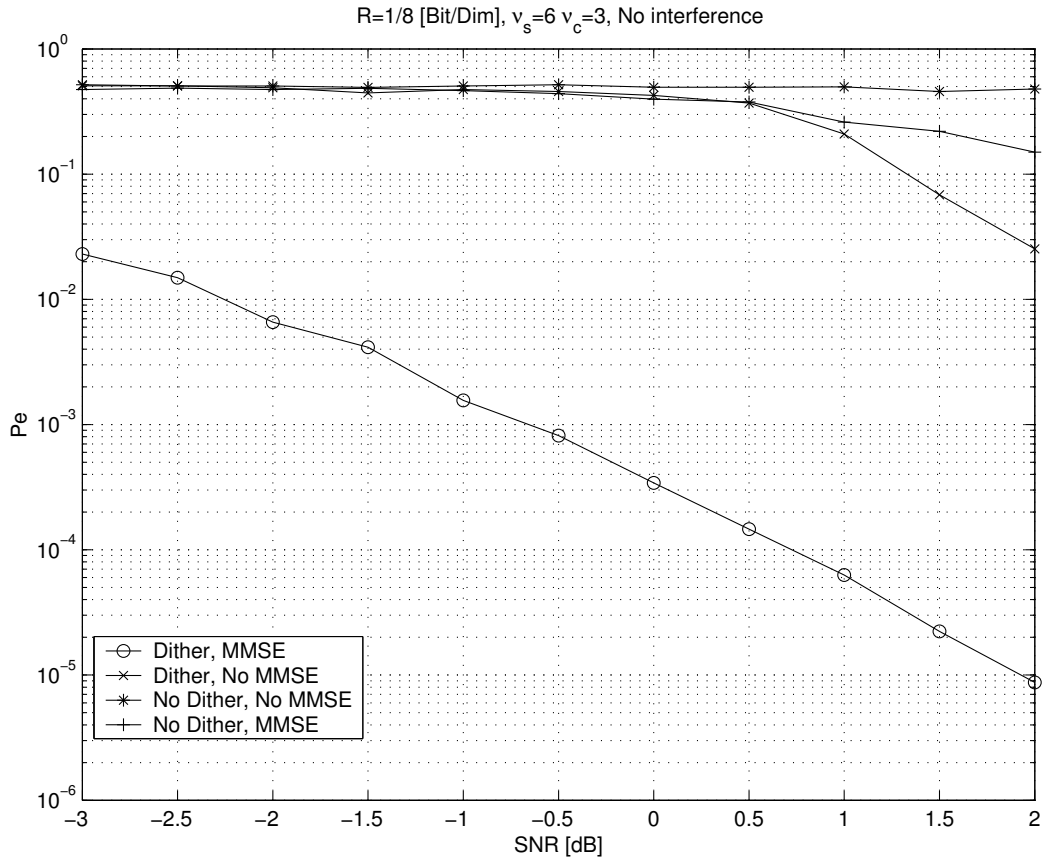

Figure 4.13: MMSE estimation and dither effect

The importance of the MMSE estimation is not doubtable, since we are at the low SNR regime, as we use MMSE-DFE instead of ZF-DFE at low SNR. The dither together with the modulo operation is used to cope with arbitrary interference. According to Figure 4.13 we need to use both the MMSE estimation and the dither, in order to have the modulo additive equivalent channel and to

minimized the effective noise. In high SNR we expect that MMSE estimation is not critical.

Achievable Rates

The shaping gain showed by the convolutional $\mathcal{C}_s$ code is quite reasonable, although convolutional code is not the optimal code for fine code. Therefore we loss rate of using convolutional code as fine code, or equivalently we do not operate near the capacity and we have a gap of from the capacity. To show the achievable rate of our scheme in this configuration referred to the capacity. We use some convolutional codes (low complexity codes) with $\nu_s = 6$, $\nu_c = 6$ and Rates $\frac{1}{2}, \frac{1}{4}, \frac{1}{6}, \frac{1}{8}, \frac{1}{12}$ bit/Dim, for $P_e = 1 \cdot 10^{-3}$, as shown in Figure 4.14.

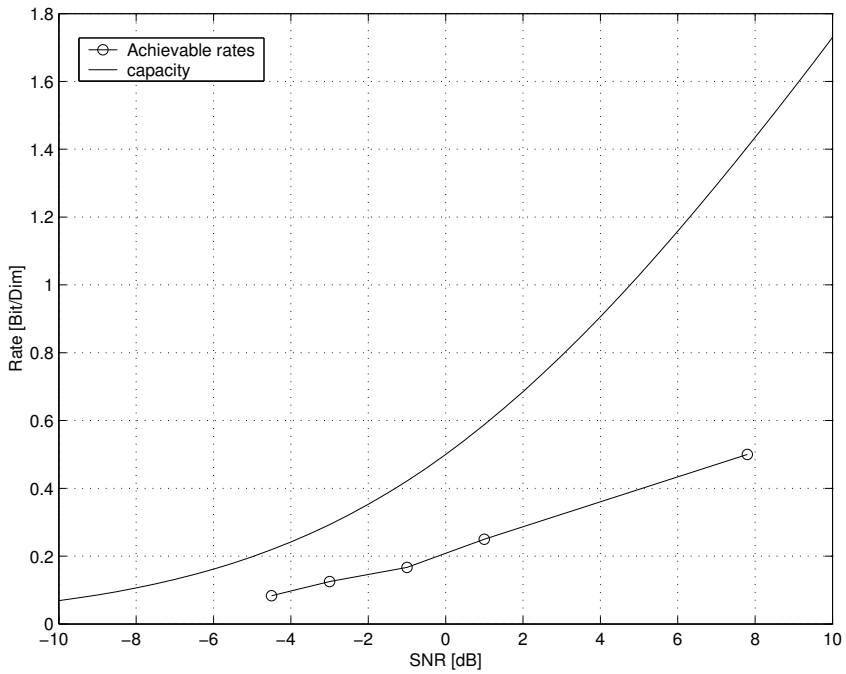

Figure 4.14: Achievable rates at $P_e = 1 \cdot 10^{-3}$

The gap from the capacity is about $4 - 4.5$ dB at around $SNR = 0$ dB.

Using powerful fine code that operates at about $3-4\ dB$ below the capacity in interference-free AWGN channel, such as Turbo codes or LDPC codes, the capacity of the interference channel can be approached. The above results are also drawn with respect to E_b/N_0, the 'x' sign at Figure 2.8 where the $4-4.5\ dB$ gap from the capacity is better shown.

Numerical Results - Summary

The following tables summarize the main results:

1. Shaping gain

The achievable shaping gains:

Shaping code	*Fine code*	*Shaping gain @* $P_e = 1 \cdot 10^{-3}$
$R_s = 1/2, \nu_s = 6$	$R_c = 1/6, \nu_c = 3$	$1.4\ dB$
$R_s = 1/2, \nu_s = 6$	$R_c = 1/4, \nu_c = 4$	$1.5\ dB$
$R_s = 1/2, \nu_s = 8$	$R_c = 1/4, \nu_c = 4$	$1.8\ dB$

2. The capacity gap

The gap from the capacity for several scenario at $P_e = 1 \cdot 10^{-3}$:

Shaping code rate	Fine code rate	Rate [bit/dim]	The gap @ $P_e = 1 \cdot 10^{-3}$
$R_s = 1/2$	no fine code	$1/2$	$7.8\ dB$
$R_s = 1/2$	$R_c = 1/2$	$1/4$	$4.6\ dB$
$R_s = 1/2$	$R_c = 1/3$	$1/6$	$4.4\ dB$
$R_s = 1/2$	$R_c = 1/4$	$1/8$	$4.3\ dB$
$R_s = 1/2$	$R_c = 1/6$	$1/12$	$4.1\ dB$

4.5 Pragmatic Decoder

The decoder can be simplified using sub optimal decoder, which decodes the shaping code and the fine code separately. Specifically, the modulo block out-

put is quantize by $\mathbb{Z}^N$ quantizer (hard decoding). The quantizer output is carried to syndrome estimation block which is performed by the shaping code parity check matrix H_s, where its output is $\hat{\mathbf{s}}_\mathbf{y}$, as shown in Figure 4.15. The $\hat{\mathbf{s}}_\mathbf{y}$ can have $\mathcal{C}_s$ syndromes which are not necessarily legitimate transmitted syndrome. Therefore, the fine code decoding search for the nearest legitimate syndrome. In case of nested trellises codes the fine code encoder is VA. Evidently, this sub-optimal decoder has inferior performance since we use the

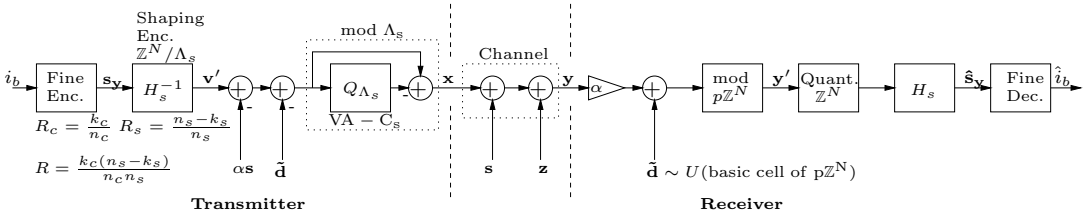

Figure 4.15: Pragmatic transmission scheme

hard decoding for $\mathcal{C}_s$ syndrome. However, any powerful codes can be incorporated as fine code, so that the total performance is improved. Turbo codes and LDPC codes, which are known as powerful codes at low SNR, can be adapted, in this way, as fine code to approach the interference channel capacity at low SNR.

Chapter 5

Summary

In this work we focused on transmission with known interference at the transmitter. Unlike for interference-free AWGN channel where the capacity can be achieved at low SNR without shaping, the interference channel does require shaping at low SNR.

The high dimensional scheme can be considered as transmission scheme for "writing on dirty paper", a problem introduced by Costa. Therefore, our scheme can be used for several communication scenarios, such as multi-terminal problems. For example, the non-degraded broadcast channel with multi antennas (MIMO channels), where single transmitter with multi antennas transmits to separate users.

The above scheme can incorporate different codes, especially Turbo codes or LDPC codes as a fine code with some decoder changes. Using these codes in our scheme enables to approach the interference channel capacity. Theoretically, for any good fine code we use, without a shaping code the capacity can not be achieved. Furthermore, at low SNR the precoding shaping-gap increases, therefore the shaping code requires specific consideration.

Further research:

• The incorporation of other powerful codes using a pragmatic decoder to improve the performance needs more detailed consideration.

• The behavior of SNR_c (the critical SNR for time sharing) as a function of the coarse lattice dimension is an interesting theoretical issue for low SNR transmission.

• The optimum MMSE factor, α, which maximizes the mutual information at a *very* low SNR as function of the coarse lattice dimension is an open issue.

Appendix A

The convolutional codes which are used for the nested trellises scheme are from [20],[19]

A.1 Generation of $\tilde{G}(D)$ Matrix - Example

For shaping code $\mathcal{C}_s$, with $R_{\mathcal{C}_s} = 1/2$ and constrain length $\nu_s = 6$, and generator matrix is

$$G_s(D) = [\ 1 + D^2 + D^4 + D^5 \quad 1 + D + D^2 + D^3 + D^5\ ]$$

the parity check matrix is given by

$$H_s(D) = [\ 1 + D + D^2 + D^3 + D^5 \quad 1 + D^2 + D^4 + D^5\ ]$$

the pseudo right inverse parity matrix is

$$H_s^{-1}(D) = [\ D \quad 1 + D\ ]^T$$

For the fine code $\mathcal{C}_c$, with rate=1/4 and constraint length $\nu_c = 3$, and generator matrix

$$G_c(D) = [\ 1 + D^2 \quad 1 + D^2 \quad 1 + D + D^2 \quad 1 + D + D^2\ ]$$

From $\tilde{G}(D)$ construction Figure 4.8, we need to represent generator matrix $G_s(D)$ as four times of the $\mathcal{C}_s$ code, meaning as code with rate 4/8 with generator matrix $G'_s(D)$. $G_s(D)$ can be written as

$$G_s(D) = [\ 1\ \ 1\ ] + [\ 0\ \ 1\ ]D + [\ 1\ \ 1\ ]D^2 + [\ 0\ \ 1\ ]D^3 + [\ 1\ \ 0\ ]D^4 + [\ 1\ \ 1\ ]D^5$$

and in binary form

$$G_s = \left(\begin{array}{cccccccc|cccccccc|cc} & & & & & & & 1 & & & & & & & & D & & & D^2 & \cdots \\ 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & \| & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & \| & 0 & 0 & \cdots \\ 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & \| & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & \| & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & \| & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & \| & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & \| & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & \| & 1 & 1 & \cdots \\ & & & & & & & \vdots & & & & & & & & \vdots & & & & \ddots \end{array}\right)$$

To have $G'_s(D)$ we need to fold all the 4×8 groups in horizontal direction with the appropriate D^i factor, thus we have

$$G'_s(D) = \begin{pmatrix} 1+D & 1 & D & 1+D & 1 & 1 & 0 & 1 \\ 0 & D & 1+D & 1 & D & 1+D & 1 & 1 \\ D & D & 0 & D & 1+D & 1 & D & 1+D \\ D^2 & D+D^2 & D & D & 0 & D & 1+D & 1 \end{pmatrix}$$

H_s^{-1} also has to represent as four times rate 1/2 encoder. In the same way we calculate $G'_s(D)$ we have that

$$(H_s'^{-1})(D) = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ D & D & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

The concatenation of the coset representative path in Figure 4.8 has the equivalent encoder

$$G_c(D)(H_s'^{-1})(D) =$$

$$= [\ D+D^2+D^3 \quad 1+D+D^3 \quad 1+D^2 \quad 0 \quad 1+D^2 \quad D \quad 1+D+D^2 \quad 0\ ]$$

Finally the $\tilde{G}(D)$ is the assembly

$$\tilde{G}(D) = \begin{pmatrix} G'_s(D) \\ G_c(D)(H'^{-1}_s)(D) \end{pmatrix} =$$

$$= \left(\begin{array}{cccc} 1+D & 1 & D & 1+D \\ 0 & D & 1+D & 1 \\ D & D & 0 & D \\ D^2 & D+D^2 & D & D \\ D+D^2+D^3 & 1+D+D^3 & 1+D^2 & 0 \end{array} \right.$$

$$\left. \begin{array}{cccc} 1 & 1 & 0 & 1 \\ D & 1+D & 1 & 1 \\ 1+D & 1 & D & 1+D \\ 0 & D & 1+D & 1 \\ 1+D^2 & D & 1+D+D^2 & 0 \end{array} \right)$$

This generator matrix, $\tilde{G}(D)$, will be used for the VA decoder, the effective constraint length for this decoder is $\nu_{dec} = 9$, i.e $2^{\nu_{dec}} = 512$ states in the trellis diagram. Generally, the pseudo right inverse matrix is not unique and in some case it has infinite memory. But for our use we can use delayed right inverse matrix so that $(H^{-1}(D))H(D) = D^i$ [21], which introduces decoding delay.

A.2 The Generator Matrix $\tilde{G}(D)$ for Simulated Codes

1. Shaping code ($\mathbf{R_{\mathcal{C}_s} = 1/2}$, $\mathbf{\nu_s = 8}$), fine code ($\mathbf{R_{\mathcal{C}_c} = 1/4}$, $\mathbf{\nu_c = 3}$)

• Shaping code - $R_{\mathcal{C}_s} = 1/2$ constraint length $\nu_s = 8$.

Generator matrix

$$G_s(D) = [\ 1 + D^3 + D^5 + D^7 \quad 1 + D + D^2 + D^3 + D^5 + D^6 + D^7\ ].$$

Parity check matrix

$$H_s(D) = [\ 1 + D + D^2 + D^3 + D^5 + D^6 + D^7 \quad 1 + D^3 + D^5 + D^7\ ].$$

pseudo right inverse parity check matrix

$$H_s^{-1}(D) = [\ 1 + D^3 + D^4 \quad D + D^2 + D^4\ ]^T.$$

• Fine code - $R_{\mathcal{C}_c} = 1/4$ constraint length $\nu_c = 3$.

Generator matrix

$$G_c(D) = [\ 1 + D^2 \quad 1 + D^2 \quad 1 + D + D^2 \quad 1 + D + D^2]\ ].$$

• $G(\tilde{D})$ matrix -

$$\left(\begin{array}{cccc} 1 & 1 & D & 1+D \\ D+D^2 & D+D^2 & 1 & 1 \\ 0 & D+D^2 & D+D^2 & D+D^2 \\ D^2 & D+D^2 & 0 & D+D^2 \\ 1+D^2 & D+D^3 & 1 & 1 \end{array}\right.$$

$$\left.\begin{array}{cccc} 0 & 1+D & 1+D & 1+D \\ D & 1+D & 0 & 1+D \\ 1 & 1 & D & 1+D \\ 1 & 1 & D & 1+D \\ 1+D+D^2 & D+D^2+D^3 & D^2+D^3 & D^2+D^3 \end{array}\right)$$

The constraint length is $\nu_{dec} = 11$ and the rate is $R = 1/8$ bit/dim.

2. Shaping code($\mathbf{R_{\mathcal{C}_s} = 1/2}$, $\mathbf{\nu_s = 6}$), fine code ($\mathbf{R_{\mathcal{C}_c} = 1/2}$, $\mathbf{\nu_c = 6}$)

• Shaping code - $R_{\mathcal{C}_s} = 1/2$ constraint length $\nu_s = 6$.

Generator matrix

$$G_s(D) = [\ 1 + D^2 + D^4 + D^5 \quad 1 + D + D^2 + D^3 + D^5\ ].$$

Parity check matrix

$$H_s(D) = [\ 1 + D + D^2 + D^3 + D^5 \quad 1 + D^2 + D^4 + D^5\ ].$$

pseudo right inverse parity check matrix

$$H_s^{-1}(D) = [\ D \quad 1 + D\ ]^T.$$

• Fine code - $R_{\mathcal{C}_c} = 1/2$ constraint length $\nu_c = 6$.

Generator matrix

$$G_c(D) = [\ 1 + D^2 + D^4 + D^5 \quad 1 + D + D^2 + D^3 + D^5\ ].$$

• $G(\tilde{D})$ matrix -

$$\begin{pmatrix} 1 + D + D^2 & 1 + D \\ D^3 & D + D^2 + D^3 \\ D + D^2 + D^3 + D^4 + D^6 & 1 + D + D^3 + D^5 + D^6 \end{pmatrix}$$

$$\begin{pmatrix} D^2 & 1 + D + D^2 \\ 1 + D + D^2 & 1 + D \\ 1 + D^2 + D^4 + D^5 & D + D^3 + D^4 \end{pmatrix}$$

The constraint length is $\nu_{dec} = 12$ and the rate is $R = 1/4$ bit/dim.

3. Shaping code ($\mathbf{R_{\mathcal{C}_s} = 1/2}$, $\mathbf{\nu_s = 8}$), fine code ($\mathbf{R_{\mathcal{C}_c} = 1/3}$, $\mathbf{\nu_c = 6}$)

• Shaping code - $R_{\mathcal{C}_s} = 1/2$ constraint length $\nu_s = 6$.

Generator matrix

$$G_s(D) = [\ 1 + D^2 + D^4 + D^5 \quad 1 + D + D^2 + D^3 + D^5\ ].$$

Parity check matrix

$$H_s(D) = [\ 1 + D + D^2 + D^3 + D^5 \quad 1 + D^2 + D^4 + D^5\ ].$$

pseudo right inverse parity check matrix

$$H_s^{-1}(D) = [\ D \quad 1 + D\ ]^T.$$

• Fine code - $R_{\mathcal{C}_c} = 1/3$ constraint length $\nu_c = 6$.

Generator matrix

$$G_c(D) = [\ 1 + D^3 + D^4 + D^5 \quad 1 + D^2 + D^4 + D^5 \quad 1 + D + D^2 + D^3 + D^5\ ].$$

• $G(\tilde{D})$ matrix -

$$\left(\begin{array}{ccc} 1 & 1+D & D \\ D+D^2 & D+D^2 & 1 \\ D^2 & D & D+D^2 \\ D+D^2+D^3+D^4+D^6 & 1+D+D^2+D^5+D^6 & 1+D^3+D^4+D^5 \end{array}\right.$$
$$\left.\begin{array}{ccc} 1 & 1+D & 1+D \\ 1+D & D & 1 \\ D+D^2 & 1 & 1+D \\ D+D^3 & 1+D^2+D^4+D^5 & D+D^3+D^4 \end{array}\right)$$

The constraint length is $\nu_{dec} = 12$ and the rate is $R = 1/6$ bit/dim.

4. Shaping code ($\mathbf{R_{\mathcal{C}_s} = 1/2}$, $\mathbf{\nu_s = 8}$), fine code ($\mathbf{R_{\mathcal{C}_c} = 1/4}$, $\mathbf{\nu_c = 5}$)

• Shaping code - $R_{\mathcal{C}_s} = 1/2$ constraint length $\nu_s = 6$.

Generator matrix

$$G_s(D) = [\ 1+D^2+D^4+D^5 \quad 1+D+D^2+D^3+D^5\ ].$$

Parity check matrix

$$H_s(D) = [\ 1+D+D^2+D^3+D^5 \quad 1+D^2+D^4+D^5\ ].$$

pseudo right inverse parity check matrix

$$H_s^{-1}(D) = [\ D \quad 1+D\ ]^T.$$

• Fine code - $R_{\mathcal{C}_c} = 1/4$ constraint length $\nu_c = 5$.

Generator matrix

$$G_c(D) = [\ 1+D^2+D^4 \quad 1+D^2+D^3+D^4 \quad 1+D+D^3+D^4 \quad 1+D+D^2+D^3+D^4\ ].$$

• $G(\tilde{D})$ matrix -

$$\left(\begin{array}{cccc} 1+D & 1 & D & 1+D \\ 0 & D & 1+D & 1 \\ D & D & 0 & D \\ D^2 & D+D^2 & D & D \\ D+D^2+D^3+D^4+D^5 & 1+D+D^3+D^5 & 1+D^2+D^4 & D^3 \end{array}\right.$$

$$\left.\begin{array}{cccc} 1 & 1 & 0 & 1 \\ D & 1+D & 1 & 1 \\ 1+D & 1 & D & 1+D \\ 0 & D & 1+D & 1 \\ 1+D^2+D^3+D^4 & D+D^2 & 1+D+D^3+D^4 & D^2 \end{array}\right)$$

The constraint length is $\nu_{dec} = 12$ and the rate is $R = 1/8$ bit/dim.

Bibliography

[1] R. J. Barron, B. Chen, and G. W. Wornell. On the duality between information embedding and source coding with side information and some applications. In *Proceedings of International Symposium on Information Theory, Washington D.C.*, page 300, June 2001.

[2] G. Caire and S. Shamai (Shitz). On the achievable throughput of a multi-antenna gaussian broadcast channel. *IEEE Trans. Information Theory*, submitted. See also, *On achievable rates in multiple-antenna broadcast downlink*, *38 Allerton Conference on Communication, Control and Computing*, Allerton House, Monticello, Illinois, October 4–6, 2000, and the *Proceedings of the 2001 IEEE International Symposium on Information Theory*, Washington, D.C., USA,p. 147, June 24–29, 2001.

[3] J. M. Cioffi, G. P. Dudevoir, M. V. Eyuboglu, and G. D. Forney. MMSE decision-feedback equalizers and coding – Part I: Equalization results, Part II: Coding results. *IEEE Trans. Communications*, 43:2582–2604, Oct. 1995.

[4] J. H. Conway and N. J. A. Sloane. *Sphere Packings, Lattices and Groups.* Springer-Verlag, New York, N.Y., 1988.

[5] M.H.M. Costa. Writing on dirty paper. *IEEE Trans. Information Theory*, IT-29:439–441, May 1983.

[6] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Wiley, New York, 1991.

[7] U. Erez, S. Litsyn, and R. Zamir. Lattices which are good for (almost) everything. *IEEE Trans. Information Theory*, in preparation.

[8] U. Erez, S. Shamai (Shitz), and R. Zamir. Capacity and lattice-strategies for cancelling known interference. *IEEE Trans. Information Theory (submitted)*, pages Also presented at the ISITA 2000, Honolulu, Hawaii, pages 681–684 (Nov. 2000).

[9] U. Erez and R. Zamir. Lattice decoding can achieve $\frac{1}{2}\log(1+SNR)$ on the AWGN channel. *IEEE Trans. Information Theory*, submitted, 2001.

[10] M. V. Eyuboglu and G. D. Forney Jr. Trellis precoding:combined coding, precoding and shaping for intersymbol interference channels. *IEEE Trans. Information Theory*, IT-38:301–314, March, 1992.

[11] L. F. Wei G. D. Forney Jr. Multidimensional constellation - part I: Introduction, Figures of Merit, and Generalized Cross Constellations. *IEEE Journal of Selected Areas in Communications*, 7,6:877–892, Aug. 1989.

[12] R. G. Gallager. *Information Theory and Reliable Communication*. Wiley, New York, N.Y., 1968.

[13] S.I. Gelfand and M. S. Pinsker. Coding for channel with random parameters. *Problemy Pered. Inform. (Problems of Inform. Trans.)*, 9, No. 1:19–31, 1980.

[14] G. Ginis and J. M. Cioffi. Vectored-DMT: A FEXT canceling modulation scheme for coordinating users. In *ICC2001, Helsinki, Finland*, volume 1, pages 305–309, June 2001.

[15] H. Harashima and H. Miyakawa. Matched-transmission technique for channels with intersymbol interference. *IEEE Trans. Communications*, COM-20:774–780, Aug. 1972.

[16] G. D. Forney Jr. Multidimensional constellation - part II: Voronoi Constellations. *IEEE Journal of Selected Areas in Communications*, 7,6:941–958, Aug. 1989.

[17] G. D. Forney Jr. Trellis shaping. *IEEE Trans. Information Theory*, IT-38:281–300, Mar., 1992.

[18] H. A. Loeliger. Averaging bounds for lattices and linear codes. *IEEE Trans. Information Theory*, 43:1767–1773, Nov. 1997.

[19] P. Orten P. Frenger and T. Ottosson. Code-spread CDMA using low-rate convolutional codes. *Proc. IEEE International Symposium on Spread Spectrum Techniques and Applications, Sun City, South Africa*, pages 374–378, Sept., 1998.

[20] T. Ottosson P. Frenger, P. Orten and A. Svensson. Multi-rate convolutional codes. *Technical Report No. 21*, April, 1998.

[21] V. S. Pless and W. C. Huffman. *Handbook of Coding Theory Volume 1.* Elsevier Science B. V., Amsterdam, Netherlands., 1998.

[22] J. G. Proakis. *Digital Communications.* McGraw-Hill, New York, 1983.

[23] C. E. Shannon. Channels with side information at the transmitter. *IBM Journal of Research and Development*, 2:289–293, Oct. 1958.

[24] S. Shamai (Shitz), S. Verdú, and R. Zamir. Systematic lossy source/channel coding. *IEEE Trans. Information Theory*, IT-44:564–579, March 1998.

[25] M. Tomlinson. New automatic equalizer employing modulo arithmetic. *Electronic Lett.*, vol. 7:138–139, Mar. 1971.

[26] G. Ungerboeck. Channel coding with multilevel/pase signals. *IEEE Trans. Information Theory*, IT-28:55–67, Jan. 1982.

[27] L. Wei. Trellis-coded modulation with multidimensional constellations. *IEEE Trans. Information Theory*, IT-33:483–501, 1987.

[28] W. Yu and J. M. Cioffi. Sum capacity of gaussian vector broadcast channels. *IEEE Trans. Information Theory*, Submitted, 2001.

[29] R. Zamir and M. Feder. On lattice quantization noise. *IEEE Trans. Information Theory*, pages 1152–1159, July 1996.

[30] R. Zamir and S. Shamai (Shitz). Nested linear / lattice codes for Wyner-Ziv encoding. In *Proc. of Info. Th. Workshop, Killarney, Ireland*, pages 92–93, Nov. 1998.

[31] R. Zamir, S. Shamai (Shitz), and U. Erez. Nested linear/lattice codes for structured multiterminal binning. *IEEE Trans. Information Theory*, IT-48:1250–1276, June 2002.