TRELLIS SHAPING

A Project Report

submitted by

N PRANITHA

in partial fulfilment of the requirements
for the award of the degree of

BACHELOR OF TECHNOLOGY

DEPARTMENT OF ELECTRICAL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY MADRAS.

May 2013

THESIS CERTIFICATE

This is to certify that the thesis titled **Trellis Shaping**, submitted by **N Pranitha**, to the Indian Institute of Technology, Madras, for the award of the degree of **Bachelor of Technology**, is a bona fide record of the research work done by her under our supervision. The contents of this thesis, in full or in parts, have not been submitted to any other Institute or University for the award of any degree or diploma.

Dr. Radhakrishna Ganti
Research Guide
Assistant Professor
Dept. of Electrical Engineering
IIT-Madras, 600 036

Place: Chennai

Date: 08th May 2013

ACKNOWLEDGEMENTS

I would like to take this opprtunity to thank my parents and family for thier support and encouragement all the time in whatever I choose to do.

I would like to thank Dr.Radhakrishna Ganti and Dr.Andrew Thangaraj for thier continual support and invaluble advice throughout the duration of the project.

I would also like to thank all Professors at IITMadras who taught me and imparted thier knowledge.

I am very thankfull to all my friends in Institute for making my stay enjoyable and helping me and being there for me. Its a big list.

I would also like to thank my labmates Arjun, Ajay, Varsha, Sudharsan, Gopal, Anjan, Debayani, Abhishek for giving very good company and creating a lively environment in lab and making it easy to work there and for helping me.

ABSTRACT

KEYWORDS: Shaping, Trellis codes, Viterbi algorithm, Constillations

Trellis shaping is a method of selecting a minimum weight sequence from an equivalent class of possible transmitted sequences by a search through the trellis diagram of shaping convolutional code C.Shaping gains of the order of 1dB may be obtained with simple 4-state shaping codes.The shaping gains obtained obtained with more complicated codes approach the ultimate shaping gain of 1.53 dB. With a feedback syndrome former for C, transmitted data can be recovered without catastrophic error propagation.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS **i**

ABSTRACT **ii**

LIST OF TABLES **iv**

LIST OF FIGURES **v**

1 INTRODUCTION **vi**

1.1 Shaping Gain vi

1.2 Genrator matrix, Parity check matrix, Syndrome sequence viii

1.3 Sign bit Shaping ix

1.4 Data Recovery using Syndrome Former x

LIST OF TABLES

1.1 Bit Mapping ix

LIST OF FIGURES

1.1 Sign bit shaping system supporting R=7 bits per symbol,using the 16x16 constellation and rate 1/2 convolutional code xiv

1.2 256-point constellation xiv

1.3 Constellation after sign bit shaping xiv

1.4 Distribution of x co-ordinates of the constellation points after sign bit shaping(Guassian) xv

1.5 Distribution of y co-ordinates of the constellation points after sign bit shaping(Gaussian) xv

1.6 Distribution of constellation points after sign bit shaping xv

1.7 SER and BER vs SNR xvi

CHAPTER 1

INTRODUCTION

Typical digital communication systems uses M-Quadrature Amplitude Modulation(QAM) to communicate through an analog channel (specifically a channel with Gaussian noise). For Higher bit rates(M) the minimum Signal to Noise ratio (SNR) required by a QAM system with Error Correcting Codes is about 1.53 dB higher than minimum SNR required by a Gaussian source(>30 percent more transmitter power) as given in Shannon Hartley theorem (Forney, 1992)

$$C = B \log_2 \left(1 + \frac{S}{N} \right)$$

where

C is the channel capacity in bits per second; B is the bandwidth of the channel in hertz; S is the total signal power over the bandwidth and N is the total noise power over the bandwidth. S/N is the signal-to-noise ratio of the communication signal to the Gaussian noise interference expressed as a straight power ratio (not as decibels).

This 1.53 dB difference is called the shaping gap. Typically digital system will encode bits with uniform probability to maximize the entropy. Shaping code act as buffer between digital sources and modulator communication system. They will receive uniformly distributed data and convert it to Gaussian like distribution before presenting to the modulator. Shaping codes are helpful in reducing transmit power and thus reduce the cost of Power amplifier and the interference caused to other users in the vicinity.

1.1 Shaping Gain

The coding gain may be viewed as the sum of two components: a coding gain due to underlying code and a shaping gain γ_s due to the choice of a particular constellation to support a finite number of bits per symbol. At high data rates coding gain and shaping gain are almost completely seperable and additive. The coding gain depends only on the shape of the constellation. The design and implementation of coding and shaping can be almost decoupled almost completely and their contributions to performance, complexity, and other code characteristics such as constellation expansion are completely independent. With conventional block constillations, every point in the constellation is equally likely. The objective of shaping is to achieve a non-uniform Guassian like distribution over a somewhat expanded constellation, so as to reduce the average signal power at the same data rate. The power reduction is called shaping gain. For simplicity, shaping is performed on a square two dimensional constellation with a binary rate-1/2 convolutional code. Trellis shaping then takes a particular simple form, which will be called "sign bit shaping". Sign bit shaping is a special case of trellis shaping. The most significant parameters of signal constellation are more generally of a shaping scheme are its datarate R in bits per two dimensions and it average energy S_x per two dimensions, relative to a scale factor such as minimum squared distance d^2_{min} between signal points. For a conventional square $M \times M$ 2D constellation such as 16×16 constillation, if $d^2_{min} = 1$, these parameters are related by

$$R = \log_2 M^2$$

$$S_x = (M^2 - 1)/6 = (2^R - 1)/6$$

For any arbitrary constellation or shaping scheme with $d^2_{min} = 1$, the data rate R and average energy S_x, the baseline average energy may be defined as

$$S_{bx}(R) = 2^R/6$$

The shaping gain is defined as

$$\gamma_s = S_{bx}(R)/S_x$$

1.2 Genrator matrix, Parity check matrix, Syndrome sequence

In coding theory, a generator matrix is a basis for a linear code, generating all possible code words. If the matrix is G and the linear code is C,

$$w = \mathbf{cG}$$

where w is the codeword of the linear code C, $\mathbf{c}$ is a row vector, and a bijection exists between w and $\mathbf{c}$. A generator matrix for a (n, k, d) code has dimensions $k \times n$. Here n is the length of the codeword, k is the number of information bits, d is the minimum distance of the code.The number of redundant bits is denoted by $r = n - k$.

The systematic form of a generator matrix is

$$G = [I_k|P]$$

where I_k is a $k \times k$ identity matrix and P is of dimension $k \times r$.

The parity check matrix for a given code can be derived from its generator matrix. For a generator matrix for an $[n, k]$-code is in standard form

$$G = [I_k|P]$$

then parity check matrix is given by

$$H = [-P^T|I_(n - k)]$$

$$GH^T = P - P = 0$$

For any (row) vector x of the ambient vector space, $s = Hx^t$ is called the syndrome of x. The vector x is a codeword if and only if $s = 0$.

1.3 Sign bit Shaping

Sign bit shaping starts with an $M \times M$ square constillation such as the 16×16 constillation shown in figure.If this constillation is scaled so that the minimum distance $d^2_{min} = 1$, then each coordinate takes on values from the 16-point PAM constillation $\{\pm 1/2, \pm 3/2, \pm 5/2, \pm 7/2, \pm 9/2, \pm 11/2, \pm 13/2, \pm 15/2\}$. The 4-bit representation is "$zabc$" where the most significant bit, z, is the sign bit. The remaining bits, abc,will be called least significant bits.

The following bit mapping is used to get a guassian distribution after shaping.

Table 1.1: Bit Mapping

coordinate value	bits
-15/2	1000
-13/2	1001
-11/2	1011
-9/2	1010
-7/2	1110
-5/2	1111
-3/2	1101
-1/2	1100
1/2	0000
3/2	0001
5/2	0011
7/2	0010
9/2	0110
11/2	0111
13/2	0101
15/2	0100

A four state rate-1/2 convolutional code C is generated using a generator matrix

$$G = [D^8 + D^5 + D^4 + D^2 + D + 1, D^8 + D^7 + D^4 + D^2 + 1]$$

where addition is modulo 2. Sign bit shaping with shaping code C can be performed as follows. Let $\alpha = \{\alpha_j\}$ be any sequence of two-dimensional points $\alpha_j = (\alpha_{1j}, \alpha_{2j})$ from the constellation, and let $z_i = \{z_i\}$ be the sequence of sign bits of α, the sequence of binary 2-tuples $z_j = (z_{1j}, z_{2j})$. The shaping operation is allowed to modify α by changing the sign bit sequence z to $z' = z \bigoplus y$, where y is any code sequence in C, and '$\bigoplus$' denotes mod-2 addition. The modification is done so as to minimize the average energy $S_x = E[||\alpha'_j||^2]$ of the modified signal points α'_j. The reduction in average energy is achieved by a viterbi algorithm search through a trellis diagram for the shaping convolutional code C, using appropriate branch metrics.

A trellis diagram for C depicts the state transition diagram of an encoder for C, with its sates σ_j at time j to σ_{j+1} at time $j+1$ corresponds to a branch in the trellis, and is labeled by the corresponding encoder output (y_{1j}, y_{2j}).The set of all paths through the trellis diagram corresponding to set of all code sequences $y \epsilon C$.

Given an original constellation point $\alpha_j = (\alpha_{1j}, \alpha_{2j})$, if $y_j = (y_{1j}, y_{2j})$ is the binary 2-tuple selected at time j in the selected code sequence $y \epsilon C$, then the modified output α'_j will be α_j with sign bits modified by y_j. Therefore, given α, we assign to any branch whose label is y_j is a branch metric $||\alpha'_j||^2$ that is equal to the Euclidean weight of the corresponding modified output α'_j. Then a search for the minimum-weight path through trellis of C is precisely equivalent to a search for $y \epsilon C$ that results in a modified sequence α' of least average energy.

1.4 Data Recovery using Syndrome Former

The modified α' is sent through a noisy channel as usual. At the receiver, if α and thus α' are uncoded, then the estimated sequence α'' can be obtained by convolutional symbol-by-symbol hard decisions. Thus shaping does not increase the complexity of transmission or detection in uncoded systems, apart from possible constellation expansion.

With or without coding, it may be assumed that the estimate α' is correct most of the time, but is subject to occasional errors. From α'', the receiver can recover an estimate z'' of the modified sign bit sequence, as well as the unmodified less significant bits. No matter what $y \epsilon C$ is choosen at the transmitter, the modified sign bit sequence $z' = z \bigoplus y$ is an element of the set $C \bigoplus z$, where z is the original sign bit sequence. Algebraically, $C \bigoplus z$ is a coset of the group code C. If z' is recovered correctly, the receiver can at least determine this coset. This may be done by passing the estimate z'' through a syndrome-former for C. In general, a syndrome-former for a rate-k/n binary linear convolutional code C with $k \times n$ generator matrix G is an n-input, $(n-k)$-output linear sequential circuit specified by an $n \times (n-k)$ transfer function matrix H^T such that

$$GH^T = 0$$

Consequently, if $y = xG$ is any code sequence in C, then $yH^T = 0$. More generally if z is any sequence of binary n-tuples, then the syndrome sequence s corresponding to z is

$$s = zH^T$$

If z is not in C, then $s \neq 0$. If $z' = z \bigoplus y$ is an element of the coset $C \bigoplus z$ of C, then z and z' have the same syndrome:

$$z'H^T = (z \bigoplus y)H^T = zH^T = s$$

conversely, if z and z' are in different cosets of C, then they have different syndromes.

It is known that every time-invarient binary linear convolutional code C has a linear, time-invarient, feedbackfree syndrome-former H^T with the same number of state as a minimal encoder G for C. For example, the code C, whose generator matrix is

$$G = [D^8 + D^5 + D^4 + D^2 + D + 1, D^8 + D^7 + D^4 + D^2 + 1]$$

, has a minimal feedbackfree syndrome-former which is specified by the polynomial matrix

$$H^T = [D^8 + D^7 + D^4 + D^2 + 1, D^8 + D^5 + D^4 + D^2 + D + 1]^T$$

It follows that, whichever $y \epsilon C$ is chosen, the receiver can recover from $z' = z \bigoplus y$ the syndrome sequence s that is associated with the initial sign bit sequence z with a syndrome-former H^T for C, since

$$z'H^T = (z \bigoplus y)H^T = zH^T = s$$

provided that $z'' = z'$. However, even if $z'' \neq z'$, occasional errors in the estimate z'' will cause only limited error propagation in the estimated syndrome sequence s', since H^T can always be chosen to be feedbackfree.

In summary, regardless of what modification is made during shaping to the original sign bit sequence z, we can recover one bit of useful information per symbol with limited error propagation, namely the syndrome bit s_j at time j in the syndrome sequence $s = zH^T$. It is, therefore, desirable to let the syndrome sequence s be part of the input data to the transmitter, and to generate from s an initial sign bit sequence z that lies in the coset of C whose syndrome is s. In general, given a syndrome-former specified by an $n \times (n-k)$ matrix H^T, any $(n-k) \times n$ left inverse $(H^{-1})^T$ for H^T may be used as a coset representative generator. From the syndrome sequence s , the transmitter generates the coset representative sequence

$$z = s(H^{-1})^T$$

then z is a sign bit sequence with syndrome s, since

$$zH^T = s(H^{-1})^T H^T = s$$

The inverse for

$$H^T = [D^8 + D^7 + D^4 + D^2 + 1, D^8 + D^5 + D^4 + D^2 + D + 1]^T$$

is

$$(H^{-1})^T = [D^7 + D^6 + D^2, D^7 + D^5 + D^4 + D^2 + D + 1]$$

In summary, sign bit shaping using rate-$1/2$ code C and the 256-point constellation may be at a data rate of $R = 7$ bits per two dimensions. Of the input data, 6 bits per symbol are used to select the less significant bits, for instance by choosing one of the 64 points in the first quadrant. (These bits may include coded bits, chosen so that the initial sequence is a sequence in a channel code C). One bit per symbol is considered to be a syndrome bit s_j. The syndrome sequence s is the input to a 1-input, 2-output coset representative generator circuit specified by $(H^{-1})^T$. The output sequence $z = s(H^{-1})^T$, a sequence of binary tuples z_j, is taken as the initial sign bit sequence. A 256-state VA decoder for C then determines the sign bit sequence z' in the coset $C \bigoplus z$ such that the modified sequence α' has minimum average energy S_x per symbol. At the receiver, a conventional symbol-by-symbol detector (or decoder for C) generates an estimated sequence α''. The less significant bits of α''_j determine the six corresponding input bits in the usual way, while the sign bit 2-tuples z''_j are passed through a feedbackfree syndrome-former H^T to produce an estimate $s' = z''H^T$ of the original syndrome sequence s.

1.5 Plots

The following are the plots when 4000 symbols i.e, 4000*7 bits are transmitted.

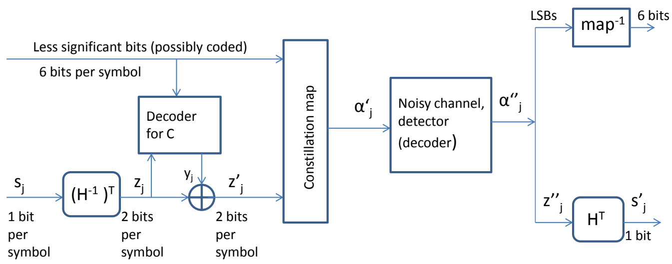

Figure 1.1: Sign bit shaping system supporting R=7 bits per symbol,using the 16x16 constellation and rate 1/2 convolutional code

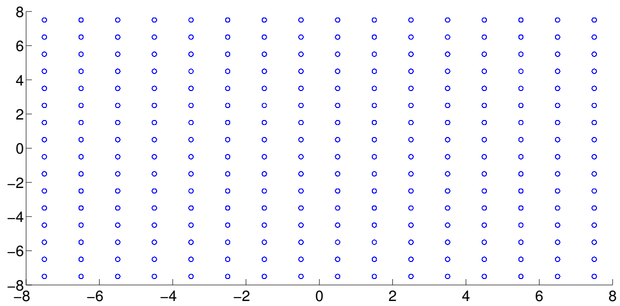

Figure 1.2: 256-point constellation

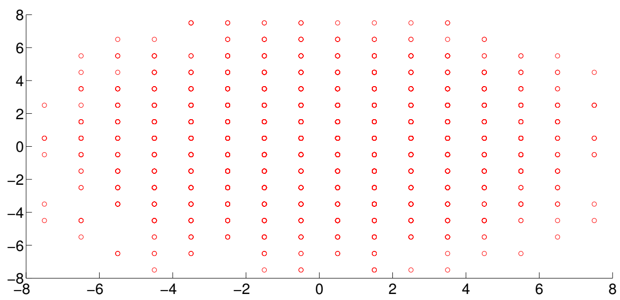

Figure 1.3: Constellation after sign bit shaping

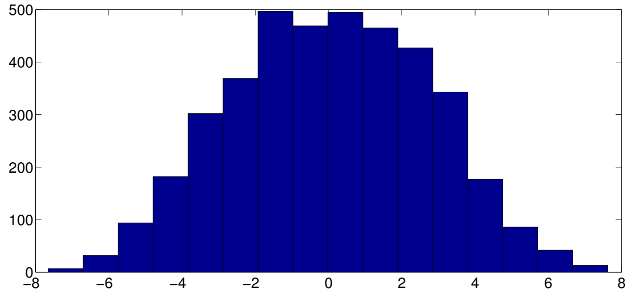

Figure 1.4: Distribution of x co-ordinates of the constellation points after sign bit shaping(Guassian)

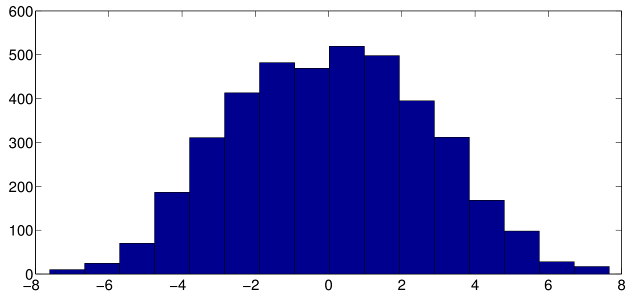

Figure 1.5: Distribution of y co-ordinates of the constellation points after sign bit shaping(Gaussian)

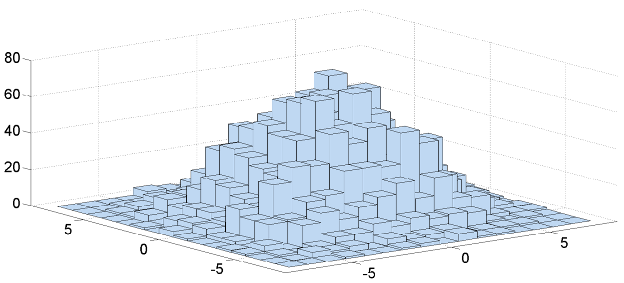

Figure 1.6: Distribution of constellation points after sign bit shaping

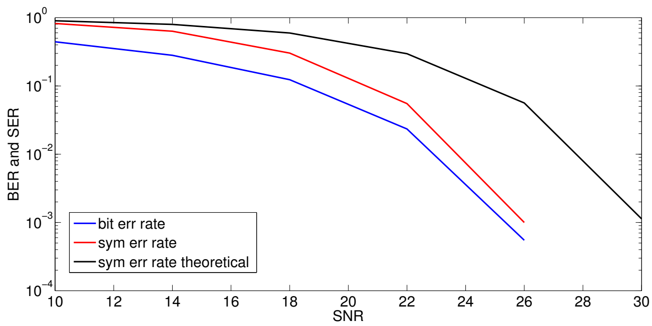

Figure 1.7: SER and BER vs SNR

REFERENCES

1. **Forney, J., G.D.** (1992). Trellis shaping. *Information Theory, IEEE Transactions on*, **38**(2), 281–300. ISSN 0018-9448.