Modulation for Interference Avoidance on the AWGN Channel

JINFENG DU

KTH Electrical Engineering

Acknowledgments

The work presented in this thesis was conducted at the Communication Theory group in the Department of Signals, Sensors and Systems (S3) of the Royal Institute of Technology (KTH) during the autumn and winter 2005–2006. Many people who have contributed to this thesis in one way or another deserve my special thanks.

Firstly I would like to thank my advisor Prof. Erik G. Larsson. Your insightful guidance, consistent help, and timely encouragement through the whole period enable me to finish this project. Every discussion with you has been of tremendous help and pleasure. And great thanks to Prof. Mikael Skoglund, your deliberate instruction and advise make me manage to keep on the right direction.

I would also like to thank the Communication Theory group for awarding me a stipend for Master's degree project which enables me to concentrate on my thesis. And many thanks to the Ph.D students Lei Bao, Xi Zhang, and Thanh Tùng Kim, all the lunch breaks and discussions with you have been pleasant time for me.

Big thanks to all the Master students in Room B512 for all the time together. The good atmosphere inside makes it a great place to work. Special thanks to Adrian Schumacher for your kindly help in C and LaTeX.

My parents deserve a special acknowledgment here. Your endless love and support accompanies me all the way. Great thanks with all my heart.

Abstract

Theoretic results have shown that the capacity of a channel does not decrease if the receiver observes the transmitted signal in the presence of interference, provided that the transmitter knows this interference non-causally. That is, if the transmitter has non-causal access to the interference, by using proper precoding this interference could be "avoided" (as if it were not present) under the same transmit power constraint. It indicates that lossless (in the sense of capacity) precoding is theoretically possible at any signal-to-noise-ratio (SNR). This is of special interest in digital watermarking, transmission for ISI channels as well as for MIMO broadcast channels. Recent research has elegantly demonstrated the (near) achievability of this "existence-type" result, while the complexity is notable. An interesting question is what one can do when very little extra complexity is permitted. This thesis treats such special cases of this problem in order to shed some light on this question. In the AWGN channel with additive interference, an optimum modulator is designed under the constraint of a binary signaling alphabet with binary interference. Tomlinson-Harashima precoding (THP), which is originally proposed for ISI channels, is improved by picking up optimized parameters and then taken as a benchmark. Simulation results show that the Optimum Modulator always outperforms the THP with optimized parameters. The difference in performance, in terms of mutual information between channel input and output as well as coded bit error rate with Turbo codes, is significant in many scenarios.

Contents

List of Figures

Chapter 1

Introduction

1.1 Background

Recently the problem of canceling known interference in noisy channels with channel state information (CSI) at the transmitter has attracted significant interest. A classic information theoretic result given by Costa [1], "dirty paper" coding (DPC), states that the capacity of a channel from A to B does not decrease if B observes the signal from A embedded in interference, provided A knows this interference non-causally. This result indicates that lossless (in the sense of capacity) precoding is theoretically possible at any signal-to-noise-ratio (SNR). Recent research [2, 3] shows that DPC can serve as a building block in architecture for both inter-symbol interference (ISI) channels and for the downlink multiuser multiple-input multiple-output (MIMO) channel. This is an "existence-type" result, but it indicates that one can find efficient downlink precoding methods. Consequently, much research has been devoted to find "practical" such methods.

1.2 Previous Works

Perhaps the simplest existing (but suboptimal) method for DPC is zero-forcing (ZF) [4], which does linear channel inversion at the transmitter. ZF precoding, however, increases the transmit power and suffers from the same serious disadvantages as ZF receivers. In the case of additive interference, ZF simply subtracts the interference and hence increases the power consumption.

Tomlinson [5] and Harashima [6] invented a nonlinear precoding method, the so called Tomlinson-Harashima precoding (THP), which introduces a modulo operation after the subtraction of the known interference in order

to maintain the transmit power constraint. Wesel and Cioffi [7] investigated the capacity loss of THP for uniformly distributed transmitting signals, given the channel impulse response, the transmit power constraint and the AWGN noise variance. Liu and Krzymień [8] applied THP to the downlink of multiple antenna multi-user systems and derived an improved THP based on a "best-first" ordering of the rows of the channel matrix. An auxiliary feedback filter was introduced by Smee and Schwartz [9] in the receiver in order to cooperate with the feedforward filter so that adaptive compensation for THP can be done in case of channel or interference variations. It also showed that the error propagation and the transient increases in mean-squared error can be avoided by adaptively updating the precoder. Liavas [10] examined the performance of THP in time-varying frequency-selective channels and proposed a robust THP suitable for low SNR scenarios with partial channel knowledge.

Another strategy for achieving capacity is known (the constructive proof is precisely described in [1]; see also [11]): First quantize the interference into a number of bins (this is essentially a source coding problem). Then, depending on which bin the interference falls into, choose an appropriate code for the encoding of the source signal. The best DPC results presented in references [12–14], which have elegantly demonstrated (near) achievability of the DPC limit, are based on this approach. However their complexity is notable.

1.3   Project Purpose and Goal

It is natural to ask what one can do about the DPC problem when permitted to add no, or very little, extra complexity to the system compared to "classical" transmission. The goal of this thesis is to shed some light on this question. More precisely, we consider the design of an optimal one-dimensional scheme which does modulation based on a source signal and the knowledge of interference so that the interference could be "avoided" (as if the interference were not present) under the same transmit power constraint. Since it is impossible to achieve the capacity with finite-dimensional modulation, we focus on the investigation of what the best one can achieve in one dimension, with low complexity.

In order to see how well the optimal scheme could perform, a good benchmark has to be introduced. The no interference cancellation scheme is acceptable but not good enough. Tomlinson-Harashima precoding seems to be

a good choice in this one-dimensional scenario, while its heuristic parameters harm the performance. A straightforward but very time-consuming way is to find the optimized parameters for THP based on an exhaustive search. This Optimal THP (THP with optimized parameters), together with the no interference case, will serve as benchmarks for our new scheme.

1.4   Outline

This thesis is divided into chapters as follows:

Chapter 2 investigates the performance of Tomlinson-Harashima precoding (THP) for one-dimensional DPC both with heuristic and optimized parameters. Both mutual information and bit error rate are used to evaluate its performance.

Chapter 3 presents an optimum modulator for a special case of the one-dimensional DPC problem: a binary signal through an AWGN channel with BPSK interference known to the transmitter. Simplification of this Optimum Modulator is also discussed.

Parts of the work in Chapter 2 and Chapter 3 resulted in a published conference paper:
[22] Jinfeng Du, Erik G. Larsson, and Mikael Skoglund, "Costa precoding in one dimension," in Proc. of *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Toulouse, France, May 2006, to appear.

Chapter 4 examines the coding gains achieved by the Optimal THP and the Optimum Modulator schemes based on the coded bit error rate when they are combined with Turbo codes.

Chapter 5 presents a summary of our work and points out some possible directions in which this one-dimensional Optimum Modulator could be extended.

1.5 Notation

Throughout this thesis the following notational conventions are used:

x	lowercase letters denote random variables.
x_i	The ith realization of the random variable x.
$P_x(\cdot)$	probability of a discrete random variable x.
$f_x(\cdot)$	probability density function of a continuous random variable x.
$F_x(\cdot)$	cumulative density function of a continuous random variable x.
$P(x \leq t)$	probability of a random variable x so that $x \leq t$.
$P(x\|y)$	conditional probability of a random variable x given y.
$\mathrm{E}[x]$	The expected value of random variable x.
$x \bmod \Lambda$	modulator operator so that the result falls inside the interval $[-\frac{\Lambda}{2}, \frac{\Lambda}{2}]$.
$\lfloor x \rfloor$	the largest integer that is smaller than x.
$\delta(\cdot)$	the Dirac function.
$n!$	n-factorial, a quantity defined as $n! = \prod_{i=1}^{n} i$.
$\log(\cdot)$	the log operator.
$\ln(\cdot)$	the log operator with base e.
$\log_2(\cdot)$	the log operator with base 2.
$\log_{10}(\cdot)$	the log operator with base 10.

1.6 Acronyms

Abbreviations used in this thesis are listed below:

- AWGN Additive White Gaussian Noise

- BER Bit Error Rate

- BPSK Binary Phase Shift Keying

- CDF Cumulative Density Function

- CSI Channel State Information

- DPC Dirty Paper Coding

- IC Interference Cancellation

- INR Interference to Noise Ratio

- ISI Inter Symbol Interference

- MAP Maximum a Posteriori

- MIMO Multiple-Input Multiple-Output

- ML Maximum Likelihood

- M-PAM M-ary Pulse Amplitude Modulation

- PDF Probability Density Function

- RSC Recursive Systematic Convolutional

- SIR Signal to Interference Ratio

- SNR Signal to Noise Ratio

- THP Tomlinson-Harashima Precoding

Chapter 2

Tomlinson-Harashima Precoding

2.1   Introduction

In this chapter, the properties of scalar THP under a power constraint is investigated in detail. The performance of THP both with heuristic and with optimized parameters is compared in terms of Mutual Information between the transmitter and the receiver. The optimal detector is derived to maximize the performance. Also efforts are spent on improving the computational efficiency of the Optimal THP method.

2.2   System Model

Figure 2.1 illustrates a general system with additive interference and AWGN noise. Interference is non-causally known at the transmitter and its probability density function (PDF) is known at the receiver.

The channel output is given by

$$y(t) = x(t) + z(t) + n(t) \tag{2.1}$$

where $x(t),\ y(t),\ z(t),\ n(t)$ denote the transmitted signal, the received signal, the interference and the AWGN noise at time instant t respectively. In the following, the time index t will be suppressed for convenience. The transmitted signal x is the output of the encoder given the information signal w and the interference z, and therefore denoted by $x(w, z)$. The information signal w is a symbol from an M-ary Pulse Amplitude Modulation (M-PAM)
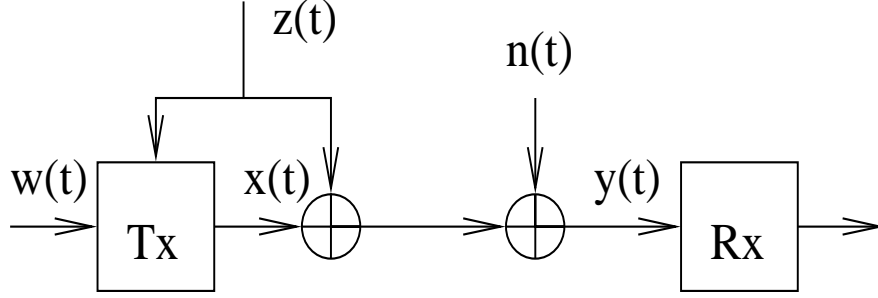
Figure 2.1: System configuration

constellation with uniform spacing Δ_w and points at

$$w_i = \left(i - \frac{M-1}{2}\right)\Delta_w, \ i = 0, 1, ..., M-1 \tag{2.2}$$

The interference z is either Gaussian or taken from a finite alphabet $\{z_n = (n - \frac{N-1}{2})\Delta_z, n = 0, 1, ..., N-1\}$, with the following distribution function

$$f_z(\tau) = \frac{1-\rho}{\sqrt{2\pi\sigma_z^2}}e^{-\frac{\tau^2}{2\sigma_z^2}} + \frac{\rho}{N}\sum_{n=0}^{N-1}\delta(\tau - z_n) \tag{2.3}$$

The CDF of z is

$$
\begin{aligned}
F_z(\tau) &= P(z \le \tau) = \int_{-\infty}^{\tau} f_z(x)dx \\
&= (1-\rho)(1-Q(\frac{\tau}{\sigma_z})) + \frac{\rho}{N}\begin{cases} 0 & \tau < -\frac{N-1}{2}\Delta_z; \\ \lfloor(\tau + \frac{N-1}{2}\Delta_z)/\Delta_z\rfloor + 1 & \text{otherwise;} \\ N & \tau \ge \frac{N-1}{2}\Delta_z. \end{cases}
\end{aligned} \tag{2.4}
$$

where $\rho \in [0, 1]$ is a relative weight which determines to what extent z is discrete or Gaussian.

Before introducing the strategy of Tomlinson-Harashima precoding, we first present some baseline schemes for this problem.

2.2.1 No Interference

If there is no interference, the best thing one can do is to transmit the signal w directly with all the available power (say, $w = \pm\sqrt{P}$ for BPSK modulated signal). The received signal is

$$y = x + n = w + n$$

Since w and n are independent, the conditional distribution function of y given w can be expressed as

$$f_y(y|w) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(y-w)^2}{2\sigma^2}} \tag{2.5}$$

where σ^2 denotes the variance of the AWGN noise n. The optimal receiver (in the minimum error-probability sense) is the one that maximizes the a posteriori probability of w when y is observed:

$$\widehat{w}_{\mathrm{MAP}} = \arg\max_w P(w|y) = \arg\max_w P_w(w)f_y(y|w) = \arg\max_w P_w(w)e^{-\frac{(y-w)^2}{2\sigma^2}}$$

2.2.2 No Interference Cancellation

If the transmitter does not know the interference but the receiver knows $f_z(\tau)$ (an assumption we do make throughout the thesis), transmitting the signal w directly with all the available power is not necessarily optimal. In our comparisons, we choose the value of P_w (subject to $P_w \leq P$) which maximizes performance. With $x = w$, the the received signal is

$$y = x + z + n = w + z + n$$

For given w, the distribution function of y can be derived from the convolution of $f_z(\tau)$ and $f_n(\tau)$ (z and n are independent)

$$f_y(y|w) = \frac{1-\rho}{\sqrt{2\pi(\sigma_z^2+\sigma^2)}}e^{-\frac{(y-w)^2}{2(\sigma_z^2+\sigma^2)}} + \frac{\rho}{N}\sum_{n=0}^{N-1}\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(y-w-z_n)^2}{2\sigma^2}} \tag{2.6}$$

The optimal receiver therefore is

$$\widehat{w}_{\mathrm{MAP}} = \arg\max_w P(w|y) = \arg\max_w P_w(w)f_y(y|w)$$

2.2.3 Interference Subtraction

If the transmitter knows the interference non-causally, the simplest way to avoid this interference is to do interference subtraction (also called ZF). That is, transmit $x = w - z$ instead of w. The received signal can be expressed as

$$y = x + z + n = (w - z) + z + n = w + n$$

which is identical with no interference. It seems like a good choice to avoid known interference at the first glance. However, this approach is not applicable unless $\mathrm{E}[x^2] = \mathrm{E}[w^2] + \mathrm{E}[z^2] \leq P$. In other words, when the power of the

interference $P_z = \mathrm{E}[z^2] \geq P - \mathrm{E}[w^2]$, this strategy becomes meaningless.

Is there anything we could do to maintain the performance of this "subtraction" strategy while largely reducing the transmit power? In early 1970s, Tomlinson [5] and Harashima [6] introduced a modulo operation after subtraction and formed the so called Tomlinson-Harashima Precoding method, which will be described in the following section.

2.3 Tomlinson-Harashima Precoding

The basic strategy of Tomlinson-Harashima precoding is to subtract the interference z from the source signal w, and then pass the resulting signal $(w - z)$ through a modulo operator. Given a real valued variable a, the modulo operation "mod Λ" outputs a new real valued variable b which falls into the region $[-\frac{\Lambda}{2}, \frac{\Lambda}{2}]$, where Λ is called the modulo range. After this modulo operation, the output signal $x = (w - z) \bmod \Lambda$ is transmitted through the noisy channel. The received signal y can be expressed by

$$y = x + z + n = (w - z) \bmod \Lambda + z + n \qquad (2.7)$$

where the modulo range Λ can be adjusted to achieve the best performance while maintaining the power constraint. The above equation can be rewritten as

$$\begin{aligned} y &= w + (w - z) \bmod \Lambda - (w - z) + n \\ &= w + \mathtt{k}\Lambda + n \end{aligned} \qquad (2.8)$$

where $\mathtt{k} = \frac{1}{\Lambda}((w - z) \bmod \Lambda - (w - z))$ is an integer with the following conditional distribution

$$\begin{aligned} P_{k|w} &= P(\mathtt{k} = k|w) = P((w - z) \bmod \Lambda - (w - z) = k\Lambda \,|w) \\ &= P(w - z \in [-(k + 1/2)\Lambda, -(k - 1/2)\Lambda] \,|w) \\ &= P(w + (k - 1/2)\Lambda \leq z \leq w + (k + 1/2)\Lambda \,|w) \\ &= F_z(w + (k + 1/2)\Lambda) - F_z(w + (k - 1/2)\Lambda) \end{aligned} \qquad (2.9)$$

where the last equality comes from the cumulative density function (CDF) of z, as shown in (2.4).

Since $\mathtt{k}\Lambda$ depends only on w, z and Λ, it is independent of noise n. According to the Bayesian rule, One can derive from (2.8) and (2.9) that

$$f_y(y|w) = \sum_{k=-\infty}^{\infty} f(y, k|w) = \sum_{k=-\infty}^{\infty} P_{k|w} f(y|w, k) = \sum_{k=-\infty}^{\infty} P_{k|w} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y - w - k\Lambda)^2}{2\sigma^2}} \quad (2.10)$$

10

where σ^2 denotes the variance of the AWGN noise n. This likelihood function can be directly used in the decoder or to calculate the mutual information, as one will see in Section 2.4. The optimal receiver is

$$\widehat{w}_{\mathrm{MAP}} = \arg \max_w P_w(w) \sum_{k=-\infty}^{\infty} P_{k|w} e^{-\frac{(y-w-k\Lambda)^2}{2\sigma^2}}$$

which differs from the heuristic (and suboptimal) detector that is given by (2.14) and usually used in papers dealing with THP.

When the interference z is known and the power of n is kept constant, the likelihood function $f_y(y|w)$ depends only on the alphabet of information signals w and the modulo range Λ. Most of the contributions related to THP link these two parameters as

$$\Lambda = \frac{3}{2} \cdot \Omega_w \tag{2.11}$$

where Ω_w is the constellation size of w. For M-PAM modulated signal w, we have

$$\Omega_w = (M - 1) \cdot \Delta_w$$

where Δ_w is the same as in (2.2). In the special case of 2-PAM (BPSK) modulated w, we have $\Omega_w = \Delta_w$. Below we show how to choose Δ_w and Λ to optimize performance. It turns out that optimal choice of (Δ_w, Λ) improves significantly over (2.11).

2.4 Performance Analysis

The performance of a communication link can either be evaluated in terms of the capacity or the bit error rate. When capacity is concerned, only the mutual information between the received signal y and the information signal w will be computed as an indicator of the capacity. When BER is computed, the optimal receiver will be used.

2.4.1 Mutual Information

According to information theory, the mutual information between y and w is

$$
\begin{aligned}
\mathbf{I}(y; w) &= \mathbf{H}(w) - \mathbf{H}(w|y) \\
&= \sum_{i=0}^{M-1} \int_{-\infty}^{\infty} P(y, w_i) \log P(w_i|y) dy - \sum_{i=0}^{M-1} P_w(w_i) \log P_w(w_i) \\
&= \sum_{i=0}^{M-1} \left[\int_{-\infty}^{\infty} f_y(y|w_i) P_w(w_i) \log \frac{f_y(y|w_i)P_w(w_i)}{f_y(y)} dy - P_w(w_i) \log P_w(w_i) \right] \\
&= \sum_{i=0}^{M-1} P_w(w_i) \int_{-\infty}^{\infty} f_y(y|w_i) \log \frac{f_y(y|w_i)}{f_y(y)} dy
\end{aligned}
\tag{2.12}
$$

where $f_y(y) = \sum_{j=0}^{M-1} f_y(y|w_j)P_w(w_j)$ and the last equality comes from the fact that $\int_{-\infty}^{\infty} f_y(y|w_i)dy = 1, \forall i$. The different $f_y(y|w)$ accompanied with different precoding schemes are shown in the following table:

Scheme	Tx(x)	Rx(y)	$f_y(y|w)$
No Interf	w	$w + n$	(2.5)
No IC	w	$w + z + n$	(2.6)
Subtract	$w - z$	$w + n$	(2.5)
THP	$(w - z) \bmod \Lambda$	$w + \mathbf{k}\Lambda + n$	(2.10)

2.4.2 Bit Error Rate

The optimal Maximum a Posteriori (MAP) receiver suitable for all strategies can be expressed as

$$\widehat{w}_{\text{MAP}} = \arg\max_w P(w|y) = \arg\max_w P_w(w)f_y(y|w) \tag{2.13}$$

For THP, there is also a suboptimal receiver which is given by

$$\widehat{w}_{\text{subopt}} = \arg\min_w \left[(y \bmod \Lambda) - w\right]^2 \tag{2.14}$$

2.5 Numerical Results

Both Mutual Information and Bit Error Rate are used as a measure of performance. If not especially mentioned, following assumptions are used in the simulations:

- The information bits are modulated by one dimensional M-PAM with equal probability $P_w = 1/M$ using Gray mapping;

- The interference z either comes from discrete symbols with equal probability or Gaussian symbols with variance σ_z^2. It is available at the transmitter and its distribution function is available at the receiver;

- A truncated sum (11 items, $k = -5, ..., 5$) is used for $f_y(y|w)$ in (2.10) and Monte-Carlo integration will be used when necessary.

When calculating the mutual information for different schemes, an exhaustive search has been used to achieve the maximum achievable mutual information for each scheme under the power constraint. For the No Interference scheme, we simply used all available power in transmission; for

other schemes, the optimal choice to achieve the maximum mutual information does not necessarily use all available power. Hence we use the transmit power "smartly" so that the system could get the best performance under the power constraint. As the Interference Cancellation (subtraction) is identical to the No Interference case except for the transmit power, it is not included in the simulations in Section 2.5.2 and Section 2.5.3 where the comparison is based on the same transmit power constraint.

Note that, strictly speaking, for a given power constraint P/σ^2, the actual SNR may be less than P/σ^2, because the optimal transmitter does not necessarily use all available power, as we mentioned earlier. Yet we refer to P/σ^2 as SNR because it facilitates a well-defined comparison with the No Interference case.

2.5.1 Investigation of THP

The performance of THP is compared with three other cases, No Interference, Interference Subtraction (Cancellation) precoding, and No Interference Cancellation. As the transmit power $P = \mathrm{E}[x^2]$ varies with different cases, adjustment of noise level has been made to ensure the same SNR level in each case.

Fig. 2.2 shows the mutual information for THP versus the modulo range Λ with SNR = 6dB together with the other three schemes. Fig. 2.3 shows the corresponding curves for the bit-error-rate with the optimal MAP detector described in (2.13). Both of them are carried out under 2-PAM modulated source signal with $\Delta_w = 2$ and equally distributed discrete interference z, which is 4-PAM modulated with $\Delta_z = 1$. Both of these two figures show that THP performs much better than the Interference Subtraction (Cancellation) scheme and the No Interference Cancellation scheme for some values of Λ. Three local minimum values of BER have been achieved at the points of modulo range 0.58, 0.79 and 1.37, which are the very points where three local maximum values of mutual information have been achieved at 0.859, 0.835 and 0.785 (bits) respectively, compared with 0.912 (bits) for the No Interference case and 0.68 (bits) for the Interference Subtraction.

Therefore in this particular case, it is possible to achieve quite good performance (comparable to the No Interference case) by properly choosing the parameters for THP. This motivates us to find the optimized parameters under a power constraint for THP and thus form the Optimal THP scheme, as described later in Section 2.5.2.
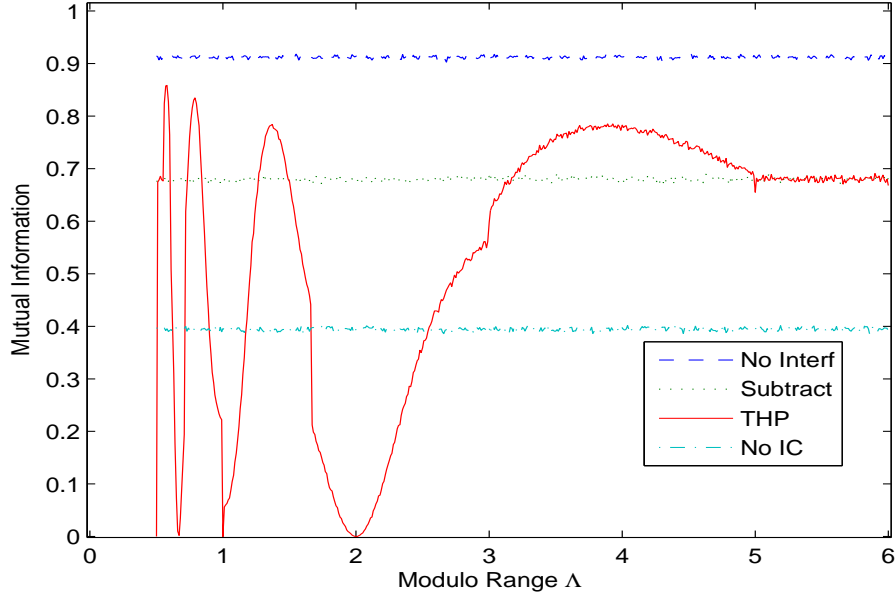
13

Figure 2.2: Mutual information vs. modulo range Λ with SNR=6dB

Those values of Λ where the THP achieves its local optimal performances, could be crudely identified from the minimum distance of the receive constellation normalized by the noise variance, as shown in Fig. 2.4 with

$$\mathrm{d}y = \min_{i \neq j} |y_i - y_j| \ \text{ where } \ y_i, y_j \in \{y|y = (w - z)\mathrm{mod}\Lambda + z\}$$

As the component with the minimum distance forms the dominating term of error probability, those with large normalized minimum distance will definitely achieve good performance. While since there are still other components that contribute to the error probability, the corresponding points in performance might vary slightly (cf. Fig. 2.4).

As indicated by Fig. 2.2 and Fig. 2.3, mutual information and bit error rate provide the same information about performance via different perspectives. In what follows we will mainly use the mutual information to measure the performance for different precoding methods.
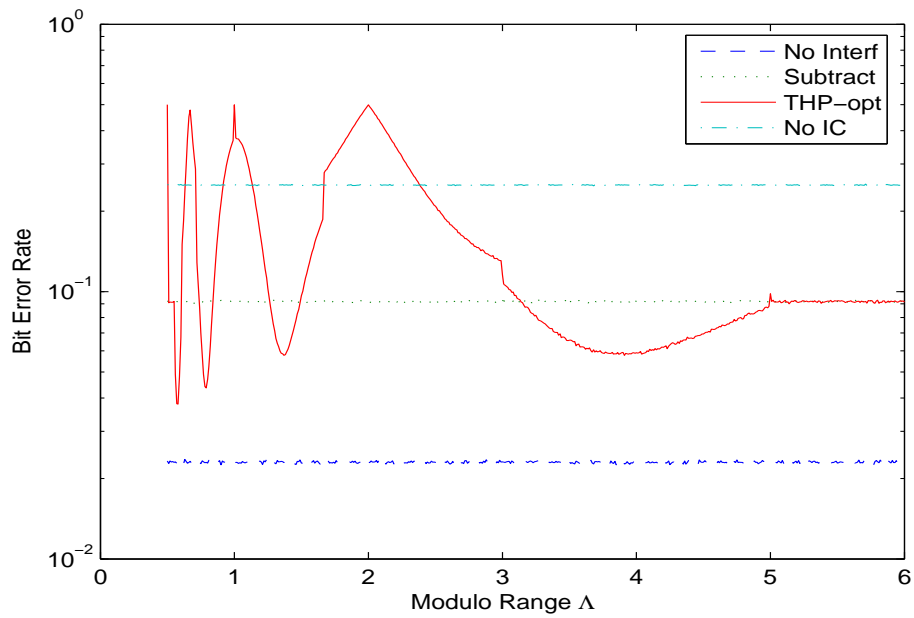
14

Figure 2.3: BER vs. modulo range Λ with fixed SNR=6dB
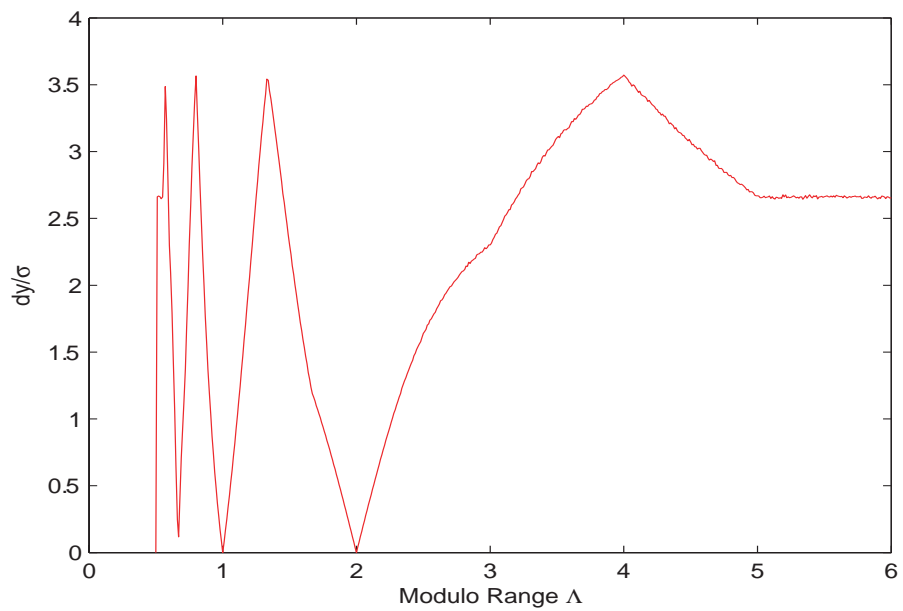

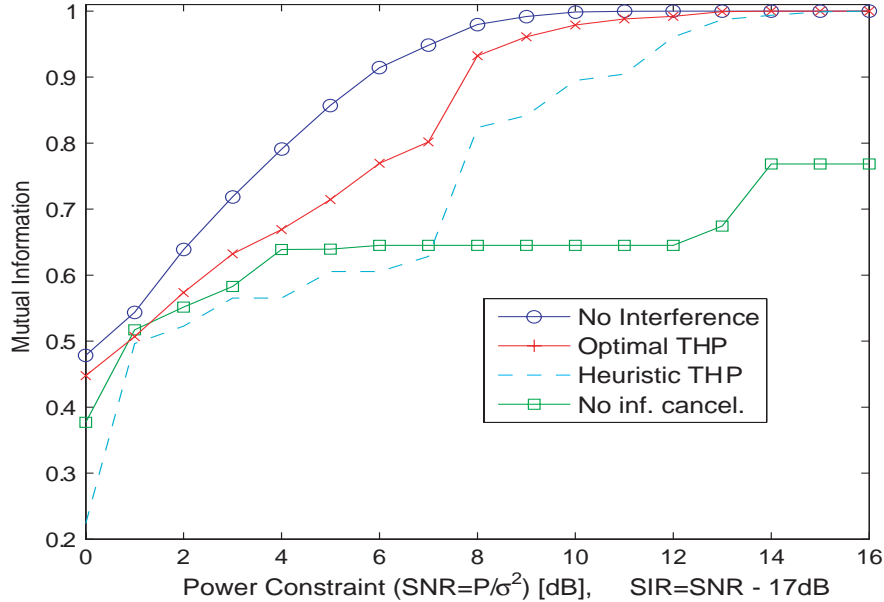Figure 2.4: Normalized minimum distance vs. modulo range Λ

Figure 2.5: Mutual information vs. power constraint, discrete interference

2.5.2 Comparison of Optimal THP and Heuristic THP

As mentioned in Section 2.3, the Heuristic THP is suboptimal. In the following simulations we investigate the performance of this Heuristic THP and an Optimal THP which uses optimized parameters by exhaustive search over all possible values of Δ_w and Λ under the power constraint.

Fig. 2.5 and 2.6 display the maximum mutual information each scheme can achieve under the same power constraint with known interference from a discrete 4-PAM constellation and Gaussian interference respectively. The source signal is 2-PAM modulated and the interference to noise ratio (INR) is 17dB. The optimal THP performs much better than the Heuristic THP both with discrete interference and with Gaussian interference.

Fig. 2.7 and 2.8 display the corresponding bit error rate each scheme achieves under the same power constraint. Here the detector used by the Heuristic THP is given by (2.14) and the detector used by the Optimal THP is given by (2.13). The suboptimal detector fails to work with some "unlucky" choices of Δ_w and Λ while the optimal detector works well regardless of Λ.
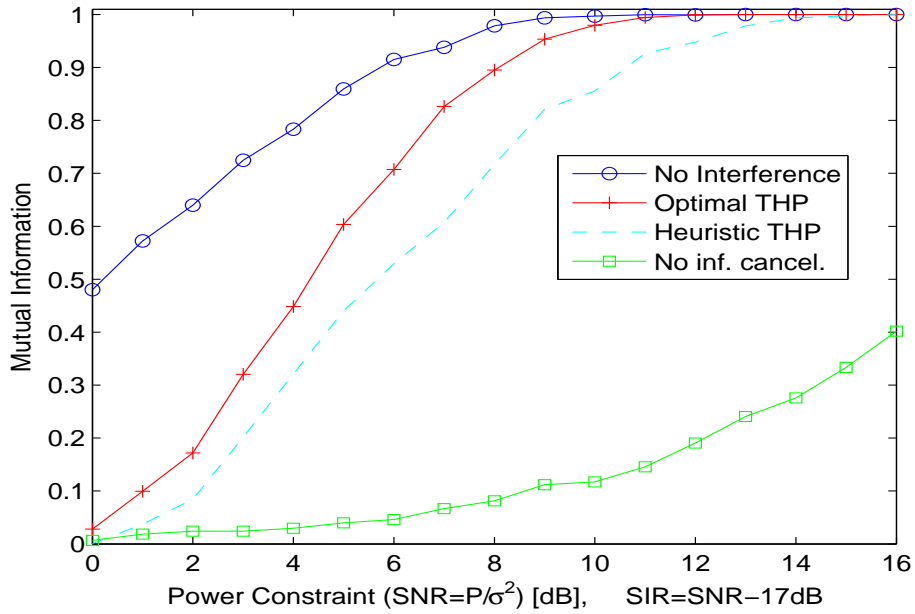
16

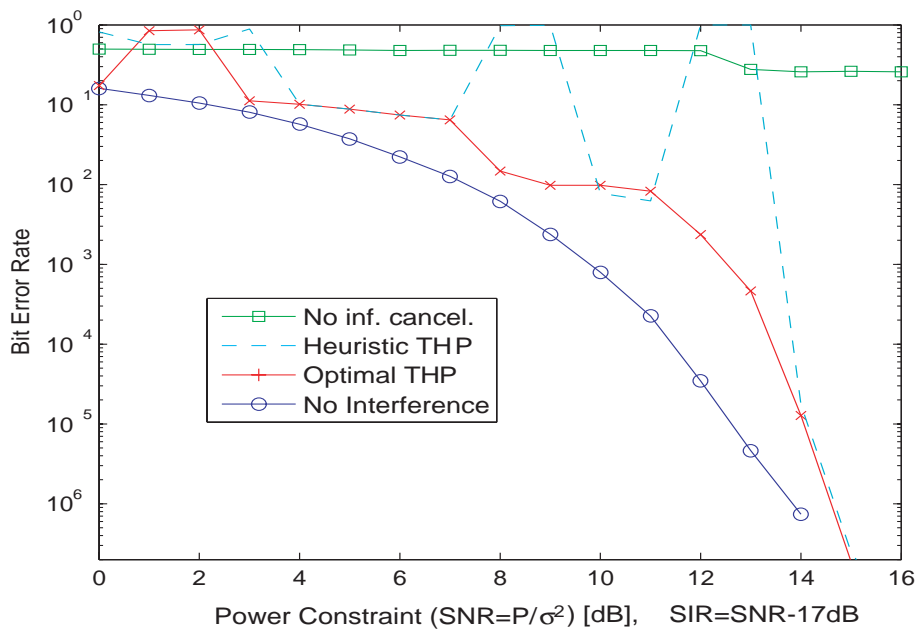Figure 2.6: Mutual information vs. power constraint, Gaussian interference

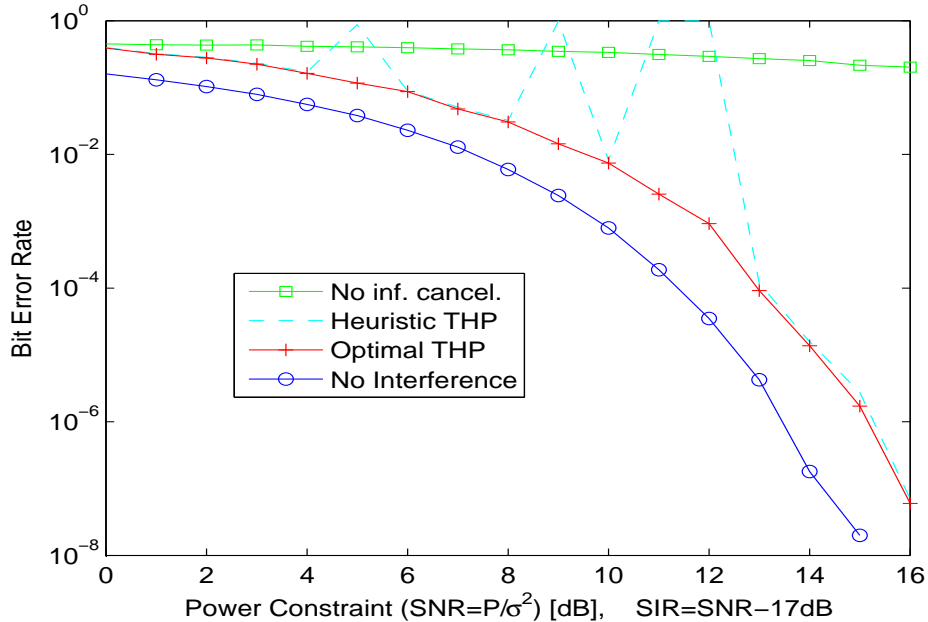Figure 2.7: Bit error rate vs. power constraint, discrete interference

Figure 2.8: Bit error rate vs. power constraint, Gaussian interference

2.5.3 Encoder Simplification

The Optimal THP works pretty well and seems more robust compared with the Heuristic THP, but finding the optimal (Δ_w, Λ) is time consuming. Is it possible to simplify the Optimal THP? In other words, is there any way to modify the Heuristic THP so that it can achieve almost the same performance as the Optimal THP?

The answer is affirmative. Let us turn to the idea when we come up with Optimal THP. We search over all possible values of Δ_w and Λ to find the best parameters. Is there any simple and straightforward relationship between these optimal parameters? First let us turn to the simple case of a 2-PAM modulated signal with 2-PAM modulated interference. Fig. 2.9 displays the relationship between the mutual information of THP and these parameters when the power of the interference is the same as the power of the noise. It is quite clear that, the line which denotes $\Lambda = \Delta_w - \Delta_z$ is where the optimal performance of THP would be achieved. (Strictly speaking, this rule does not work when $\Lambda < \Delta_z$, as shown in Fig. 2.9. Thus we choose $\Lambda = \Delta_z$ instead in such scenario.) This claim is further supported by Fig. 2.10, which shows the same relationship when the power of the interference is two times of the power of the noise.
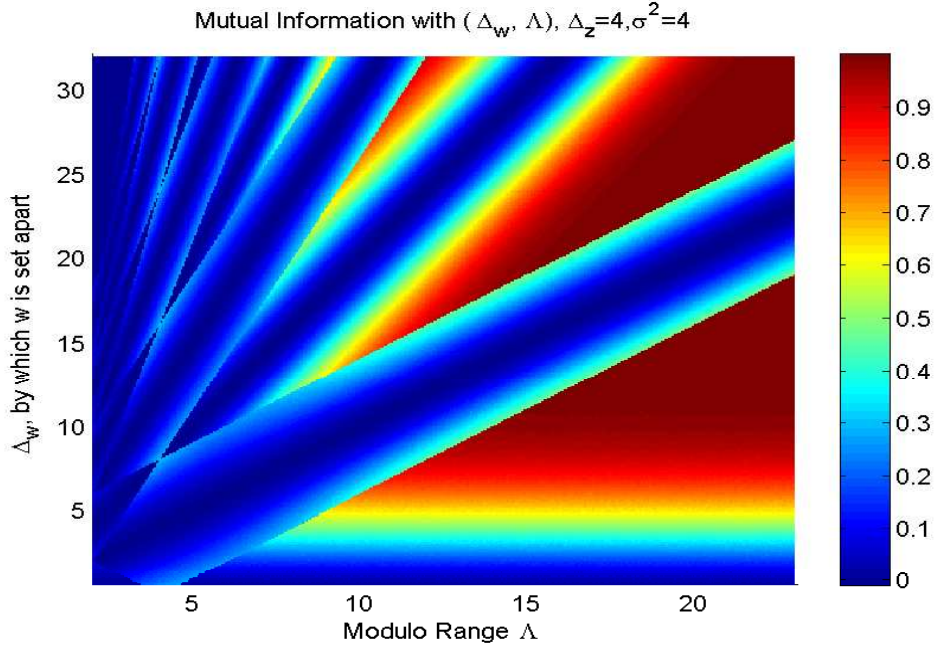
18

Figure 2.9: Mutual information vs. Δ_w and Λ, with SNR = SIR

Now it is quite simple to apply this rule to the Heuristic THP to achieve better performance without increasing the complexity. However the rule that $\Lambda = \Delta_w - \Delta_z$ is not exact at all, it is simply deduced from the figure and thus rather heuristic. Let us resort to simulation again.

Fig. 2.11 and Fig. 2.12 display the mutual information for all different schemes versus the power constraint in two different scenarios (INR = 3dB, 0dB respectively). The Heuristic THP uses the parameter $\Lambda = \Delta_w - \Delta_z$ (strictly speaking, $\Lambda = \max\{\Delta_w - \Delta_z, \Delta_z\}$), as mentioned earlier. It is the same for the following cases when using this rule, and hence we will not explain it in the following paragraphs. With this rule applied, the Heuristic THP performs almost as good as Optimal THP in the high SNR region when the power of the interference is two times as large as the power of noise (INR = 3dB), as shown in Fig. 2.11. In a different scenario, however, it does not work (Fig. 2.12 with INR = 0dB). Although slightly modification of this rule regarding of specific scenarios is possible to make it work, it is not easy find a general expression of this rule that is suitable for all scenarios.

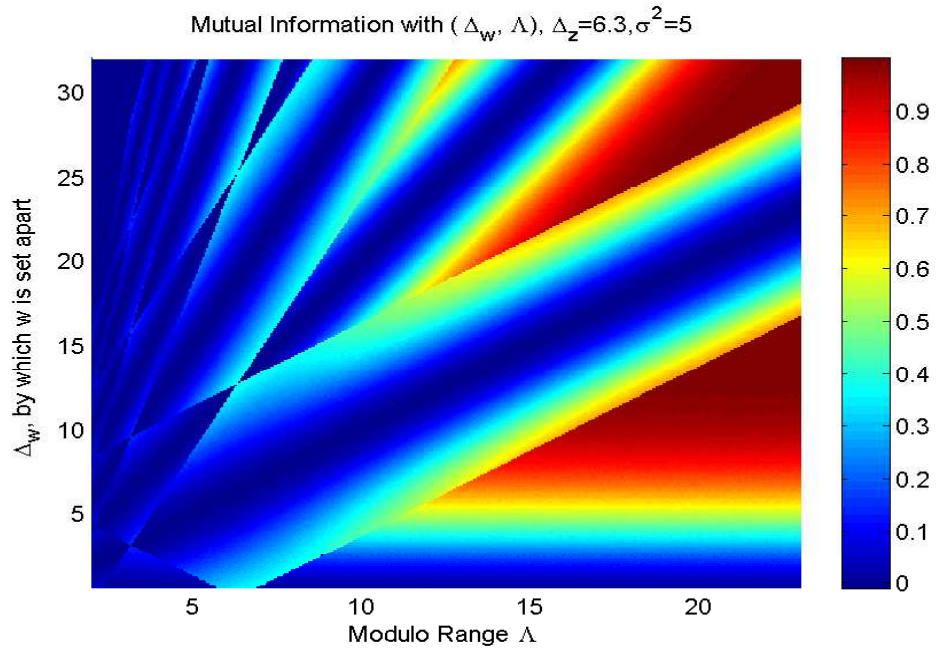In more complicated cases, say 2-PAM modulated signal with 4-PAM

Figure 2.10: Mutual information vs. Δ_w and Λ, with SNR = SIR + 3dB
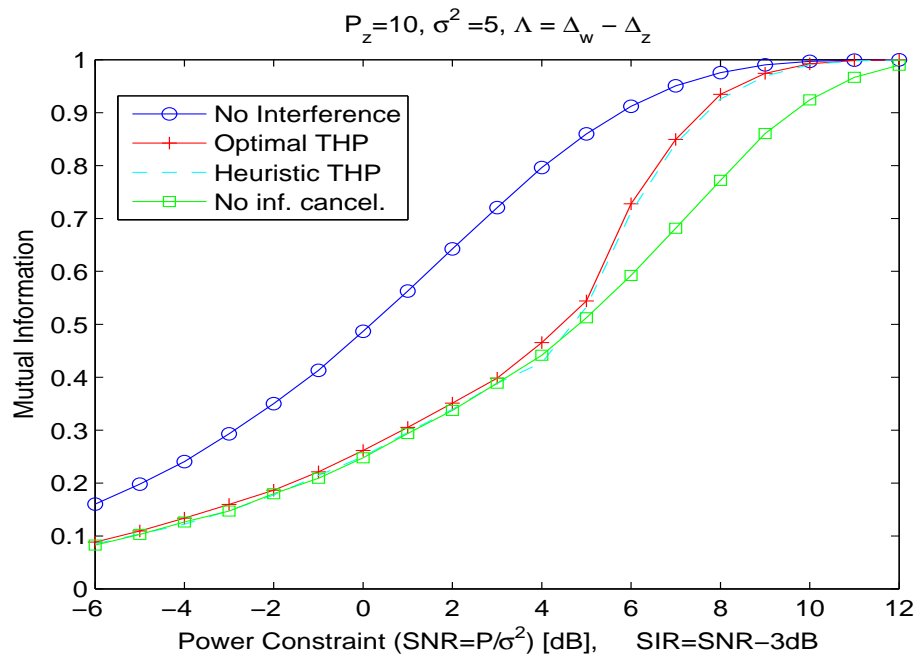

Figure 2.11: Mutual information vs. power constraint, with $\Lambda = \Delta_w - \Delta_z$
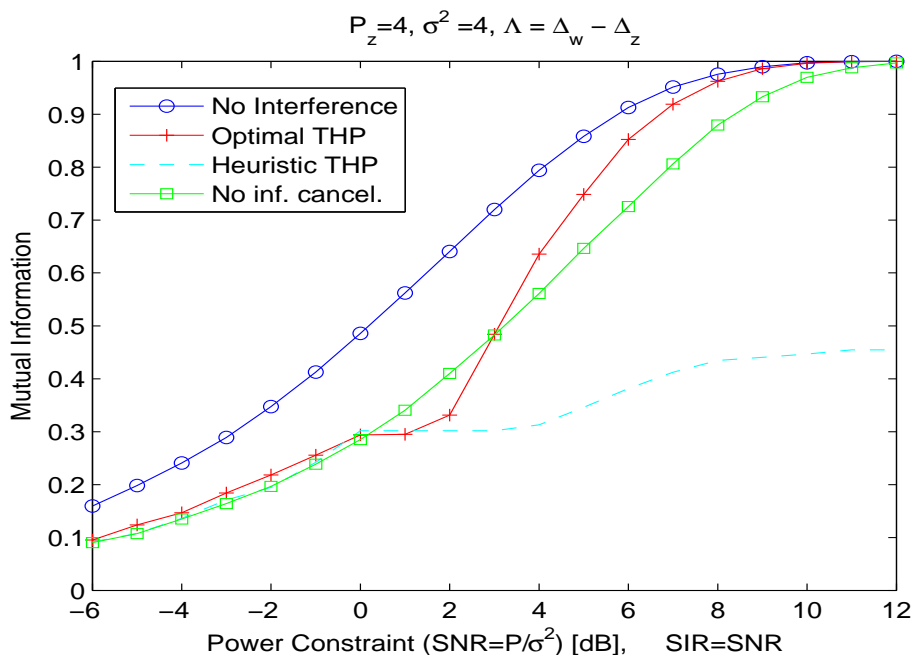
Figure 2.12: Mutual information vs. power constraint, with $\Lambda = \Delta_w - \Delta_z$

modulated interference, there also seems to be similar relationships between these optimal parameters Λ, Δ_w, and Δ_z, as shown in Fig. 2.13. This completes our discussion of THP.

2.6 Summary

Tomlinson-Harashima precoding could largely decrease the transmit power while maintaining the same communication quality, both in the mutual information sense and in terms of bit error rate. With optimized parameters, THP can eliminate most of the effects brought in by the known interference, regardless of whether it consists of discrete symbols or Gaussian components. The optimal MAP detector works better than the suboptimal detector.

Note that, all the deductions in this chapter are valid without any assumptions on the PDF of the information signal. And the distribution function for the interference could also be changed to a more general one. Such kind of changes may affect the results we have presented here while the way to investigate the performance remains the same.

However, there are also some inherent shortcomings in the construction
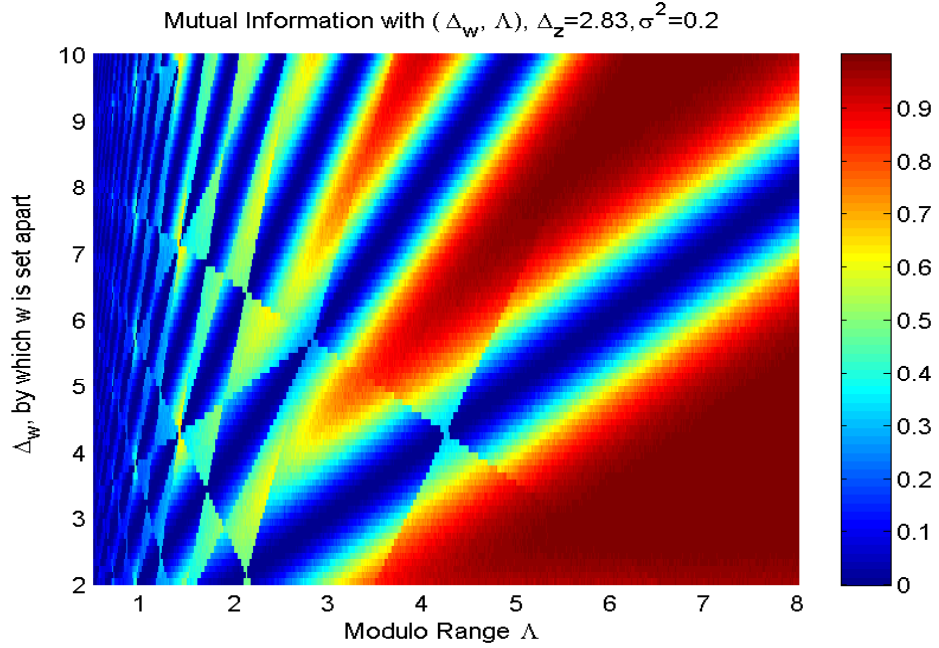
Figure 2.13: Mutual information vs. Δ_w and Λ, with SNR = SIR + 17dB

of THP. Subtraction of interference before the modulo operation reduces the freedom of arranging transmit constellation and hence may cause performance loss in some special cases (say, the power constraint range between 0dB and 3dB in Fig. 2.12). Additionally, the symmetry in the constellations of w and z as well as the modulo operation introduces extra constraints on the mapping of $x(w, z)$. For example, $x(w_1, z_1) = -x(w_0, z_0)$ and $x(w_1, z_0) = -x(w_0, w_1)$ always hold in case of a binary signaling alphabet with binary interference. This extra constraint is unnecessary from the point of view of modulation design. Could one find a better precoding (rather, modulation) scheme so that such shortcomings could be conquered? This motivation inspires us to examine what is the best one can do for this problem. It consequently results in an Optimum Modulator which will be displayed in the next chapter.

Chapter 3

Modulator Optimization

3.1 Introduction

In this chapter we turn to a more specific case of Dirty Paper coding, namely for BPSK signals with BPSK Interference. An Optimum Modulator is proposed to mitigate the interference and in some cases it shows a performance close to that of the No Interference system. It generally outperforms the Optimal Tomlinson-Harashima precoding(THP). Simulation results show that, the Optimum Modulator suffers at most 1.5dB loss against the No Interference system.

3.2 System Model

In the special case under study, both w and z are BPSK modulated with following distribution functions

$$f_w(\tau) = \alpha\delta(\tau - w_0) + (1-\alpha)\delta(\tau - w_1),\ 0 < \alpha < 1 \tag{3.1}$$

$$f_z(\tau) = \beta\delta(\tau - z_0) + (1-\beta)\delta(\tau - z_1),\ 0 < \beta < 1 \tag{3.2}$$

where

$$w_i = \left(i - \frac{1}{2}\right)\Delta_w,\ i = 0, 1$$

$$z_j = \left(j - \frac{1}{2}\right)\Delta_z,\ j = 0, 1$$

Then the CDF of z described in (2.4) could be rewritten as

$$F_z(\tau) = \begin{cases} 0 & \tau < -\frac{\Delta_z}{2}; \\ \beta & \text{otherwise}; \\ 1 & \tau \geq \frac{\Delta_z}{2}. \end{cases} \tag{3.3}$$

23

In the following discussions, both the signal w and interference z are supposed to be evenly distributed on their own alphabet ($\alpha = \beta = \frac{1}{2}$), which is a reasonable assumption in communication systems.

3.3 Optimum Modulator

3.3.1 Constellation Design

For different combinations of w and z, four different values of x are possible, as shown in Fig. 3.1.
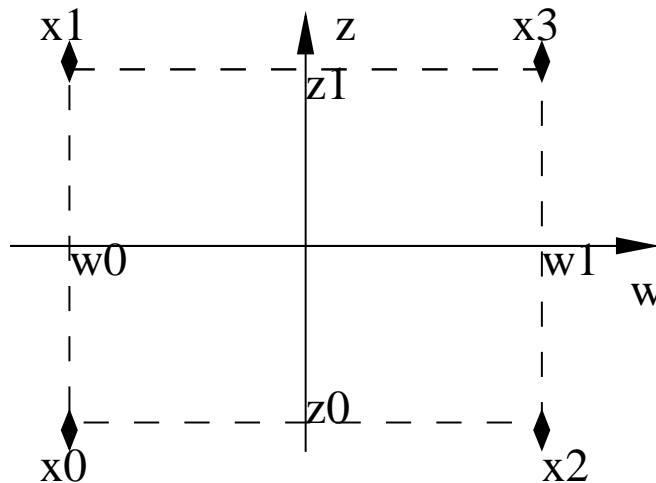
Figure 3.1: Constellation for optimum modulator

The symmetry in the BPSK constellations for w and z decreases the degree of freedom for x. We can assume that x comes from the following finite alphabet

$$x \in \{-b, -a, a, b\}, \text{ where } \mathrm{E}[x^2] = \frac{a^2 + b^2}{2} \leq P$$

for some positive constants a, b. With no constraint on the ordering of a and b, there are totally $4! = 24$ possibilities, of which 12 are redundant (because a and b are not ordered). The set of possible mappings to be considered

24

therefore is

$$
\begin{array}{llllll}
\text{(I)} & x_0 = a, & x_1 = -a, & x_2 = b, & x_3 = -b \\
\text{(II)} & x_0 = a, & x_1 = -b, & x_2 = b, & x_3 = -a \\
\text{(III)} & x_0 = -b, & x_1 = -a, & x_2 = a, & x_3 = b \\
\text{(IV)} & x_0 = -a, & x_1 = -b, & x_2 = a, & x_3 = b \\
\text{(V)} & x_0 = -b, & x_1 = a, & x_2 = b, & x_3 = -a \\
\text{(VI)} & x_0 = -a, & x_1 = a, & x_2 = b, & x_3 = -b \\
\text{(VII)} & x_0 = -a, & x_1 = a, & x_2 = -b, & x_3 = b \\
\text{(VIII)} & x_0 = -b, & x_1 = a, & x_2 = -a, & x_3 = b \\
\text{(IX)} & x_0 = a, & x_1 = b, & x_2 = -b, & x_3 = -a \\
\text{(X)} & x_0 = a, & x_1 = b, & x_2 = -a, & x_3 = -b \\
\text{(XI)} & x_0 = a, & x_1 = -b, & x_2 = -a, & x_3 = b \\
\text{(XII)} & x_0 = a, & x_1 = -a, & x_2 = -b, & x_3 = b
\end{array}
\tag{3.4}
$$

A more intuitive description for the 12 mappings is displayed in Fig. 3.2
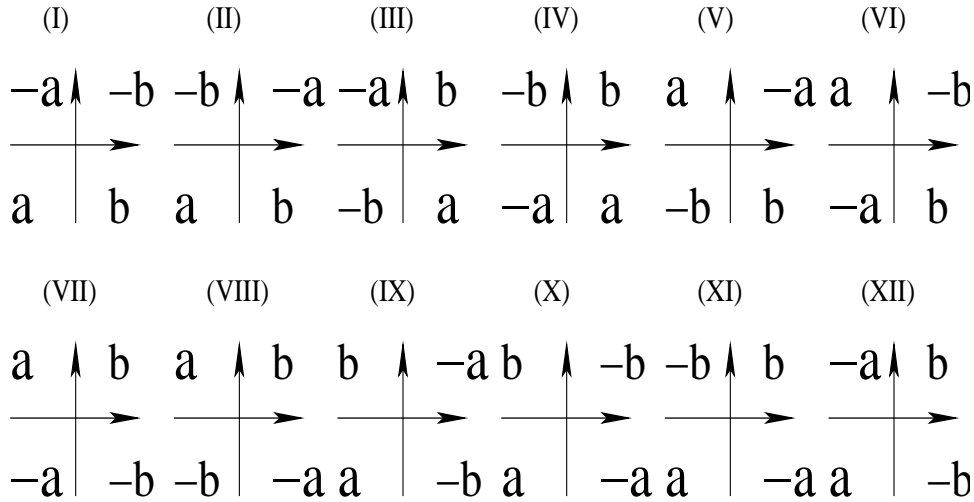
Figure 3.2: The 12 possible mappings for the Optimum Modulator

3.3.2 Strategy for Optimal Mapping

The Optimum modulator works as follows: For a given power constraint, interference and noise level, it first searches over different values of a, b. For each pair of (a,b), every mapping in the 12 combinations is used to calculate the mutual information. The one with the largest mutual information is recorded as the optimal constellation associated with that point (a,b). Among these recorded points, the one with the maximum mutual information and with power less than or equal to the power constraint is selected as

25

the optimal mapping for this given scenario.

As one can see from the above 12 different mappings, there is a lot of symmetry. This set actually could be further reduced to 4, or even 2 without an apparent loss of performance. This topic will be further discussed in Section 3.5.2.

3.3.3   Conditional Probability

The independence of z and n gives the distribution function of y conditioned on w as

$$
\begin{aligned}
f_y(y|w) &= \sum_{i=0}^{1} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-z_i-x(w,z_i))^2}{2\sigma^2}} P_z(z=z_i) \\
&= \frac{1}{2\sqrt{2\pi\sigma^2}} \left(e^{-\frac{(y-z_0-x(w,z_0))^2}{2\sigma^2}} + e^{-\frac{(y-z_1-x(w,z_1))^2}{2\sigma^2}} \right)
\end{aligned}
\tag{3.5}
$$

This results in a very simple form of the optimal receiver

$$
\widehat{w}_{\mathrm{MAP}} = \arg\max_w \left(e^{-\frac{(y-z_0-x(w,z_0))^2}{2\sigma^2}} + e^{-\frac{(y-z_1-x(w,z_1))^2}{2\sigma^2}} \right)
$$

3.4   Mutual Information

The Mutual information between the received signal y and the information signal w will be computed as an indicator of the performance of the communication system. The achievable mutual information for different methods will be studied under a power constraint to examine how much gain the Optimum Modulator could achieve over the Optimal THP.

According to information theory, the mutual information between y and w is

$$
\begin{aligned}
\mathbf{I}(y;w) &= \mathbf{H}(w) - \mathbf{H}(w|y) \\
&= \sum_{i=0}^{1} \int_{-\infty}^{\infty} P(y,w_i) \log P(w_i|y) dy - \sum_{i=0}^{1} P_w(w_i) \log P_w(w_i) \\
&= \sum_{i=0}^{1} \left[\int_{-\infty}^{\infty} f_y(y|w_i) P_w(w_i) \log \frac{f_y(y|w_i) P_w(w_i)}{f_y(y)} dy - P_w(w_i) \log P_w(w_i) \right] \\
&= \sum_{i=0}^{1} P_w(w_i) \int_{-\infty}^{\infty} f_y(y|w_i) \log \frac{f_y(y|w_i)}{f_y(y)} dy
\end{aligned}
\tag{3.6}
$$

where

$$
f_y(y) = \sum_{j=0}^{1} f_y(y|w_j) P_w(w_j)
$$

26

and the last equality comes from the fact that

$$\int_{-\infty}^{\infty} f_y(y|w_i)dy = 1, \ i = 0, 1$$

The conditional PDF $f_y(y|w)$ for the Optimal THP and the Optimum Modulator are given by (2.10) and (3.5) respectively.

3.5   Numerical Results

In the numerical simulation, we assume that both w and z are equally distributed. That is, both α and β equal to 0.5. Mutual information is calculated by Monte-Carlo integration and the infinite summation in (2.10) is truncated. A transmit power constraint is applied everywhere and expressed via the signal to interference ratio (SIR) or the signal to noise ratio (SNR).

3.5.1   Mutual Information Simulation

Fig. 3.3 examines how the gap between the Optimum Modulator system and the No Interference system varies when the SIR increases while the SNR is kept constant. We fix $P_z = 4$, simultaneously vary P and σ^2 while keeping the ratio P/σ^2 (SNR) constant and equal to 1dB (marker "Δ"), 3dB (marker "O") and 6dB (marker "$*$") respectively. The No Interference Cancellation (IC) case is also included as a reference. The largest loss the Optimum Modulator suffers is 1.5dB (at least for the SNR values shown in Fig. 3.3).

Fig. 3.4 illustrates how the Optimum Modulator relates to the No Interference system and the Optimal THP precoding. The Heuristic THP utilizes the modulo range as $\Lambda = 1.5 \cdot \Delta_w$. The No Interference Cancellation (IC) case is also included as a baseline. The interference to noise ratio (INR) equals to 3dB. The gap between the Optimum Modulator and the No Interference system is relatively small while the gain over the Optimal THP is pretty large.

Fig. 3.5 shows the performance comparison when both the power of the noise and that of the interference are constant and equal to each other. The largest loss for the Optimal Modulator against the No Interference case is no larger than 1.5dB for all values of SNR. The Heuristic THP only works well in the high SNR region while the Optimal THP is pretty good at high SNR but suffers at SNR between 0 to 3dB.
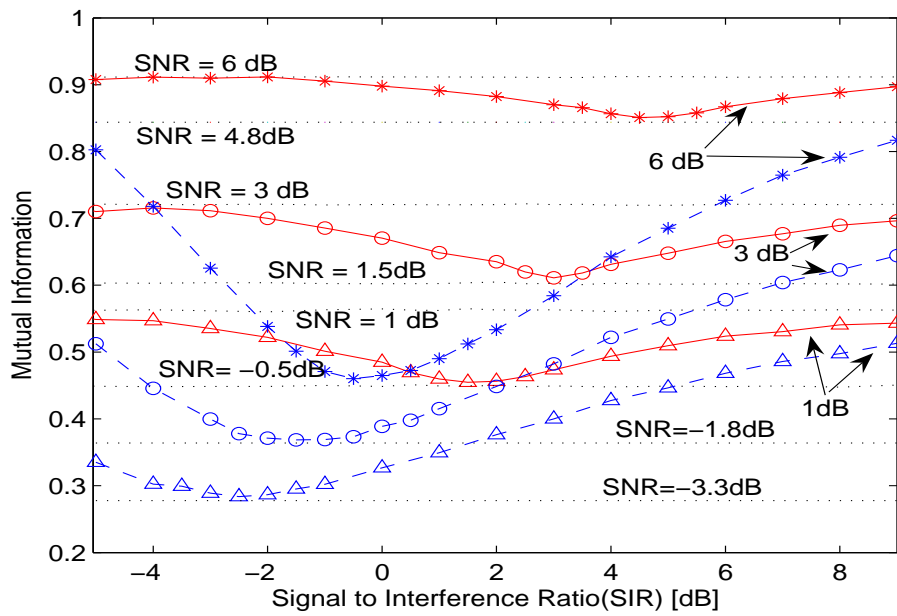
Figure 3.3: Mutual information for Optimal Modulator (solid lines), No IC (dashed lines)and No Interference cases (dotted lines), with $P_z = 4$.
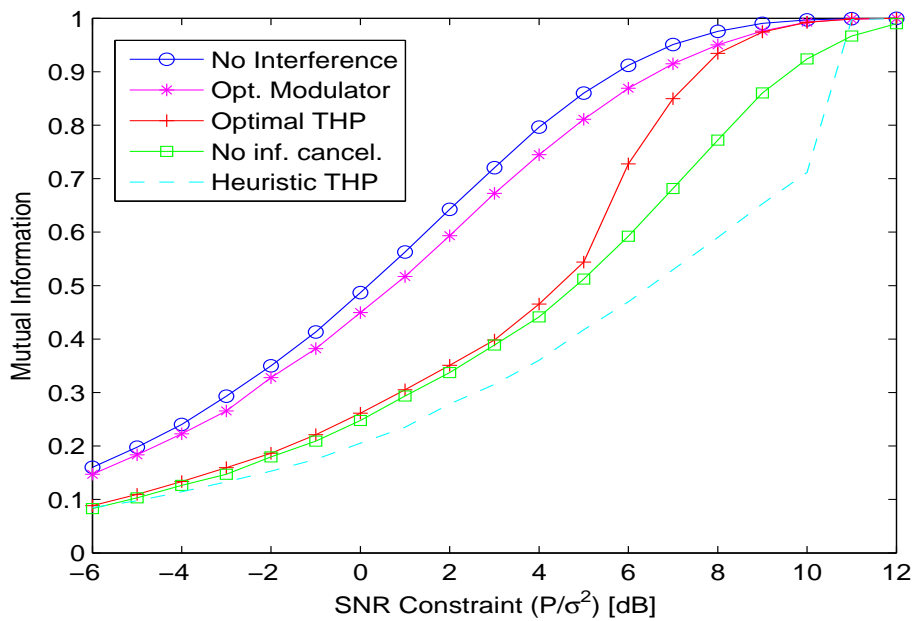
Figure 3.4: Mutual information vs. power constraint, with INR = 3dB.

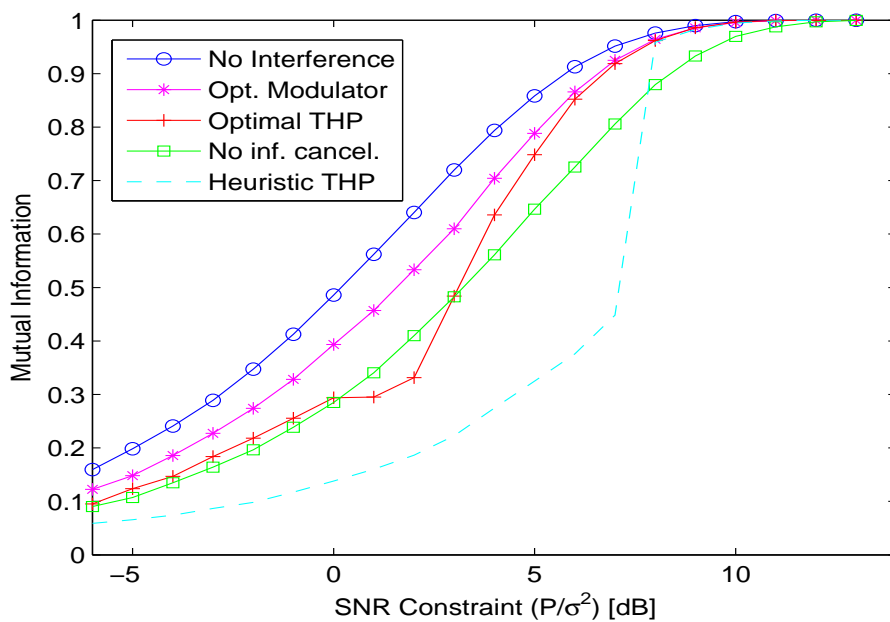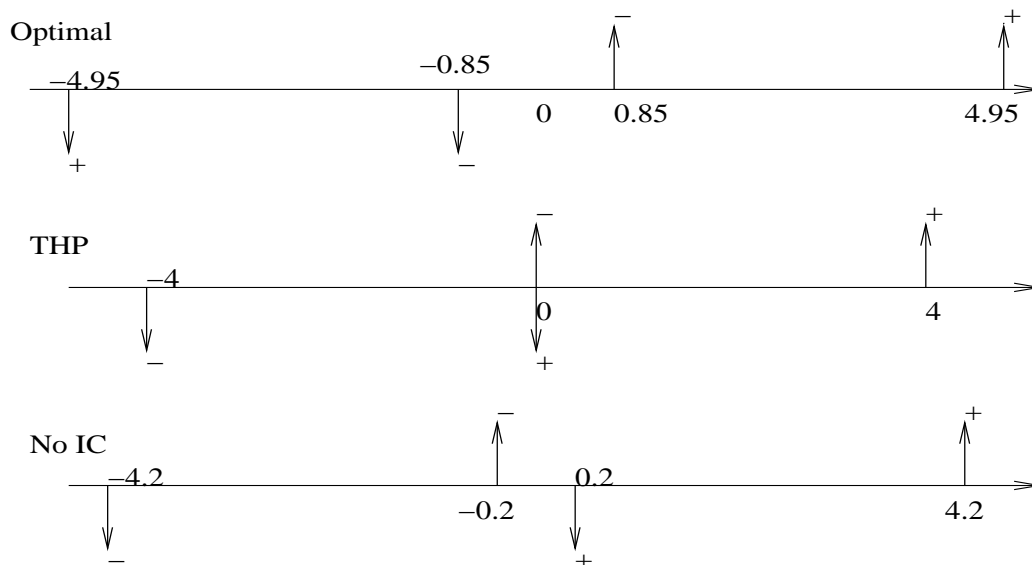Figure 3.5: Mutual information vs. power constraint, with SIR = SNR.

Figure 3.6: Received constellation regardless of noise, SNR = 1dB. Arrows above axis stand for z_1, otherwise for z_0.

An explanation for why the THP even with optimized parameters still can perform worse than the No Interference Cancellation case is that, the inherent structure of THP (subtract the interference and then do modulo operation) confines its performance in some special cases. The No Interference Cancellation scheme we used here is actually not as simple as its name indicates: we use the available transmit power "smartly" as indicated in Section 2.5. This improves the performance of the No IC scheme and makes it possible for some cases to outperform the THP scheme even with optimized parameters. Fig. 3.6 tries to give an intuitive but crude explanation for why the Optimal THP in some special case is worse than the No IC by examining the minimum distance in the received constellation with SNR = 1dB. The minimum distance between w_0 ("-") and w_1 ("+") dominates the performance.

3.5.2 Constellation Simplification

As mentioned previously in Section 3.3.2, the set of constellations could be reduced. A statistical result that shows how frequently each of the 12 combinations are recorded when searching over (a,b) and how frequently each of the 12 combinations are selected as the optimal constellation is shown in the table bellow:

Combinations	(I)	(II)	(III)	(IV)	(V)	(VI)
Recorded. No	5	76	14988	137	133	8438
Frequency	0.0001	0.0016	0.3145	0.0029	0.0028	0.1770
Selected No	0	0	21	0	0	26
Frequency	0	0	0.2234	0	0	0.2766
Combinations	(VII)	(VIII)	(IX)	(X)	(XI)	(XII)
Recorded. No	269	7	14941	124	92	8451
Frequency	0.0056	0.0001	0.3135	0.0026	0.0019	0.1773
Selected No	0	0	22	0	0	25
Frequency	0	0	0.2340	0	0	0.2660

Based on this we believe that the combinations (III), (VI), (IX), and (XII) should be enough to be used by this Optimum Modulator. The set could be reduced to 4, or even 2, due to symmetry between (III) and (IX), (VI) and (XII). The statistical result for how frequently each mapping would be recorded seems to converge to some constant values as more and more simulations are carried out under different conditions. The most important result derived from this statistics is that four combinations seem to be sufficient for Optimum Modulator. This opinion is also supported by numerous simulations which have not been displayed here. It is meaningless, however, to check the specific percentage that indicates how frequently each of these four mappings are selected as the optimal constellation. According to simulation records, the Optimum Modulator will pick up (VI) or (XII) as the optimal constellation in the low SNR region and (III) or (IX) in the high SNR region. (More often than not, shown by simulations, 3dB serves as a good dividing line between low and high SNR regions in the scenario with INR = 0dB.) Therefore as the number of simulation trials increases, the percentage for (VI) and (XII) will become identical. It is the same case for (III) and (IX).

An intuitive explanation for why the Optimum Modulator performs better than other schemes in the presence of interference is given in Fig. 3.7, which displays the received constellations without noise with fixed $P_z = \sigma^2 = 4$ at SNR = 0dB ($a < b$, (XII) selected), 1dB ($a > b$, (VI) selected), 4dB ($a > b$, (IX) selected) and 6dB ($a > b$, (III) selected) respectively. The Optimal Modulator utilizes the information of the interference z in a "smart" way. It also shows clearly the symmetry between (III) and (IX), (VI) and (XII).

Three different simulations have been carried out with different sets consisting of 12 (dashed line), 4 (solid line) and 2 (dotted line) combinations
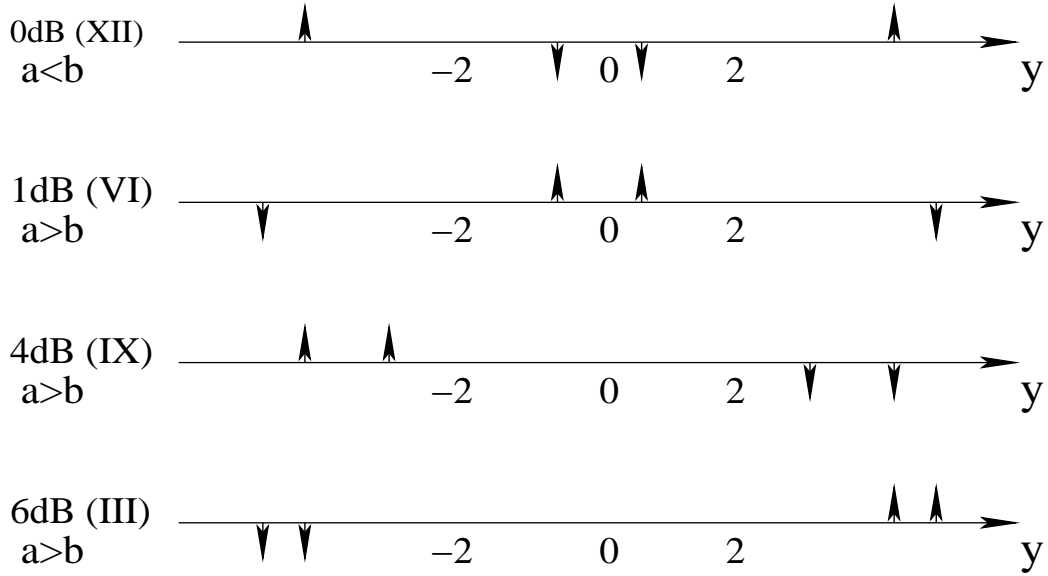
Figure 3.7: Receive constellations regardless of noise, $P_z = 4$, $\sigma^2 = 4$. Arrows above axis stand for w_1, otherwise for w_0.

respectively, as shown in Fig. 3.8 and Fig. 3.9. The set of 2 combinations could either be {(III),(VI)} or {(IX),(XII)}. In Fig. 3.8, we fix $P_z = 9$ and simultaneously vary P and σ^2 while keeping the ratio P/σ^2 (SNR) constant and equal to certain values (1, 3 and 6dB). The difference between set size 12 and 4 is too small to be noticed, and the difference between set size 4 and 2 is mainly caused by the Monte-Carlo simulation.

In Fig. 3.9, we fix $P_z = 10$, $\sigma^2 = 5$ (INR = 3dB, marker "∇") or 10 (INR = 0dB, marker "O") for two different cases. There is no difference between set size 12 and 4, and the difference between set size 4 and 2 is small. The redundancy of mapping sets could be removed to increase the Monte-Carlo simulation efficiency with little or almost no performance loss.

Another straightforward and attractive question is whether it is possible to get rid of the exhaustive search over (a, b). That is to say, if there is any simple relationship between the optimal (a, b) and Δ_z. A mathematical statement of this question would be like:
Given Δ_z and σ^2 in this Optimum Modulator system, for any $P > 0$, is there a positive real number a and a real-valued function $f(a, \Delta_z)$, so that

$$\begin{cases} E[x^2(a, f)] \triangleq (a^2 + f^2(a, \Delta_z))/2 \leq P \\ \mathbf{I}(a, f) = \max_{E[x^2] \leq P} \mathbf{I}(y; w) \end{cases} \quad \text{always hold?}$$

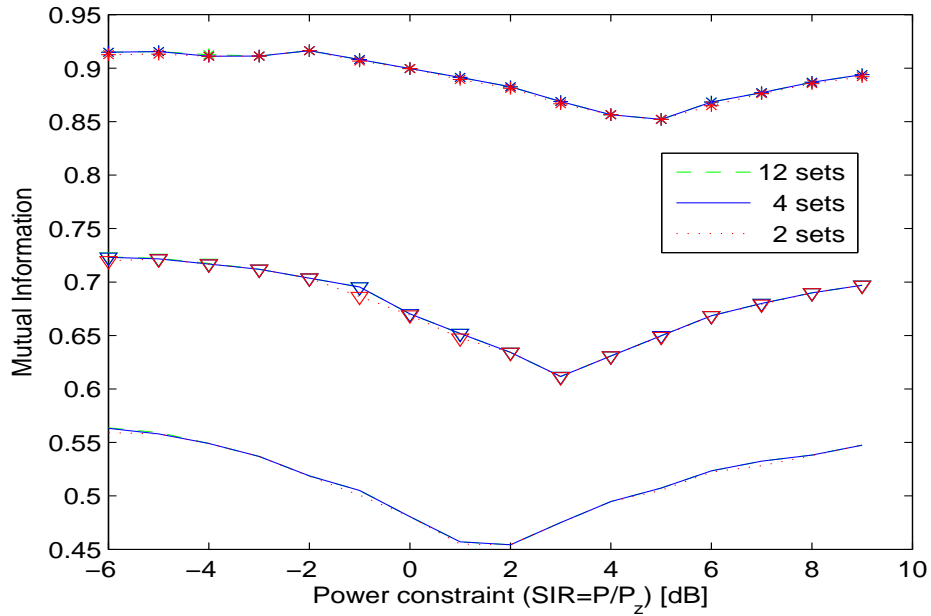Figure 3.8: Mutual information versus power constraint ($\text{SIR} = P/P_z$), with fixed SNR = 1dB, 3dB (marker "∇") and 6dB (marker "$*$") respectively.
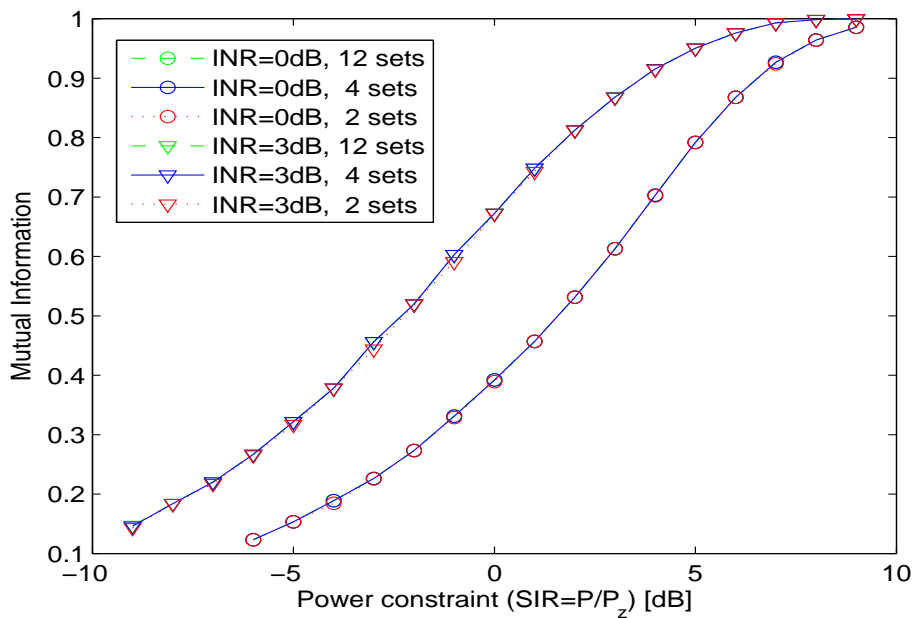

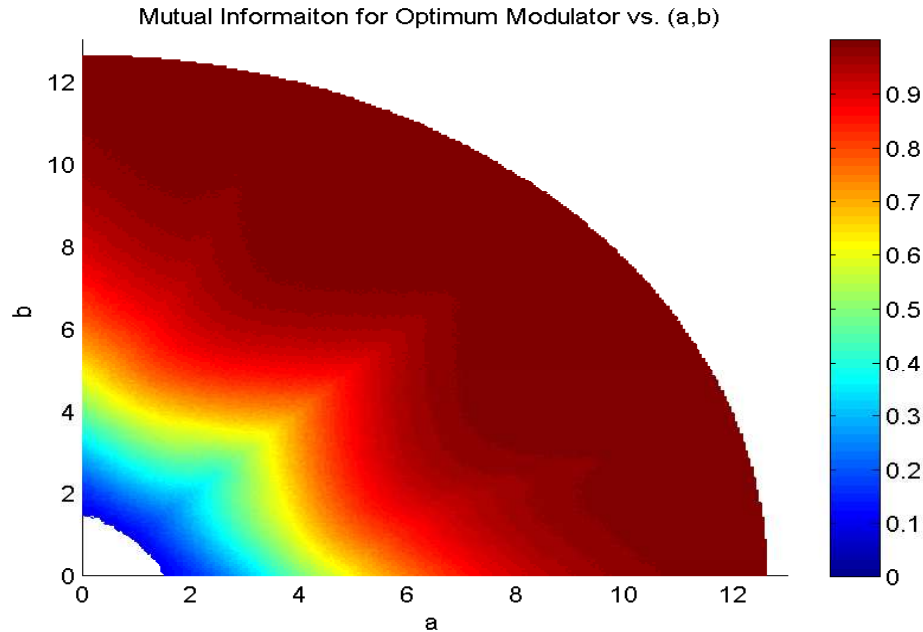Figure 3.9: Mutual information versus power constraint ($\text{SIR} = P/P_z$)

Figure 3.10: Recorded mutual information versus (a,b), with $\Delta_z = 4$, $\sigma^2 = 4$.

This is a difficult problem. The simulation provides some hints, as shown in Fig. 3.10 which displays the mutual information for every point (a, b) when doing the exhaustive search. It does not seem possible to find a simple function $b = f(a, \Delta_z)$ to achieve the optimum.

3.6 Summary

The Optimum Modulator for BPSK signals with BPSK interference can come close in performance to the No Interference case. The largest loss the Optimum Modulator suffers compared with the No Interference system is less than 1.5dB. Simulations show that the algorithm efficiency could be possibly increased by reducing the set of combinations. And it is possible to further reduce the complexity. Although this problem has not been solved in this thesis due to time limitations, it could be a good topic for further work.

Note that, the Optimal THP is a special case of the Optimum Modulator:

- for THP: $x(w, z) = (w - z) \bmod \Lambda$
  The symmetry of w and z as well as the modulo operation confines the constellation choices of x. Equality $x(w_1, z_1) = -x(w_0, z_0)$ and

$x(w_1, z_0) = -x(w_0, w_1)$ always holds and this places an extra constraint on THP.

- for Optimal Modulator: $x(w, z) \in \{-b, -a, a, b\}$, a, b are positive constants. The constraint of THP does not exist here.

It is now clear that the Optimum Modulator has a larger number of degrees of freedom in the constellation design than THP, and this is the inherent reason why the Optimal Modulator always outperforms the THP. This concludes our discussion of the Optimum Modulator. In the next section, we will apply this scheme to the system with a Turbo code as the outer code and examine the gain it could achieve in a coded bit error rate sense.

Chapter 4

Combination with Turbo Coding

4.1 System Configuration

A DPC communication system with a Turbo code as the outer code is depicted in Fig. 4.1. This system is used in all the simulations in this chapter to compute the coded bit error rate for different precoding schemes which have been mentioned in earlier chapters.

4.2 Turbo Coding

Turbo code, invented by C. Berrou, A. Glavieux, and P. Thitimajshima in [15] and [16], is well known for its capability to approach the Shannon limit. From then on, numerous efforts have been spent on reducing the complexity of the decoder. Those results and their byproducts have made Turbo code the most popular outer code in many communication systems. A comprehensive description of the Turbo Code can be found in [17, 18].
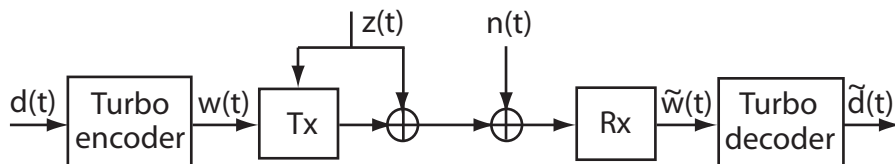

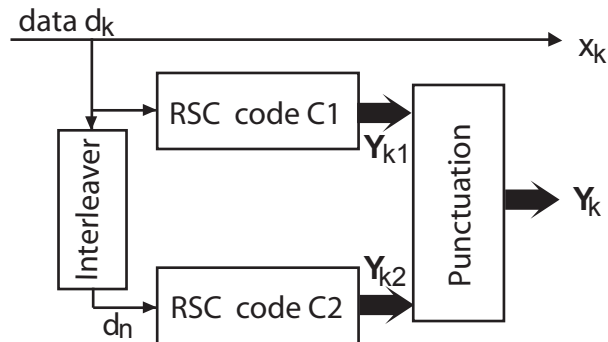Figure 4.1: System configuration with Turbo code as the outer code

Figure 4.2: Turbo encoder

4.2.1 Turbo Encoder

A Turbo code is formed by a parallel concatenation of two Recursive Systematic Convolutional (RSC) codes, linked together by an interleaver, as shown in Fig. 4.2.

The Turbo encoding process is carried out block by block. For each block of transmitted bits, the data flow (d_k at time k) goes directly to the RSC encoder C1 and feeds another RSC encoder C2 (by d_n at time k) after interleaving, where C1 and C2 are generally different. The data d_k are systematically transmitted as symbols x_k. The parity symbols Y_{k1} generated by encoder C1 and Y_{k2} by C2 are then combined (after punctuation when necessary) to form the transmitted parity symbols Y_k associated with x_k at time k. The aim of introducing punctuation is to increase the effective data rate. Generally, punctuation patterns must be chosen with care.

4.2.2 Iterative Turbo Decoder

The common structure of an iterative decoder for Turbo codes is shown in Fig. 4.3. For each received block, the input data are firstly separated into systematic symbols y_k, parity symbols P_{k1} associated with encoder C1, and parity symbols P_{k2} associated with C2. Zeros will be inserted into the parity symbol sequences according to the punctuation operation rule in the encoder if necessary.

For component decoder 1 (decoder 2), information about the transmitted data d_k (d_n after interleaving) is collected from three different sources: 1) the systematic symbols y_k (y_n); 2) parity symbols P_{k1} (P_{k2}) associated with

corresponding encoder C1 (C2); 3) the a-priori information $v_k(u_n)$ about the transmitted data y_k (y_n) provided by decoder 2 (decoder 1) after de-interleaving (interleaving). Interleaving and De-interleaving are exploited properly to make sure that different information about the same bit shall appear in the same position.

Let's take a closer look at how this iterative decoder works. In the first iteration, the component decoder 1 starts without any a-priori information from 3). Based on y_k and P_{k1}, decoder 1 gives out a soft output. Then decoder 2 starts to work based on y_d and P_{k2}, as well as the information from 3) provided by decoder 1. As this additional information from 3) is independent from 1) and 2), it provides an "second opinion" about the transmitted bits and hence could help to improve the performance. After decoder 2 gives out the a-posteriori information about the transmitted bits, the first iteration ends. Then the second iteration starts. This time decoder 1 gets information from 1), 2) and 3) which is provided by decoder 2 in the first iteration and hence it gives a better result compared with the first iteration and feeds it to decoder 2. Then decoder 2 will also be able to give a better result compared with the first iteration due to this improved information.

This iteration continues. In each iteration, information from 1) and 2) is constant while information from 3) is updated. Hence we manage to get a better understanding about the block of transmitted bits after each iteration. It is important in this iterative decoding to make sure that the same kind of information shall be used only one time in each iteration. That's why the information from 1) and 3) is removed from the soft output before bing fed to another decoder as the a-priori information. After several iterations, the soft output z_n from component decoder 2 is passed to make hard decision after de-interleaving.

4.2.3 Turbo Decode Algorithm

Two kinds of decoders are used in the iterative decoding of Turbo codes: the Soft Output Viterbi Algorithm (SOVA) decoder and the Maximum a Posteriori (MAP) decoder. As their names indicate, the SOVA decoder minimizes the sequence error probability whereas the MAP decoder minimizes the coded BER. However, the MAP decoder has higher complexity.

In order to be used in the iterative decoding, both of these component decoders (see Fig. 4.3) should output soft decisions in terms of so-called Log Likelihood Ratios (LLRs): the magnitude gives the sign of the bit while
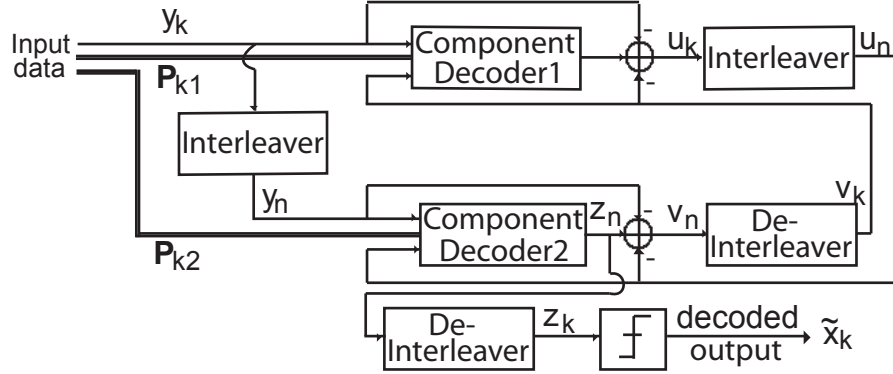
39

Figure 4.3: Iterative decoder

the amplitude indicates the probability of a correct decision. Working with the logarithm of the likelihood functions makes it possible to reduce the complexity of the MAP decoder. According to the Jacobian logarithm

$$\ln(e^{x_1} + e^{x_2}) = \max(x_1, x_2) + \ln(1 + e^{-|x_1 - x_2|}) \tag{4.1}$$

where the second part on the right side can be treated as a correction term. When this correction term is omitted in approximation, the resulting decoder is named as "Max-Log-MAP" decoder [19, 20]. When the exact rather than approximate value of the correction item is calculated, it will result in the "Log-MAP" decoder [21]. One can also build up a lookup table to find the approximate value of the correction term instead of computing the accurate one to improve the algorithm efficiency [21]. For a detailed comparison of different component decoders, see [18].

4.2.4 Recursive Systematic Convolutional Codes

Two Recursive Systematic Convolutional (RSC) encoders with constraint length $K = 3$ (memory length $v = K - 1$), rate $R = 1/2$, and generator polynomials $(G0 = 7, G1 = 5)$ in octal format or $(G0 = 111, G1 = 101)$ in binary format are depicted in Fig. 4.4. $G0$ and $G1$ stand for the feedback loop and the parity output branch respectively.

At time instant k, the input to the RSC encoder is a bit d_k and the corresponding output consists of the systematic transmitted bit x_k and the parity bit y_k as follows

$$\begin{cases} x_k = d_k \\ y_k = \sum_{i=0}^{v} g_{1i} a_{k-i} \quad g_{1i} = 0, 1 \end{cases} \tag{4.2}$$
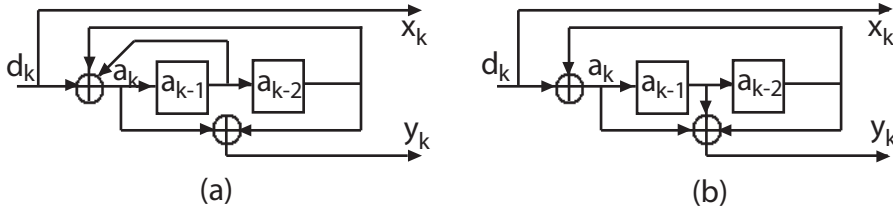
40

Figure 4.4: RSC encoders with constraint length $K = 3$ and different generator polynomials: (a) $G0 = 7$, $G1 = 5$ (b) $G0 = 5$, $G1 = 7$

where g_{ni} is the ith bit of Gn in the binary format and a_k is the input of the shift register given by

$$a_k = d_k + \sum_{i=1}^{v} g_{0i}a_{k-i} \quad g_{0i} = 0, 1 \tag{4.3}$$

Setting $g_{00} = 1$, rewriting (4.3) based on binary operation and putting it together with (4.2), we get

$$\begin{cases} x_k = d_k \\ y_k = \sum_{i=0}^{v} g_{1i}a_{k-i} & g_{1i} = 0, 1 \\ a_k = \sum_{i=0}^{v} g_{0i}a_{k-i} & g_{0i} = 0, 1 \end{cases} \tag{4.4}$$

Hence the generator polynomials $(G0, G1)$ are sufficient to describe the RSC encoders.

4.2.5 Interleaving Sequence

Interleaving plays an important role in Turbo codes. Large efforts have been spent on this topic. Detailed information about the effect of the interleaving sequences can be found in [18].

4.3 Numerical Results

In the following simulations, a Turbo code frequently appearing in many papers, consisting of two identical RSC encoders with constraint length $K = 3$ and generator polynomial $(G0 = 7, G1 = 5)$, is used. The number of iterations in the Turbo decoder is set to 8 and two MAP algorithms, "Log-MAP" and "Max-Log-MAP", are used in the component decoders. Both random interleaving sequences and WCDMA interleaving sequences are used with a sequence length 1000 bits. The coded bit rate is either 1/2 with

41

punctuation or 1/3 without. When rate $R = 1/2$ is exploited, parity symbols from the RSC encoder C1 and C2 are alternatively punctured so that the redundancy information for the original bits and the interleaved bits is kept evenly. Interference $z(t)$ comes from a binary alphabet $\{\sqrt{P_z}, -\sqrt{P_z}\}$ with P_z as its power level and the power of the AWGN noise $n(t)$ is expressed as σ^2. When the curves of coded BER are drawn, it is E_b/N_0 rather than SNR that would be more reasonable to be used for comparison between different rates. With $P = E_b \cdot R$ and $N_0/2 = \sigma^2$, we get

$$\text{SNR} = P/\sigma^2 = \frac{E_b \cdot R}{N_0/2} = E_b/N_0 + 10 \cdot \log_{10}(2 \cdot R) \text{ (dB)} \qquad (4.5)$$

To measure the coded BER, 100M bits Monte-Carlo simulation is used with a minimum number of 200 bit errors recorded for each scenario.

Fig. 4.5 displays the coded BER curves versus E_b/N_0 for the different schemes. The "Log-MAP" algorithm and random interleaving sequences are used. With bit rate $R = 1/2$, we get SNR $= E_b/N_0$ as indicated in (4.5). With interference power $P_z = 4$ and noise power $\sigma^2 = 2$, we have INR = 3dB. At BER $= 10^{-4}$, the Optimum Modulator suffers about 0.7dB loss against the No Interference case and achieves about 2.9dB gain over the Optimal THP, just as indicated by Fig. 3.4 in the sense of mutual information at about 0.6 bits.

All the above conclusions are still kept when the "Max-Log-MAP" algorithm is used, as shown in Fig. 4.6. The only difference is that the absolute performance for each scheme is slightly better (0.1dB) when using the "Log-MAP" decoder than that when using the "Max-Log-MAP" decoder. It is the relative gain between different schemes rather than their absolute performance that is of more interest to us. Meanwhile, the "Max-Log-MAP" algorithm is much more time efficient than the "Log-MAP" algorithm. Thus only the "Max-Log-MAP" algorithm will be exploited in the following simulations.

Fig. 4.7 displays the coded BER with a 1/2 rate Turbo code. The "Max-Log-MAP" algorithm and random interleaving sequences are used with INR = 0dB. At BER $= 10^{-4}$, the Optimum Modulator suffers about 1.4dB loss against the the No Interference case and achieves about 0.8dB gain over the Optimal THP. It is the same gain as displayed in Fig. 3.5 with mutual information at about 0.6 bits. The Optimal THP performs worse than the No Interference Cancellation (No IC) scheme when E_b/N_0 (SNR) is less than
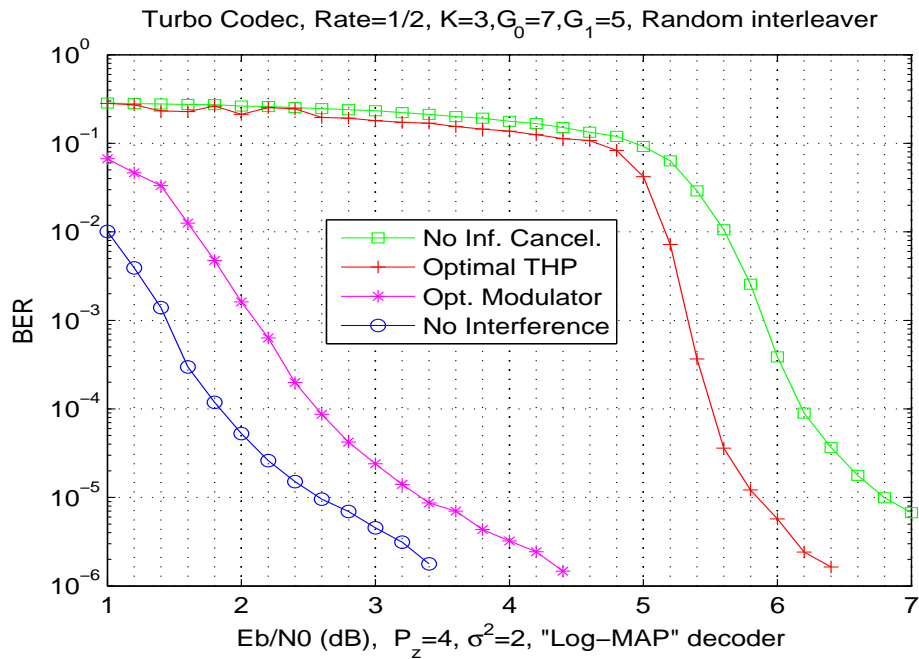
Figure 4.5: Coded BER with Turbo code vs. E_b/N_0, INR = 3dB, "Log-MAP" decoder, bit rate $R = 1/2$, random interleaving.
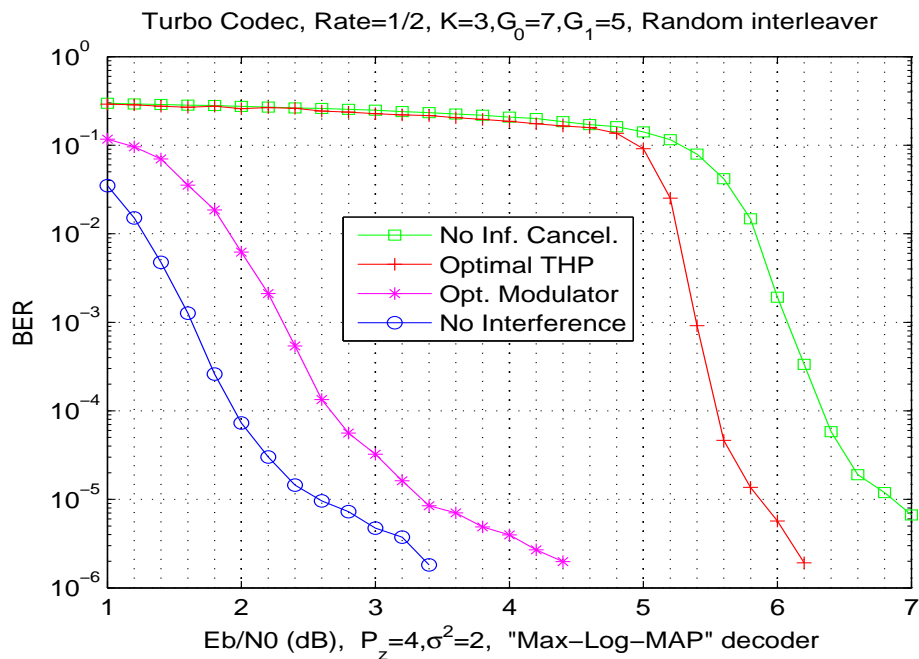

Figure 4.6: Coded BER with Turbo code vs. E_b/N_0, INR = 3dB, "Max-Log-MAP" decoder, bit rate $R = 1/2$, random interleaving.
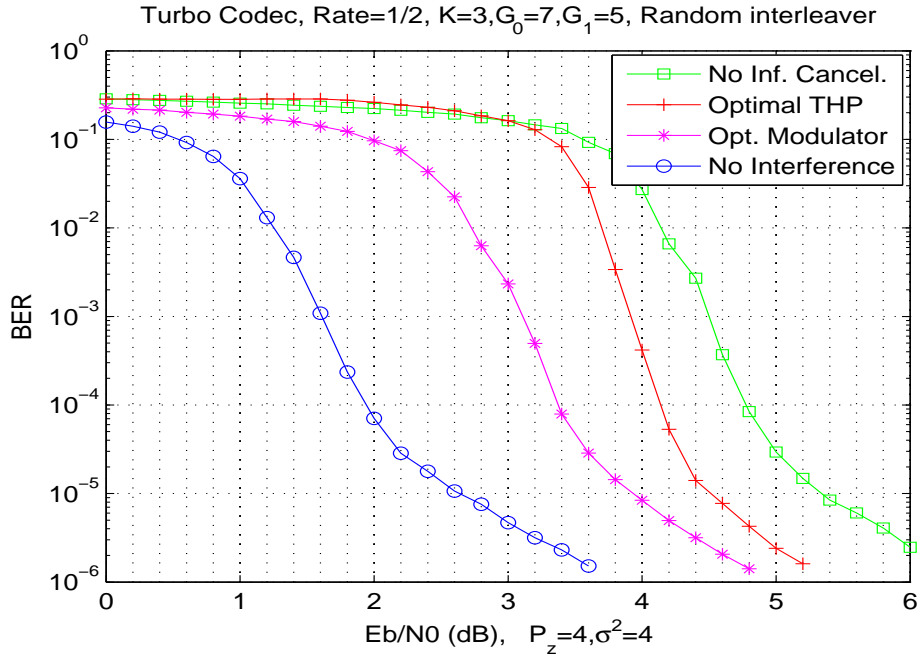
Figure 4.7: Coded BER with Turbo code vs. E_b/N_0, INR = 0dB, "Max-Log-MAP" decoder, bit rate $R = 1/2$, random interleaving.

3dB, as shown before in Fig. 3.5.

Fig. 4.8 displays the coded BER with a 1/2 rate Turbo code. The "Max-Log-MAP" algorithm and WCDMA interleaving sequences are used with INR = 0dB. At BER = 10^{-4}, the Optimum Modulator suffers about 1.5dB loss against the No Interference case and achieves about 0.8dB gain over the Optimal THP. The benefit of using WCDMA interleaving sequences is obvious in high E_b/N_0 (SNR) region.

Fig. 4.9 displays the coded BER with a 1/3 rate Turbo code (without punctuation). According to (4.5) we get SNR = $E_b/N_0 - 1.76$dB. The "Max-Log-MAP" algorithm and random interleaving sequences are used with INR = 0dB. As this 1/3 rate Turbo code is very strong, only the part with SNR < 3.8dB (or $E_b/N_0 < 5.56$dB in equivalent) is shown. As indicated already by Fig. 4.7, the Optimal THP performs worse than the No Interference Cancellation scheme in the SNR region of 0.2dB–3dB, corresponding to the E_b/N_0 region of 1.96dB–4.76dB in Fig. 4.9. This corresponds well to the conclusion drawn in Fig. 3.5 in Chapter 3.5.1. At BER = 10^{-4}, the Optimum Modulator suffers about 1.5dB loss against the No Interference case and achieves about 1.8dB gain over the Optimal THP.
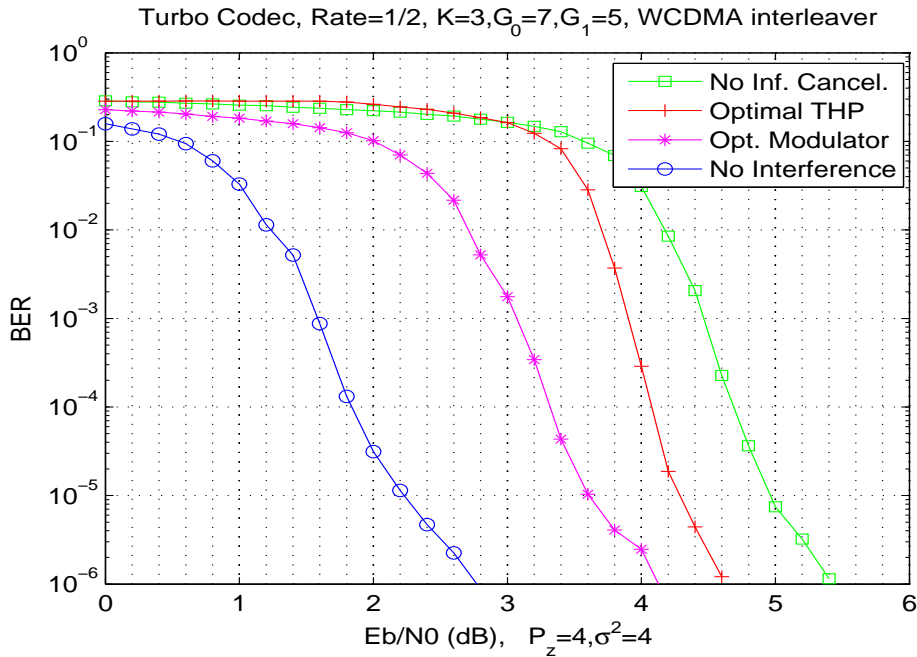
44

Figure 4.8: Coded BER with Turbo code vs. E_b/N_0, INR = 0dB, "Max-Log-MAP" decoder, bit rate $R = 1/2$, WCDMA interleaving.

Fig. 4.10 displays the coded BER with a 1/3 rate Turbo code. The "Max-Log-MAP" algorithm and WCDMA interleaving sequences are used with INR = 0dB. The upper part with BER < 10^{-4} is almost identical to Fig. 4.9 while the lower part shows the gain of introducing WCDMA interleaving sequences over random interleaving sequences. At BER = 10^{-4}, the Optimum Modulator displays the same gain as in Fig. 4.9.

4.4   Summary

Simulations in this chapter show that the capacity gains that the Optimum Modulator achieves in different scenarios over the Optimal THP in the sense of coded BER correspond to the gains in mutual information found in Chapter 3. The Optimum Modulator always outperforms the Optimal THP. Parameter adjustments inside the Turbo code does not change the difference in performance significantly.
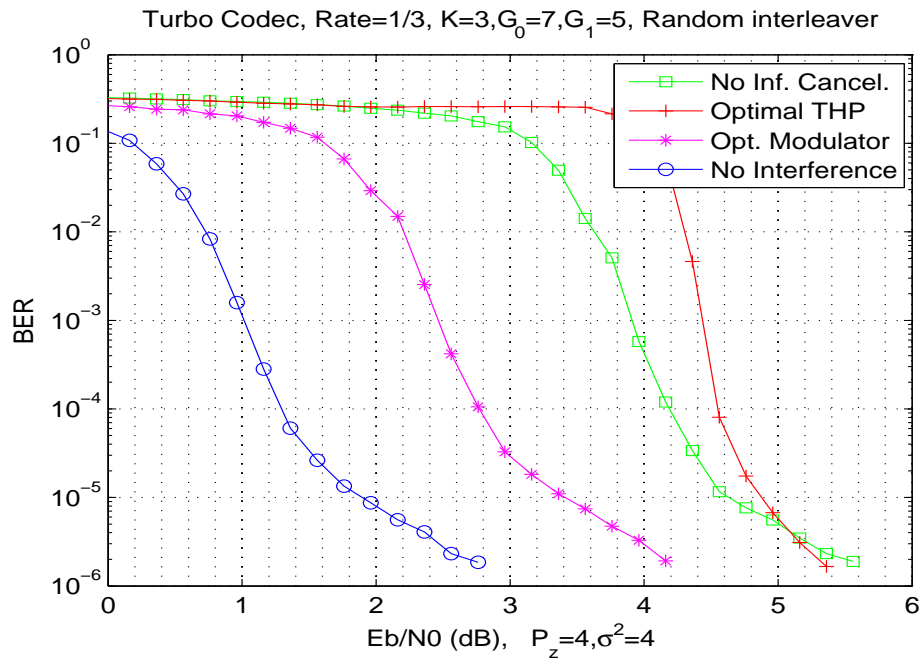
Figure 4.9: Coded BER with Turbo code vs. E_b/N_0, INR = 0dB, "Max-Log-MAP" decoder, bit rate $R = 1/3$, random interleaving.
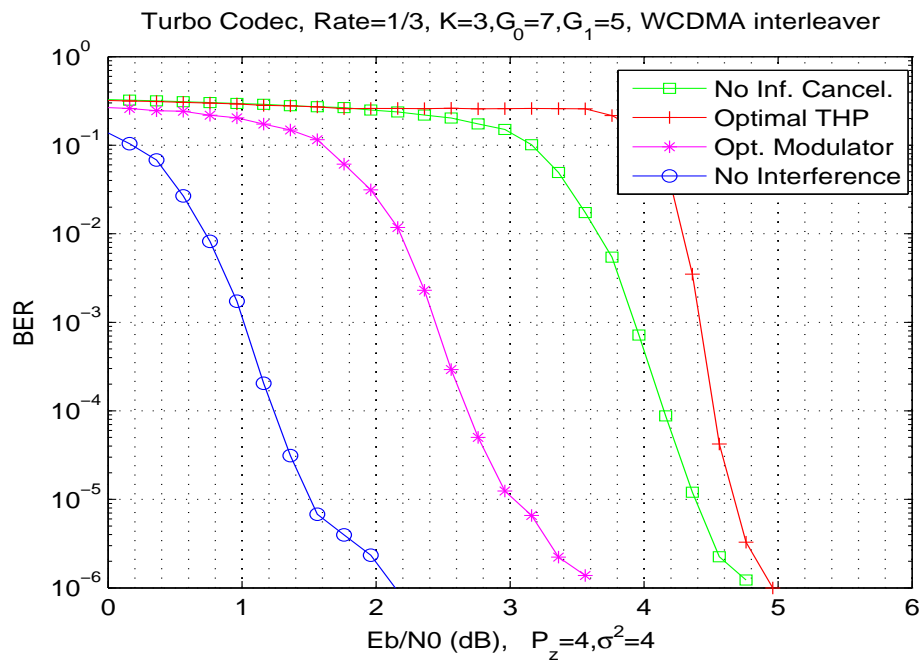

Figure 4.10: Coded BER with Turbo code vs. E_b/N_0, INR = 0dB, "Max-Log-MAP" decoder, bit rate $R = 1/3$, WCDMA interleaving.

Chapter 5

Conclusion and Future Work

5.1   Conclusion

In this thesis, we discussed the problem of modulation for interference avoidance on the AWGN Channel, more precisely, in one dimension. Firstly the performance of scalar Tomlinson-Harashima precoding (THP) in one-dimensional both with heuristic and optimized parameters was investigated. Both mutual information and bit error rate are exploited to evaluate its performance. Simulation shows that THP with optimized parameters always outperforms the Heuristic THP. Next we propose an Optimum Modulator based on exhaustive constellation search to further improve the coding gain. Numerous efforts have been spent on reducing the complexity of finding the optimized parameters for THP and the optimized constellation for the Optimum Modulator. Simulation results show that the complexity could be greatly reduced without performance loss. The Mutual information between the transmitter and the receiver as well as the Bit Error Rate have been used to compare the performance with several other methods. At the end, the coded Bit Error Rate (BER) with a Turbo Code as the outer code was examined to display the coding gain achieved by these two new precoding methods. Simulations show that the capacity gains that the Optimum Modulator achieves in different scenarios over the Optimal THP in the sense of coded BER correspond to those predicted by the mutual information.

5.2   Future Work

All conclusions we have drawn under the assumption of binary constellations do not necessarily translate (at least not quantitatively) to the case of larger constellations. However, our study does indicate that rather impressive inter-

ference avoidance performance can be achieved in a single dimension. This result serves as motivation to continue studying low-complexity approaches to the Costa problem.

The work can be extended in several directions. First, the constraint of binary signal constellations may be relaxed. In this case, it is not clear how the resulting optimization problem can be solved: an exhaustive search over the mapping $x(w, z)$ does not seem feasible. Second, one may attempt to extend our strategy to a higher (but small) dimension; that is, let x, w, z, n, y be vectors and work with a multivariate mapping $\mathbf{x}(\mathbf{w}, \mathbf{z})$. An implementation of Costa precoding in practice will likely rely on operations in a space of small dimension, so the problems outlined here would be of much interest.

Bibliography

[1] M. Costa, "Writing on dirty paper," *IEEE Transactions on Information Theory*, vol. 29, pp. 439-441, May 1983.

[2] G. Caire, S. Shamai, "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Transactions on Information Theory*, vol. 49, pp. 1691-1706, July 2003.

[3] D. Tse, P. Viswanath, *Fundamentals of Wireless Communications*, Cambridge University Press, 2005.

[4] J. G. Proakis, M. Salehi,*Communications Systems Engineering, second edtion*, Prentice-Hall, Inc, 2002.

[5] M. Tomlinson, "New automatic equalizer employing modulo arithmetic" *Electronics Letters*, pp. 138-139, March 1971.

[6] H. Harashima, H. Miyakawa, "Matched-transmission technique for channels with intersymbol interference," *IEEE Transactions on Communications*, vol. COM-20 No. 4, pp. 774-780, August 1972.

[7] R. D. Wesel, J. M. Cioffi, "Achievable rates for Tomlinson-Harashima precoding," *IEEE Transactions on Information Theory*, vol. 44 No. 2, pp. 824-831, March 1998.

[8] J. Liu, W. A. Krzymień, "Improved Tomlinson-Harashima precoding for the downlink of multiple antenna multi-user systems," in Proc. of *IEEE Wireless Communications and Networking Conference, 2005*, vol. 1, pp. 466-472, March 2005.

[9] J. E. Smee, S. C. Schwartz, "Adaptive compensation techniques for communications systems with Tomlinson-Harashima precoding," *IEEE Transactions on Communications*, vol. 51 No. 6, pp. 865-869, June 2003

[10] A. P. Liavas, "Tomlinson-Harashima precoding with partial channel knowledge," *IEEE Transactions on Communications*, vol. 53 No. 1, pp. 5-9, January 2005.

[11] R. Zamir, S. Shamai, and U. Erez, "Nested linear/lattice codes for structured multiterminal binning," *IEEE Transactions of Information Theory*, vol. 48 No.6, pp. 1250-1276, June 2002.

[12] A. Bennatan, D. Buhrstein, G. Caire, and S. Shamai, "Superposition coding for side-information channels," Preprint.

[13] U. Erez, S. ten Brink, "Approaching the dirty paper limit for cancelling known interference," in Proc. of *Allerton Conference on Communications, Control and Computing*, pp. 799-808, October 2003.

[14] U. Erez, S. ten Brink, "A close-to-capacity dirty paper coding scheme," *IEEE Transactions on Information Theory*, vol. 51 No. 10,pp. 3417-3432, October 2005.

[15] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo-codes," in Proc. of *IEEE International Conference on Communications (ICC)*, Geneva, Switzerland, May 1993, pp. 1064-1070.

[16] C. Berrou, A. Glavieux, "Near optimum error correcting coding and decoding: Turbo-codes," *IEEE Transaction on Communications*, vol. 44 No. 10, pp. 1262-1271, October 1996.

[17] B. Sklar, "A primer on Turbo code concepts," *IEEE Communications Magazine*, pp. 94-102, December 1997.

[18] J. P. Woodard, L. Hanzo, "Comparative study of Turbo decoding techniques: an overview," *IEEE Transactions on Vehicular Technology*, vol. 49 No. 6, pp. 2208-2233, November 2000.

[19] W. Koch, A. Baier, "Optimum and sub-optimum detection of coded data disturbed by time-varying inter-symbol interference," *IEEE Globecom*, pp. 1679-1684, December 1990.

[20] J. A. Erfanian, S. Pasupathy, and G. Gulak, "Reduced complexity symbol detectors with parallel structures for ISI channels," *IEEE Transaction on Communications*, vol. 42, No.2-4, pp. 1661-1671, February 1994.

[21] P. Robertson, E. Villebrun, and P. Hoeher, "A comparison of optimal and sub-optimal MAP decoding algorithms operating in the log domain," in Proc. of *IEEE International Conference on Communications (ICC)*, Seattle, WA, June 1995, pp. 1009-1013.

[22] J. Du, E. G. Larsson, and M. Skoglund, "Costa precoding in one dimension," in Proc. of *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Toulouse, France, May 2006, to appear.