國 立 交 通 大 學

電信工程研究所

碩 士 論 文

有通道估計誤差下之多天線信號偵測

MIMO Signal Detection in the Presence of Channel Estimation Errors

研 究 生：林玉婷

指導教授：蘇育德 博士

中 華 民 國 九十九 年 十 月

有通道估計誤差下之多天線信號偵測

MIMO Signal Detection in the Presence of Channel Estimation Errors

研究生：林玉婷　　Student : Yu-Ting Lin

指導教授：蘇育德 博士　　Advisor : Dr. Yu T. Su

國 立 交 通 大 學

電 信 工 程 研 究 所

碩 士 論 文

A Thesis

Submitted to Institute of Communications Engineering

College of Electrical and Computer Engineering

National Chiao Tung University

in Partial Fulfillment of the Requirements

for the Degree of Master of Science

in Communications Engineering

October 2010

Hsinchu, Taiwan, Republic of China

中 華 民 國 九 十 九 年 十 月

有通道估計誤差下之多天線信號偵測

研究生：林玉婷　　指導教授:蘇育德 博士

國立交通大學電信工程學系碩士班

中文摘要

相較於傳統的單輸入單輸出(SISO)系統，使用多傳送與接收天線的多輸入多輸出(MIMO)的無線通訊系統可大幅的增加系統容量(capacity)，因此在過去十年被廣泛的研究，相關技術並已應用於實際系統成為國際無線通訊標準規格不可或缺的一部份。大多數對於MIMO訊號偵測的研究，都假設接收端有百分之百正確的通道資訊(CSI)，但實際上，這是不可能的。通道資訊需透過某種估計演法得到，誤差在所難免，而基於此不正確CSI所偵測的數據之可靠性也必須打折扣。事實上許多研究都顯示接受器的效能將有顯著的降低。因此，如何在設計MIMO接收器時將通道估計誤差一併考慮以盡量減少因上述CSI不匹配的因素造成的效能損失是一項相當迫切的課題，也是本文的主要研究動機。我們首先探討通道估計誤差對兩種MIMO信號偵測法–即粒子群趨動交叉熵偵測法及QRD-*M*偵測法–的性能之影響。

粒子群趨動交叉熵法（Particle-Swarm-Driven Cross-Entropy）偵測法是在給定接收向量後，試圖尋找傳送信號位置的事後機率分布。這個機率分布是藉由在接收向量附近取樣並且反覆的更新使得機率分佈擁有最小交叉熵（相較於現有的機率分佈）。為了改善效能，粒子群趨動的概念則被引入來提供一組新的機率分布。另一方面，QRD-*M*則是一個以樹狀結構為基礎的低複雜度偵測法。在少許效能損失下，利用樹狀結構搜尋並且減少每一層的存活路徑。

由於最小歐式距離偵測法在有通道估計誤差的情形下對於前述兩個偵測器並不適用，我們提出考量通道估計誤差效應的新型最佳偵測結構，提出了新的解碼度量(decoding metric)。模擬結果顯示新法雖然增加了些許計算複雜度但卻有顯著的效能提昇。我們也發現將同樣的度量應用在有時空編碼的MIMO系統亦可得到類似的改善。

MIMO Signal Detection in the presence of Channel Estimation Errors

Student : Yu-Ting Lin Advisor : Yu T. Su

Institute of Communications Engineering

National Chiao Tung University

Abstract

The multiple-input multiple output (MIMO) technology promises significant capacity increase over conventional single-input single-output systems. Most investigations on MIMO signal detection, however, assume perfect channel state information (CSI), which is difficult, if not impossible, to realize in practice. Earlier studies have shown that that the performance of MIMO detection schemes will suffer from severe degradation in the presence of channel estimation errors.

In this thesis, the effects of imperfect CSI on two MIMO signal detectors, namely, the Particle-Swarm-Driven Cross-Entropy (PSD-CE) based detector and the QRD-M detector, are studied, and new detector structures that take into account the CSI error are proposed.

The PSD-CE detector tries to estimate and refine the 'a posteriori probability distribution' of the transmitted signal location given the received vector. The distribution is estimated by sampling over the neighborhood of the received vector and is iteratively updated to the one which has the minimum cross entropy with respect to the current distribution. It is further modified by applying the concept of Particle Swarm Optimization to render a mixture of probability distribution. QRD-M, on the other hand, is an efficient tree-search based detector. It prunes the search tree to reduce the number of surviving paths with minimum performance loss.

Since the minimum Euclidean distance criterion is no longer suitable for both detectors in the presence of channel estimation errors, the proposed MIMO detectors take into account the imperfect CSI effect by averaging the estimation errors to obtain a new decoding metric. Numerical examples are given to demonstrate the performance improvement which is attained with insignificant complexity increase.

Acknowlegements

I am truly grateful to my advisor, Professor Y. T.Su. This thesis would not have been possible without his guidance and support. His wisdom and encouragements guided me not only in the research for this thesis, but also in my daily life. It is really an honor to be his student.

I am greatly indebted to my seniors, Jen-Yang and Chang-Ming. They helped me in times of trouble and gave me many suggestions during the years. The knowledge I gained from the many discussions we had over the past two years is invaluable.

I would also like to thank the members of lab 811, they have always been there for me, giving me encouragements and advises in times of need. I will always remember the sharing of our week in the meetings.

Finally, I wish to express my heartfelt gratitude to my dearest family for their spiritual support and unconditional love.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Multiple-input and multiple-output (MIMO) technologies have gained enormous popularity in the recent years because the significant communication capacity enhancement they promised to deliver [1], [3]. A MIMO system uses multiple antennas at both the transmit and the receive ends. In a rich-scattering MIMO channel that can be characterized by a channel matrix with independent and identically distributed (i.i.d.) Rayleigh entries, the associated capacity can increase linearly with the minimum of the number of transmit and receive antennas [2].

However, such a capacity increase can only be realized if the channel matrix is perfectly known. The more practical concern of detecting MIMO signals has been answered by many alternatives, e.g. the Zero Forcing (ZF) [4] detector and V-BLAST [6] detector. Like the capacity issue, these detection schemes perform well only if the channel state information (CSI), i.e., the channel matrix, is perfectly known at the receiver [7]. In practical situations, the knowledge of the channel is never perfect and is attained by some channel estimation scheme at the receiver. The most popular and practical way to obtain the CSI at the receiver (CSIR), is to insert known sequences of symbols, often referred to as pilot signals, within the transmit data stream to aid channel estimation [8], [9]. Information symbols are then detected as though the estimated channel matrix is perfect. Due to the finite number of pilot symbols and various sources of noise and interference, the receiver can only obtain imperfect estimates of the channel. As a result,

such a *mismatched* detector is a sub-optimal approach and its performance is severely degraded in the presence of channel estimation errors as shown in [8] and [10].

The optimum detection method in MIMO systems is the Maximum Likelihood (ML) detector which often calls for exhaustive search over the entire set of possible transmitted symbol vectors. Apparently, the major drawback of ML detector is its high complexity that increases exponentially with the number of transmit antennas and modulation size of the constellation. On the other hand, when perfect CSI is not available, the minimum Euclidean distance criterion is no longer an ML one. To find an ML detection metric, many investigations [20], [22], assume that the channel (matrix) is uncorrelated with the channel estimation error (matrix). This is not appropriate if we recall that both the least-squares (LS) and minimum mean square error (MMSE) channel estimators follows the orthogonal principle which says that the estimation error should be orthogonal to the estimator in some sense. Based on this observation we adopt a different model and derive a new detection metric based on the latter assumption.

In our work, we consider two MIMO detection schemes: the Particle-Swarm-Driven Cross-Entropy (PSD-CE) [11] method and the M-algorithm combined with QR decomposition (QRD-M) [12], [13]. Both schemes have comparatively low complexities compared with the ML detector.

The PSD-CE is a Monte Carlo based stochastic approach that attempts to find the transmitted symbol. It is an iterative method that tries to approach the 'optimal probability distribution' of the transmitted signal. It generates samples over the entire neighborhood of the received vector with uniform distribution. Then the probability distribution of the samples is updated with the 'better' samples. After some iterations, the probability distribution should close in on the optimal distribution in the Cross-Entropy (CE) sense [14]. To improve the performance further, the concept of Particle Swarm Optimization (PSO) [16], [17] is added.

In order to accelerate the convergence speed of this stochastic approach, we make

some changes to the original PSD-CE method. Going over the steps of PSD-CE, we realize that initializing the probability distribution with uniform distribution means beginning the search without any information. Instead of starting the search in such an inefficient manner, we could try to use some information that can be easily obtained by some other simple methods, Zero Forcing for example. By changing the initial distribution so that it contains some useful information, the modified PSD-CE method becomes more efficient, bearing better performance than the conventional PSD-CE with lower computational complexity.

Unlike the PSD-CE method, QRD-M is a deterministic detection algorithm. Different from the QR detection [21], QRD-M is a tree search algorithm [13], [21]. The concept of QRD-M is to apply the tree search to detect symbols in a sequential manner. Instead of making a decision for the symbols at each detection level as in the QR detection method, QRD-M retains M reliable paths. Decision is made only after all the layers have been processed. It can achieve near-ML performance with relatively low but fixed complexity. There is a trade-off between the complexity and performance on the selection of M. With a large M, better performance is achieved at the cost of higher complexity. Usually, M is set to be the size of the modulation constellation.

The new optimization criterion mentioned earlier is used in place of the original criterion $\|\mathbf{y} - \mathbf{Hx}\|^2$ in the PSD-CE method. However, it cannot be directly applied to the QRD-M algorithm because it can no longer be decomposed into the sum of separate branch metrics, as will be shown in Chapter 5. We figured out a way such that the new criterion becomes the lower bound and can be used in the QRD-M method to achieve better performance in the presence of channel estimation errors.

By employing some simple space-time block codes (STBC) [18], [19], the effects of the gain in the modified ML criterion is apparent. G. Taricco *et al.* [20] examines space-time decoding with imperfect channel estimations, but their analysis is based on the assumption that the channel estimation error is orthogonal to the channel.

The rest of this thesis is organized as follows: in Chapter 2 we describe the system model used and the main assumptions concerning channel estimation. In Chapter 3 we review some existing MIMO detection schemes that are of interest to our study. Chapter 4 gives a detailed description of the PSD-CE algorithm, as well as the modifications made to improve its performance. We then propose improved detectors in the presence of channel estimation errors in Chapter 5. Space-Time codes are introduced in Chapter 6. Simulation results are given in Chapter 7 and finally, Chapter 8 gives the conclusion.

The following notations are used throughout the thesis: upper case bold symbols denote matrices and lower case bold symbols denote vectors. $\mathbf{I}_N$ is a $N \times N$ identity matrix. The superscripts $(\cdot)^T, (\cdot)^H$ and $(\cdot)^\dagger$ represent the transpose, Hermitian transpose and the pseudo-inverse, respectively. $\mathbb{E}\{\cdot\}$ denotes the statistical expectation, $tr(\cdot)$ is the trace of a square matrix.

Chapter 2

System Model

2.1 MIMO Systems with Perfect Channel Estimation

We consider a MIMO system [1] with N_T transmit antennas and N_R receive antennas. As shown in Fig. 2.1, data is demultiplexed into N_T data substreams, then the substreams are mapped onto sequences of M-QAM symbols where the signal constellation is denoted by $\mathcal{A}_M$. These symbol sequences are transmitted over the N_T antennas simultaneously.

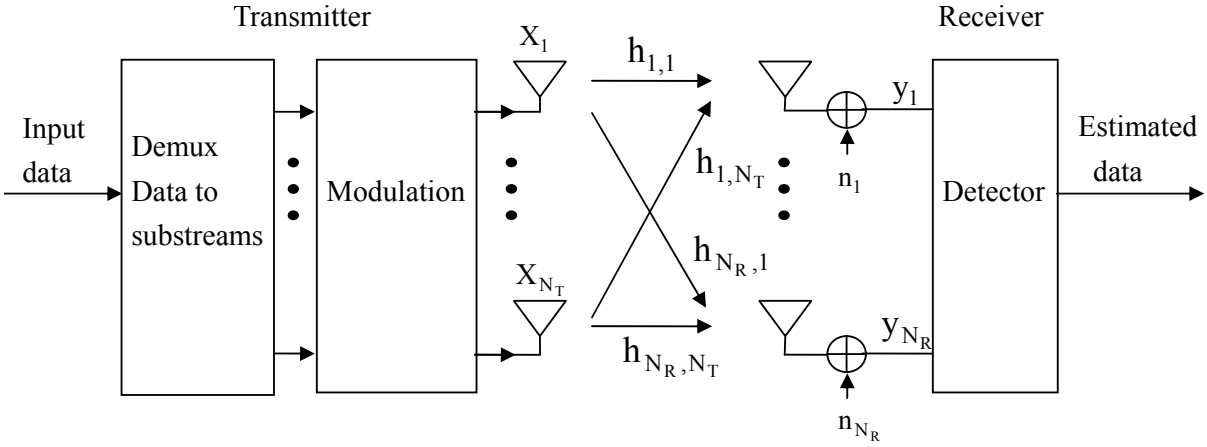

Figure 2.1: A MIMO system model.

If T consecutive signal vectors are transmitted, the transmitted signal can be represented as an $N_T \times T$ matrix $\mathbf{X}$, where $\mathbf{X} \triangleq [\mathbf{x}_1, \ldots \mathbf{x}_T]$ and $\mathbf{x}_t$ is the signal vector transmitted at the tth time slot. The signal vector transmitted at the tth time slot $\mathbf{x}_t$ can be represented as

$$\mathbf{x}_t = [x_{1,t}, \ldots, x_{N_T,t}]^T \in \mathcal{A}_M^{N_T} \tag{2.1}$$

Similarly, assuming perfect timing receiver and quasi-static channel , the corresponding $N_R \times T$ received signal matrix, $\mathbf{Y}$, can be expressed as

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{Z} \tag{2.2}$$

where

$$\mathbf{Y} = [\mathbf{y}_1, \ldots, \mathbf{y}_T] \tag{2.3}$$

$$\mathbf{Z} = [\mathbf{z}_1, \ldots, \mathbf{z}_T] \tag{2.4}$$

and

$$\mathbf{y}_t = [y_{1,t}, \ldots, y_{N_R,t}]^T \in \mathcal{C}^{N_R} \tag{2.5}$$

$$\mathbf{z}_t = [z_{1,t}, \ldots, z_{N_R,t}]^T \in \mathcal{C}^{N_R} \tag{2.6}$$

$$\mathbf{H} = \begin{bmatrix} h_{1,1} & h_{1,2} & \cdots & h_{1,N_T} \\ h_{2,1} & h_{2,2} & \cdots & h_{2,N_T} \\ \vdots & \vdots & \ddots & \vdots \\ h_{N_R,1} & h_{N_R,2} & \cdots & h_{N_R,N_T} \end{bmatrix} \in \mathcal{C}^{N_R \times N_T} \tag{2.7}$$

Here, $\mathcal{C}$ denotes the complex-valued domain. $\mathbf{H}$ describes the overall $N_R \times N_T$ channel matrix, the (j,i)th element of $\mathbf{H}$, $h_{j,i}$, is the channel response between the ith transmit antenna and the jth receive antenna. In this thesis, the elements of $\mathbf{H}$ are independently and identically distributed (i.i.d.) zero-mean complex Gaussian random variables with unit variance. $\mathbf{Z}$ is the $N_R \times T$ additive white Gaussian noise (AWGN) matrix observed at receiver. Each element of $\mathbf{Z}$ has zero mean and variance σ_z^2, while the average transmit power of each antenna is normalized to one. Hence, we have

$$\mathrm{E}_s \triangleq \frac{1}{N_T\, T}\, \mathbb{E}\left[\, \| \mathbf{X} \|^2 \,\right] = \frac{1}{N_T\, T}\, \mathbb{E}\left[\, tr\left\{ \mathbf{X}\mathbf{X}^H \right\} \right] = 1 \tag{2.8}$$

and

$$\mathbb{E}\{\, \mathbf{Z}_t \, \mathbf{Z}_t^H \} = \sigma_z^2 \, \mathbf{I}_{N_R}. \tag{2.9}$$

where tr denotes the trace operation, $\mathbf{I}_N$ is an identity matrix of size $N \times N$ and $t = 1, \cdots, T$.

2.2 Pilot-Based Channel Estimation

The channel information is estimated at receiver and it is often assumed to be perfectly estimated in a lot of literature. However, due to the time-varying channels and quantization errors, only an estimate of the channel matrix that is corrupted by estimation errors is available at the receiver side and the assumption about perfect estimate is not rational in a practical system.

For pilot-based channel estimation scheme, a pilot matrix, $\mathbf{X}_P$, is sent and the receiver observes

$$\mathbf{Y}_P = \mathbf{H}\mathbf{X}_P + \mathbf{Z}_P \tag{2.10}$$

where $\mathbf{Z}_P$ is the noise matrix affecting the transmission of the pilot symbols. $\mathbf{X}_P$ is a known $N_T \times P$ matrix at receiver with average pilot symbol energy

$$\mathrm{E}_p \triangleq \frac{1}{N_T \, P} \parallel \mathbf{X}_P \parallel^2 = \frac{1}{N_T \, P} \; tr \, \{ \, \mathbf{X}_P \, \mathbf{X}_P^H \} \tag{2.11}$$

Since the number of the elements of H is $N_R N_T$, at least $N_R N_T$ independent measurements are needed to estimate the $N_R \times N_T$ channel matrix. To satisfy this condition, N_R measurements are observed in each time slot and we require P to be larger than or equal to N_T. Furthermore, in order to yield independent measurements, $\mathbf{X}_P$ must have rank N_T.

Several channel estimation methods have been proposed [25]. The property of channel estimation error depends on the adopted channel estimation method. A channel estimation method, called Least Square (LS) method, is a common technique used to estimate the channel $\mathbf{H}$ and is considered in this thesis.

The LS estimate of $\mathbf{H}$ can be found by

$$\hat{\mathbf{H}} = \mathbf{Y}_P \, \mathbf{X}_P^H \, \left(\mathbf{X}_P \mathbf{X}_P^H\right)^{-1} \tag{2.12}$$

Moreover, it can be expressed in terms of $\mathbf{H}$ as

$$\mathbf{H} = \hat{\mathbf{H}} + \mathbf{E} \tag{2.13}$$

where $\mathbf{E}$ is the channel error matrix. Due to the orthogonal property of LS method, $\mathbf{E}$ and $\hat{\mathbf{H}}$ are uncorrelated. In addition, because the elements of $\hat{\mathbf{H}}$ and $\mathbf{H}$ are complex Gaussian random variable, the elements of $\mathbf{E}$ are also. Hence, $\hat{\mathbf{H}}$ and $\mathbf{E}$ are independent.

From (2.10), (2.12) and (2.13), we can show that

$$\begin{aligned} \mathbf{H} &= \mathbf{Y}_P \, \mathbf{X}_P^H \, \left(\mathbf{X}_P \mathbf{X}_P^H\right)^{-1} + \mathbf{E} \\ &= \left(\mathbf{H}\mathbf{X}_P + \mathbf{Z}\right) \mathbf{X}_P^H \, \left(\mathbf{X}_P \mathbf{X}_P^H\right)^{-1} + \mathbf{E} \\ \mathbf{E} &= -\mathbf{Z}_P \, \mathbf{X}_P^H \, \left(\mathbf{X}_P \mathbf{X}_P^H\right)^{-1} \end{aligned} \tag{2.14}$$

From (2.14), we also have

$$\mathbf{E}_i = -(\mathbf{Z}_P)_i \, \mathbf{X}_P^H \, (\mathbf{X}_P \mathbf{X}_P^H)^{-1} \tag{2.15}$$

$(\cdot)_i$ denoting the ith row of the matrix $(\cdot)$. We use orthogonal pilot matrices which can be generated from a perfect root-of-unity sequence (PRUS) [26], then the covariance matrix of $\mathbf{E}_i$ is

$$\mathbb{E}\left[\, \mathbf{E}_i^H \, \mathbf{E}_i \,\right] = \sigma_z^2 (\mathbf{X}_P \mathbf{X}_P^H)^{-1} \tag{2.16}$$

Since the pilot matrices are orthogonal, with (2.11), the covariance matrix becomes

$$\mathbb{E}\left[\, \mathbf{E}_i^H \, \mathbf{E}_i \,\right] = \sigma_e^2 \, \mathbf{I}, \qquad \sigma_e^2 = \frac{\sigma_z^2}{(P \mathrm{E}_P)} \tag{2.17}$$

Considering one time slot, we rewrite (2.10) with (2.13) as:

$$\begin{aligned} \mathbf{Y}_t &= \mathbf{H}\mathbf{X}_t + \mathbf{Z}_t \\ &= \left(\hat{\mathbf{H}} + \mathbf{E}\right) \mathbf{X}_t + \mathbf{Z}_t \\ &= \hat{\mathbf{H}} \, \mathbf{X}_t + \mathbf{E} \, \mathbf{X}_t + \mathbf{Z}_t \\ &= \hat{\mathbf{H}} \, \mathbf{X}_t + \mathbf{Z'}_t \end{aligned} \tag{2.18}$$

When the channel estimate $\hat{\mathbf{H}}$ is used in place of the perfect channel matrix $\mathbf{H}$, the corresponding noise term becomes $\mathbf{Z}'_t$. It consists of the additive noise component $\mathbf{Z}_t$ as well as the influence caused by the channel estimation error $\mathbf{EX}_t$. Thus, the power of $\mathbf{Z}'_t$ becomes [22]

$$\begin{aligned} \sigma^2_{Z'_t} &= \tfrac{1}{M} tr\{\mathbb{E}\{\mathbf{Z}'_t \mathbf{Z}'^H_t\}\} \\ &= \tfrac{1}{M} tr\{\mathbb{E}\{(\mathbf{Z}_t + \mathbf{EX}_t)(\mathbf{Z}_t + \mathbf{EX}_t)^H\}\} \\ &= \tfrac{1}{M} tr\{\mathbb{E}\{\mathbf{EE}^H\} + \sigma^2_Z\} \end{aligned} \tag{2.19}$$

where we assume that $\mathbf{E}_t$ is uncorrelated with $\mathbf{Z}_t$. From (2.19), we can see that the channel estimation error causes the overall noise power to increase, resulting in performance degradation.

Chapter 3

A Review of MIMO Detectors

In this chapter, we give a brief overview of the detectors that are of interest to our investigation, assuming, for the time being, that the channel is perfectly known at the receiver.

3.1 Maximum Likelihood Detector

The Maximum Likelihood (ML) detector performs a search over the entire set of possible candidates and chooses the one that minimizes

$$\hat{\mathbf{X}}_{ML} = \arg\min_{\mathbf{X}} ||\mathbf{Y} - \mathbf{HX}||^2. \tag{3.1}$$

The ML search is often exhaustive whence very time-consuming, especially for systems with a large number of transmit antennas or a high-order modulation. Even though the ML detector has the best performance among all MIMO detectors, its complexity makes it infeasible in practice.

3.2 Zero-Forcing (ZF) Detector

Linear detector is a class of simple approaches to recover the transmitted signal matrix, $\mathbf{X}$, from the received interference-corrupted signal matrix, $\mathbf{Y}$, by using an $N_T \times N_R$ weight matrix $\mathbf{P}$ to linearly combine (filter) the elements of $\mathbf{Y}$. The zero-forcing

(ZF) detector belongs to this class. By multiplying the received signal matrix $\mathbf{Y}$ with the Moore-Penrose pseudo-inverse [5] of the channel matrix $\mathbf{H}$, it attempts to null out the influence introduced by the channel. Thus, the ZF linear filtering (weight) matrix is

$$\mathbf{P}_{ZF} = \mathbf{H}^{\dagger} = (\mathbf{H}^H\mathbf{H})^{-1}\mathbf{H}^H. \tag{3.2}$$

Each element of the output matrix, $\mathbf{P}_{ZF}\mathbf{Y}$, is quantized to the nearest symbol in the constellation, $\mathcal{A}_M$, to obtain an estimate of the transmitted signal vector $\hat{\mathbf{X}}$, i.e.

$$\hat{\mathbf{X}}_{ZF} = \mathcal{Q}\{\,\mathbf{H}^{\dagger}\,\mathbf{Y}\,\} = \mathcal{Q}\{\,\mathbf{X} + (\mathbf{H}^H\,\mathbf{H})^{-1}\,\mathbf{H}^H\,\mathbf{n}\,\}. \tag{3.3}$$

where $\mathcal{Q}$ denotes the quantization operator.

The main advantage of a ZF detector is its low implementation complexity. However while the detector nulls out the spatial interference, it fails to take the noise into consideration and causes the power of the noise to boost up significantly, resulting in performance degradation. Only in the case of an orthonormal channel matrix, is the performance of the ZF detector identical to that of the optimum ML detector. Otherwise, ZF generally leads to noise enhancement, thus does not provide satisfactory performance.

3.3 The QRD-M Detector

The tree search based QRD-M algorithm attracts special attention as it achieves near-ML performance, while requiring substantially low complexity in comparison with the ML detector. The QRD-M algorithm selects only M candidate nodes with the smallest accumulated metrics at each level of the search tree, hence reduces the search complexity.

The first step of the QRD-M algorithm is to perform QR-decomposition on the channel matrix $\mathbf{H}$, in doing so we obtain

$$\mathbf{H} = \mathbf{Q}\hat{\mathbf{R}} \tag{3.4}$$

where $\mathbf{Q}$ is an $N_R \times N_R$ unitary matrix, and

$$\hat{\mathbf{R}} = \begin{bmatrix} \mathbf{R} \\ \mathbf{0}_{(N_R - N_T) \times N_T} \end{bmatrix} \tag{3.5}$$

$\mathbf{R}$ is an $N_T \times N_T$ upper triangular matrix and $\mathbf{0}_{(N_R - N_T) \times N_T}$ is a zero matrix of size $(N_R - N_T) \times N_T$. Note that $\mathbf{Q}^H \mathbf{Q} = \mathbf{I}$. Ignoring the Gaussian noise for a moment and pre-multiplying the equation $\mathbf{y} = \mathbf{H}\mathbf{x}$, where $\mathbf{x} = [x_1, \ldots, x_{N_T}]$ is the transmit signal vector, by $\mathbf{Q}^H$, we obtain

$$\tilde{\mathbf{y}} = \mathbf{R}\mathbf{x} + \tilde{\mathbf{z}} \tag{3.6}$$

where $\tilde{\mathbf{y}}$ is the first N_T rows of $\mathbf{Q}^H \mathbf{y}$ and $\tilde{\mathbf{z}}$ is the first N_T rows of $\mathbf{Q}^H \mathbf{z}$.

Based on (3.6), we can span a tree-like structure with depth N_T, as shown in Fig. (3.1). The process starts from the last element of $\mathbf{x}$, i.e., $x_i, i = N_T$, because from equation (3.6) we have $y_{N_T} = r_{N_T,N_T} x_{N_T} + n_{N_T}$, where $r_{i,i}$ is the (i,i)th element of $\mathbf{R}$, suggesting that it has no interference from the other antennas. The algorithm calculates the metric for all possible values of x_{N_T} from the constellation set of size $\mathcal{C}$ using [13]

$$| \tilde{y}_{N_T - i + 1} - R_{N_T - i + 1} \hat{\mathbf{x}}_i |^2 \tag{3.7}$$

where $\tilde{y}_i$ is the ith element of $\tilde{\mathbf{y}}$, $\mathbf{R}_i$ is the ith row of $\mathbf{R}$ and $\hat{\mathbf{x}}_i = [\hat{x}_{i+1}, \hat{x}_{i+2}, \ldots, \hat{x}_{N_T}]$ is the vector of estimated symbols of the specific survivor path. Only M branches with the smallest metrics are retained and the rest of the list is discharged. This procedure is applied to the nodes of the next level, and is repeated until a tree depth of N_T is reached, $i = 1$.

The general detection process is described in Table 3.1 [13].

Table 3.1: The QRD-M detection algorithm.

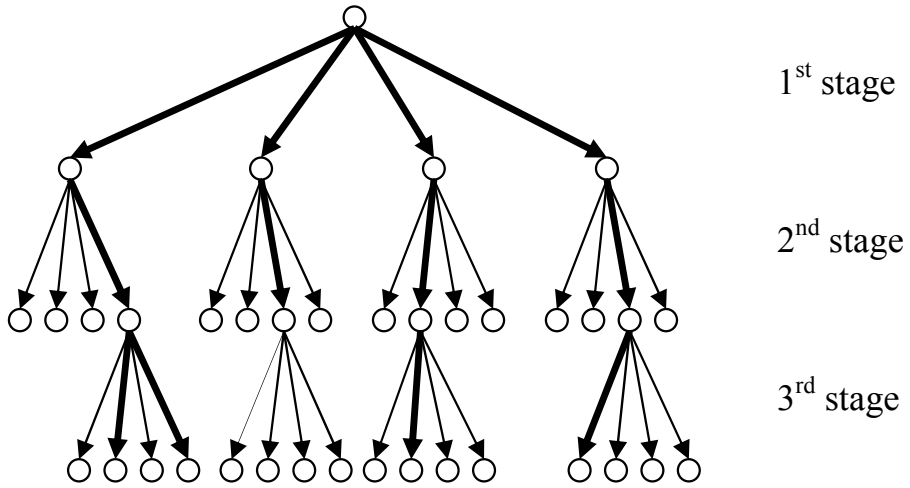

Figure 3.1: Tree structure of QRD-M ($M = 4$), $N_T = 3$, $N_R = 3$, with 4-QAM modulation.

Step 1 : **Perform QR-decomposition on the channel matrix H.**

Step 2 : **Premultiply the received vector y with $\mathbf{Q}^H$.**

Step 3 : **For every retained node, extend all branches to $\mathcal{C}$ nodes.**

Step 4 : **Calculate the branch metrics using equation (3.7).**

Step 5 : **Retain only M branches with the smallest metrics and delete the rest of the list.**

Step 6 : **Go to next level and return to Step 3.**

Chapter 4

Particle-Swarm-Driven Cross-Entropy MIMO Detector

4.1 The Cross-Entropy Method

In this section, we give a detailed description of the concepts used in the Particle-Swarm-Driven Cross-Entropy (PSD-CE) method. We consider a single time slot only; the extension to T time slots is straightforward. The PSD-CE method is a Monte Carlo based stochastic approach to solve

$$\arg \min_{\mathbf{x} \in \mathcal{A}_M^{N_T}} \| \mathbf{y} - \mathbf{H}\mathbf{x} \|^2. \tag{4.1}$$

As the name suggests, this method combines the ideas of both Particle Swarm Optimization (PSO) and the Cross-Entropy (CE) method. Let us first begin with the latter approach.

Assuming perfect channel knowledge at the receive end, we define the score function

$$\mathcal{S}(\mathbf{x}) = \| \mathbf{y} - \mathbf{H}\mathbf{x} \|^2. \tag{4.2}$$

In this situation, the optimal importance distribution $g^*(\mathbf{x})$ should be a peak at $\hat{\mathbf{x}}_{ML}$, i.e. $g^*(\mathbf{x}) = \delta(\mathbf{x} - \hat{\mathbf{x}}_{ML})$. Since we do not know $g^*(\mathbf{x})$, we want to find the parameter $\mathbf{v}$ of $f(\mathbf{x}; \mathbf{v})$ such that the cross-entropy between $g^*(\mathbf{x})$ and $f(\mathbf{x}; \mathbf{v})$ is minimum. The

cross-entropy, a convenient measure of distance between two distributions, e.g. $g(\mathbf{x})$ and $f(\mathbf{x})$, is defined as [14]

$$\mathcal{D}(g, f) = \mathbb{E}_g \ln \frac{g(\mathbf{X})}{f(\mathbf{X})} = \int g(\mathbf{x}) \ln g(\mathbf{x}) d\mathbf{x} - \int g(\mathbf{x}) \ln f(\mathbf{x}) d\mathbf{x} \tag{4.3}$$

Thus, our goal becomes

$$\arg\max_{\mathbf{v}} \left[\int g^*(\mathbf{x}) \ln f(\mathbf{x}; \mathbf{v}) d\mathbf{x} \right], \tag{4.4}$$

where $\{f(\,\cdot\,; \mathbf{v})\}$ is a family of importance distributions on $\mathcal{A}_M^{N_T}$. To solve the above problem, we first use another distribution $g'(\mathbf{x}; \mathbf{w})$ to replace $g^*(\mathbf{x})$, where

$$g^*(\mathbf{x}; \mathbf{w}) = \frac{f(\mathbf{x}; \mathbf{w})\, \mathbf{I}\{\mathcal{S}(\mathbf{x}) \leq \gamma\}}{c} \tag{4.5}$$

$\mathcal{S}(\mathbf{x}) = \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2$, $\mathbf{I}\{\,\cdot\}$ is an indicator function and c is a constant for normalization. Notice that if $f(\hat{\mathbf{x}}; \mathbf{w}) \neq 0$ and $\gamma = \mathcal{S}(\hat{\mathbf{x}})$, $g'(\mathbf{x}; \mathbf{w}) = g^*(\mathbf{x})$. Let γ be the smallest value such that, under the distribution $f(\mathbf{x}; \mathbf{w})$, $P_{\mathbf{w}}(\{\mathcal{S}(\mathbf{x}) \leq \gamma\}) \geq \rho$, where ρ is a predetermined parameter.

To find the smallest γ such that $P_{\mathbf{w}}(\{\mathcal{S}(\mathbf{x}) \leq \gamma\}) \geq \rho$ can be done using the Monte Carlo method. First draw U random samples, $\mathbf{x}_1, \ldots, \mathbf{x}_U$, from $f(\mathbf{x}; \mathbf{u})$ and evaluate their scores by $\mathcal{S}(\mathbf{x})$ respectively. Let γ be the $\lceil \rho U \rceil$th smallest score among all scores, then

$$P_{\mathbf{w}}(\{S(\mathbf{x}) \leq \gamma\}) \approx \frac{1}{U} \sum_{i=1}^{U} I\{S(\mathbf{x}_i) \leq \gamma\} = \frac{\lceil \rho U \rceil}{U} \geq \rho \tag{4.6}$$

Substitute $g'(\mathbf{x}; \mathbf{w})$ into (4.4), we obtain

$$\arg\max_{\mathbf{v}} \left[\int \frac{f(\mathbf{x}; \mathbf{w}) \mathbf{I}\{S(\mathbf{x}) \leq \gamma\}}{c} \ln f(\mathbf{x}; \mathbf{v}) d\mathbf{x} \right] \tag{4.7}$$

By using the U random samples drawn from $f(\mathbf{x}; \mathbf{w})$ to estimate (4.7),

$$\arg\max_{\mathbf{v}} \frac{1}{U} \sum_{i=1}^{U} \mathbf{I}\{S(\mathbf{x}_i) \leq \gamma\} \ln f(\mathbf{x}_i; \mathbf{v}) \tag{4.8}$$

we can obtain $\hat{\mathbf{v}}$ by solving

$$\frac{1}{U} \sum_{i=1}^{U} \mathbf{I}\{S(\mathbf{x}_i) \leq \gamma\} \frac{\partial}{\partial \mathbf{v}} \ln f(\mathbf{x}_i; \mathbf{v}) = 0 \tag{4.9}$$

Since the initial guess $f(\mathbf{x};\mathbf{w})$ may not be a well-approximated $g^*(\mathbf{x})$, we use $f(\mathbf{x};\hat{\mathbf{v}})$ to approximate $g^*(\mathbf{x})$ again. The above process is repeated to estimate another $\hat{\mathbf{v}}'$ and γ'. In fact, we will iteratively estimate the parameter $\mathbf{v}^{(k)}$, $\gamma^{(k)}$ and approximate the $g^*(\mathbf{x})$ by $f(\mathbf{x};\mathbf{v}^{(k)})$, where k is the iteration index. It is shown in [15] that the CE method converges, and if $\gamma^{(k)}$ converges to $\gamma^* = S(\hat{\mathbf{x}})$ in the fixed number of iterations, we can get an exact solution.

4.2 Particle Swarm Optimization

In [11], it was shown that the performance of a CE-based MIMO detection algorithm exhibits an error floor in the high SNR region. Obviously, to remove the error floor and improve the detector performance with small sample size, other elements need to be included.

Particle Swarm Optimization (PSO) is an optimization technique inspired by birds-flocking. A swarm algorithm consists of a number of possible 'particles' (or samples) that move through the feasible solution space to explore the optimal solution. Every particle keeps a record of the position of its best performance, called the individual best position. The position of the best performance among all particles is the global best position. The PSO concept consists of changing the velocity of each particle towards its individual best position and the global best position, in every iteration. This concept is added into the previously mentioned algorithm. In every iteration, the best sample is recorded (relates to the individual best position), as well as the best sample over all the iterations (the global best position), and their influences are added when updating the new distribution. The evolutional concept acts as a driving force, pulling out the samples that have sunk into the undesirable local minima and pushing them towards the global minimum.

The complete algorithm is listed below:

Step 1 : **Initialize the distribution $f^{(k)}(x_i)$ with uniform distribution for $i = 1, \cdots, N_T$, respectively. Set $k = 0$.**

Step 2 : **Generate U samples $x_{i,u}^k$ from $f^{(k)}(x_i)$ for $u = 1, \cdots, U$ and construct the set $\{\mathbf{x}_u^k\}_{u=1}^U$ where $\mathbf{x}_u^k = [x_{1,u}^k, \cdots, x_{i,u}^k, \cdots, x_{N_T,u}^k]^T$.**

Step 3 : **Calculate the set of scores $\{S(\mathbf{x}_u^k)\}_{u=1}^U$ according to $S(\mathbf{x}_u^k) = ||\mathbf{y} - \mathbf{H}\mathbf{x}_u^k||^2$.**

Step 4 : **Find a $\gamma^{(k)}$ satisfying $\gamma^{(k)} = \arg\min_{\gamma} P(S(\mathbf{x}) \leq \gamma) \geq \rho$ for $\mathbf{x} \in \{\mathbf{x}_u^k\}_{u=1}^U$. Then an elite set can be defined as $\{\mathbf{x}|S(\mathbf{x}) \leq \gamma^{(k)}, \mathbf{x} \in \{\mathbf{x}_u^k\}_{u=1}^U\}$.**

Step 5 : **Calculate the distribution of samples in the elite set:**

$$f_s^{(k)}(x_i = a) = \frac{\sum_{u=1}^U \mathbf{I}\{S(\mathbf{x}_u^k) \leq \gamma^{(k)}\}\mathbf{I}\{x_{i,u}^k = a\}}{\sum_{u=1}^U \mathbf{I}\{S(\mathbf{x}_u^k) \leq \gamma^{(k)}\}}, \tag{4.10}$$

where $a \in \mathcal{A}_M$ for $i = 1, \cdots, N_T$.

Step 6 : **Update the sample vector with the overall best score (from the 1st iteration to the kth iteration),$\mathbf{x}_{g(1)}^k$, and the best sample vector in the current iteration, $\mathbf{x}_{p(1)}^k$.**

Step 7 : **Update the distribution according to**

$$\begin{aligned} f^{(k+1)}(x_i = a) = \; & \alpha_1 f_s^{(k)}(x_i = a) + \alpha_2 f_{g(1)}^{(k)}(x_i = a) + \alpha_3 f_{p(1)}^{(k)}(x_i = a) \\ & + \left(1 - \textstyle\sum_{i=1}^3 \alpha_i\right) f^{(k)}(x_i = a) \end{aligned} \tag{4.11}$$

where α is the weighting factor and $0 \leq \alpha < 1$.

Step 8 : **Stop at iteration $k = K$ if the pre-defined stopping criterion is met; otherwise, let $k = k + 1$ and go back to Step 2.**

In (4.11), the weighting factors α_i are smoothing factors that take the value between 0 and 1. The best values are found empirically, so that the PSD-CE method achieves a balance between exploration and exploitation. $f_{g(1)}$ and $f_{p(1)}$ are distributions given in [11], based on $\mathbf{x}_{g(1)}$ and $\mathbf{x}_{p(1)}$, respectively.

4.3 Improving the PSD-CE Detector

Even with inclusion of a PSO element, the convergence speed of the PSD-CE method is still not fast enough. Thus, we propose a way to improve PSD-CE. In step 1 of the

algorithm above, the initial distribution, $f^{(0)}(x_i)$, is set to be uniformly distributed. This is the main reason of the slow convergence since the algorithm will need more iterations and samples to roam the entire solution space and converge at the best solution. A good initial distribution is crucial if a better performance or faster convergence is required.

Perform the Zero Forcing (ZF) method to find an initial solution, $\hat{\mathbf{x}}_{ZF}$. The reason we choose ZF is its low complexity. Although the ZF solution is usually not a very 'good' one, it still contains some information that can greatly improve the performance of the PSD-CE method.

The next problem is how to use the information drawn from $\hat{\mathbf{x}}_{ZF}$. The $\hat{\mathbf{x}}_{ZF}$ gives a rough idea about the whereabouts of the $\hat{\mathbf{x}}_{ML}$ so it would be a good decision to emphasize the search more in that area. Given the jth element of $\hat{\mathbf{x}}_{ZF}$, $\hat{x}_{ZF,j}$, where $j = 1, \ldots, N_T$, the new initial distribution is designed as

$$f_j^{(0)}(x_i) = \frac{D_i}{\sum_{k=1}^{M} D_k} \tag{4.12}$$

where

$$x_i \in \mathcal{A}_M, i = 1, \ldots, M \tag{4.13}$$

and

$$D_i = \frac{1}{\|\hat{x}_{ZF,j} - x_i\|^2}. \tag{4.14}$$

Since the summation of D_i for $i = 1, \ldots, M$ does not add up to 1, the denominator of (4.12) normalizes the probability distribution.

From (4.12), the probability of each entry x_i in the constellation is given as the reciprocal of the distance between $\hat{x}_{ZF,j}$ and x_i. In other words, the entries nearer to $\hat{x}_{ZF,j}$ are set to have larger probabilities, while those further have smaller probabilities. Simulation results show that with this new initial distribution, the performance of the PSD-CE method has been greatly improved.

The additional computational complexity required to set the new initial distribution to all the elements of the constellation is $(M-1)N_T$ additions and $17MN_T$ multiplica-

tions (here we assume that the complexity of division is equal to that of 8 mulplications), where M is the constellation size. The computational complexity needed to compute the Euclidean distance of a sample is $N_R(N_T + 1) - 1$ additions and $(N_T + 1)N_R$ multiplications. Taking a 4×4 MIMO system with 16-QAM for example, the new initial distribution requires 64 additions and 1088 multiplications, while calculating the Euclidean distance of one sample requires 19 additions and 20 multiplications. Changing the initial distribution is equivalent to adding about 55 samples only! Hence, this small additional complexity results in a great improvement in the performance, as can be seen in the simulation results in Chapter 7.

Chapter 5

Signal Detection With CSI Uncertainty

5.1 Effects of Channel Estimation Error

In the previous sections, the channel state information is assumed to be perfectly known at the receiver. However, in practical situations, pilot symbols are often employed to estimate the channel. Due to the finite number of pilots and noise, channel estimation errors are prone to exist. Hence, minimization of the Euclidean distance criterion is no longer optimum. Consideration of the errors of channel estimation has to be included in the optimal criterion. In most of the studies up until today, systems with channel estimation errors are modeled as

$$\hat{\mathbf{H}} = \mathbf{H} + \mathbf{E} \tag{5.1}$$

where $\mathbf{E}$ is the channel estimation error. However, according to the least squares method [25], the estimated channel should be orthogonal to the estimate errors. Hence, according to least squares, the channel should be modeled as

$$\mathbf{H} = \hat{\mathbf{H}} + \mathbf{E} \tag{5.2}$$

Under this assumption, we will derive the optimal criterion for the system in the presence of channel estimation errors.

Theorem 5.1.1. *[27] Denote* $\mathbf{z}_1$ *and* $\mathbf{z}_2$ *circularly symmetric complex Gaussian random vectors with zero means and full-rand covariance matrices* $\mathbf{\Sigma}_{ij} \triangleq \mathbb{E}[\mathbf{z}_i \mathbf{z}_j^\dagger]$. *Then, conditionally on* $\mathbf{z}_2$, $\mathbf{z}_1$ *is circularly symmetric complex Gaussian with mean* $\mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{z}_2$ *and covariance matrix* $\mathbf{\Sigma}_{11} - \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21}$.

We want to compute $p(\mathbf{Y}|\hat{\mathbf{H}}, \mathbf{X})$, and we know that

$$p(\mathbf{Y}|\hat{\mathbf{H}}, \mathbf{X}) = \prod_{i=1}^{N_R} p(\mathbf{Y}_i|\mathbf{X}, \hat{\mathbf{H}}_i). \tag{5.3}$$

Let $\mathbf{z}_1 = \mathbf{Y}_i^\dagger = \mathbf{X}^\dagger \mathbf{H}_i^\dagger + \mathbf{Z}_i^\dagger = \mathbf{X}^\dagger(\hat{\mathbf{H}}_i + \mathbf{E}_i)^\dagger + \mathbf{Z}_i^\dagger$ and $\mathbf{z}_2 = \hat{\mathbf{H}}_i^\dagger$, then apply Theorem 5.1.1 to get

$$\mathbf{\Sigma}_{11} = \mathbf{X}^\dagger\mathbf{X} + \sigma_z^2\mathbf{I}_T \tag{5.4}$$

$$\mathbf{\Sigma}_{12} = \mathbf{X}^\dagger\left(\mathbf{I}_{N_T} - \sigma_e^2\mathbf{I}_{N_T}\right) \tag{5.5}$$

$$\mathbf{\Sigma}_{22} = \mathbf{I}_{N_T} - \sigma_e^2\mathbf{I}_{N_T} \tag{5.6}$$

Therefore, $p(\mathbf{Y}_i^\dagger|\mathbf{X}, \hat{\mathbf{H}}^\dagger{}_i)$ is circularly symmetric complex Gaussian distributed with

$$\text{mean} = \mathbf{X}^\dagger\hat{\mathbf{H}}_i^\dagger \tag{5.7}$$

and

$$\text{covariance matrix} = \sigma_z^2\mathbf{I}_T + \sigma_e^2\mathbf{X}^\dagger\mathbf{X} \tag{5.8}$$

Finally, be multiplying $p(\mathbf{Y}_i|\mathbf{X}, \hat{\mathbf{H}}_i)$ from $i = 1, \ldots, N_R$ as in (5.3), we obtain

$$p\left(\mathbf{Y}|\,\mathbf{X}, \hat{\mathbf{H}}\right) = \frac{\exp\left\{tr\left\{-\left(\mathbf{Y} - \hat{\mathbf{H}}\mathbf{X}\right)\left[\,\sigma_z^2\,\mathbf{I}_T + \sigma_e^2\mathbf{X}^H\mathbf{X}\,\right]^{-1}\left(\mathbf{Y} - \hat{\mathbf{H}}\mathbf{X}\right)^H\right\}\right\}}{\det\left\{\pi\left[\sigma_z^2\mathbf{I}_T + \sigma_e^2\mathbf{X}^H\mathbf{X}\right]\right\}^{N_R}} \tag{5.9}$$

Take the logarithm of (5.9) and drop the constant terms to achieve

$$\begin{aligned}\hat{\mathbf{X}}_{ML} = \arg\min_{\mathbf{X}} \quad & N_R \ln\left\{\det\left\{\sigma_z^2\mathbf{I}_T + \sigma_e^2\mathbf{X}^H\mathbf{X}\right\}\right\} + \\ & tr\left\{\left(\mathbf{Y} - \hat{\mathbf{H}}\mathbf{X}\right)\left[\sigma_z^2\mathbf{I}_T + \sigma_e^2\mathbf{X}^H\mathbf{X}\right]^{-1}\left(\mathbf{Y} - \hat{\mathbf{H}}\mathbf{X}\right)^H\right\}\end{aligned} \tag{5.10}$$

Note that with perfect CSI, $P\mathrm{E}_P/\sigma_z^2 \to \infty$, $\sigma_e^2 \to 0$, we will get the original metric, i.e. $\|\mathbf{Y} - \hat{\mathbf{H}}\mathbf{X}\|^2$.

5.2 QRD-M detection with imperfect CSI

Consider the case of one time slot only (generalization to T time slots is straightforward), equation (5.10) will simplify to

$$\hat{\mathbf{x}}_{ML} = \arg \min_{\mathbf{x} \in \mathcal{A}_M^{N_T}} \left[N_R \ln \left(\sigma_z^2 + \sigma_e^2 \|\mathbf{x}\|^2 \right) + \frac{\|\mathbf{y} - \hat{\mathbf{H}}\mathbf{x}\|^2}{(\sigma_n^2 + \sigma_e^2 \|\mathbf{x}\|^2)} \right]. \tag{5.11}$$

In the QRD-M algorithm, we start the detection from the last layer, i.e., $i = N_T$. This is viable because $\mathbf{R}$ is an upper-triangular matrix hence the last element of the vector $\tilde{\mathbf{y}}$ can be expressed as

$$\tilde{y}_{N_T} = r_{N_T,N_T} x_{N_T} + \tilde{n}_{N_T} \tag{5.12}$$

where $r_{i,j}$ is the (i,j)th element of $\mathbf{R}$. $\tilde{y}_{N_T}$ does not contain interference from other antennas, so no information of the elements from the upper layers are needed. When detecting upper layers, only the elements of the previously extended paths are used.

However, in (5.11), the term $\|\mathbf{x}\|^2$ in both the logarithm and denominator of the fraction causes a problem when we want to apply QRD-M under imperfect channel state information. This is because information of the elements of all the other layers will be needed at every detection layer.

In this thesis, we provide a way to solve this problem. Observe that the QR-decomposition can be performed on $\hat{\mathbf{H}}$, i.e., $\hat{\mathbf{H}} = \hat{\mathbf{Q}}\hat{\mathbf{R}}$. Keeping in mind that $\hat{\mathbf{Q}}$ is a unitary matrix and $\hat{\mathbf{R}}$ is an upper-triangular matrix, the term $\|\mathbf{y} - \hat{\mathbf{H}}\mathbf{x}\|^2$ can still be expanded into a tree structure.

$$\begin{aligned} \|\mathbf{y} - \hat{\mathbf{H}}\mathbf{x}\|^2 &= \|\mathbf{y} - \hat{\mathbf{Q}}\hat{\mathbf{R}}\mathbf{x}\|^2 \\ &= \|\hat{\mathbf{Q}}^H\mathbf{y} - \hat{\mathbf{R}}\mathbf{x}\|^2 \end{aligned} \tag{5.13}$$

As the exact metric for a node cannot be obtained, we modify the optimization criterion in (5.11) by using a lower bound of the exact one.

Denote the metric to be minimized by $\Xi(\mathbf{x})$, i.e.

$$\Xi(\mathbf{x}) = N_R \ln(\sigma_n^2 + \sigma_E^2 \|\mathbf{x}\|^2) + \frac{\|\mathbf{y} - \hat{\mathbf{H}}\mathbf{x}\|^2}{(\sigma_n^2 + \sigma_E^2 \|\mathbf{x}\|^2)}. \tag{5.14}$$

Suppose now we are at the ith detection layer, and $\hat{\mathbf{x}}_i = [\hat{x}_i, \hat{x}_{i+1}, \ldots, \hat{x}_{N_T}]^T$, $i = 1, \ldots, N_T$, are the estimated symbols in the current and previous layers of the path in consideration. The lower bound for the metric in equation (5.14) in the ith detection layer is given as

$$LB(\Xi(\hat{\mathbf{x}}_i)) = N \ln(\sigma_n^2 + \sigma_E^2(\|\hat{\mathbf{x}}_i\|^2 + (i-1)A_{min})) + \frac{\|\mathbf{y} - \boldsymbol{\mu}(\hat{\mathbf{x}}_i)\|^2}{(\sigma_n^2 + \sigma_E^2(\|\hat{\mathbf{x}}_i\|^2 + (i-1)A_{max}))} \tag{5.15}$$

where A_{min} and A_{max} are the smallest power and the largest power of the symbol vectors among the constellation alphabets, taking 16-QAM for example, $A_{min} = 2$ and $A_{max} = 18$. $\boldsymbol{\mu}(\hat{\mathbf{x}}_i)$ is the multiplication of the corresponding $\hat{\mathbf{R}}$ and $\hat{\mathbf{x}}_i$. Think of equation (5.15) as equation (5.14) under the best conditions in every detection layer.

The new QRD-M based detector begins from the first detection layer, $i = N_T$. The metrics for all possible values of $\hat{x}_i$ are calculated with equation (5.15). Only M nodes with the smallest metrics are kept, and the rest of the list is discarded. The same steps are applied to the nodes of the next layer, and is done iteratively until $i = 1$. Finally, the path with the smallest metric is the estimated signal vector.

Chapter 6

Space Time Code

6.1 Background

In wireless communication systems with multiple antennas, we use space time codes (STC) to improve the transmission reliability, obtaining both diversity and coding gains. STC is based on joint encoding across space and time domains. STCs transmit redundant copies of the data stream to the receiver through multiple transmit antennas (transmit diversity). The transmitted signals will traverse different environments with scattering, reflection, and so on, resulting in different copies of the data at the receiver end. Some of the copies are better than others, meaning that they are less faded and thus provide more reliable information. In fact, the receiver combines all copies of the received signal in an optimal way to extract as much information from each of them as possible.

6.2 Alamouti Space-Time Block Code

In this section, we introduce the Alamouti space-time block code [19]. It is a simple orthogonal space-time block code with a maximum-likelihood decoding algorithm detector. It is shown in [19] that with 2 transmit antennas, the scheme provides a diversity order of $2N_R$.

Consider a system with $N_T = N_R = 2$ employing the transmit-diversity scheme of

Alamouti. The Alamouti codeword is given by

$$\mathbf{C} = \begin{bmatrix} s_1 & -s_2^* \\ s_2 & s_1^* \end{bmatrix} \tag{6.1}$$

In the case of BPSK, $s_1, s_2 \in \{-1, 1\}$. This scheme requires two signaling periods to convey a pair of symbols s_1 and s_2. During the first time slot, the symbols s_1 and s_2 are simultaneously transmitted from the first and second transmit antennas, respectively; then, in the next time slot, $-s_2^*$ and s_1^* are transmitted from the respective transmit antennas.

The received signal matrix will be

$$\begin{bmatrix} y_{1,1} & y_{1,2} \\ y_{2,1} & y_{2,2} \end{bmatrix} = \begin{bmatrix} h_{1,1} & h_{1,2} \\ h_{2,1} & h_{2,2} \end{bmatrix} \begin{bmatrix} s_1 & -s_2^* \\ s_2 & s_1^* \end{bmatrix} + \mathbf{Z} \tag{6.2}$$

where

$$\begin{aligned} y_{1,1} &= h_{1,1}s_1 + h_{1,2}s_2 + z_{1,1} \\ y_{1,2} &= h_{1,2}s_1^* - h_{1,1}s_2^* + z_{1,2} \\ y_{2,1} &= h_{2,1}s_1 + h_{2,2}s_2 + z_{2,1} \\ y_{2,2} &= h_{2,2}s_1^* - h_{2,1}s_2^* + z_{2,2} \end{aligned}$$

For operation convenience, we rewrite (6.2) as

$$\begin{bmatrix} y_{1,1} \\ y_{1,2}^* \\ y_{2,1} \\ y_{2,2}^* \end{bmatrix} = \begin{bmatrix} h_{1,1} & h_{1,2} \\ h_{1,2}^* & -h_{1,1}^* \\ h_{2,1} & h_{2,2} \\ h_{2,2}^* & -h_{2,1}^* \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} + \mathbf{Z}' \tag{6.3}$$

where it can be written with the more common notation

$$\mathbf{Y}_{equi} = \mathbf{H}_{equi}\mathbf{X}_{equi} + \mathbf{Z}' \tag{6.4}$$

Under the circumstances that the channel $\mathbf{H}$ is perfectly known at the receiver, the maximum likelihood (ML) detector can be derived to be

$$\begin{bmatrix} h_{1,1} & h_{1,2} \\ h_{1,2}^* & -h_{1,1}^* \\ h_{2,1} & h_{2,2} \\ h_{2,2}^* & -h_{2,1}^* \end{bmatrix}^\dagger \begin{bmatrix} y_{1,1} \\ y_{1,2}^* \\ y_{2,1} \\ y_{2,2}^* \end{bmatrix} = \begin{bmatrix} hs & 0 \\ 0 & hs \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} + \begin{bmatrix} h_{1,1} & h_{1,2} \\ h_{1,2}^* & -h_{1,1}^* \\ h_{2,1} & h_{2,2} \\ h_{2,2}^* & -h_{2,1}^* \end{bmatrix}^\dagger \mathbf{Z}' \tag{6.5}$$

where $hs = |h_{1,1}|^2 + |h_{1,2}|^2 + |h_{2,1}|^2 + |h_{2,2}|^2$. Note that this is also a zero-forcing (ZF) detector, hence one of the many advantages of Alamouti coding is its easy decoding.

It is interesting that when employing Alamouti codes with any constant envelope modulation, such as phase shift keying (PSK), the *mismatched* detector is exactly (6.5). The *mismatched* detector is expressed as

$$\begin{aligned}
&\arg\min_{\mathbf{X}} \left\| \mathbf{Y}_{equi} - \hat{\mathbf{H}}_{equi}\mathbf{X}_{equi} \right\|^2 \\
&= \arg\min_{\mathbf{X}} \left(\mathbf{Y}_{equi} - \hat{\mathbf{H}}_{equi}\mathbf{X}_{equi} \right)^* \left(\mathbf{Y}_{equi} - \hat{\mathbf{H}}_{equi}\mathbf{X}_{equi} \right) \\
&= \arg\min_{\mathbf{X}} \mathbf{Y}_{equi}^*\mathbf{Y}_{equi} + \left(\mathbf{X}_{equi}\hat{\mathbf{H}}_{equi} \right)^* \hat{\mathbf{H}}_{equi}\mathbf{X}_{equi} - 2\mathrm{Re}\left[\mathbf{Y}_{equi}^* \left(\mathbf{H}_{equi}\mathbf{X}_{equi} \right) \right]
\end{aligned} \tag{6.6}$$

Since $\mathbf{Y}_{equi}^*\mathbf{Y}_{equi}$ is a constant term, when trying to find the argument that minimizes a function, it can be omitted. The second term of the equation can be calculated as

$$\begin{aligned}
\left(\mathbf{X}_{equi}\hat{\mathbf{H}}_{equi} \right)^* \hat{\mathbf{H}}_{equi}\mathbf{X}_{equi} &= \mathbf{X}_{equi}^*\hat{\mathbf{H}}_{equi}^*\hat{\mathbf{H}}_{equi}\mathbf{X}_{equi} \\
&= hs\mathbf{X}_{equi}^*\mathbf{X}_{equi} \\
&= c\,\mathbf{I}
\end{aligned} \tag{6.7}$$

where c is a constant, therefore, the second term in (6.6) can also be ignored. That leaves us with the last term in the equation. Denote $\mathbf{X}_{ZF}$ the solution found using the ZF method, i.e. $\mathbf{X}_{ZF} = \mathbf{H}_{equi}^*\mathbf{Y}_{equi}$.

$$\begin{aligned}
\arg\min_{\mathbf{X}} -2\mathrm{Re}\left[\mathbf{Y}_{equi}^* \left(\mathbf{H}_{equi}\mathbf{X}_{equi} \right) \right] &= \arg\max_{\mathbf{X}} \mathrm{Re}\left[\left(\mathbf{H}_{equi}^*\mathbf{Y}_{equi} \right)^* \mathbf{X}_{equi} \right] \\
&= \arg\max_{\mathbf{X}} \mathrm{Re}\left[\mathbf{X}_{ZF}^*\mathbf{X}_{equi} \right]
\end{aligned} \tag{6.8}$$

Take QPSK symbols for example,

$$\mathrm{Re}\left[x_{ZF,i}^* x_{equi,i} \right] = \mathrm{Re}[x_{ZF,i}^*]\mathrm{Re}[x_{equi,i}] + \mathrm{Im}[x_{ZF,i}^*]\mathrm{Im}[x_{equi,i}], \quad i = 1, \ldots, N_T \tag{6.9}$$

In order to find the solution, the multiplication of the real parts has to be positive. The

same is for the multiplication of the imaginary parts. We get the following equations:

$$\text{If } \mathrm{Re}[x_{ZF,i}] = \quad \mathrm{Re}[x^*_{ZF,i}] > 0, \quad \mathrm{Re}[x_{equi,i}] = 1 \tag{6.10}$$

$$\text{If } \mathrm{Re}[x_{ZF,i}] = \quad \mathrm{Re}[x^*_{ZF,i}] < 0, \quad \mathrm{Re}[x_{equi,i}] = -1 \tag{6.11}$$

$$\text{If } \mathrm{Im}[x_{ZF,i}] = -\mathrm{Im}[x^*_{ZF,i}] > 0, \quad \mathrm{Im}[x_{equi,i}] = 1 \tag{6.12}$$

$$\text{If } \mathrm{Im}[x_{ZF,i}] = -\mathrm{Im}[x^*_{ZF,i}] < 0, \quad \mathrm{Im}[x_{equi,i}] = -1 \tag{6.13}$$

The above equations are exactly the ZF solutions! Thus, it is proved that when using the Alamouti code with QPSK (and can be generalized to any constant envelope modulations), the *mismatched* detector is equivalent to the ML detector.

However, when the modulation does not have a constant envelope, such as $16-QAM$, the *mismatched* detector of (6.5) is no longer optimal. In such cases, the ML decoding criterion is still given by (5.10).

6.3 Hamming based Space-Time Block Code

Richard Hamming proposed an error-correcting code that could correct a 1-bit error and detect 2-bit errors. This code, a binary linear error-correcting code, is named Hamming code [24] after its inventor. Let n be the length of the codeword, and k be the length of data bit, then it is apparent that $n-k$ redundant bits are used for error checking. This is often referred to as the (n,k) code.

Denote the $k \times n$ matrix $\mathbf{G}$ the generator matrix of a linear (n,k) code and let $\mathbf{H}$ denote the $(n-k) \times n$ parity check matrix. $\mathbf{G}$ and $\mathbf{H}$ for linear block codes must satisfy

$$\mathbf{H}\mathbf{G}^H = \mathbf{0}. \tag{6.14}$$

For Hamming code, to encode a k-bit data $\mathbf{u}$, the corresponding n-bit codeword $\mathbf{x}$ can be obtained by

$$\mathbf{x} = \mathbf{u}\mathbf{G} \tag{6.15}$$

In this thesis, we consider the Hamming based space-time block code. The rationality behind choosing the Hamming based space-time block code is mentioned in [20].

The generating Hamming based space-time block code composed of two steps. First, the Hamming codeword is generated. Then, the codeword is multiplexed into two streams. The performance of Hamming based space-time block code depends on the multiplexing scheme and the optimal multiplexing scheme can be found through simulations.

Specifically, we consider a (8,4) binary Hamming based space-time block code. The generator matrix of (8,4) binary Hamming code is

$$\mathbf{G} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \end{bmatrix} \tag{6.16}$$

After encoding the data into a 8-bit codeword, $\mathbf{x} = [x_1, \ldots, x_8]$, we multiplex the 8 bits into two streams. A possible multiplexing scheme is to fill in the first row and then the second row. Hence, the generated codeword is

$$\mathbf{X} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \\ x_5 & x_6 & x_7 & x_8 \end{bmatrix} \tag{6.17}$$

However, we can also fill in the column first and the generated codeword is

$$\mathbf{X} = \begin{bmatrix} x_1 & x_3 & x_5 & x_7 \\ x_2 & x_4 & x_6 & x_8 \end{bmatrix} \tag{6.18}$$

The performance of these two multiplexing schemes and related comparison will be provided in Chapter **??**.

Chapter 7

Simulation Results

This chapter is divided into two parts: In the first part, we present computer-simulated performance some conventional MIMO detectors as well as the PSDCE and the modified PSDCE detectors for a system with $N_T = 4$ transmit antennas, $N_R = 4$ receive antennas and perfect CSI. In the second part, simulation results for the performance of detectors with modified ML criterion based on the nonzero channel estimation error are presented. We consider spatial multiplexing, Alamouti-coded and Hamming-coded MIMO systems.

7.1 Performance Comparison with Perfect CSI

In Fig. 7.1, the BER performance comparison of some known detectors, such as ZF, V-BLAST, PSD-CE and modified PSD-CE, is shown. In this simulation, we generated 60 samples in every iteration for only 3 iterations, for both the PSD-CE detector and the modified PSD-CE detector using 4-QAM. We can see that the performance of the conventional PSD-CE is better than that of ZF and V-BLAST. It is also shown that the performance is further improved by the modified PSD-CE detector. Comparison between the BER performance of the conventional PSD-CE and the modified PSD-CE with 16-QAM is shown in Fig. 7.2. The solid lines are performances of the cases that run 5 iterations, and in every iteration 500 samples are generated. The proposed algorithm outperforms the conventional method by more than 5dB at BER= 10^{-2}. The dotted line

shows the performance of the modified PSD-CE with only 3 iterations, also generating 500 samples in each iteration. We can see that the performance of the modified PSD-CE even outperforms the conventional PSD-CE with a larger sample size! Therefore, it is obvious that the modified PSD-CE improves the performance, as well as decreases the complexity of the original PSD-CE method.

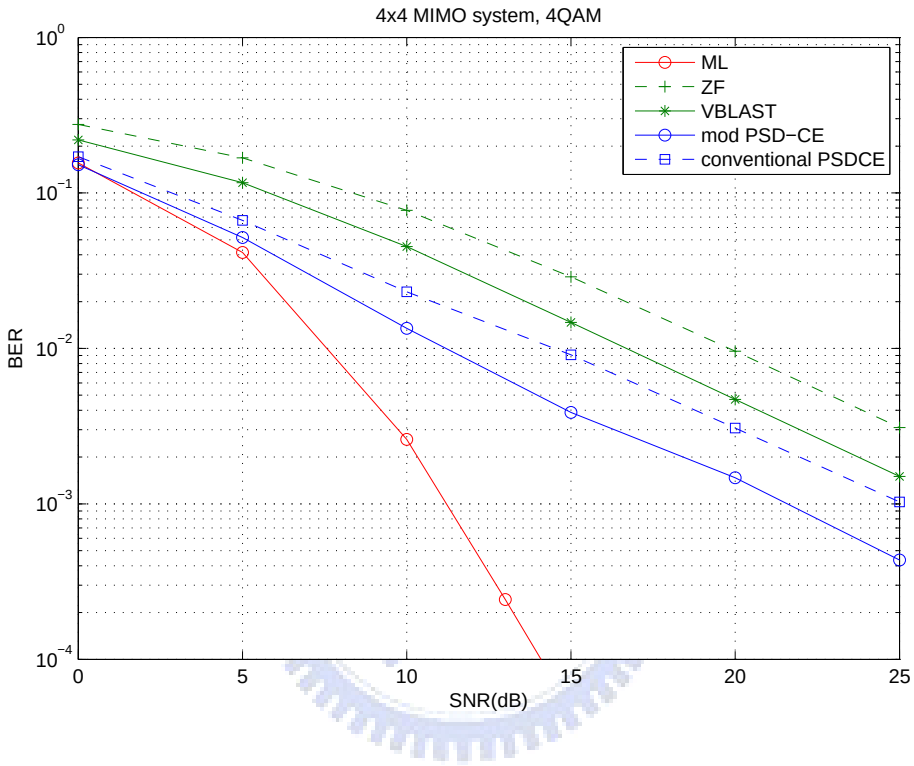

Figure 7.1: Comparison of the BER performance of the modified PSD-CE detector (3 iterations and 60 samples/iteration) and other known detectors in a MIMO system with $N_T = N_R = 4$ using 4-QAM.

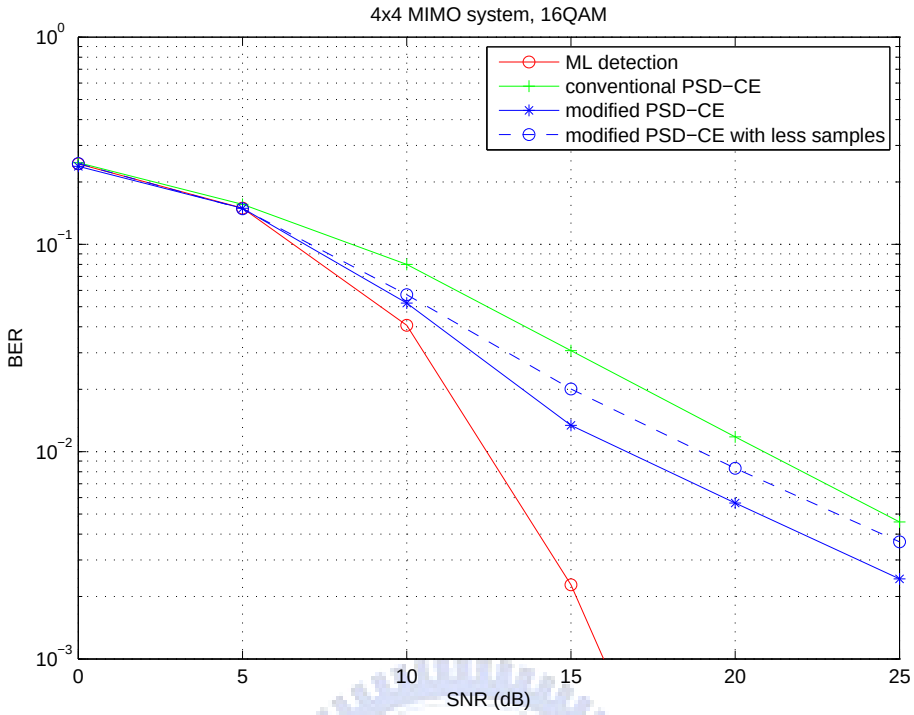

Figure 7.2: Comparison of the BER performance of the conventional PSD-CE detector and the modified PSD-CE in a MIMO system with $N_T = N_R = 4$ using 16-QAM. solid line : 5 iterations and 500 samples/iteration; dashed line : 3iterations and 500 samples/iteration.

7.2 Performance in the Presence of CSI Error

In this part of the simulations, we examine the influence of a system in the absence of perfect CSI at the receiver. Consider a MIMO system with $N_T = N_R = 4$ and 16-QAM modulation. Figure 7.3 shows the frame error rate (FER) versus SNR in dB for different values of $(S/N)_P \triangleq PE_P/\sigma_z^2$ with the *mismatched* metric, (2.19).

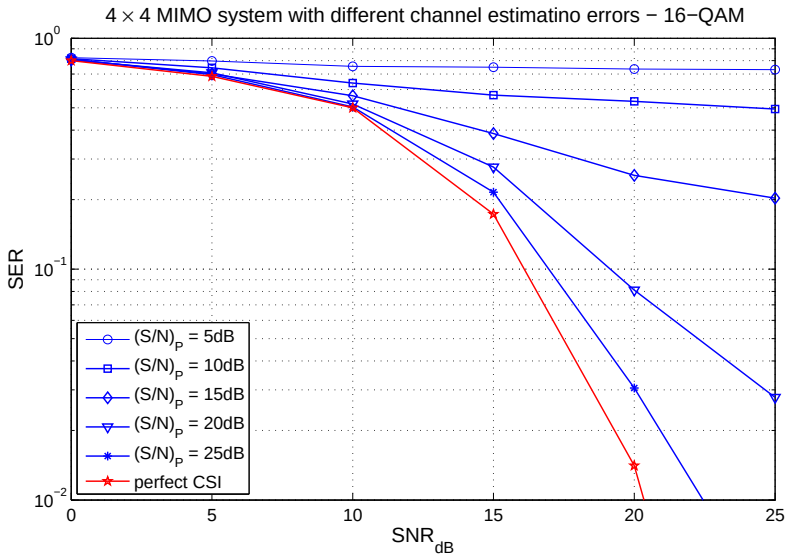

Figure 7.3: Effect of channel estimation errors on the system performance.

7.2.1 Spatial multiplexing system

In this section, we consider the performance of spatial multiplexing (SM) system which uses two transmit antennas to transmit two independent data streams. For a 2×2 MIMO system, instead of a $N_R \times 1$ vector, let $\mathbf{X}$ be a $N_T \times T$ matrix. Figures 7.4 and 7.5show the performances of $T = 2$ and 4 using the modified ML detector (5.10), respectively. The rate-2 code here implies that there is no actual 'coding', as it does not provide error correcting or detecting functions.

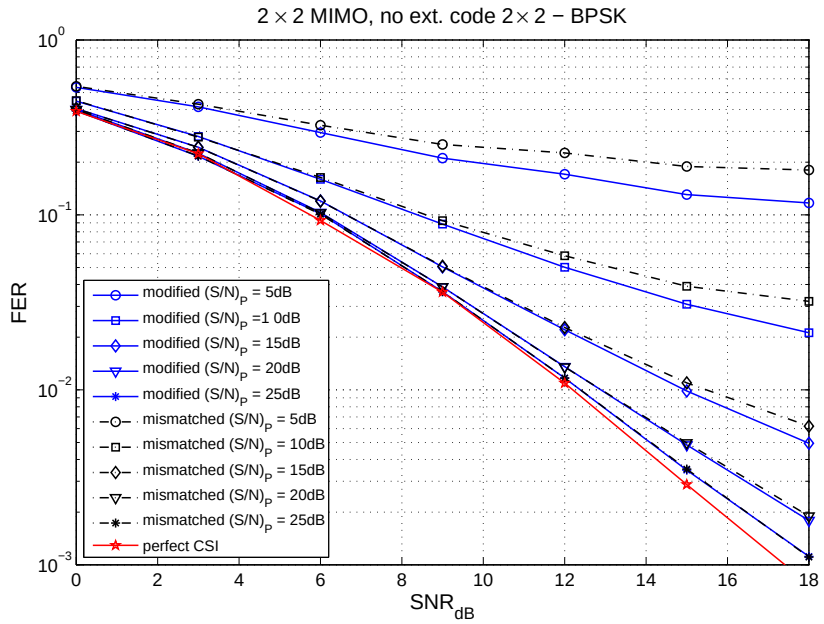

Figure 7.4: Effect of channel estimation errors on a 2×2 MIMO with $T = 2$ using BPSK.

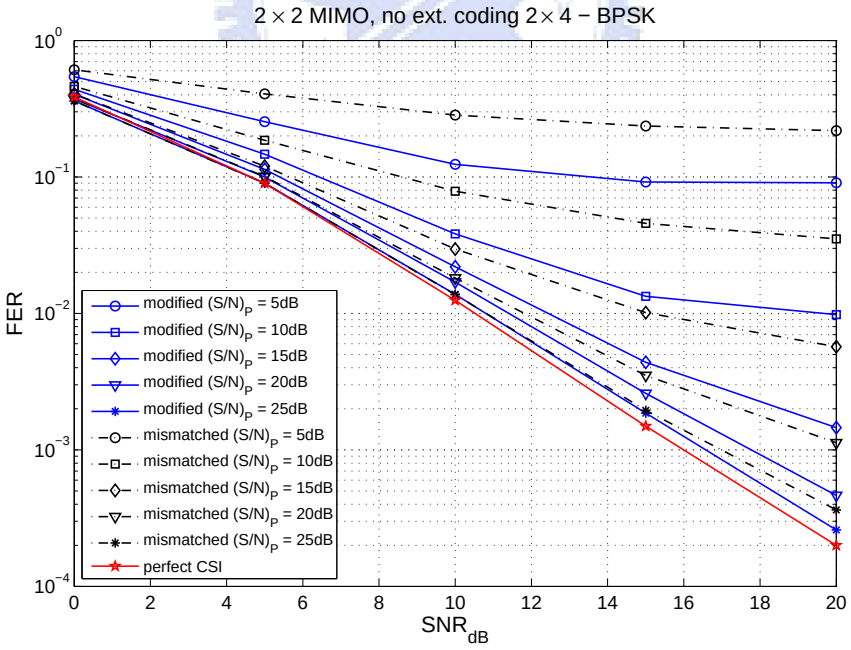

Figure 7.5: Effect of channel estimation errors on a 2×2 MIMO with $T = 4$ using BPSK.

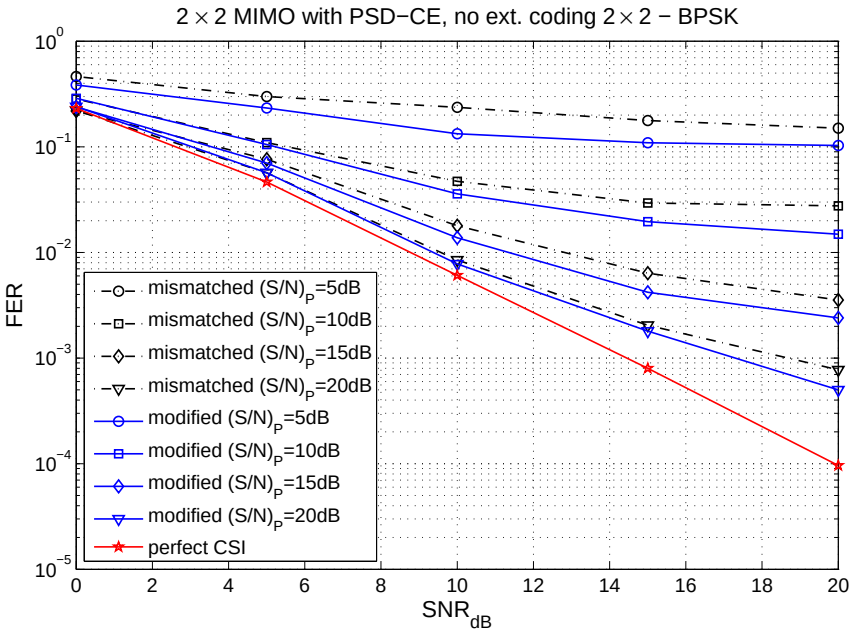

Figure 7.6: The use of PSDCE detector under channel estimation errors in a 2×2 MIMO with $T = 2$ using BPSK.

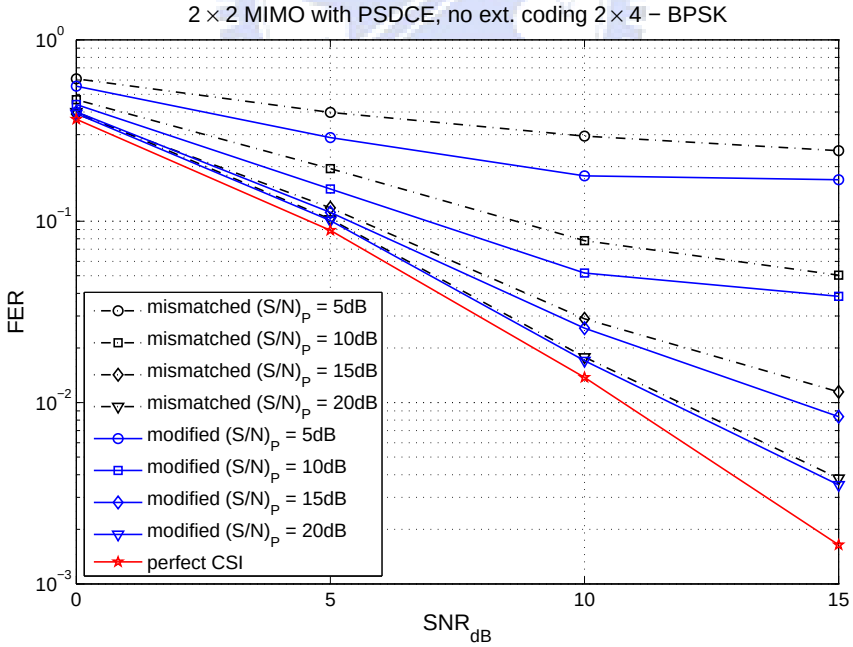

Figure 7.7: The use of PSDCE detector under channel estimation errors in a 2×2 MIMO with $T = 4$ using BPSK.

7.2.2 Alamouti space-time coded system

In this section, Alamouti code is employed in a 2×2 MIMO system using 16-QAM, i.e. $T = 2$. Recall that the ML performance of the system using constant envelope constellation is equal to that of the *mismatched* detector. Figure 7.8 shows comparison between the performances of the previously described system and the *mismatched* system. The gain is small but certain.

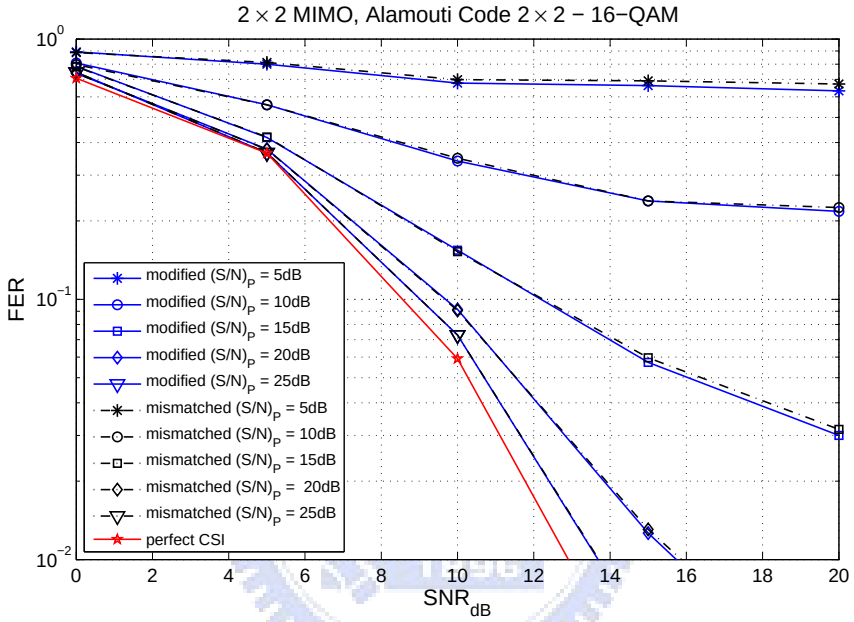

Figure 7.8: Employing Alamouti code in a 2×2 systme with estimation errors with $T = 4$ using 16-QAM.

7.2.3 Hamming coded system

Here, we consider a space-time block code obtained by mapping a $(8, 4)$ binary Hamming code to 2×4 BPSK codewords. The 8-bit Hamming codeword $\mathbf{c} = [c_1, \ldots, c_8]$ is mapped to the BPSK symbols like so

$$\mathbf{x} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \\ x_5 & x_6 & x_7 & x_8 \end{bmatrix} \tag{7.1}$$

where $x_i = (-1)^{c_i}$. In the first time slot, x_1 and x_5 are transmitted, then x_2 and x_6 are transmitted in the second time slot and so on. The results obtained is shown in figure

7.9.

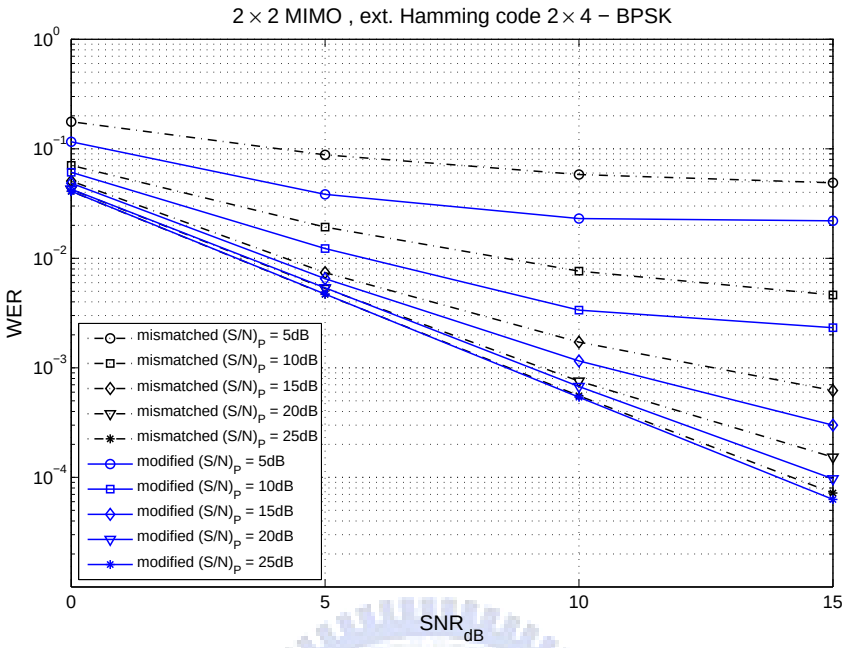

Figure 7.9: Employing Alamouti code in a 2×2 systme with estimation errors with $T = 4$ using 16-QAM.

Chapter 8

Conclusion

In this thesis, we investigate the optimum criterion of MIMO signal detection in the presence of channel estimation errors. Most existing studies on this subject assumed that the true channel matrix is independent of the associated error matrix. However, for least square or minimum mean square error channel estimators, the orthogonal principle implies that the channel estimator should be uncorrelated with the estimation error. Therefore, it is more appropriate to use such an assumption accordingly.

Although the optimal (ML)@detector can be derived, its high complexity makes it infeasible in practice. To overcomes this difficulty, we consider two sub-optimal detector structures that take imperfect CSI into account. The first suboptimal detector, referred to as Particle-Swarm-Driven Cross-Entropy (PSD-CE) detector, is a stochastic search based detection scheme. Its performance depends on the setting of the initial distribution and uniform distribution is often used for lack of a priori information. We propose a method to obtain the initial distribution which then leads to performance and complexity improvements. The other suboptimal detector is the modified QRD-M detector.

Since our design criterion is no longer equivalent to the minimum Euclidean distance criterion, the original QRD-M method cannot be used directly. We propose a low-complexity QRD-M detection method, using a new decoding metric which is derived from our assumption on the uncorrelatedness between the estimated channel matrix and the estimation error matrix. Finally, we extend our study to space-time block

coded MIMO systems and examine their performance using the new decoding metric.

In all cases under investigation, we show by computer simulated numerical examples that the proposed decoding metric does offer performance improvement over that achieved by using conventional metric that does not consider the imperfect CSI effect.

Bibliography

[1] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*, Cambridge University Press, 2005.

[2] E. Biglieri, R. Calderbank, A. Constantinides, A. Goldsmith, A. Paulraj and H. V. Poor, *MIMO Wireless Communications*, Cambridge University Press, 2007.

[3] I. Telatar, "Capacity of multiple antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, no. 6, pp. 585-595, Nov/Dec, 1999.

[4] R. Xu, F.C.M. Lau, "Performance analysis for MIMO systems using zero forcing detector over fading channels," *IEE Proc. Commun.*, Vol. 153, No. 1, February 2006.

[5] G. H. Golub and C. F. Van Loan, *Matrix Computations*, Johns Hopkins University Press, Baltimore, MD, 1983.

[6] P. W. Wolniansky, G. J. Foschini, G. D. Golden and R. A. Valenzuela, "V-BLAST: An architecture for realizing very high data rates over the rich-scattering wireless channel," *Proc. of IEEE Int. Symposium on Signals, Systems and Electronics (ISSSE'98)*, Pisa, Italy, Sep. 1998.

[7] R. Narasimhan, "Error propagation analysis of VVBLAST with channel estimation errors," *IEEE Trans. Commun.*, vol. 53, no. 1, pp. 27-31, Jan. 2005.

[8] V.Tarokh, A.Naguib, N.Seshadri, and A.R.Calderbank, "Space-time codes for high data rate wireless communication: Performance criteria in the presence of channel

estimation errors, mobility, and multiple paths," *IEEE Trans. Commun.*, vol.47, no.2,pp.199-207,Feb.1999.

[9] P. Hoeher and F. Tufvesson, "Channel estimation with superimposed pilot sequence," in *Proc. IEEE GLOBECOM'99*, vol. 4, pp. 2164-2166, 1999.

[10] M. Biguesh and A. B. Gershman, "Training based MIMO channel estimation: A study of estimator tradeoffs and optimal training signals," *IEEE Trans. Sig. Proc.*, vol.54, no.3, pp. 884-893, March 2006.

[11] C. L. Wang, Y. T. Lin and Y. T. Su, "Particle-Swarm-Driven Cross-Entropy Method for MIMO Signal Detection," *IEEE WCNC 2010.*

[12] J.Yue, K.J.Kim, J.D.Gibson, and R.A.Iltis, "Channel Estimation and Data Detection for MIMO-OFDM systems," in Proc, GLOBECOM 2003, pp.581-585,2003.

[13] Chin,W.H. "QRD based tree search data detection for MIMO communication systems", *2005 IEEE 61st Vehicular Technology Conference*, Spring 2005.

[14] R. Y. Rubinstein and D. P. Kroese, *The Cross-Entropy Method*, Springer, 2004.

[15] L. Margolin, "On the convergence of the cross-entropy method", *Annals of Operations Research*, vol. 134, no.1, pp. 201-214, 2005.

[16] J. Kennedy and R. C. Eberhart, "Particle swarm optimization", *Proc. of the IEEE Int. Joint Conf. on Neural Networks,* pp. 1942-1948, 1995.

[17] F. Heppner and U. Grenander, "A stochastic nonlinear model for coordinated bird flocks", *S. Krasner, editor, The Ubiquity of Chaos,* AAAS Publications, Washington, DC, 1990.

[18] V. Tarokh, N. Sheshadri and A. R. Calderbank, "Space-time codes for high data rate wireless communications: Performance criteria and code construction," *IEEE Trans. Inf. Theory*, vol. 44, pp. 744-765, Mar 1998.

[19] S. M. Alamouti, "A Simple Transmit Diversity Technique for Wireless Communications", *IEEE Journal on Select Areas in Commun.*, vol. 16, no. 8, Oct 1998.

[20] G. Taricco and E. Biglieri, "Space-Time Decoding With Imperfect Channel Estimation", *IEEE Trans. on Wireless Communications*, vol. 4, no. 4, July 2005.

[21] B. Kim and K. Choi, "A Very Low Complexity QRD-M Algorithm Based on Limited Tree Search for MIMO Systems", *2008 IEEE 67th Vehicular Technology Conference*, Marina Bay, Singapore, pp. 1246-1250, May 2008.

[22] A. A. Farhoodi and M. Fazaelifar, "Sphere Detection in MIMO Communication Systems with Imperfect Channel State Information," *Proceedings of the Communication Networks and Services Research Conference*, pp. 228-233, May 2008.

[23] C. E. Shannon, "A Mathematical Theory of Communication", *Bell System Technical Journal*, pp. 379-423(Part 1); pp. 623-656(Part 2), July 1948.

[24] R. W. Hamming, "Error Detecting and Error Correcting Codes", *Bell System Technical Journal*, pp. 147-161, April 1950.

[25] S. M. Kay, *Fundamentals of Statistical Signal Processing : Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993, vol. I.

[26] O. Weikert, U. Zölzer, "Efficient MIMO channel estimation with optimal training sequences," *Proc. of 1st Workshop on Commercial MIMO-Components and-Systems (CMCS 2007),* Duisburg, Germany, Sep. 2007.

[27] M. Bilodeau and D. Brenner, *Theory of Multivariate Statistics*. New York: Springer, 1999.

作者簡歷

林玉婷
民國七十四年十二月二十三日出生。
從小因為父親工作的關係，全家搬到馬來西亞。
民國九十三年畢業於檳城台灣學校，擔任多次班長、風紀股長、衛生股長等，永遠考前五名因為班上只有５人。
民國九十七年畢業於國立交通大學電信工程學系，一年級加入交大羽球校隊，從此便愛上了羽球。經過苦練，四年級之後成為大專盃團體賽上場球員及梅竹上場球員，多次代表學校比賽，感到無比光榮。
同年推甄進入國立交通大學電信工程研究所碩士班，師蘇育德博士，兩年來學到了很多，對老師的感激無法言喻。
國九十九年十月獲得碩士學位。

Graduate Courses：

1. Coding Theory(I)
2. Coding Theory(II)
3. Wireless Communications
4. Digital Communications
5. Digital Signal Processing
6. Adaptive Signal Processing
7. Information Theorem
8. Stochastic Processes