

國立交通大學

電信工程學系

碩士論文

盲蔽式量子濾波器等化技術在頻率選擇性
通道下之演算法與效能分析

Particle Filter Based Blind Equalization for
Frequency Selective Channels

— Algorithms and Performance Analysis

研究生：何恩慶

指導教授：紀翔峰 博士

中華民國九十五年九月

盲蔽式量子濾波器等化技術在頻率選擇性通道下之

演算法與效能分析

**PARTICLE FILTER BASED BLIND EQUALIZATION FOR
FREQUENCY SELECTIVE CHANNELS**

— ALGORITHMS AND PERFORMANCE ANALYSIS

研究生：何恩慶

Student：An-Ching Ho

指導教授：紀翔峰博士

Advisor：Dr. Hsiang-Feng Chi

國立交通大學

電信工程學系碩士班

碩士論文

1896

A Thesis

Submitted to Department of Communication 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 Communication Engineering

September 2006

Hsinchu, Taiwan, Republic of China

中華民國九十五年九月

盲蔽式量子濾波器等化技術在頻率選擇性通道下之 演算法與效能分析

研究生：何恩慶

指導教授：紀翔峰 博士

國立交通大學

電信工程學系碩士班



摘要:

近年來，有關於量子濾波器 (Particle Filter) 在系統通道等化器之應用問題已經在很多論文當中引起廣泛的討論。如同參考資料[3]中所提到，當通道有一個很微弱的直線傳輸路徑(light of sight, LOS)時，連續關鍵取樣(Sequential Importance Sampling, SIS)等化器的系統效能會被嚴重的影響而衰退。在參考資料[3]中同時也提出了一種延遲連續關鍵取樣演算法(Delayed SIS Algorithm)來試圖解決這個問題。然而這種演算法需要大量的運算複雜度來提升系統的效能，因此在本篇論文中我們提出一種新的盲蔽式(blind)連續關鍵取樣演算法，該演算法不會受通道直線傳輸路徑的強弱影響而衰退。

在論文中，我們首先簡單的介紹量子濾波器的理論，並建立起量子濾波等化器的系統架構。接著透過錯誤率(bit error rate, BER)的數學分析，我們可以了解到一個擁有微弱的直線傳輸路徑的通道是如何的影響連續關鍵取樣等化器的效能。為了克服這個問題，我們利用最小相位濾波器(minimum phase filter)的概

念來轉化原來的通道，使轉化後的等效通道的直線傳輸路徑增強。這種方式稱做連續關鍵取樣決策回授等化器(SIS decision feedback equalizer, SIS DFE)。在通道已知的情況下，最小相位濾波器可以用回授等化器來實現，其中濾波器的係數可以根據(zero-forcing, ZF)或最小平方平均誤差(minimum mean square error, MMSE)準則來求出。而在通道係數未知以及時變通道的系統中，我們提出適應性盲蔽式連續關鍵取樣等化器，利用適應性濾波器演算法如最小平均平方(least mean-square, LMS)或遞迴最小平方(recursive least squares, RLS)來調出最小相位濾波器的係數。針對上述兩種情況，我們都提出電腦模擬結果來驗證這種連續關鍵取樣決策回授等化器可以改善系統的效能。另外我們也比較連續關鍵取樣決策回授等化器與延遲連續關鍵取樣演算法的效能，並證明我們提出的方法不論在錯誤率或系統複雜度方面都有更好的表現。

此外，為了更進一步簡化這種適應性連續關鍵取樣等化器的複雜度，我們提出最大比重適應性盲蔽式連續關鍵取樣決策回授等化器(Max-Weight blind SIS DFE)。透過電腦模擬結果，我們證明這種節省系統資源的演算法可以大幅減少系統的複雜度，卻可以提供幾乎不遜於原本的適應性盲蔽式連續關鍵取樣決策回授等化器的效能。

PARTICLE FILTER BASED BLIND EQUALIZATION FOR
FREQUENCY SELECTIVE CHANNELS
— ALGORITHMS AND PERFORMANCE ANALYSIS

Student: An-Ching Ho

Advisor: Dr. Hsiang-Feng Chi

Department of Communication Engineering

National Chiao Tung University

Hsinchu, Taiwan



ABSTRACT:

The use of particle filter on the channel equalization problem has been studied by many researches for years. As mentioned in [3], the weak light-of-sight (LOS) channel is one of the problems limiting the performance of the sequential importance-sampling (SIS) based equalization. In [3], the delayed-SIS (D-SIS) algorithm was proposed to solve this problem at the expense of high computation complexity. In this thesis we introduce a new class of blind SIS equalization algorithms for the frequency selective channels no matter how the first impulse of CIR (Channel Impulse Response) is.

We begin with a brief review of the particle filtering theory and establish the model of the particle filter equalizer. After the mathematical analysis of the BER in

the particle filter based channel equalization systems, we can know how the performance of the SIS equalization algorithm is affected by the channel with an attenuated LOS. To overcome this problem and improve the performance, we use the idea of *minimum-phase pre-filtering* to maximize the LOS of the equivalent channel and propose the SIS decision feedback equalization algorithms. In the case when the channel state information (CSI) is known, this minimum-phase pre-filtering can be implemented with the decision feedback equalizers (DFE), whose coefficients are computed based on either the zero-forcing (ZF) or minimum mean square error (MMSE) criteria. In the case of unknown CSI and time-varying channel, the proposed adaptive blind SIS equalization algorithm pre-filters the receiver input by using the adaptive filters such as the least mean-square (LMS) or the recursive least squares (RLS). In both cases, we conduct the computer simulations to illustrate how the proposed SIS DFE algorithms improve the performance. We compare the performance of the D-SIS equalization and the proposed SIS decision feedback equalization and find that our approach outperforms the D-SIS on both the BER performance and the computation complexity.

Moreover, to save the computation of the adaptive SIS equalization algorithms, we propose a simplified scheme of the adaptive blind SIS DFE, named the Max-Weight blind SIS DFE algorithm. We show that the proposed cost-effective algorithm can provide the performance almost the same as the original adaptive SIS DFE algorithm from the computer simulation results.

誌謝

時光飛逝，短短兩年碩士班的研究生涯，以此論文之告成，即將劃上句點。在這兩年裡，首當感謝指導教授 紀翔峰老師於研究以及論文上給予充分且詳實的指導與教誨鼓勵，令學生受益匪淺。

其次感謝交通大學電信系教授 蘇育德老師以及教授 黃家齊老師擔任口試委員的指導與建議，在此也表達由衷的感謝。並感謝系上的其他教授，在課堂上授予我專業知識與技能。

衷心的感謝實驗室的學長、同學及學弟們給予研究上的幫助，有學長們的指導，同學們的互相切磋，以及學弟們的鼓勵打氣，我才能順利完成碩士生涯，也在實驗室度過許多快樂的時光，更有難忘的回憶，謝謝你們。

最後，要特別感謝我的父母以及家人還有女友，特別是父母的辛苦栽培以及默默付出，還有女友平時的照顧，使得我得以全心投入學業。再次衷心的謝謝以上諸位，由於你們的幫忙，此論文得以完成，衷心感謝。

僅以此篇論文獻給對我付出關心的你們。

民國九十五年九月

研究生何恩慶謹識於交通大學

TABLE OF CONTENTS

中文摘要	1
英文摘要	3
誌謝.....	5
TABLE OF CONTENTS	6
LIST OF FIGURES	8
1 INTRODUCTION TO PARTICLE FILTERING THEORY ...	10
1.1 INTRODUCTION.....	10
1.2 MOTIVATION.....	10
1.3 FUNDAMENTALS OF THE PARTICLE FILTERING THEORY	11
1.4 RESAMPLING	14
1.5 PRACTICAL APPLICAIONS AND LIMITATIONS OF PARTICLE FILTERING	16
1.6 THESIS ORGANIZATION	17
2 PARTICLE FILTERING FOR CHANNEL EQUALIZATION	18
2.1 SYSTEM MODEL OF THE PARTICLE FILTER EQUALIZER.....	18
2.2 THE SIS-BASED EQUALIZATION	20
2.2.1 <i>Log SIS</i>	22
2.2.2 <i>Max-Log SIS</i>	23
2.3 PRACTICAL ISSUE — PROBLEM AND SOLUTIONS OF SIS EQ UNDER WEAK LOS CHANNELS	27
2.3.1 <i>Delayed Sampling</i>	28
2.3.2 <i>Minimum Phase filter Solutions</i>	29
2.4 CHAPTER SUMMARY	31
3 PERFORMANCE ANALYSIS OF MAX-LOG SIS	
EQUALIZATION ALGORITHMS	32
3.1 DERIVATION OF THE BIT ERROR PROBABILITY.....	32
3.2 CONVERGENCE BEHAVIOR OF THE AVERAGE BER OF A PARTICLE.....	35
3.3 ANALYSIS OF SIS DECISION FEEDBACK EQUALIZATION.....	38
3.3.1 <i>Zero-Forcing DFE (ZF-DFE)</i>	39
3.3.2 <i>Minimum Mean Square Error DFE (MMSE-DFE)</i>	43
3.4 CHAPTER SUMMARY	47

4 BLIND AND ADAPTIVE PARTICLE FILTERING

EQUALIZATIONS.....49

4.1	ADAPTIVE DECISION FEEDBACK EQUALIZATIONS.....	49
4.2	ADAPTIVE BLIND SIS EQUALIZATIONS.....	53
4.3	COMPLEXITY ISSUES AND THE MAX-WEIGHT SIS DFE.....	56
4.4	CHAPTER SUMMARY.....	58

5 COMPUTER SIMULATIONS & OBSERVATIONS.....59

5.1	PERFECT CHANNEL STATE INFORMATION.....	59
5.1.1	<i>Channel with a strong LOS</i>	60
5.1.2	<i>Channel with a weak LOS</i>	61
5.2	BLIND AND ADAPTIVE EQUALIZATIONS.....	65
5.2.1	<i>Channel with a strong LOS</i>	66
5.2.2	<i>Channel with a weak LOS</i>	67
5.3	CHAPTER SUMMARY.....	70

6 CONCLUSIONS AND PROSPECTIVE.....71

6.1	CONCLUSION.....	71
6.2	PROSPECTIVE AND THE FUTURE WORKS.....	72

REFERENCES.....74

APPENDIX.....76

A.	JACOBIAN APPROXIMATION.....	76
B.	DERIVATION OF THE EXPECTED VALUES IN THE BER DERIVATION (3-1).....	77
C.	PROOF OF THE BER AS A MONOTONICALLY DECREASING FUNCTION.....	80
D.	PRE-FILTER COEFFICIENTS.....	83

LIST OF FIGURES

CHAPTER 1

Fig. 1-1 Illustration of resampling (when $N_s = 100$)..... 15

CHAPTER 2

Fig. 2-1 The system model..... 18

Fig. 2-2 Illustration of Log SIS sampling mechanism 23

Fig. 2-3 Illustration of Max-Log SIS approximating Log SIS with different γ 25

Fig. 2-4 Illustration of line of sight in communication systems 27

Fig. 2-5 Illustration of the idea of delayed-sampling 28

Fig. 2-6 Illustration of the idea of minimum phase filtering method 30

Fig. 2-7 System diagram of the SIS based decision feedback equalization 30

CHAPTER 3

Fig. 3-1 BER $P_{err,k}^{(i)}$ versus SNR ζ with different $|h_0|^2$ as $\lambda_k^{(i)} = 10^{-3}$ 34

Fig. 3-2 BER $P_{err,k}^{(i)}$ versus SNR ζ with different $|h_0|^2$ as $\lambda_k^{(i)} = 0.1$ 35

Fig. 3-3 Convergence trajectory of $P_{err,k}^{(i)}$ as $|h_0|^2 = 0.2$, SNR = 15 dB. 37

Fig. 3-4 Convergence trajectory of $P_{err,k}^{(i)}$ as $|h_0|^2 = 0.6$, SNR = 15 dB..... 37

Fig. 3-5 Convergence trajectory of $P_{err,k}^{(i)}$ as $|h_0|^2 = 0.36$, SNR = 15 dB..... 38

Fig. 3-6 System model of the ZF-DFE..... 39

Fig. 3-7 System model of the MMSE-DFE 44

CHAPTER 4

Fig. 4-1 Block diagram of the adaptive filter algorithm..... 50

Fig. 4-2 System diagram of the SIS based adaptive DFE. 52

Fig. 4-3 Practical implementation of the SIS based blind DFE. 54

Fig. 4-4	Illustration of information transmission between i-th PF and adaptive filter. ...	55
Fig. 4-5	Implementation of the Max-Weight blind PF DFE.	57

CHAPTER 5

Fig. 5-1	BER versus SNR plots of the three equalizers under channel with a strong LOS. (Perfect CSI)	61
Fig. 5-2	BER versus SNR plots of the SIS EQ under channel with different LOS. (Perfect CSI)	62
Fig. 5-3	BER versus SNR plots of the three equalizers under channel with weak LOS. (Perfect CSI)	63
Fig. 5-4	Comparison of BER versus SNR plots of the D-SIS EQ with different delay d under channel with weak LOS. (Perfect CSI)	64
Fig. 5-5	BER versus SNR plots of the different blind equalizations under channel with strong LOS	67
Fig. 5-6	BER versus SNR plots of the three blind equalizations under channel with weak LOS.	67
Fig. 5-7	BER versus SNR plots of the four different blind equalizations under channel with weak LOS	69

1 INTRODUCTION TO PARTICLE FILTERING THEORY

1.1 INTRODUCTION

Particle filter is a novel optimal filtering algorithm based on the Bayesian formulation and the sequential importance sampling (SIS) technique. It is basically one of the sequential Monte Carlo (MC) methodologies. The particle filter employs discrete random measures to recursively compute and to approximate the desired probability distributions. As a simulation-based algorithm, particle filtering is particularly useful in dealing with nonlinear and non-Gaussian problems, in which it is typically impossible to obtain the desired statistical estimates or distributions by using the traditional methods.

The sequential Monte Carlo methods like particle filtering were invented decades ago. However, they did not cause much attention because of the high computation complexity. The recent advances in digital circuit technologies make high-bandwidth computation possible and are able to put the MC methods into practice. Therefore the research of these topics has become active and attractive again recently. The SIS algorithm and the particle filtering theory have been discussed in detail in some previously published papers, as references [1] [2]. In this chapter, we will first bring out the motivation of this thesis, and then give a brief introduction on the particle filtering theory. The applications of equalizations using particle filtering are main topics of this thesis and will be discussed in the next chapters.

1.2 MOTIVATION

There have been many articles discussing about the particle filtering equalization. However, rare of them addressed about the mathematical analysis of its bit error rate (BER) performance. In sight of this, we would like to derive the BER of a particle as

a mathematic equation, and then analyze its behavior in the first part of this thesis.

Besides, followed by the BER analysis of the particle filter equalizer, we can soon discover that its performance decays greatly under channels with an attenuated line of sight (LOS). As also discussed in the reference [3], this problem can be alleviated by the method of the *delayed sequential importance sampling* (D-SIS). The D-SIS algorithm, however, has some limitations and the notorious complexity issue in practical. This gives us some motivation to bring up an alternative solution that utilizes the *minimum-phase pre-filtering* to maximize the LOS of the equivalent channel.

1.3 FUNDAMENTALS OF THE PARTICLE FILTERING THEORY

The particle filtering is basically a suboptimal approach to perform *Bayesian filtering*. Consider the following state-space model,

$$\mathbf{x}_k = f_k(\mathbf{x}_{k-1}, \mathbf{u}_k) \quad (1-1)$$

$$\mathbf{y}_k = h_k(\mathbf{x}_k, \mathbf{n}_k) \quad (1-2)$$

where \mathbf{u}_k is the input vector, \mathbf{y}_k is a vector of observations, \mathbf{x}_k is the state vector, $f_k(\cdot)$ is the system transition function, $h_k(\cdot)$ is the measurement function, and \mathbf{n}_k is the noise vector. Equation (1-1) is known as the state equation, and equation (1-2) is the measurement equation. Note that $f_k(\cdot)$ and $h_k(\cdot)$ are possibly nonlinear functions. In many signal processing applications, we are interested in the way of *recursively computing* the posteriori distribution $p(\mathbf{x}_{k:0} | \mathbf{y}_{k:0})$ ¹. To recursively compute the posteriori probability, we can consider the following decomposition (see [2]):

$$p(\mathbf{x}_{k:0} | \mathbf{y}_{k:0}) \propto p(\mathbf{x}_k | \mathbf{x}_{k-1}) \cdot p(\mathbf{y}_k | \mathbf{x}_{k:0}) \cdot p(\mathbf{x}_{k-1:0} | \mathbf{y}_{k-1:0}) \quad (1-3)$$

¹ Here $\mathbf{y}_{k:0} = \{\mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_k\}$ is a set of observation vectors. We would use this form to represent a sequential set of symbols or vectors in different time indices throughout this article.

The Bayesian filtering is stated as the ways of analytically computing these kinds of distributions in the close forms. Unfortunately, the analytic solutions (known as Kalman filter [1]) to the transition from $p(\mathbf{x}_{k-1:0}|\mathbf{y}_{k-1:0})$ to $p(\mathbf{x}_{k:0}|\mathbf{y}_{k:0})$ exist only in a very restricted set of cases (e.g. when the systems are *linear* and the noises are *Gaussian* distributed). Therefore the particle filtering method has become an important alternative when we are solving nonlinear and non-Gaussian problems. When the particle filter is used, the distributions are approximated by discrete random measures via the Monte Carlo method. The random measures are composed of numerous particles which are samples drawn based on an estimated probability density function of the particles. The state space, and the weights are computed by using the Bayes theory. They are used to represent the “probability mass” of the particles. By the concept of Monte Carlo, the interested distribution $p(x)$ can be approximated as

$$p(x) \approx \hat{p}(x) = \sum_{i=1}^{N_s} w^{(i)} \cdot \delta(x - x^{(i)}) \quad (1-4)$$

where $x^{(i)}$ is the i -th particle, $w^{(i)}$ is its weight, and N_s is the number of the particles used in the approximation. It can be shown that as $N_s \rightarrow \infty$, the approximation approaches $p(x)$. $\delta(\cdot)$ is the Dirac delta function, i.e. $\delta(x - x^{(i)}) = 1$ if $x = x^{(i)}$, and $\delta(x - x^{(i)}) = 0$, otherwise. As shown in [9], this kind of Monte Carlo approximation has a certain advantage: the computations of expectations involving complicated integrations are now simplified as summations. Particle filtering is simply a simulation-based Monte Carlo method to solve the *sequential* Bayesian problems by the approximation in (1-4).

The key concept in constructing the weights and the particles of particle filtering is the *importance sampling* principle. In the process of drawing particles, it is often

not possible to draw the samples directly from the desired distribution $p(x)$. Therefore we have to generate particles from another distribution $q(x)$, which is also known as the *importance function*. The un-normalized weights are assigned as follows,

$$\tilde{w}_k^{(i)} = \frac{p(x)}{q(x)} \quad (1-5)$$

The weights are obtained after normalization:

$$w_k^{(i)} = \frac{\tilde{w}_k^{(i)}}{\sum_{j=1}^N \tilde{w}_k^{(j)}}, \forall i = 1, 2, \dots, N_s \quad (1-6)$$

The selection of the importance function $q(x)$ plays a crucial role in determining the performance of the particle filter. In general, the closer the importance function is selected to the desired distribution, the better the performance would be. A good choice of importance function can also reduce the effect of the degeneracy problem, which will be discussed in the next section. There are two most frequently used importance functions in the particle filtering applications: the *priori* importance function and the *optimal* importance function. The *priori* one is much easier for implementation while the *optimal* importance function minimizes the variance of the importance weights conditional on the trajectory of transmitted signal and the observations. The detail of the selection of the importance function is discussed in [8]. We would adopt the optimal importance function for analysis and simulations throughout this thesis.

To sequentially and recursively compute the importance functions and the particle weights, we need to update the weights and the importance function of each particle at every time slot through the equations as (1-3). The resulting method is referenced to the *sequential importance sampling* (SIS) algorithm. The detailed theory of the SIS algorithm can be found in many articles like [1] and [8]. We will discuss some important issues about the particle filtering in the following sections.

1.4 RESAMPLING

One major problem of the particle filtering is the *degeneracy* problem, which has been addressed in many previous articles like [1], [2] and [8]. It has been shown (as in [11] and [12]) that the variance of the importance weights can only increase over time due to propagation. Most of the particles will have negligible weights (very close to zero) after a few iterations. That is, much computation would be devoted in vain, The performance of the particle filter will become worse after a few interations. However, this problem can be alleviated by a good choice of the importance sampling function and *resampling*.

The idea of resampling is to elimiate the particles with negilible weights and replicate the largely-weighted particles. During the resampling, the larger weight a particle has, the more propable its particle (trajectory) would be replicated and overwrites the others. Each particles, after being re-allocated , would be given the same weight and then proceed the particle filtering process. In this article, the resampling is conducted whenever the *effective sample size* of the particle filter goes below a threshold ϵ . The effective sample size is defined as:

$$N_{eff} = \frac{N_s}{1 + \text{var}(w_k^{(i)})} \leq N_s$$

In practical, we estimate N_{eff} with

$$N_{eff} \cong \frac{1}{\sum_{i=1}^{N_s} (w_k^{(i)})^2} \leq N_s$$

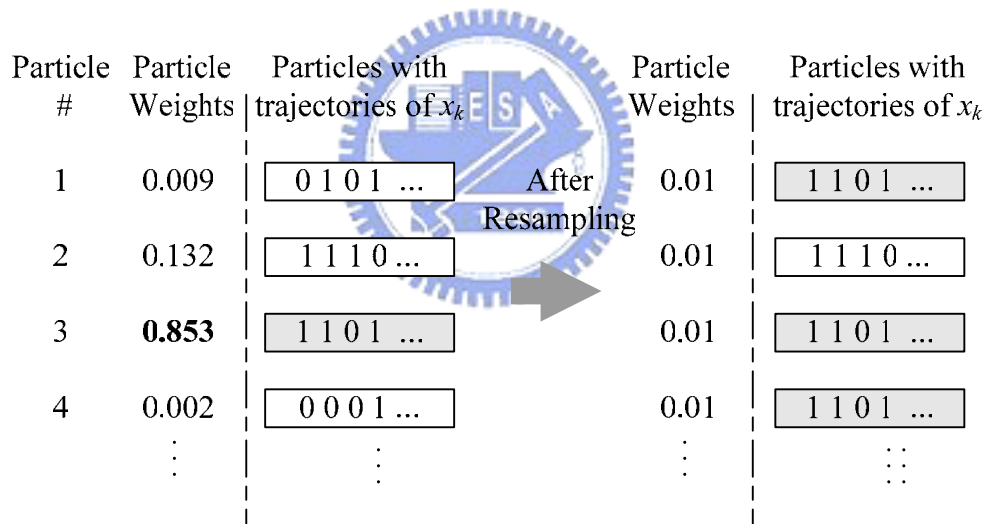
The effective sample size can roughly represent the number of particles with weights that are not negilible (or, say, the particles that are still effective). The procedure of resampling is summarized as the following:

if ($N_{eff} \leq \epsilon$), then

1. Replicate a new set of particles $\tilde{x}_k^{(i)} = x_k^{(i)}$, for all i .
2. For each i , the i -th particle is assigned to the value of particle j with probability $w_k^{(j)}$, i.e. $x_k^{(i)} = \tilde{x}_k^{(j)}$ with probability $w_k^{(j)}$.
3. Normalize/re-assign the particle weights $w_k^{(i)} = \frac{1}{N_s}$, for each i .

end if

Finally we give an example in Fig. 1-1, illustrating how the resampling works during the particle filtering.



- Every particle has a probability of 0.853 to be overwritten with the trajectory of the 3rd particle. (and 0.09 to be written with the 1st particle ..., etc.)

Fig. 1-1 Illustration of resampling (when $N_s = 100$).

1.5 PRACTICAL APPLICATIONS AND LIMITATIONS OF PARTICLE FILTERING

Particle filtering has been demonstrated as a powerful and promising methodology with a great range of applications in science and engineering. Within the field of communications, particle filtering has been widely applied to solve the channel equalization problems, including blind equalization, blind detection over flat fading channel (as addressed in [2]), and multi-user detections (MIMO) [13][14].

Nevertheless, as proposed in many articles like [3], particle filtering equalization suffers from performance decay under channels with a much attenuated line of sight (LOS), i.e. the first channel impulse response (CIR) is weak. To improve the performance under these channels, some solutions have been proposed. One of them is the method of *delayed sequential importance sampling with resampling* (D-SIR) proposed in [3], which incorporates the future observations to compute each particle at present. This method, however, requires the knowledge of the position of the largest impulse in CIR. Furthermore, the performance improvement of the D-SIR is achieved at the expense of the computational complexity, which is *exponentially increased* in proportional to the delay d . A large delay would lead to a great computational complexity which is nearly impractical.

Hence in this thesis we come out with a novel method that employs the idea of *minimum phase filter* implemented with the decision feedback equalizations (DFE) scheme. Instead of the idea to incorporate the future observations, we are considering to *move* the largest pulse in the CIR to the first one by employing the *minimum energy-delay property* of the minimum phase filter. This can be done by pre-calculated feedforward and feedback filter coefficients when the channel is known or by utilizing the adaptive filtering techniques (the least-mean-square (LMS) or the recursive least-squares (RLS) algorithms) when the channel is unknown.

1.6 THESIS ORGANIZATION

In Chapter 2 and Chapter 3, we would propose a particle filtering equalization scheme and analyze its bit error rate (BER) performance mathematically in the first part of this paper. And in the second part we come out with a new idea to keep good performance of the particle filtering under channel with attenuated LOS by applying *minimum phase* pre-filter followed by the particle filter. We verify the improvement of this method with mathematical proves and give some comments at the end of Chapter 3.

Next we put this idea into practice in Chapter 4 as a realization of particle filtering blind equalizations. We propose the system diagram of the SIS-based blind and adaptive equalizers. In sight of the complexity of the blind particle filter equalizer (PF EQ), we further induce the method of Max-Weight PF EQ, which can greatly reduce the computation and system complexity.

In Chapter 5 we utilize the computer simulations to verify the inference and conclusions in Chapters 3 and 4. Finally we make the thesis conclusion and present the future works in Chapter 6. The appendixes and the references are provided at the end of this thesis for further information and consultation.

2 PARTICLE FILTERING FOR CHANNEL EQUALIZATION

The main topics of this thesis are the performance and problems of the particle filter applied to equalization in the communication systems. In this chapter we first introduce the signal model and the SIS algorithm for equalizations. Then we discuss the problems and solutions of particle filter when the fading channels have a weak light-of-sight transmission path.

2.1 SYSTEM MODEL OF THE PARTICLE FILTER EQUALIZER

Consider a digital communication system where the BPSK symbols $b_k \in \{\pm 1\}, k = 0, 1, 2, \dots$ are transmitted to a frequency-selective multi-path channel. The channel impulse response (CIR) is assumed to be time invariant within the frame duration (i.e. the coherence time of the fading channel is much longer than the frame duration.). The system model is shown in Fig. 2-1.

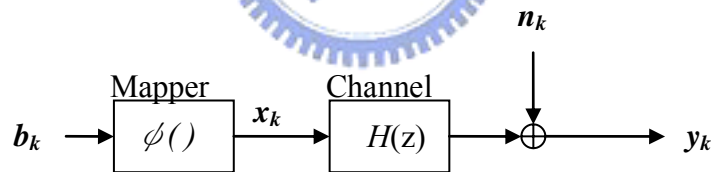


Fig. 2-1 The system model.

The transmitted bit stream b_k is mapped into x_k by using the BPSK mapping function $\phi()$ as illustrated in Table. 2-1.:

b_k	0	1
$x_k = \phi(b_k)$	+A	-A

Table2-1.BPSK mapping function $\phi()$

The CIR is assumed to be finite-length and has M tap weights: $\{h_0, h_1, \dots, h_{M-1}\}$. For simplicity, the CIR is normalized to have unity power, i.e. $\sum_{l=0}^{M-1} |h_l|^2 = 1$. The channel noise n_k is assumed to be additive white Gaussian noise (AWGN) with zero mean and the variance σ_n^2 , i.e. $n_k \sim N(0, \sigma_n^2)$. Thus the discrete-time sequence of observations can be written as:

$$y_k = \sum_{l=0}^{M-1} x_{k-l} h_l + n_k \quad (2-1)$$

We can now model the transmission by modifying the state-space formulation in (1-1) and (1-2) as follows:

$$\text{(state equation)} \quad \mathbf{x}_k = \mathbf{G}\mathbf{x}_{k-1} + \mathbf{u}_k \quad (2-2)$$

$$\text{(observation equation)} \quad y_k = \mathbf{h}^H \mathbf{x}_k + n_k \quad (2-3)$$

where the $M \times 1$ vector $\mathbf{x}_k = [x_k, x_{k-1}, \dots, x_{k-M+1}]^T$ is the system state at symbol k , and $\mathbf{h} = [h_0, h_1, \dots, h_{M-1}]^T : M \times 1$ is the CIR vector. \mathbf{G} is the $M \times M$ state-transition matrix:

$$\mathbf{G} = \begin{bmatrix} \mathbf{0}_{1 \times M-1} & 0 \\ \mathbf{I}_{M-1} & \mathbf{0}_{M-1 \times 1} \end{bmatrix}$$

such that $\mathbf{G} \cdot [x_{k-1}, x_{k-2}, \dots, x_{k-M}]^T = [0, x_{k-1}, \dots, x_{k-M+1}]^T$, and $\mathbf{u}_k = [x_k, 0, \dots, 0]^T$ is an $M \times 1$ vector of inputs. Our goal is to estimate the transmitted symbols \mathbf{x}_k recursively from the observations \mathbf{y}_k . The optimum estimate of \mathbf{x}_k is given by the maximum likelihood (ML) estimate (or MAP, if \mathbf{x}_k is sent equally likely). However, it is not possible to compute the posteriori probability $p(\mathbf{x}_{k:0} | \mathbf{y}_{k:0})$ in the close form if the noise is non-Gaussian or the system is non-linear. Instead of using the complicated optimal solution, we are seeking the sub-optimal way, in which the approximation of posteriori probability $p(\mathbf{x}_{k:0} | \mathbf{y}_{k:0})$ is recursively computed with the

aid of sequential importance sampling (SIS) algorithm.

2.2 THE SIS-BASED EQUALIZATION

The optimal blind equalization is achieved by MAP detection of the transmitted sequence $\mathbf{x}_{k:0}$, given the observations $\mathbf{y}_{k:0}$. Let $p(\mathbf{x}_{k:0}|\mathbf{y}_{k:0})$ denote the posteriori *probability mess function* (p.m.f.) of the data sequence given the observations, The MAP algorithm is to find the estimated data sequence that maximizes this posteriori probability. The most straightforward way of finding the MAP solution is to exhaustedly compute the probabilities of all 2^{k+1} possible $\hat{\mathbf{x}}_{k:0}$ and select the largest one. However, this is impractical due to the high computation complexity. It is preferable to find the estimate $\hat{\mathbf{x}}_{k:0}^{MAP}$ that maximizes the posteriori p.m.f. from the previous estimate $\hat{\mathbf{x}}_{k-1:0}^{MAP}$ recursively even though the final solution is sub-optimal. To recursively compute the posteriori probability, we utilize the decomposition in (1-3) of Chapter 1:

$$p(\mathbf{x}_{k:0}|\mathbf{y}_{k:0}) \propto p(x_k|x_{k-1}) \cdot p(y_k|\mathbf{x}_{k:0}) \cdot p(\mathbf{x}_{k-1:0}|\mathbf{y}_{k-1:0}) \quad (1-3)$$

Assume that the data bit b_k is uncorrelated with the bit b_l at the different time l , $k \neq l$, i.e. $p(x_k|x_{k-1}) = p(x_k)$. Using the similar derivation in [3], we can obtain the likelihood function:

$$p(y_k|\mathbf{x}_{k:0}) = \frac{1}{\sqrt{2\pi}\sigma_n} \exp\left(\frac{-|y_k - \mathbf{h}^H \mathbf{x}_k|^2}{2\sigma_n^2}\right) \quad (2-4)$$

From the equations (1-3) and (2-4), we can find a recursive equalization algorithm by computing the posteriori p.m.f. $p(\mathbf{x}_{k:0}|\mathbf{y}_{k:0})$ sequentially.

According to the particle filtering theory in Chapter 1 and [1, 2], we want to draw N_s particles $x_k^{(i)}, i=1,2,\dots,N_s$ from an importance function $q(\mathbf{x}_{k:0}|\mathbf{y}_{k:0})$ and

give them the weights based on the true posteriori probability $p(\mathbf{x}_{k:0}|\mathbf{y}_{k:0})$:

$$x_k^{(i)} \sim q(x_k^{(i)}|\mathbf{x}_{k-1:0}, y_k) \quad (2-5)$$

$$w_k^{(i)} \propto w_{k-1}^{(i)} \cdot \frac{p(x_k^{(i)}|x_{k-1}^{(i)}) \cdot p(y_k|\mathbf{x}_{k:0}^{(i)})}{q(x_k^{(i)}|\mathbf{x}_{k-1:0}, y_k)} \quad (2-6)$$

The importance function $q(\mathbf{x}_{k:0}|\mathbf{y}_{k:0})$ is the trial p.m.f. which approximates $p(\mathbf{x}_{k:0}|\mathbf{y}_{k:0})$. It is easier to draw samples from $q(\mathbf{x}_{k:0}|\mathbf{y}_{k:0})$ than $p(\mathbf{x}_{k:0}|\mathbf{y}_{k:0})$. The adaptation of the normalized weight (2-6) is derived from (1-3) and the following factorization of importance function:

$$q(\mathbf{x}_{k:0}|\mathbf{y}_{k:0}) \propto q(x_k|\mathbf{x}_{k-1:0}, y_k) \cdot q(\mathbf{x}_{k-1:0}|\mathbf{y}_{k-1:0}) \quad (2-7)$$

The N_s particles constitute a Monte Carlo smoother (MC smoother) that approximates the true posteriori p.m.f. for each time slot k :

$$p(\mathbf{x}_{k:0}|\mathbf{y}_{k:0}) \approx \hat{p}(\mathbf{x}_{k:0}|\mathbf{y}_{k:0}) = \sum_{i=1}^{N_s} w_k^{(i)} \cdot \delta(\mathbf{x}_{k:0} - \mathbf{x}_{k:0}^{(i)}) \quad (2-8)$$

where $\mathbf{x}_{k:0}^{(i)}$ are the particles, $w_k^{(i)}$ are their weights and $\delta(\cdot)$ is the Dirac delta function.

There are many choices to decide the importance sampling function $q(\mathbf{x}_{k:0}|\mathbf{y}_{k:0})$. Here we choose the optimal importance function that collects all available information up to time k .

$$\begin{aligned} q(x_k|\mathbf{x}_{k-1:0}, y_k) &= p(x_k|\mathbf{x}_{k-1:0}, y_k) \\ &= \frac{p(y_k|x_k, \mathbf{x}_{k-1:0}) \cdot p(x_k|\mathbf{x}_{k-1:0})}{p(y_k|\mathbf{x}_{k-1:0})} \end{aligned} \quad (2-9)$$

Substituting (2-9) into (2-5) and (2-6), we have

$$x_k^{(i)} \sim P_k^{(i)}(\chi) \equiv \frac{p(y_k|x_k = \chi, \mathbf{x}_{k-1:0}^{(i)}) \cdot p(x_k = \chi)}{\sum_{\substack{\forall x \in \\ \text{constellation}}} p(y_k|x_k = x, \mathbf{x}_{k-1:0}^{(i)}) p(x_k = x)} \quad (2-10)$$

$$w_k^{(i)} \propto w_{k-1}^{(i)} \cdot p(y_k | \mathbf{x}_{k-1:0}^{(i)}) = w_{k-1}^{(i)} \cdot \sum_{x_k} p(y_k | x_k, \mathbf{x}_{k-1:0}^{(i)}) \cdot p(x_k) \quad (2-11)$$

If the channel is known, the estimate of the likelihood function for the i -th particle at time k is

$$\hat{p}\left(y_k | x_k = \chi, \mathbf{x}_{k-1:0}^{(i)}\right) = \frac{1}{\sqrt{2\pi}\sigma_n} \exp\left(\frac{-|y_k - \mathbf{h}^H \tilde{\mathbf{x}}_k^{(i)}|^2}{2\sigma_n^2}\right) \quad (2-12)$$

where $\tilde{\mathbf{x}}_k^{(i)} = (\chi, x_{k-1}^{(i)}, x_{k-2}^{(i)}, \dots, x_{k-M+1}^{(i)})^T$, and χ is the drawn sample for testing $x_k^{(i)}$.

To draw samples for a particle according to the likelihood function (2-10) and put the SIS algorithm into practice, we propose two algorithms here for sample drawing of each particle.

2.2.1 Log SIS

From the definition of $P_k^{(i)}(\chi)$ in (2-10) and the fact that χ could be either $+A$ or $-A$. We can determine a sample drawing mechanism that employs a uniformly distributed random variable (r.v.) $u \sim U(0,1)$ to draw samples for $x_k^{(i)}$. Since the denominator of $P_k^{(i)}(\chi)$ is the same when $\chi = +A$ or $-A$, we consider only the numerator. Define $\rho_{k,+A}^{(i)}$ & $\rho_{k,-A}^{(i)}$ as the numerator of $P_k^{(i)}(\chi)$ when $\chi = +A$ and $-A$, respectively:

$$\begin{aligned} \rho_{k,+A}^{(i)} &\equiv p(y_k | x_k = +A, \mathbf{x}_{k-1:0}^{(i)}) \cdot p(x_k = +A) \\ &= \frac{1}{\sqrt{2\pi}\sigma_n} \exp\left(\frac{-|y_k - A \cdot h_0 - \sum_{l=1}^{M-1} h_l \cdot x_{k-l}^{(i)}|^2}{2\sigma_n^2}\right) \cdot p(x_k = +A) \end{aligned} \quad (2-13)$$

and similarly

$$\rho_{k,-A}^{(i)} = \frac{1}{\sqrt{2\pi}\sigma_n} \exp\left(\frac{-|y_k + A \cdot h_0 - \sum_{l=1}^{M-1} h_l \cdot x_{k-l}^{(i)}|^2}{2\sigma_n^2}\right) \cdot p(x_k = -A) \quad (2-14)$$

By normalizing $\rho_{k,+A}^{(i)}$ & $\rho_{k,-A}^{(i)}$, we can determine a threshold $\zeta_{k,+A}^{(i)}$

$$\zeta_{k,+A}^{(i)} \equiv \rho_{k,+A}^{(i)} / (\rho_{k,+A}^{(i)} + \rho_{k,-A}^{(i)}) = P_k^{(i)}(+A) \quad (2-15)$$

so that if the uniform r.v. u falls in the interval $[0, \zeta_{k,+A}^{(i)}]$, then the particle $x_k^{(i)}$ is assigned as $+A$, and vice versa. The way of generating the i -th particle $x_k^{(i)}$ at time k is illustrated in Fig. 2-2.

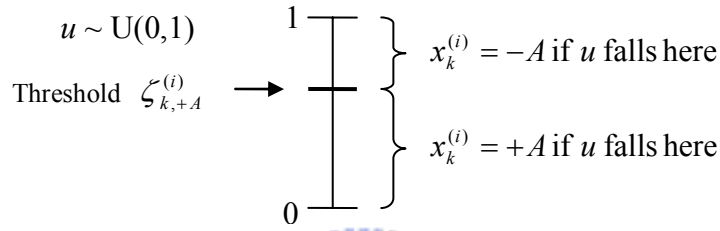


Fig. 2-2 Illustration of Log SIS sampling mechanism

The update of the weight for the i -th particle is (from (2-11))

$$\tilde{w}_k^{(i)} \equiv w_{k-1}^{(i)} (\rho_{k,+A}^{(i)} + \rho_{k,-A}^{(i)}) \propto w_k^{(i)} \quad (2-16)$$

where $\tilde{w}_k^{(i)}$ is the weight before normalization, and the new weight $w_k^{(i)}$ of i -th particle at time k can be obtained after the normalization:

$$w_k^{(i)} = \frac{\tilde{w}_k^{(i)}}{\sum_{j=1}^{N_s} \tilde{w}_k^{(j)}}, \forall i = 1, 2, \dots, N_s \quad (2-17)$$

2.2.2 Max-Log SIS

In this section we want to further simplify the computation of the threshold by using some approximations. Let us look at the definition of $\zeta_{k,+A}^{(i)}$ in (2-15). Since $\rho_{k,+A}^{(i)}$ & $\rho_{k,-A}^{(i)}$ are composed of exponential terms, we can obtain the new threshold $\psi_{k,+A}^{(i)}$ of Max-Log SIS by applying the Jacobian approximation (see Appendix A) on

the denominator of the original threshold:

$$\zeta_{k,+A}^{(i)} \equiv \frac{\rho_{k,+A}^{(i)}}{(\rho_{k,+A}^{(i)} + \rho_{k,-A}^{(i)})}$$

$$= \begin{cases} 1 - \frac{\exp(-\alpha_{k,-A}^{(i)})}{\exp(-\alpha_{k,+A}^{(i)}) + \exp(-\alpha_{k,-A}^{(i)})} = 1 - \exp(-(\alpha_{k,-A}^{(i)} - \alpha_{k,+A}^{(i)})) , & \text{if } \rho_{k,+A}^{(i)} > \rho_{k,-A}^{(i)} \\ \frac{\exp(-\alpha_{k,+A}^{(i)})}{\exp(-\alpha_{k,+A}^{(i)}) + \exp(-\alpha_{k,-A}^{(i)})} = \exp(-(\alpha_{k,-A}^{(i)} - \alpha_{k,+A}^{(i)})) , & \text{if } \rho_{k,+A}^{(i)} \leq \rho_{k,-A}^{(i)} \end{cases}$$

Thus we have simplified the computation of the threshold as

$$\zeta_{k,+A}^{(i)} \equiv \begin{cases} 1 - \frac{\rho_{k,-A}^{(i)}}{\rho_{k,+A}^{(i)}} = 1 - \exp(-|\alpha_{k,+A}^{(i)} - \alpha_{k,-A}^{(i)}|) , & \text{if } \rho_{k,+A}^{(i)} > \rho_{k,-A}^{(i)} \\ \frac{\rho_{k,+A}^{(i)}}{\rho_{k,-A}^{(i)}} = \exp(-|\alpha_{k,+A}^{(i)} - \alpha_{k,-A}^{(i)}|) , & \text{if } \rho_{k,+A}^{(i)} \leq \rho_{k,-A}^{(i)} \end{cases}$$

where $\alpha_{k,+A}^{(i)} = -\log(\rho_{k,+A}^{(i)})$, $\alpha_{k,-A}^{(i)} = -\log(\rho_{k,-A}^{(i)})$. If $p(x_k = +A) = p(x_k = -A)$, then

$\alpha_{k,+A}^{(i)}$ and $\alpha_{k,-A}^{(i)}$ can be simplified as the exponent term of $\rho_{k,+A}^{(i)}$ and $\rho_{k,-A}^{(i)}$:

$$\alpha_{k,+A}^{(i)} = \left| y_k - A \cdot h_0 - \sum_{l=1}^{M-1} h_l \cdot x_{k-l}^{(i)} \right|^2$$

$$\alpha_{k,-A}^{(i)} = \left| y_k + A \cdot h_0 - \sum_{l=1}^{M-1} h_l \cdot x_{k-l}^{(i)} \right|^2$$

To reduce the error produced by this approximation and to make the new threshold closer to the original one, we generalize the ratio of $\rho_{k,+A}^{(i)}/\rho_{k,-A}^{(i)}$ or $\rho_{k,-A}^{(i)}/\rho_{k,+A}^{(i)}$ as

$$\rho_k^{(i)} = \frac{1}{2} \exp(-\gamma |\alpha_{k,+A}^{(i)} - \alpha_{k,-A}^{(i)}|), \quad \gamma > 0$$

with the parameter γ to give some exponential weighting on it. The effects of this parameter would be discussed later in this section. Now we can define the threshold

$\psi_{k,+A}^{(i)}$ for the Max-Log SIS as:

$$\psi_{k,+A}^{(i)} \equiv \begin{cases} 1 - \rho_k^{(i)} , & \text{if } \rho_{k,+A}^{(i)} > \rho_{k,-A}^{(i)} \\ \rho_k^{(i)} , & \text{if } \rho_{k,+A}^{(i)} \leq \rho_{k,-A}^{(i)} \end{cases} \quad (2-18)$$

We can see that in both cases, $\psi_{k,+A}^{(i)}$ would approach to $\zeta_{k,+A}^{(i)}$ if γ is chosen

properly. To further illustrate that the Log SIS algorithm can be approximated by the Max-log SIS, we plot $\zeta_{k+A}^{(i)}$ and $\psi_{k+A}^{(i)}$ as functions of $(y_k - \sum_{l=1}^{M-1} h_l \cdot x_{k-l}^{(i)})$ in Fig. 2-3. with different values of γ .

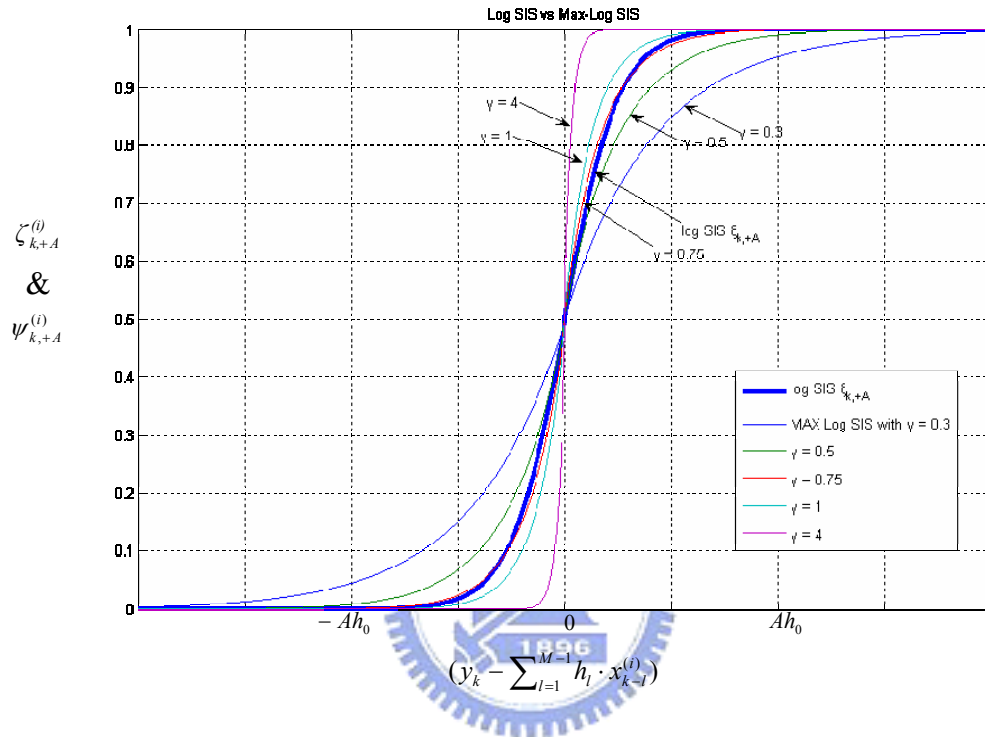


Fig. 2-3 Illustration of Max-Log SIS approximating Log SIS with different γ

We can observe that $\psi_{k+A}^{(i)}$ approaches $\zeta_{k+A}^{(i)}$ as γ is around 0.75. As γ increases, the characteristic function of $\psi_{k+A}^{(i)}$ would approach to a step function, which implies that the *hard decision* scheme is applied to draw sample for $x_k^{(i)}$. In this case, most of the particles would be drawn to be the same value at each time slot, and the particle filter can no longer compute the estimation of the desired posteriori probability. On the other hand, if we choose γ to be a very small value (e.g. $\gamma < 0.5$), we can observe from Fig. 2-3 that even when $(y_k - \sum_{l=1}^{M-1} h_l \cdot x_{k-l}^{(i)}) > A \cdot h_0$, the threshold is still not very close to 1. That is, there

is still quite a little probability that the particles would be drawn wrongly, and the error probability would increase.

Similarly, if we define $W_k^{(i)} = \log(w_k^{(i)})$ and $\tilde{W}_k^{(i)} = \log(\tilde{w}_k^{(i)})$, we can take logarithm and apply Jacobian approximation to the weight $w_k^{(i)}$ (2-16) and (2-17). We can obtain the following weight adaptation for Max-Log SIS:

$$\begin{aligned}
\tilde{W}_k^{(i)} &\equiv \log(w_{k-1}^{(i)}) + \log(\rho_{k,+A}^{(i)} + \rho_{k,-A}^{(i)}) \\
&= \tilde{W}_{k-1}^{(i)} + \log(\max(\rho_{k,+A}^{(i)}, \rho_{k,-A}^{(i)})) \\
&= \tilde{W}_{k-1}^{(i)} - \min(\alpha_{k,+A}^{(i)}, \alpha_{k,-A}^{(i)})
\end{aligned} \tag{2-19}$$

Similar to the Log-SIS case, the normalized weights would be:

$$W_k^{(i)} = \tilde{W}_k^{(i)} - \log(\sum_{j=1}^N \tilde{w}_k^{(j)}), \forall i = 1, 2, \dots, N \tag{2-20}$$

These logarithm SIS algorithms have an addition advantage on the implementation. Consider the case when the likelihood function is too small. The value may be truncated and may not be stored accurately due to the finite precision problem caused by limiting number of bits in a word. By utilizing the logarithms we can avoid these undesired truncations caused by insufficient bits.

2.3 PRACTICAL ISSUE — PROBLEM AND SOLUTIONS OF SIS EQ UNDER WEAK LOS CHANNELS

As we have mentioned in Chapter 1, the particle filter based equalization encounters a performance loss under the channel with a much attenuated line of sight (LOS)².

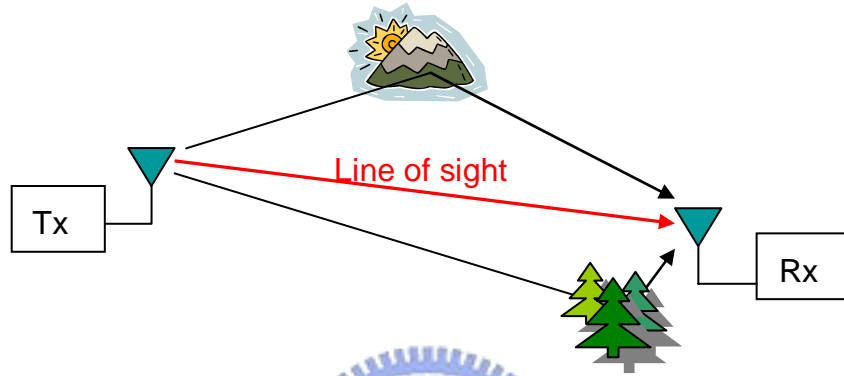


Fig. 2-4 Illustration of line of sight in communication systems

In this case, the first impulse of the CIR (h_0) would be very small. The value of the likelihood

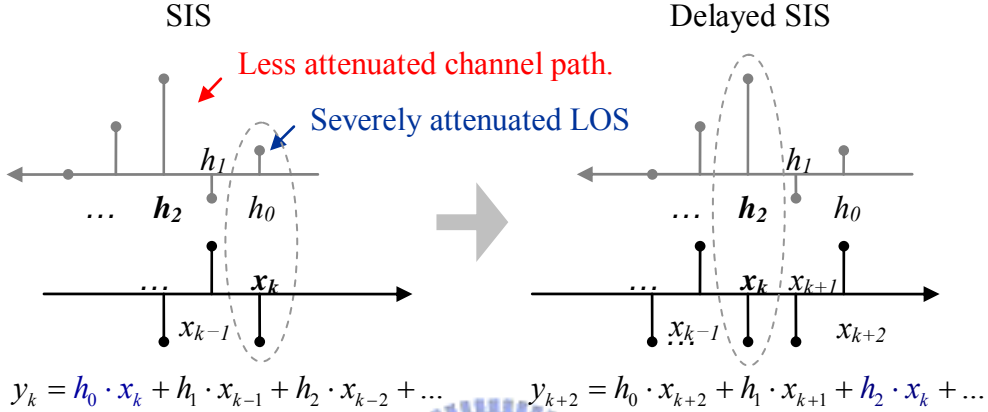
$$\hat{p}(y_k | x_k = \chi, \mathbf{x}_{k-1:0}^{(i)}) = \frac{1}{\sqrt{2\pi}\sigma_n} \exp\left(-\frac{|y_k - h_0 \cdot \chi - \sum_{l=1}^{M-1} h_l \cdot x_{k-l}^{(i)}|^2}{2\sigma_n^2}\right) \quad (2-12)$$

for all possible outcomes, $\chi = +A$ and $-A$, would be very close. Assume that symbols are sent equally likely, the difference between $\rho_{k,+A}^{(i)}$ and $\rho_{k,-A}^{(i)}$, in (2-13) and (2-14), would be small so that the SIS algorithm can hardly determine (or *draw out*) the desired particle. Hence the symbols will be erroneously decided in the receiver, and the performance of the particle filter decreases.

² Line of sight (LOS) is commonly used to refer to telecommunication links that rely on a line of sight directly between the transmitting antenna and the receiving antenna (as illustrated in Fig. 2-4), and its gain would usually be valued as the first impulse response of the CIR.

2.3.1 Delayed Sampling

One strategy to overcome this problem is the *delayed-sampling* technique, proposed in [3]. The basic idea of the *delayed-sampling* algorithm is to utilize the possible large (un-attenuated) and lagged CIR pulses and enumerate the future observations to determine the present particle, as illustrated in Fig. 2-5.



- Delayed SIS incorporates future observation y_{k+2} because in which x_k is multiplied with larger channel impulse h_2

Fig. 2-5 Illustration of the idea of delayed-sampling

Note that the CIR in this figure is placed in reversed order to illustrate how the computations work in the convolution of \mathbf{h} and \mathbf{x} .

More specifically, the sampling of a particle $x_k^{(i)}$ is delayed until y_{k+d} is observed,

$$x_k^{(i)} \sim q\left(x_k^{(i)} \mid \mathbf{x}_{k-1:0}^{(i)}, \mathbf{y}_{k:k+d}\right) \quad (2-21)$$

$$w_k^{(i)} \propto w_{k-1}^{(i)} \cdot \frac{p\left(x_k^{(i)} \mid x_{k-1}^{(i)}\right) \cdot p\left(\mathbf{y}_{k:k+d} \mid \mathbf{x}_{k:0}^{(i)}\right)}{q\left(x_k^{(i)} \mid \mathbf{x}_{k-1:0}^{(i)}, \mathbf{y}_{k:k+d}\right)} \quad (2-22)$$

Compared with the original SIS algorithm (2-10) and (2-11), the particles are now drawn from the delayed importance p.m.f. and the weights are updated accordingly.

The likelihood function now becomes

$$\begin{aligned} \hat{p}(\mathbf{y}_{k:k+d} | x_k = \mathcal{X}, \mathbf{x}_{k-1:0}^{(i)}) &= \sum_{\tilde{\mathbf{s}}_{t+1:t+d} \in \{\pm 1\}^d} \hat{p}(\mathbf{y}_{k:k+d}, \tilde{\mathbf{x}}_{t+1:t+d} | x_k = \mathcal{X}, \mathbf{x}_{k-1:0}^{(i)}) \\ &\propto \sum_{\tilde{\mathbf{s}}_{t+1:t+d} \in \{\pm 1\}^d} \hat{p}(\mathbf{y}_{k:k+d} | \tilde{\mathbf{x}}_{t+1:t+d}, x_k = \mathcal{X}, \mathbf{x}_{k-1:0}^{(i)}) \end{aligned} \quad (2-23)$$

Again, the proportionality comes from the assumption that the transmitted symbols are sent equally likely.

The delayed-SIS algorithm, however, has some limitations. First, as depicted in Fig. 2-5, it is straightforward that the *delay* d needs to be large enough to cover the less attenuated CIR. This requires the knowledge of the channel delay, which is usually unknown in the blind equalization scenarios. In addition, the computational complexity of the delayed-SIS algorithm has been increased *exponentially* because of the marginalization of all possible outcomes as shown in (2-23). Furthermore, the performance of the delayed-SIS method is not necessarily growing with the *delay* d (we will illustrate this in Chapter 5). In other words, it is possible that the performance decays when the delay d increases. The computation spent for the extra delay would be in vain. This usually happens when the delay d is larger than the channel length. In this situation, all of the information given in (2-23) would be the future symbols $\tilde{\mathbf{x}}_{t+1:t+d}$, and none of the previous particles would be applied in this computation. The particle filter loses its function to storage the discrete measures.

2.3.2 Minimum Phase filter Solutions

As discussed in the previous section, the delayed-SIS has some limitations and high computation complexity. Here we will propose a novel particle filtering scheme, which utilizes the *minimum phase filters* implemented in the form of decision feedback equalizations (DFE). As introduced in many *digital signal processing* textbooks, the minimum phase system has the property that the partial energy is most concentrated at the first impulse of the impulse response $h(n)$ (the *minimum energy delay* property [sec. 5.6.3 [7]]). The idea of this method is illustrated in Fig. 2-6.

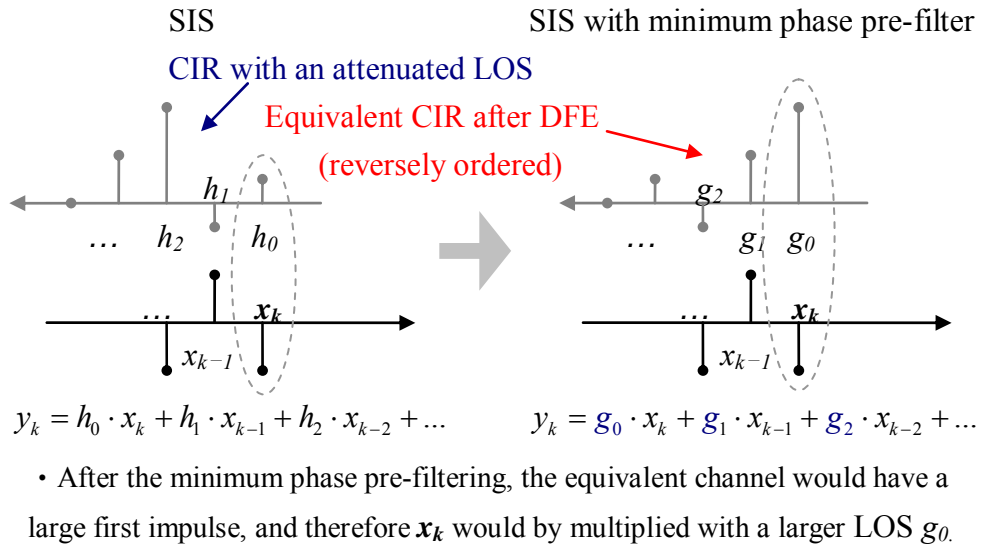


Fig. 2-6 Illustration of the idea of minimum phase filtering method

To combine the minimum phase system with particle filters, we propose the following system diagram of the *SIS decision feedback equalization* (SIS DFE), as shown in Fig. 2-7.

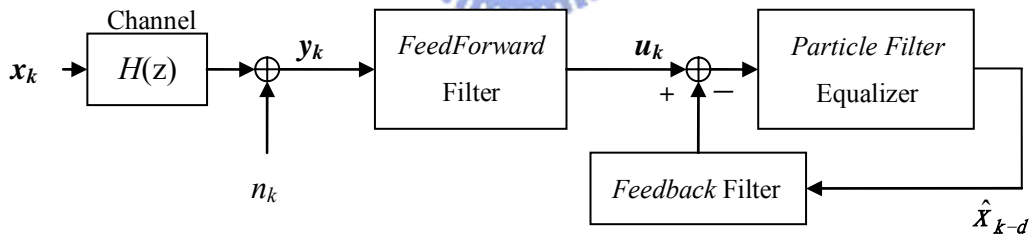


Fig. 2-7 System diagram of the SIS decision feedback equalization

The match filter cascaded with a noise whitening filter forms a feedforward filter (FFF). The feedback filter (FBF) is a minimum phase filter. In the case of known channel state information (CSI) in the time-invariant system, the coefficients of the FFF and FBF can be pre-calculated under the criteria of *zero-forcing* (ZF) or *minimum mean square error* (MMSE). When the channel information is unknown or varies with time, the adaptive filtering techniques such as RLS or LMS can be applied to fulfill

the operation of blind equalization. The details of the SIS based ZF and MMSE DFE would be introduced in the next chapter, and the implementation of blind equalization with adaptive filtering would be discussed in Chapter 4.

2.4 CHAPTER SUMMARY

In this chapter, we have briefly introduced the particle filtering equalization structure, and conceptually explained why channels with a weak LOS would reduce the performance of a particle filtering equalizer (PF EQ). To solve the weak LOS problem, we propose the SIS decision feedback equalization structure. Based on the defined notations and models, we will provide the detailed mathematical analysis of its bit error rate in the next chapter. We will show the effects of the weak LOS problem and how the proposed SIS decision feedback equalization can solve this problem.



3 PERFORMANCE ANALYSIS OF MAX-LOG SIS EQUALIZATION ALGORITHMS

In this chapter, we attempt to analyze the BER of the equalization system based on the Max-Log SIS algorithm. We will first derive the BER in 3.1 and further analyze the convergence behavior in 3.2. Then, we will focus on the performance analysis of the SIS decision feedback equalizers proposed in Chapter 2 and illustrate how they improve the performance (BER). Finally we will have a chapter summary in the end of the chapter.

3.1 DERIVATION OF THE BIT ERROR PROBABILITY

In the following analysis, we consider only the i -th particle and analyze the bit error rate (BER) of this particle. Assume that the data sequence b_k are sent equally likely, i.e. the priori probability $p(b_k = 0) = p(b_k = 1) = 1/2$. According to the Max-Log SIS algorithm, we can calculate the error probability of the i -th particle as:

$$\begin{aligned}
 P_{err,k}^{(i)} &= 1 - p(x_k^{(i)} = +A | x_k = +A) \cdot p(x_k = +A) - p(x_k^{(i)} = -A | x_k = -A) \cdot p(x_k = -A) \\
 &= 1 - \frac{1}{2} \cdot \int_0^1 p(u \leq \psi_{k,+A}^{(i)} | \psi_{k,+A}^{(i)} = \varphi, x_k = +A) \cdot p(\psi_{k,+A}^{(i)} = \varphi | x_k = +A) d\varphi \\
 &\quad - \frac{1}{2} \cdot \int_0^1 p(u \leq \psi_{k,-A}^{(i)} | \psi_{k,-A}^{(i)} = \varphi, x_k = -A) \cdot p(\psi_{k,-A}^{(i)} = \varphi | x_k = -A) d\varphi \\
 &= 1 - \frac{1}{2} \cdot \int_0^1 \varphi \cdot p(\psi_{k,+A}^{(i)} = \varphi | x_k = +A) d\varphi - \frac{1}{2} \cdot \int_0^1 \varphi \cdot p(\psi_{k,-A}^{(i)} = \varphi | x_k = -A) d\varphi \\
 &= 1 - \frac{1}{2} E[\psi_{k,+A}^{(i)} | x_k = +A] - \frac{1}{2} E[\psi_{k,-A}^{(i)} | x_k = -A] \tag{3-1}
 \end{aligned}$$

where $u \sim U(0,1)$ is a uniform distributed r.v. with the range $[0,1]$.

Based on the calculation of $E[\psi_{k,+A}^{(i)} | b_k = 0]$ and $E[\psi_{k,-A}^{(i)} | b_k = 1]$ (the details are in Appendix B), we obtain the error probability $P_{err,k}^{(i)}$ as:

$$\boxed{P_{err,k}^{(i)} = F_0(\lambda_k^{(i)}, |h_0|^2, \zeta) + F_1(\lambda_k^{(i)}, |h_0|^2, \zeta, \gamma)} \quad (3-2)$$

where

$$F_0(\lambda_k^{(i)}, |h_0|^2, \zeta) = Q\left(\sqrt{\frac{h_0^2}{4\lambda_k^{(i)} + 1/\zeta}}\right)$$

$$F_1(\lambda_k^{(i)}, |h_0|^2, \zeta, \gamma) = \frac{1}{2} \cdot \exp(2|h_0|^2 \zeta \gamma^2 (4\zeta\lambda_k^{(i)} + 1 - 1/\gamma)) Q\left(\gamma \left(8\zeta\lambda_k^{(i)} + 2 - \frac{1}{\gamma}\right) \sqrt{\frac{h_0^2}{4\lambda_k^{(i)} + 1/\zeta}}\right) - \frac{1}{2} \cdot \exp(2|h_0|^2 \zeta \gamma^2 (4\zeta\lambda_k^{(i)} + 1 + 1/\gamma)) Q\left(\gamma \left(8\zeta\lambda_k^{(i)} + 2 + \frac{1}{\gamma}\right) \sqrt{\frac{h_0^2}{4\lambda_k^{(i)} + 1/\zeta}}\right)$$

where $\zeta = A^2 (\sum_{l=0}^{M-1} |h_l|^2) / \sigma_n^2 = A^2 / \sigma_n^2$ (recall that $\sum_{l=0}^{M-1} |h_l|^2 = 1$) is the SNR, and the Q -function is defined as following:

$$Q(x) \equiv \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{t^2}{2}} dt.$$

We define the *error propagation factor* as

$$\lambda_k^{(i)} = \sum_{l=1}^{M-1} |h_l|^2 P_{err,k-l}^{(i)}$$

This factor indicates the error contribution of the previous BERs after the channel responses are applied.

In the rest of this section, we will observe the influence of all parameters on F_0 and F_1 . The BER of the i -th particle at time k $P_{err,k}^{(i)}$ is a function of the *error propagation factor* $\lambda_k^{(i)}$, the power of the first CIR $|h_0|^2$, the SNR ζ , and the Max-Log SIS parameter γ . To see the effect $|h_0|^2$ on the performance, we choose $\gamma = 0.75$ and fix $\lambda_k^{(i)}$ to observe the BER versus SNR graph with different values of $|h_0|^2$.

Fig. 3-1 is the plot of $P_{err,k}^{(i)}$ versus the SNR ζ according to (3-2) with several

values of $|h_0|^2$ with $\lambda_k^{(i)} = 10^{-3}$.

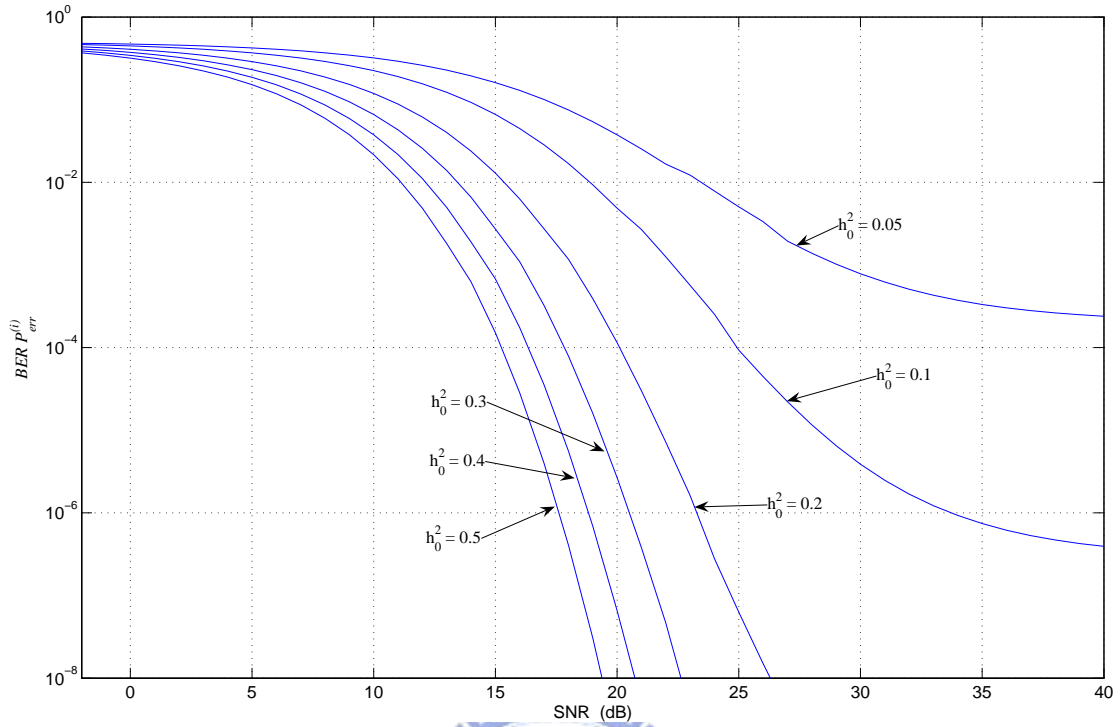


Fig. 3-1 BER $P_{err,k}^{(i)}$ versus SNR ζ with different $|h_0|^2$ with $\lambda_k^{(i)} = 10^{-3}$

As shown in the figure, we can find that the double of $|h_0|^2$ is equivalent to a 6 dB increase in SNR when $|h_0|^2$ is sufficiently large. Therefore it is essential to make $|h_0|^2$ large enough to obtain acceptable low bit error rate. Note that this BER versus SNR plot is drawn under the condition that the error propagation factor $\lambda_k^{(i)}$ is fixed. In fact, when $|h_0|^2$ is small, the value of $\lambda_k^{(i)}$ would be increased because the error propagation may be enlarged by the rest impulses of the CIR. When $\lambda_k^{(i)}$ is increased to 0.1, the BER versus SNR plot would be as shown in Fig. 3-2.

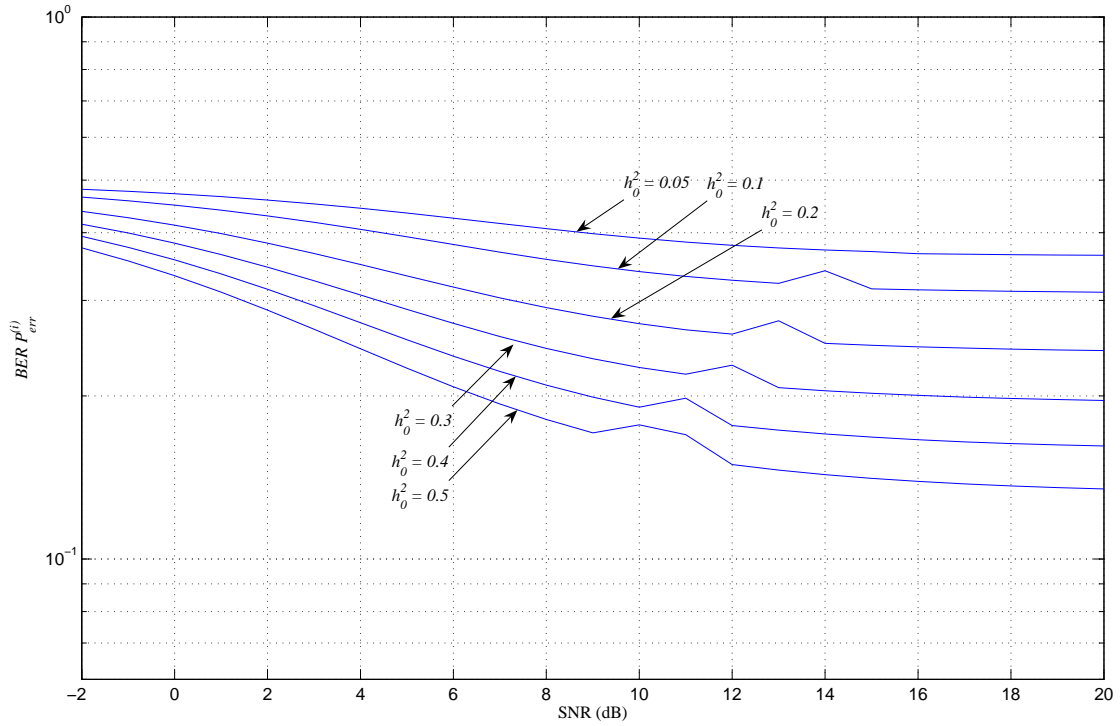


Fig. 3-2 BER $P_{err,k}^{(i)}$ versus SNR ζ with different $|h_0|^2$ as $\lambda_k^{(i)} = 0.1$

To evaluate the effects of $\lambda_k^{(i)}$ on the performance is difficult because $\lambda_k^{(i)}$ varies with time. However, our simulation results show the BER curves are similar to the lower curves in Fig. 3-1 when $|h_0|^2$ is large and similar to the upper ones in Fig. 3-2 as $|h_0|^2$ is small.

3.2 CONVERGENCE BEHAVIOR OF THE AVERAGE BER OF A PARTICLE

From (3-2) and the definition of $\lambda_k^{(i)}$, we can observe that $P_{err,k}^{(i)}$ is a function of the previous $M-1$ averaged bit error probabilities $\{P_{err,k-l}^{(i)} | l = 1, 2, \dots, M-1\}$. That is, the averaged bit error probability of a particle may change with time. It is not easy to analyze the convergence behavior of $P_{err,k}^{(i)}$ when M is large. In general $P_{err,k}^{(i)}$ will converge to stable equilibrium points, which are determined by the crossing points of a plane and a line in the hyper space $(p_0, p_1, \dots, p_{M-1})$ as the following.

The hyper plane:

$$p_0 = P_{err,0}^{(i)}(\lambda_k^{(i)}, |h_0|^2, \zeta, \gamma)$$

and the hyper lines:

$$\begin{cases} p_1 = p_0 \\ p_2 = p_1 \\ \vdots \\ p_{M-1} = p_{M-2} \end{cases}$$

The negative going cross points correspond to stable equilibrium points whereas the positive going cross points correspond to unstable equilibrium points.

To illustrate the convergence of $P_{err,k}^{(i)}$, we observe the case $M = 2$ with different values of $|h_0|^2$ and plot the convergence trajectories of the $P_{err,k}^{(i)}$ as in Fig. 3-3, Fig. 3-4, and Fig. 3-5. The arrowed lines represent the convergence trajectories of $P_{err,k}^{(i)}$ as the time index k increases ($k = 0, 1, 2, \dots$). We can see that in Fig. 3-3 and Fig. 3-4, $P_{err,k}^{(i)}$ will eventually converge to single stable equilibrium point E. In Fig. 3-5, when $|h_0|^2 = 0.36$ (in the middle range), there are two stable equilibrium points (E0 and E1) and an unstable equilibrium point (U). If $P_{err,0}^{(i)}$ is initially located in the range I0, $P_{err,k}^{(i)}$ would eventually converge to E0. On the other hand, if $P_{err,0}^{(i)}$ is initially located in the range I1, $P_{err,k}^{(i)}$ would eventually converge to E1, as shown in Fig.3-3. The steady state $P_{err,k}^{(i)}$ at E0 is much smaller than at E1. We call the initial point within the interval I0 a *good initial*. How to make $P_{err,0}^{(i)}$ fall in good initials is one of the important factors of obtaining good performance. That is actually the research topics we intend to study in the future. In addition, the unstable equilibrium point separates the convergence intervals I0 and I1.

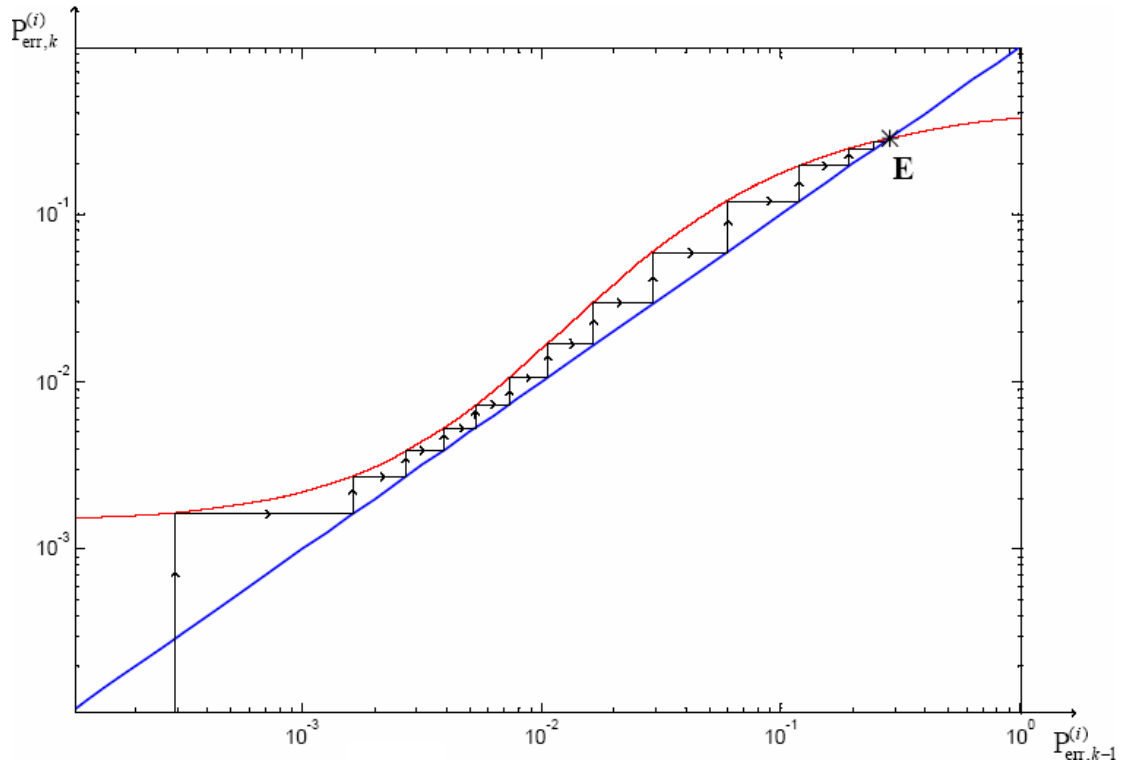


Fig. 3-3 Convergence trajectory of $P_{err,k}^{(i)}$ as $|h_0|^2 = 0.2$, SNR = 15 dB.

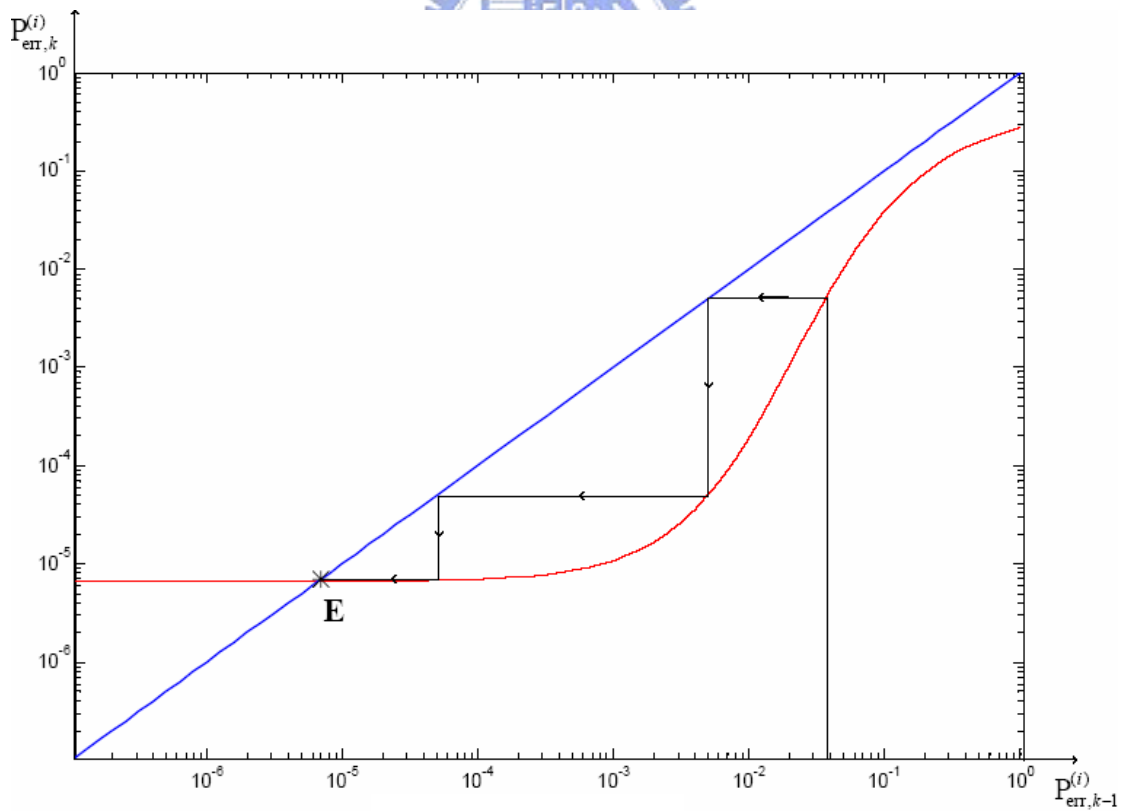


Fig. 3-4 Convergence trajectory of $P_{err,k}^{(i)}$ as $|h_0|^2 = 0.6$, SNR = 15 dB.

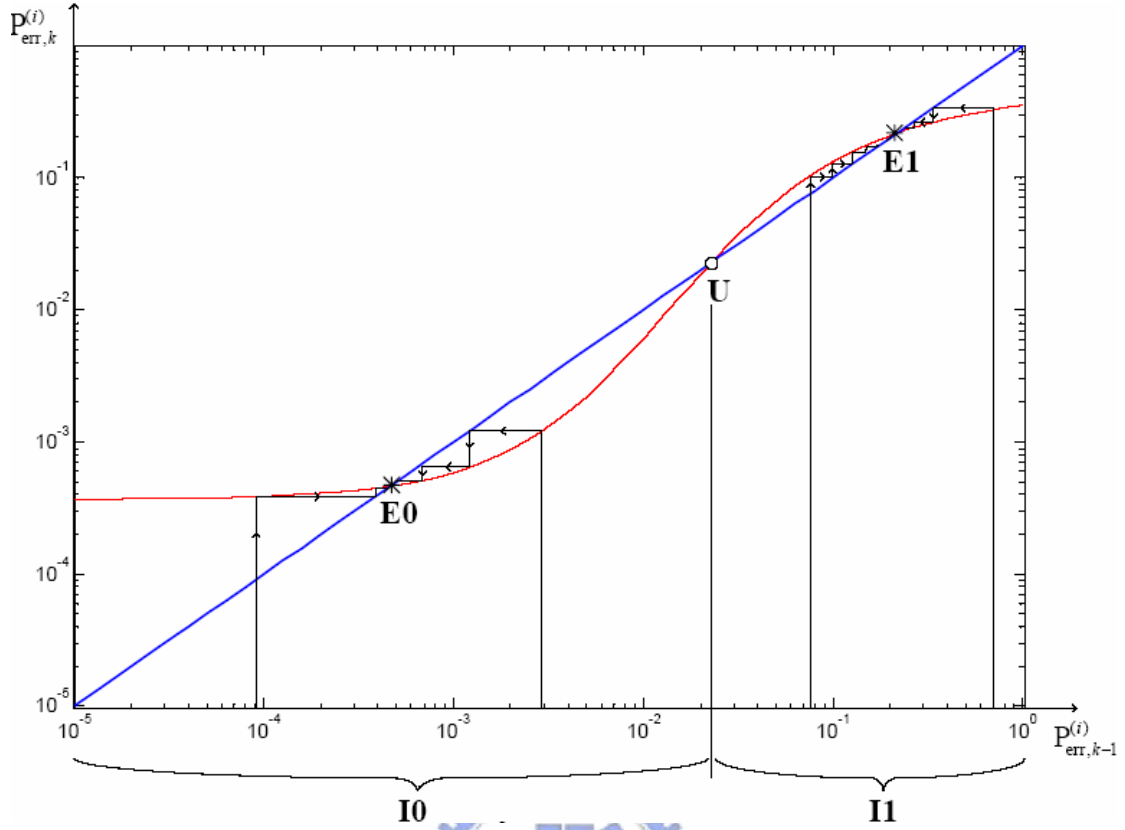


Fig. 3-5 Convergence trajectory of $P_{err,k}^{(i)}$ as $|h_0|^2 = 0.36$, SNR = 15 dB.

From the above observations, we conclude that the averaged bit error probability is likely to converge to a higher value (poor performance) if $|h_0|^2$ is not high enough. This situation usually happens when the channel has a weak LOS.

3.3 ANALYSIS OF SIS DECISION FEEDBACK EQUALIZATION

From the previous sections, we have proven that the power of the first impulse response of the CIR plays an important role in the SIS equalization. In this section, we will see how the SIS decision feedback equalization scheme proposed in Chapter 2 improves the system performance. In the following analysis, we assume the channel impulse response is known. First we will derive the coefficients of the FFF and FBF under the criteria of *zero-forcing* (ZF) and *minimum mean square error* (MMSE), respectively. Then we analyze the BER performance of SIS decision feedback

equalization with the coefficients calculated from these two criteria.

3.3.1 Zero-Forcing DFE (ZF-DFE)

We employ the technique of zero-forcing decision-feedback equalization and modify the system model in Fig. 2-7 into the ZF-DFE scheme, as shown in Fig. 3-6. The original receiver input y_k is filtered by a *feedforward filter* (FFF) with the response $F_{ZF}(z)$, which is composed of the match filter and the error whitening filter. The estimated sequence \hat{x}_{k-d} is filtered by a *feedback filter* (FBF) with the response $B_{ZF}(z)$, as shown in Fig. 3-6.

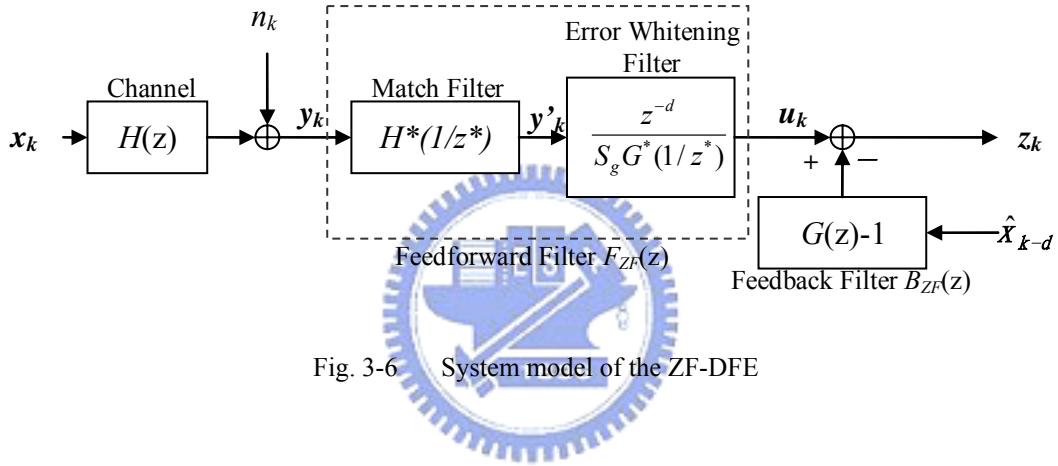


Fig. 3-6 System model of the ZF-DFE

The intermediate output y'_k after the match filter can be written as:

$$Y'_k(z) = X(z) \cdot R_{hh}(z) + N'(z) \quad (3-3)$$

where the noise sequence $N'(z)$ is a complex Gaussian sequence with autocorrelation function $R_{n'n'}(z) = E[N'(z)N'^*(1/z^*)] = R_{hh}(z) \cdot \sigma_n^2$. The average noise energy is $\sigma_{n'}^2 = R_{hh}(0) \cdot \sigma_n^2$.

By applying *spectral factorization* in [4], we can factorize the autocorrelation function of the CIR as (If it is factorizable³):

³ An autocorrelation function $R(z)$ is said to be *factorizable* if its Fourier transform, the *power spectrum density* (PSD) $|S(\theta)|^2$ and its logarithm $\log|S(\theta)|^2$ are both integrable over $-\pi < \theta \leq \pi$.

$$\boxed{R_{hh}(z) \equiv H(z) \cdot H^*(1/z^*) = S_g G(z) \cdot G^*(1/z^*)} \quad (3-4)$$

where $S_g = \exp\left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \log(|H(e^{j\omega})|^2) d\omega\right)$.

$G(z)$ is a *canonical* filter response of the length N_g , i.e. it is *causal* ($g_k = 0$ if $k < 0$), *monic* ($g_0 = 1$), and *minimum-phase* (all poles and zeros are fallen inside the unit circle)[4]. In addition, from the unity power property of CIR ($\sum_{l=0}^{M-1} |h_l|^2 = 1$) and the relation in (3-4), the power of the filter response g_k would be equal to $\sum_{l=0}^{M-1} |g_l|^2 = 1/S_g$.

Based on the derivations in [5], we have the following ZF-DFE solutions :

$$F_{ZF}(z) = \frac{z^{-d} H^*(1/z^*)}{S_g G^*(1/z^*)} \quad (3-5)$$

Therefore the equalized output is:

$$U(z) = G(z)X(z) + V(z)$$

or equivalently,

$$u_k = \sum_{i=0}^{N_g-1} g_i \cdot x_{k-d-i} + v_k \quad (3-6)$$

where v_k is the whitened Gaussian noise signal with the variance $\sigma_v^2 = \sigma_n^2/S_g$ [5], and the *feedforward* filter $F_{ZF}(z)$ is called the *whitened match filter* (WMF) because it whitens the noise signal after the match filter $H^*(1/z^*)$. Note that the $F_{ZF}(z)$ is an anti-causal IIR filter so that the delay d is usually induced for the implementation purpose.

From (3-6), we can treat u_k as the convolution result of the input x_k and the minimum-phase response \mathbf{g} plus the whitened noise v_k . In other words, u_k is the equivalent channel output when the input x_k is transmitted through the *equivalent*

minimum-phase channel \mathbf{g} with the noise v_k .

Comparing the linear convolution formula in (3-6) with (2-1), we can apply the BER analysis on the SIS-based ZF-DFE using the same procedure as in 2.2 (simply replacing y_k with u_k) and obtain:

$$x_{k-d}^{(i)} \sim P_{k-d}^{(i)}(\chi) \equiv \frac{p(u_k | x_{k-d} = \chi, \mathbf{x}_{k-d-1:0}^{(i)}) \cdot p(x_{k-d} = \chi)}{\sum_{\substack{\forall x \in \\ \text{constellation}}} p(u_k | x_{k-d} = x, \mathbf{x}_{k-d-1:0}^{(i)}) p(x_{k-d} = x)} \quad (3-7)$$

$$w_{k-d}^{(i)} \propto w_{k-d-1}^{(i)} \cdot p(u_k | \mathbf{x}_{k-d-1:0}^{(i)}) = w_{k-1}^{(i)} \cdot \sum_{x_{k-d}} p(u_k | x_{k-d}, \mathbf{x}_{k-d-1:0}^{(i)}) \cdot p(x_{k-d}) \quad (3-8)$$

And the estimate of the likelihood function is calculated as

$$\hat{p}(u_k | x_{k-d} = \chi, \mathbf{x}_{k-d-1:0}^{(i)}) = \frac{1}{\sqrt{2\pi}\sigma_v} \exp\left(-\frac{|u_k - \mathbf{g}^H \tilde{\mathbf{x}}_{k-d}^{(i)}|^2}{2\sigma_v^2}\right) \quad (3-9)$$

where $\mathbf{g} = [g_0, g_1, \dots, g_{N_g-1}]^T$ is the coefficients of $G(z)$, and

$$\tilde{\mathbf{x}}_{k-d}^{(i)} = [\chi, x_{k-d-1}^{(i)}, x_{k-d-2}^{(i)}, \dots, x_{k-d-M+1}^{(i)}]^T.$$

We can see that the equations (2-10)~(2-12) and the equations (3-7)~(3-9) have the same mathematical form except for the delay d . Likewise, by following the same derivations as in 3.1 and Appendix B, we can obtain the averaged bit error probability of the i -th particle in the SIS-based ZF-DFE as follows (using the same parameter γ):

$$\boxed{P_{err,k,ZF}^{(i)} = F_0(\lambda_{k,ZF}^{(i)}, |g_0|^2, \zeta_{ZF}) + F_1(\lambda_{k,ZF}^{(i)}, |g_0|^2, \zeta_{ZF}, \gamma)} \quad (3-10)$$

To compare $P_{err,k,ZF}^{(i)}$ in (3-10) and $P_{err}^{(i)}$ in equation (3-2), we could express the parameters $\lambda_{k,ZF}^{(i)}$, $|g_0|^2$, and ζ_{ZF} in terms of $\lambda_k^{(i)}$, $|h_0|^2$, and ζ . Then we compare the difference between (3-10) and (3-2). Unfortunately, since the BER $P_{err}^{(i)}$ is complicated, and the parameters are mutually dependent on each other, it is very difficult to have a close form comparison of these two BERs. To simplify the comparison, we assume that the N_g particles before the time k are all detected

correctly, i.e. $P_{err,k-l}^{(i)} = 0$, for $l = 1, 2, \dots, N_g - 1$.

Thus

$$\lambda_{k,ZF}^{(i)} = \sum_{l=1}^{M-1} |g_l|^2 P_{err,k-l}^{(i)} = 0$$

We can simplify $P_{err,k}^{(i)}(\lambda_k^{(i)}, |h_0|^2, \zeta, \gamma)$ as

$$P_{err,k}^{(i)}(\lambda_k^{(i)}, |h_0|^2, \zeta, \gamma) \Big|_{\lambda_k^{(i)}=0} = P_{err,k}^{(i)}(|h_0|^2 \cdot \zeta, \gamma) = P_{err,k}^{(i)}(\eta, \gamma)$$

where $\eta \equiv (|h_0|^2 \cdot \zeta)$. We can show (in Appendix C) that $P_{err,k}^{(i)}(\eta, \gamma)$ is a monotonically decreasing function of η when γ is within $[0.5, 1]$. Thus, we can show that $P_{err,k,ZF}^{(i)} \Big|_{\lambda_{k,ZF}^{(i)}=0} = P_{err,k}^{(i)}(\eta_{ZF}, \gamma)$ is smaller than $P_{err,k}^{(i)}(\eta, \gamma)$ if we can prove that $\eta_{ZF} \geq \eta$. The proof is as follows.

From the definition of η_{ZF} :

$$\eta_{ZF} \equiv |g_0|^2 \cdot \zeta_{ZF} = 1 \cdot A^2 / \sigma_v^2 = S_g \cdot \zeta \quad (3-11)$$

where $g_0 = 1$ ($G(z)$ is *monic*), and $\zeta_{ZF} = A^2 / \sigma_v^2 = A^2 / (\sigma_n^2 / S_g) = S_g \cdot \zeta$. Let $G'(z)$ be the minimum phase response having the same magnitude response as $H(z)$, i.e.

$$|H(z)|^2 = |G'(z)|^2$$

Compare this with (3-4),

$$\boxed{R_{hh}(z) \equiv H(z) \cdot H^*(1/z^*) = S_g G(z) \cdot G^*(1/z^*)} \quad (3-4)$$

it can be easy to obtain that $G'(z) = \sqrt{S_g} \cdot G(z)$. Since $G(z)$ is monic with $g_0 = 1$, the first impulse response of $G'(z)$ would be $g'_0 = \sqrt{S_g}$ and its power is $|g'_0|^2 = S_g$. From the *Komogoroff's error formula*⁴ [sec. 13-A, [6]] and the *minimum*

⁴ Let $H(z)$ be a minimal-phase response, it can be proven (in [6]) that

$\log h_0^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log |H(e^{j\omega})|^2 d\omega$, where h_0 is the magnitude of the first impulse response of $h(n)$.

energy-delay property⁵ of the minimum phase filter [sec. 5.6.3 [7]], we can know

$S_g = |g'_0|^2 \geq |h_0|^2$. Hence we have proven

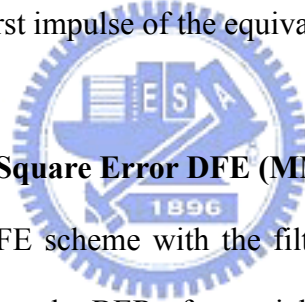
$$\begin{aligned}\eta_{ZF} &= S_g \cdot \zeta \geq |h_0|^2 \cdot \zeta = \eta \\ \Rightarrow \eta_{ZF} &\geq \eta\end{aligned}\quad (3-12)$$

It follows:

$$\boxed{P_{err,k,ZF}^{(i)} = P_{err,k}^{(i)}(\eta_{ZF}, \gamma) \leq P_{err,k}^{(i)}(\eta, \gamma)} \quad (3-13)$$

Note that the equality holds when the CIR $h(n)$ itself is a minimum phase response.

In conclusion, when the previous particles are drawn correctly and γ is chosen properly, the SIS based ZF-DFE scheme indeed improves the performance (i.e. reduces the averaged bit error probability) of the particle filter because it enhances the amplitude (or power) of the first impulse of the equivalent CIR $G(z)$.



3.3.2 Minimum Mean Square Error DFE (MMSE-DFE)

Now we consider the DFE scheme with the filter coefficients computed under the MMSE criterion. To analyze the BER of a particle in the SIS based MMSE-DFE scheme, we consider the equivalent discrete system model as shown in Fig. 3-7. The receiver's input signal y_k is filtered by a *feedforward filter* (FFF) with the response $F_{MMSE}(z)$. The estimated sequence \hat{x}_{k-d} is filtered by a *feedback filter* (FBF) with the response $B_{MMSE}(z)$.

⁵ Among all impulse responses $h(n)$ belonging to the same family of systems that have the same magnitude response, that is,

$$|H(e^{j\omega})| = |H_{\min}(e^{j\omega})|, \text{ for every } H(e^{j\omega}),$$

the partial energy of a minimum phase system $h_{\min}(n)$ would be most concentrated around its first impulse $h_{\min}(0)$:

$$\sum_{k=0}^n |h(k)|^2 \leq \sum_{k=0}^n |h_{\min}(k)|^2, \forall n \geq 0$$

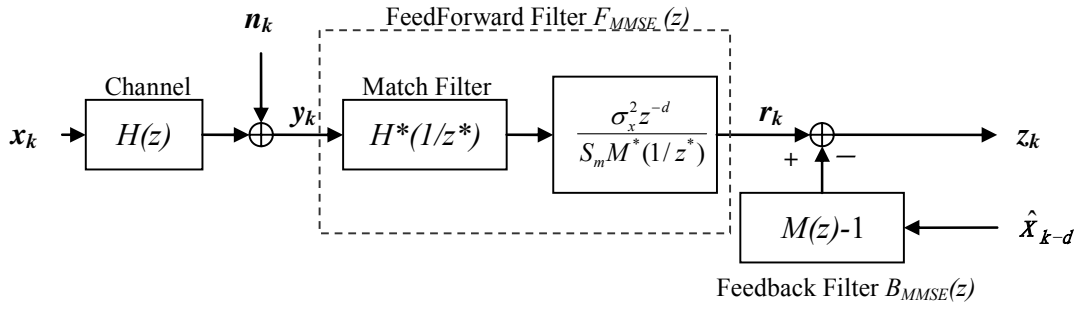


Fig. 3-7 System model of the MMSE-DFE

According to the theory of *linear prediction* and *spectral factorization* ([4] and [5]), we can write down the *key equation*:

$$\boxed{\sigma_x^2 H(z) H^*(1/z^*) + \sigma_n^2 = S_m M(z) M^*(1/z^*)} \quad (3-14)$$

where $S_m = \exp\left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \log(\sigma_x^2 |H(e^{j\omega})|^2 + \sigma_n^2) d\omega\right)$.

The left hand side of the equation is known as the *system autocorrelation function*, and the right hand side is the *spectral factorization*. Note that $M(z)$ is also a *canonical* filter response with the length N_m . From [5], the MMSE-DFE solutions of the FFF and FBF are

$$F_{MMSE}(z) = \frac{\sigma_x^2 z^{-d} H^*(1/z^*)}{S_m M^*(1/z^*)} \quad (3-15)$$

$$B_{MMSE}(z) = M(z) - 1 \quad (3-16)$$

and the equalized output becomes

$$R(z) = M(z)X(z) + E(z)$$

or equivalently,

$$r_k = \sum_{i=0}^{N_m-1} m_i \cdot x_{k-d-i} + e_k \quad (3-17)$$

The *feedforward* filter $F_{MMSE}(z)$ is called the *mean-square whitened match filter* (MS-WMF) because it whitens the filtered output noise e_k , which is the combination

of both the noise and the inter-symbol interference (ISI) sequences. After the MS-WMF, e_k would have the variance $\sigma_e^2 = (\sigma_x^2/S_m) \cdot \sigma_n^2$ [5]. Although e_k is white, it is not Gaussian distributed in general. That is because e_k is dependent on the data sequence x_k . However, from the Central Limit Theorem, the error signal e_k can still be approximated to be Gaussian.

Following the same procedure as used in the previous sections, we proceed the BER analysis of this MMSE-DFE scheme by writing:

$$x_{k-d}^{(i)} \sim P_{k-d}^{(i)}(\chi) \equiv \frac{p(r_k | x_{k-d} = \chi, \mathbf{x}_{k-d-1:0}^{(i)}) \cdot p(x_{k-d} = \chi)}{\sum_{\substack{\forall x \in \\ \text{constellation}}} p(r_k | x_{k-d} = x, \mathbf{x}_{k-d-1:0}^{(i)}) p(x_{k-d} = x)} \quad (3-18)$$

$$w_{k-d}^{(i)} \propto w_{k-d-1}^{(i)} \cdot p(r_k | \mathbf{x}_{k-d-1:0}^{(i)}) = w_{k-1}^{(i)} \cdot \sum_{x_{k-d}} p(r_k | x_{k-d}, \mathbf{x}_{k-d-1:0}^{(i)}) \cdot p(x_{k-d}) \quad (3-19)$$

The estimate of the likelihood function is calculated as

$$\hat{p}(r_k | x_{k-d} = \chi, \mathbf{x}_{k-d-1:0}^{(i)}) = \frac{1}{\sqrt{2\pi}\sigma_e} \exp\left(-\frac{|r_k - \mathbf{m}^H \tilde{\mathbf{x}}_{k-d}^{(i)}|^2}{2\sigma_e^2}\right) \quad (3-20)$$

where $\mathbf{m} = [m_0, m_1, \dots, m_{N_m-1}]^T$ is the coefficients of $M(z)$, and

$$\tilde{\mathbf{x}}_{k-d}^{(i)} = [\chi, x_{k-d-1}^{(i)}, x_{k-d-2}^{(i)}, \dots, x_{k-d-M+1}^{(i)}]^T.$$

Similar to the ZF-DFE, we can see that the equations (2-10)~(2-12) and the equations (3-18)~(3-20) again have the same mathematical form except for the delay d . Hence we can obtain the averaged bit error probability of a particle in the MMSE-DFE scheme as

$$P_{err,k,MMSE}^{(i)} = F_0(\lambda_{k,MMSE}^{(i)}, |m_0|^2, \zeta_{MMSE}) + F_1(\lambda_{k,MMSE}^{(i)}, |m_0|^2, \zeta_{MMSE}, \gamma) \quad (3-21)$$

Similar to the BER analysis in the case of ZF-DFE, we simplify $P_{err,k,MMSE}^{(i)}$ as a function of η_{MMSE} and γ by assuming that $\lambda_{k,MMSE}^{(i)} = 0$:

$$P_{err,k,MMSE}^{(i)} \Big|_{\lambda_{k,MMSE}^{(i)}=0} = P_{err,k}^{(i)}(\eta_{MMSE}, \gamma)$$

where $\eta_{MMSE} \equiv |m_0|^2 \cdot \zeta_{MMSE} = 1 \cdot A^2 / \sigma_e^2 = S_m / \sigma_n^2$, and

$$S_m = \exp\left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \log(A^2 |H(e^{j\omega})|^2 + \sigma_n^2) d\omega\right).$$

If $\sigma_n^2 = 0$, we can obtain

$$\begin{aligned} S_m \Big|_{\sigma_n^2=0} &= \exp\left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \log(A^2 |H(e^{j\omega})|^2) d\omega\right) \\ &= \exp\left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \log(A^2) + \log(|H(e^{j\omega})|^2) d\omega\right) \\ &= \exp\left(\log(A^2) + \frac{1}{2\pi} \int_{-\pi}^{\pi} \log(|H(e^{j\omega})|^2) d\omega\right) \\ &= A^2 \cdot S_g \end{aligned}$$

where $S_g = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log(|H(e^{j\omega})|^2) d\omega$ is the same as (3-4) in ZF-DFE scheme, and

A^2 is the symbol power.

Since σ_n^2 is always positive and the logarithm function is also a monotonic increasing function, we can conclude that S_m is always larger than $A^2 S_g$.

$$S_m \geq A^2 \cdot S_g$$

Utilizing the property $S_g \geq |h_0|^2$, as mentioned in the previous section, we have:

$$S_m \geq A^2 \cdot S_g \geq A^2 \cdot |h_0|^2 \quad (3-22)$$

Thus:
$$\eta_{MMSE} = \frac{S_m}{\sigma_n^2} \geq \frac{A^2 \cdot |h_0|^2}{\sigma_n^2} = |h_0|^2 \cdot \zeta \equiv \eta$$

$$\Rightarrow \eta_{MMSE} \geq \eta \quad (3-23)$$

We have finally proved that

$$\boxed{P_{err,k,MMSE}^{(i)} = P_{err,k}^{(i)}(\eta_{MMSE}, \gamma) \leq P_{err,k}^{(i)}(\eta, \gamma)} \quad (3-24)$$

Note that the equality holds when the both equalities in (3-22) maintain. The first

equality in (3-22) holds when $\sigma_n^2 = 0$. In this situation, the MMSE-DFE solution becomes the ZF-DFE solution. And the second equality in (3-22) maintain as the CIR $h(n)$ is a minimum phase system.

In sum, we have shown that when the particle filter operates with correct previous bit decisions and with a properly chosen γ , the MMSE-DFE scheme also improves the performance of the SIS-based equalization. The improvement is caused by enlarging the amplitude (or power) of the LOS of the equivalent channel impulse response $M(z)$.

Furthermore, from the first inequality in (3-22), we can have the following relationship between η_{MMSE} and η_{ZF}

$$\eta_{MMSE} = \frac{S_m}{\sigma_n^2} \geq \frac{A^2 \cdot S_g}{\sigma_n^2} = S_g \cdot \zeta \equiv \eta_{ZF}$$

The equality holds when $\sigma_n^2 = 0$. Therefore we can conclude that as noise variance approaches to zero ($\sigma_n^2 \rightarrow 0$), or signal-to-noise ratio (SNR) is high enough, the MMSE-DFE scheme would approaches to the ZF-DFE. In general, the SIS-based MMSE-DFE schemes would generally have a better performance than the ZF-DFE schemes, i.e.

$$P_{err,k,MMSE}^{(i)} \leq P_{err,k,ZF}^{(i)}$$

3.4 CHAPTER SUMMARY

In this chapter we have conducted the analyses of the averaged bit error probability of a particle in the SIS-based equalization schemes. During the analysis, we have found that the performance is unacceptable when the power of the first impulse of the channel impulse response (CIR) is low, namely, the weak LOS

problem. To combat this effect, we have proposed the SIS-based decision feedback equalization (DFE) schemes (either based on ZF or MMSE criterion). We have proven that by employing the DFE schemes, the BERs are lower than that of the original SIS-based equalization. In the next chapter, we would focus on the unknown CSI case and propose the SIS-based *blind* decision feedback equalization algorithms. The simulation results based on all the proposed schemes would be shown in Chapter 5.



4 BLIND AND ADAPTIVE PARTICLE FILTERING EQUALIZATIONS

In Chapter 3, we have introduced the SIS decision feedback equalization based on ZF and MMSE criteria. We have confirmed that they can improve the performance when the channel LOS is weak. Nevertheless, the coefficients of the FFF and FBF must be calculated in advance. That means the channel impulse response must be known and not vary with time. In this chapter, we will propose an *adaptive blind* SIS decision feedback equalization for the situation that the CIR is unknown or varies with time.

4.1 ADAPTIVE DECISION FEEDBACK EQUALIZATIONS

According to the *adaptive filter theory* [10], the adaptive filtering algorithms can be divided into two categories: the *stochastic gradient approach* and the *least square estimate*. The former algorithm is designed to minimize the mean-square error (MMSE) $E[e^2(n)]$ by solving the Wiener-Hopf equations for stationary inputs (or observations) in the sense of “ensemble average”. The later algorithm is to minimize the sum of error squares $\sum_{n=n_1}^{n_2} e^2(n)$ by solving the normal equation with “time averages”. The stochastic gradient algorithm is optimal in the probability sense while the method of least squares involves the time averages and is the optimal solution based on the deterministic data.

In practical situations, when the statistical information of the channel is unknown, the representative recursive solutions of these two methods are the *Least-Mean-Square* (LMS) adaptive filter and *Recursive Least Squares* (RLS) adaptive filter. Compared with the RLS algorithm, the LMS algorithm is much simpler in complexity. But the LMS algorithm has a slower convergence and is more sensitive to the

*eigenvalue spread*⁶ of the correlation matrix than the RLS algorithm.

Both of the two adaptive filtering algorithms can be applied in the blind SIS-based decision feedback equalization. Here we will adopt the *RLS* algorithm as an example.

First we have a quick overview of the RLS algorithm for the linear equalization filter (without feedback), as given in [10]:

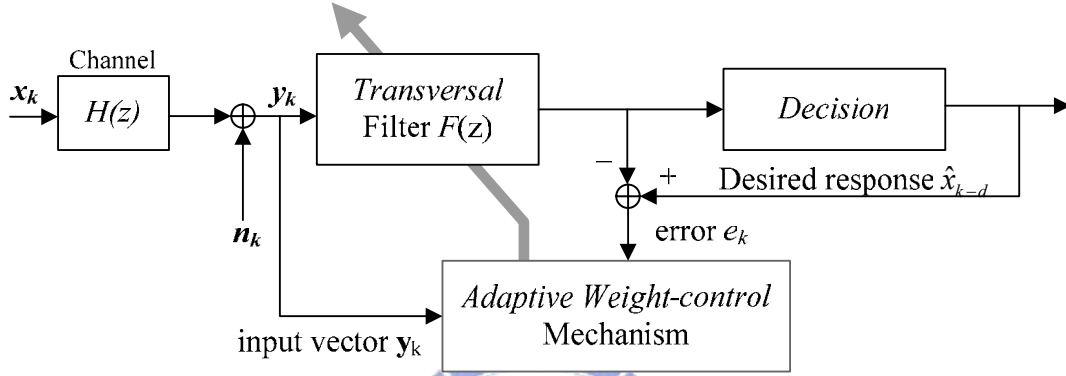


Fig. 4-1 Block diagram of the adaptive filter algorithm

With the system block diagram and the notations shown in Fig. 4-1, the method of least squares is to minimize the sum of error squares with respect to the filter coefficient vector $\mathbf{f}(n) = [f_0(n), f_1(n), \dots, f_{N_f-1}(n)]^T$ at time n . i.e.

$$\arg \left\{ \min_{\mathbf{f}} \left(\sum_{k=n_1}^{n_2} |e(k)|^2 \right) \right\} \quad (4-1)$$

where

$$e(k) = \hat{x}(k) - \mathbf{f}^H(n) \mathbf{y}(k) \quad (4-2)$$

, and $\mathbf{y}(k)$ (or \mathbf{y}_k) is the *tap-input vector at time k*:

$\mathbf{y}(k) = [y(k), y(k-1), \dots, y(k-N_f+1)]^T$. From [10], the cost function

$\varepsilon(\mathbf{f}) \equiv \sum_{k=n_1}^{n_2} e(k) \cdot e^*(k)$ is minimized for a particular estimation error $e_{\min}(k)$ such

⁶ The eigenvalue spread is the ratio of the largest over the smallest eigenvalue. When the eigenvalue spread of a channel correlation matrix is large, the convergence of the learning curve would be slow and the misadjustment would be relatively large [10].

that $\nabla \varepsilon = 0$ or, equivalently,

$$\sum_{k=n_1}^{n_2} y(k-i) \cdot e_{\min}^*(k) = 0, \quad \text{for } i = 0, 1, \dots, N_f - 1 \quad (4-3)$$

Let $\hat{\mathbf{f}}(n) = [\hat{f}_0(n), \hat{f}_1(n), \dots, \hat{f}_{N_f-1}(n)]^T$ be the optimal filter coefficients such that

$$e_{\min}(k) = \hat{x}(k) - \hat{\mathbf{f}}^H(n) \mathbf{y}(k) = \hat{x}(k) - \sum_{t=0}^{N_f-1} \hat{f}_t^*(n) \cdot y(k-t) \quad (4-4)$$

By substituting (4-4) into (4-3), we obtain the system of *normal equations*:

$$\sum_{t=0}^{N_f-1} \hat{f}_t(n) \cdot \sum_{k=n_1}^{n_2} y(k-i) \cdot y^*(k-t) = \sum_{k=n_1}^{n_2} y(k-i) \cdot \hat{x}^*(k)$$

$$\Rightarrow \boxed{\sum_{t=0}^{N_f-1} \hat{f}_t(n) \cdot \Phi(t, i) = z(-i), \text{ for } i = 0, 1, \dots, N_f - 1} \quad (4-5)$$

where $\Phi(t, i) = \sum_{k=n_1}^{n_2} y(k-i) \cdot y^*(k-t)$ and $Z(-i) = \sum_{k=n_1}^{n_2} y(k-i) \cdot \hat{x}^*(k)$. Please note that $\Phi(t, i)$ and $Z(-i)$ represent the time-averaged autocorrelation of $\mathbf{y}(k)$ and the cross-correlation between $\mathbf{y}(k)$ and $\hat{\mathbf{x}}(k) = [\hat{x}(k), \hat{x}(k-1), \dots, \hat{x}(k-N_f+1)]^T$, respectively. If we rewrite (4-5) in matrix form:

$$\boxed{\mathbf{\Phi} \cdot \hat{\mathbf{f}} = \mathbf{z}, \text{ or } \hat{\mathbf{f}} = \mathbf{\Phi}^{-1} \cdot \mathbf{z}} \quad (4-6)$$

where $\mathbf{z} = \mathbf{y}(k) \cdot \hat{\mathbf{x}}^*(k) = [z(0), z(-1), \dots, z(-N_f+1)]^T$,

$$\mathbf{\Phi} = \mathbf{y}(k) \cdot \mathbf{y}^H(k) = \begin{pmatrix} \phi(0, 0) & \dots & \phi(N_f-1, 0) \\ \vdots & \ddots & \vdots \\ \phi(0, N_f-1) & \dots & \phi(N_f-1, N_f-1) \end{pmatrix}_{N_f \times N_f}$$

Thus we have derived the least squares solution of the filter $\hat{\mathbf{f}}$. The inversion lemma is applied to obtain the way of recursively computing the inverse of the autocorrelation matrix $\mathbf{\Phi}^{-1}(n)$ from $\mathbf{\Phi}^{-1}(n-1)$ without direct matrix inversion [Chapter 9, [10]]. The resulting algorithm is known as the *Recursive Least-Squares* (RLS) algorithm. We will generalize this algorithm form linear equalization filter to *decision feedback* equalization filters.

Consider the following conceptual diagram of adaptive filtering generalized from Fig. 2-7 and Fig. 4-1:

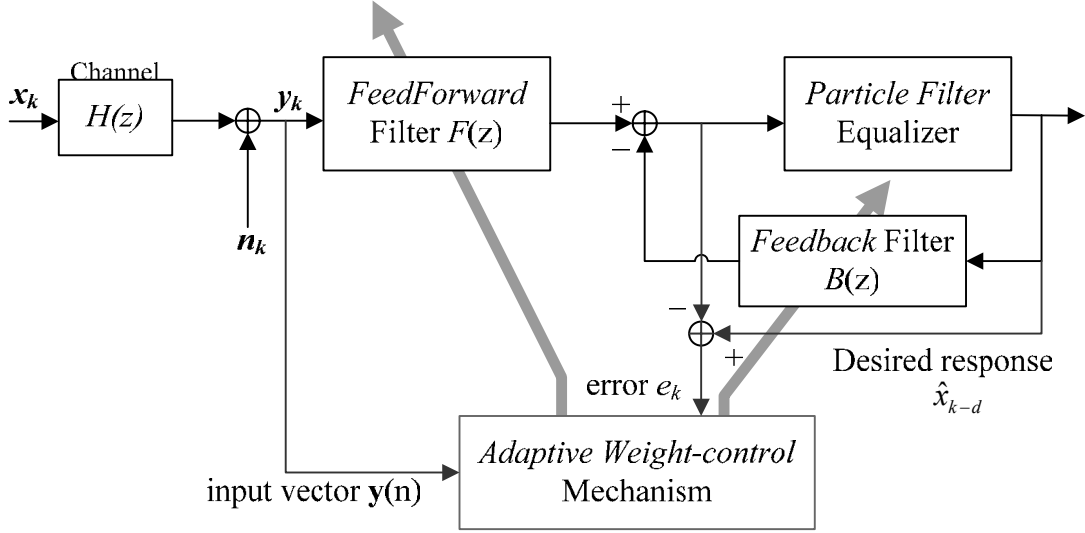


Fig. 4-2 System diagram of the SIS adaptive DFE.

The major difference between linear equalizer and the decision feedback equalizers is the error term. As shown in Fig. 4-2, the error term now can be written as:

$$\begin{aligned} e(k) &= \hat{x}(k-d) - (\mathbf{f}^H(n)\mathbf{y}(k) - \mathbf{b}^H(n)\hat{\mathbf{x}}(k-d)) \\ &\equiv \hat{x}(k-d) - \mathbf{w}^H(n) \cdot \mathbf{u}(k) \end{aligned} \quad (4-7)$$

$$\begin{aligned} \text{where } \mathbf{w}(n) &= [\mathbf{f}^T(n), -\mathbf{b}^T(n)]^T \\ &= [f_0(n), f_1(n), \dots, f_{N_f-1}(n), -b_0(n), -b_1(n), \dots, -b_{N_b-1}(n)]^T \\ \mathbf{u}(k) &= [\mathbf{y}^T(k), \hat{\mathbf{x}}^T(k-d)]^T \\ &= [y(k), y(k-1), \dots, y(k-N_f+1), \hat{x}(k-d), \hat{x}(k-d-1), \dots, \hat{x}(k-d-N_b+1)]^T \end{aligned}$$

Note that the delay d is added for the practical considerations. In this way, (4-7) has the same mathematical form as (4-2). Using the similar derivations, we can obtain

$$\tilde{\Phi} \cdot \hat{\mathbf{w}} = \tilde{\mathbf{z}}, \text{ or } \hat{\mathbf{w}} = \tilde{\Phi}^{-1} \cdot \tilde{\mathbf{z}} \quad (4-8)$$

with $\tilde{\mathbf{z}} = \mathbf{u}(k) \cdot \hat{\mathbf{x}}^*(k-d) = [\tilde{z}(0), \tilde{z}(-1), \dots, \tilde{z}(-N_f+1)]^T$,

$$\tilde{\Phi} = \mathbf{u}(k) \cdot \mathbf{u}^H(k) = \begin{pmatrix} \Phi_{yy} & \Phi_{yx} \\ \Phi_{xy} & \Phi_{xx} \end{pmatrix}_{(N_f+N_b) \times (N_f+N_b)}$$

where $\Phi_{yy} = \Phi$ is the time-averaged autocorrelation matrix of $\mathbf{y}(k)$, Φ_{yx} is the time-averaged cross-correlation matrix between $\mathbf{y}(k)$ and $\hat{\mathbf{x}}(k)$, and so forth. Similarly, we can derive the recursive computation of the inverse $\tilde{\Phi}^{-1}$ and the RLS algorithm simply by replacing $\mathbf{f}(n)$ with $\mathbf{w}(n)$, and $\mathbf{y}(k)$ with $\mathbf{u}(k)$.

The coefficients of the LMS based adaptive DFE can be derived likewise by solving the *Wiener-Hopf equation* instead of the normal equation (4-5).

4.2 ADAPTIVE BLIND SIS EQUALIZATIONS

To conduct the SIS algorithm based on the DEF structure, we have to consider a number of particles during the equalization process. Each of the particles has its own filter coefficients (FFF+FBF), weight, and the drawn particle trajectory $\hat{x}_{k,0}$. The drawn samples $\hat{x}_{k,0}^{(i)}$ and the filter coefficients ($\mathbf{f}_k^{(i)}$ and $\mathbf{b}_k^{(i)}$) of the i -th particle are used to update the filter coefficients and obtain $\mathbf{f}_{k+1}^{(i)}$ and $\mathbf{b}_{k+1}^{(i)}$. The final decision bit sequence would be determined when k reaches the end of a time frame. Unlike the structure in Fig. 4-2, each particle of the adaptive SIS-based DFE has the equivalent minimum-phase channel response estimated based on its own drawn particle trajectory by applying adaptive filtering algorithms (LMS, RLS, etc), as illustrated in Fig. 4-3.

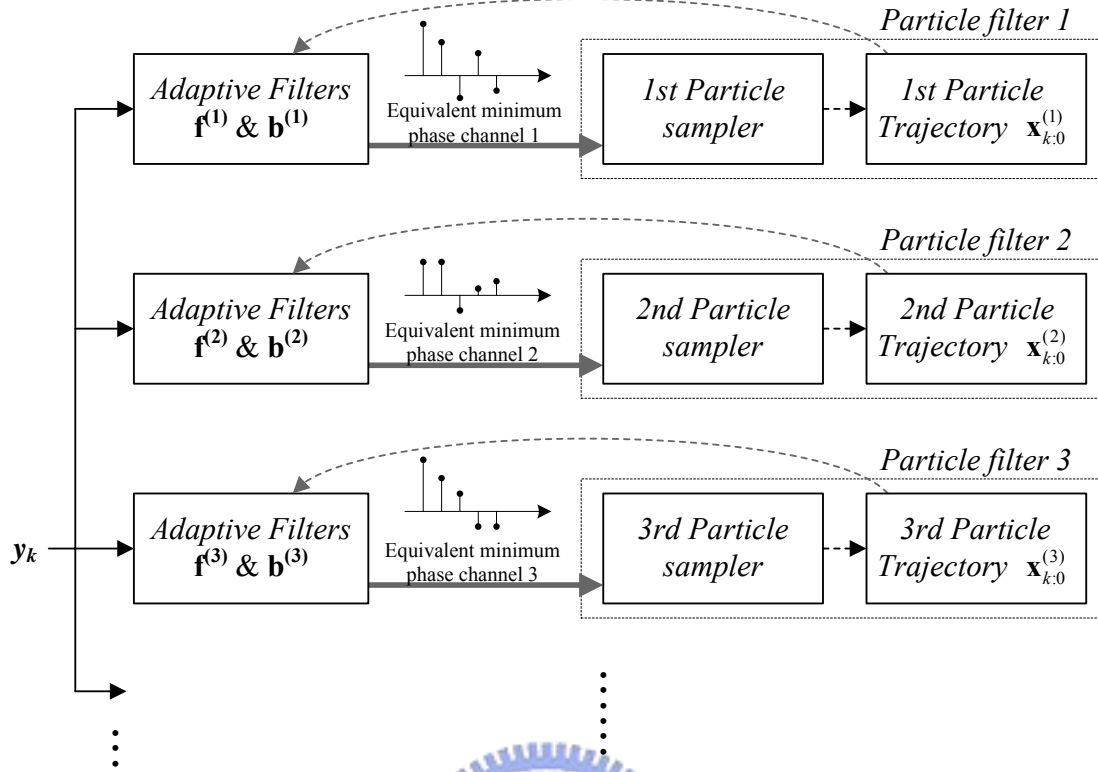


Fig. 4-3 Practical implementation of the SIS blind DFE.

We apply the RLS algorithm to the adaptive SIS-based DFE as an example. At an iteration, the i -th particle filter uses the trajectory $\hat{\mathbf{x}}_{k:0}^{(i)}$ to calculate the coefficient vector (4-9) and acquires the FFF and FBF coefficients (from $\hat{\mathbf{w}}^{(i)}$) to calculate the likelihood functions (4-10).

$$\hat{\mathbf{w}}^{(i)} = \mathbf{P}^{(i)} \cdot \tilde{\mathbf{z}}^{(i)}, \text{ where } \mathbf{P}^{(i)} = \left(\Phi^{(i)} \right)^{-1} \quad (4-9)$$

with $\tilde{\mathbf{z}}^{(i)} = \sum_{k=n_1}^{n_2} \mathbf{u}^{(i)}(k) \cdot x^{(i)*}(k-d) = [\tilde{z}^{(i)}(0), \tilde{z}^{(i)}(-1), \dots, \tilde{z}^{(i)}(-N_f+1)]^T$,

$$\tilde{\Phi}^{(i)} = \sum_{k=n_1}^{n_2} \mathbf{u}^{(i)}(k) \cdot \left(\mathbf{u}^{(i)}(k) \right)^H,$$

$$\begin{aligned} \mathbf{u}^{(i)}(k) &= \left[\mathbf{y}^T(k), \left(\mathbf{x}^{(i)}(k-d) \right)^T \right]^T \\ &= [y(k), y(k-1), \dots, y(k-N_f+1), x^{(i)}(k-d), x^{(i)}(k-d-1), \dots, x^{(i)}(k-d-N_b+1)]^T \end{aligned}$$

$$\hat{p}(r_k | x_{k-d}^{(i)} = \chi, \mathbf{x}_{k-d-1:0}^{(i)}) = \frac{1}{\sqrt{2\pi}\sigma_e} \exp\left(\frac{-|r_k^{(i)} - \mathbf{m}^H \mathbf{x}_{k-d}^{(i)}|^2}{2\sigma_e^2}\right) \quad (3-20)$$

$$= \frac{1}{\sqrt{2\pi}\sigma_e} \exp\left(\frac{-|\mathbf{f}^{(i)H} \cdot \mathbf{y}_k^{(i)} - \mathbf{b}^{(i)H} \mathbf{x}_{k-d-1}^{(i)} - \chi|^2}{2\sigma_e^2}\right)$$

$$= \frac{1}{\sqrt{2\pi}\sigma_e} \exp\left(\frac{-|(\mathbf{w}^{(i)})^H \cdot \mathbf{u}_k^{(i)} - \chi|^2}{2\sigma_e^2}\right) \quad (4-10)$$

where $r_k^{(i)} = \mathbf{f}^{(i)H} \cdot \mathbf{y}_k$, here we have chosen $\mathbf{b} = \mathbf{m} - [1, 0, \dots, 0]^T$ (since the first impulse response of the canonical response $M(z)$ would always be 1) The function of the RLS adaptive filter is to calculate the minimum phase solution, namely, the coefficients FFF and FBF for the PF whereas the PF is responsible for the decision of the next data samples being used for the RLS adaptation. As shown in the conceptual block diagram in Fig.4-4, they provide information to each other. In the SIS algorithm, the particle in which the RLS has the good estimate of the channel would be given a large weight and is likely to be preserved (via the particle weight adaptations and resampling⁷). We will see this example in the next chapter talking about simulations.

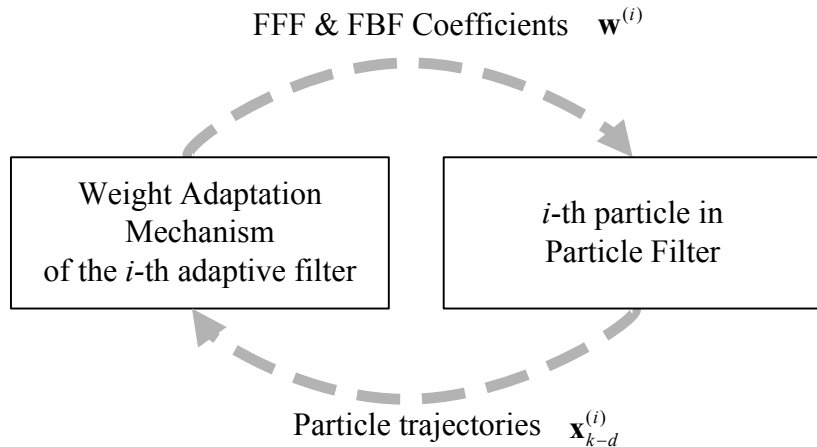


Fig. 4-4 Illustration of information transmission between i -th PF and adaptive filter.

⁷ Notably, the coefficients of the filters $\mathbf{w}^{(i)}$ and the matrix $\mathbf{P}^{(i)}$ should be copied and replaced together with the particle trajectories during the resampling process.

4.3 COMPLEXITY ISSUES AND THE MAX-WEIGHT SIS DFE

It is realizable that the complexity of the blind SIS DFE (as well as the blind delayed-SIS equalizer (blind D-SIS EQ)) is very high. Every particle filter has its own set of filter coefficients adapted individually for the blind SIS equalization. Take RLS as an example; each adaptive filter requires a $(N_f + N_b)$ by $(N_f + N_b)$ matrix $\mathbf{P}^{(i)}$ and several $(N_f + N_b)$ by 1 vectors and updates the coefficients with 8 matrix multiplications and 2 matrix additions at every time iteration for each particle. Obviously, the computation and storage requirement would be huge. To meet the broadband applications, the enormous hardware cost is required in practice. Another important task of this thesis is to reduce the number of adaptive filter coefficient adaptations to minimize the computation complexity and the storage required for the SIS equalization.

From the particle filtering theory, the particle trajectories with large weights are more likely to be preserved than those with small weights. A particle with the large weight will be replicated and replace the others in the procedure of resampling. Normally, the particle trajectory of the maximum weight would become the final decision sequence of the particle filter. We could use this feature to save the computation complexity. We use only *a single set* of adaptive filters and update the coefficients according to the trajectory of the *maximally weighted particle*. Thus, we can greatly reduce the complexity of the blind SIS DFE without suffering from great performance loss. The idea of this Max-Weight blind SIS DFE is illustrated in Fig. 4-5.

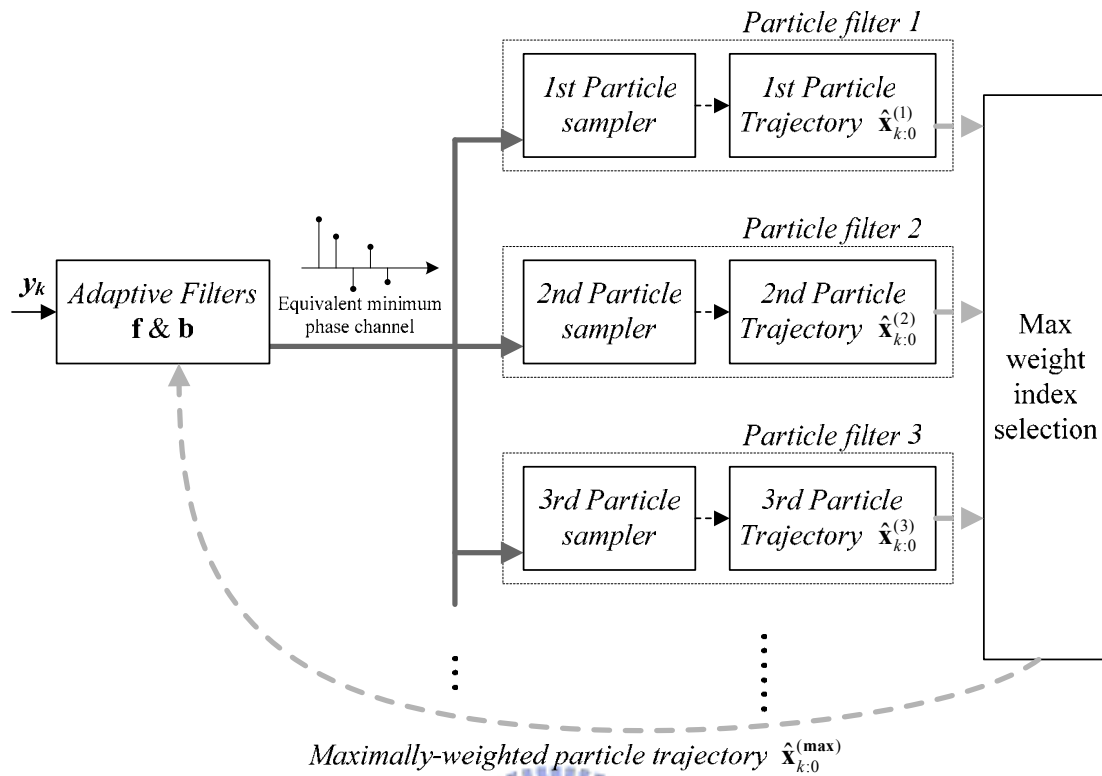


Fig. 4-5 Implementation of the Max-Weight blind PF DFE.

Compared with Fig. 4-3, this implementation requires only one set of adaptive filter so that the computation complexity is greatly reduced. The performance comparison of the Max-Weight blind SIS DFE and the original blind SIS DFE would be discussed in Chapter 5.

4.4 CHAPTER SUMMARY

In this chapter, we have proposed an adaptive blind SIS decision feedback equalizer (blind SIS DFE). Unlike the algorithms introduced in Chapter 3, this method does not need the knowledge of the CSI in advance in order to applying the minimum phase pre-filtering before the SIS equalization. But it requires N_s sets of adaptive filters and the adaptation processes for N_s individual particles of the blind SIS equalization. The high computation requirement can be alleviated by the Max-Weight SIS DFE algorithm proposed in 4.3. The proposed Max-Weight blind SIS DFE algorithm indeed outperform the original SIS equalization proposed in [3] especially in the environment with the weak LOS.

Compared with the method of blind delayed-SIS equalization, which was also proposed in [3] to solve the weak LOS problem, the Max-Weight blind SIS equalization algorithm can provide better performance at lower computation complexity requirement. As shown in (2-23) of 2.3.1, the computational complexity of the delayed-SIS grows *exponentially* with the delay d . In addition, the selection of the delay d in the blind delayed SIS algorithm has the big impact on the performance. Unfortunately, increasing the delay d does not necessarily improve the performance. This effect and the performance these algorithms will be observed by using computer simulation in the next chapter.

5 COMPUTER SIMULATIONS & OBSERVATIONS

In this chapter, we will show the computer simulation results based on the aforementioned algorithms. The purpose is to confirm the conclusions we obtain from the mathematical performance analysis in the previous chapters. We first performed the simulations by assuming that the CIR is known. The simulation results coincide with the analysis results in Chapter 3. We also do the performance comparison of the regular SIS equalizer [3], the minimum-phase SIS DFE, and the delayed-SIS EQ. In the second section, we provide the simulation results of adaptive *blind* SIS equalization algorithms as mentioned in Chapter 4. We compare the performance of the blind SIS EQ, the adaptive minimum-phase SIS DFE and the blind Delayed-SIS EQ with different delays d . Finally we make some simple conclusion.

5.1 PERFECT CHANNEL STATE INFORMATION

In this section, we examine the results we obtain in the BER analysis of SIS equalization in Chapter 3 under the assumption that the channel state information (CSI) is perfectly known. In this case, the fixed CIR is used to calculate the minimum-phase decision feedback filter coefficients according to the method in Appendix D.

In the simulation, we normalize the CIR to have the unity power, i.e. $\sum_{l=0}^{M-1} |h_l|^2 = 1$. The number of the particles N_s is 100. The Max-Log SIS weighting parameter is set as $\gamma = 0.75$. The signals are transmitted equally likely with the BPSK modulation. In this section, we consider three types of equalizers during the simulations:

- (1). *The SIS equalizer* (SIS EQ): This is the original particle filter using the Max-Log SIS algorithm as proposed in 2.2.2. The BER and the convergence

behaviors are analyzed in 3.1 and 3.2.

(2). *The (minimum-phase) SIS decision feedback equalizer (SIS DFE)*: This algorithm is proposed in 2.3.2. The coefficients are calculated based on the given CIR according to two criteria – *zero forcing (ZF)* and *minimum mean square error (MMSE)*. The structures of this algorithm based on these two criteria are illustrated in Fig. 3-6 and Fig. 3-7.

(3). *The delayed-SIS equalizer (D-SIS EQ)*: The delayed-SIS algorithm is proposed in [3] and described in 2.3.1 of this thesis. Different from [3], the channel is assumed to be known here. We will see the effect of the delay d by conducting the simulation with different values of d .

To observe how the channel LOS amplitude affects the performance, we consider the following two cases of CIR separately: the first case is the CIR with a large LOS and the second one is the opposite case which has an attenuated LOS.

5.1.1 Channel with a strong LOS

We begin with a channel with a strong LOS. As shown above in Fig. 5-1, there is not much difference between the SIS EQ and the SIS DFE when the first impulse of the channel response is large. That is because the most energy of this channel is concentrated at the first channel impulse. The minimum phase DFE has very limited effect on this kind of channel, as described in 3.3.2. Nevertheless, we can find that the D-SIS EQ behaves well as $d = 0$. However, when the delay $d = 2$, which is larger than the actual channel delay (in this case, channel delay = 0), the performance deteriorates greatly in spite of spending extra computation. This problem can be further testified in the case when channel LOS is weak.

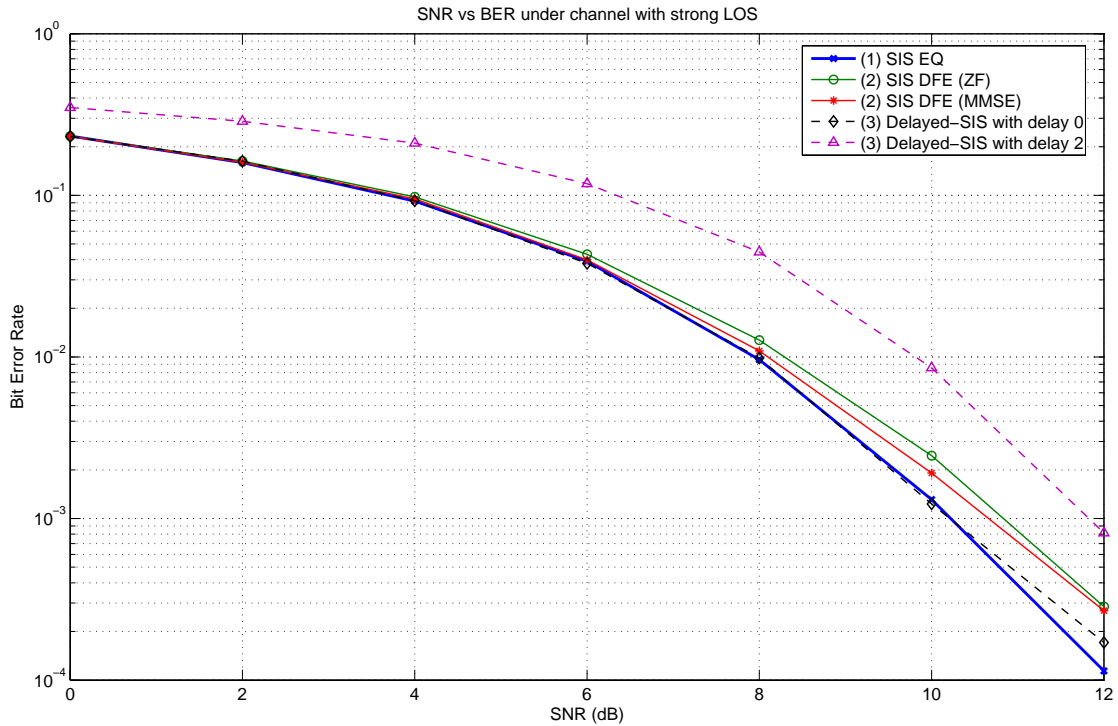


Fig. 5-1 BER versus SNR plots of the three equalizers under channel with a strong LOS. (Perfect CSI)

5.1.2 Channel with a weak LOS

In this subsection, we will provide the simulation results to see how well the minimum-phase SIS DFE and the D-SIS EQ improve the system performance when the channel LOS is weak. One of the simulations is purposed to compare the performance of the original SIS EQ under channels with different values of LOS, as shown in Fig. 5-2. We can see that the SIS EQ operating with weak LOS channels suffers from error propagation and has a relatively high BER in the high SNR (as SNR = 8~12) whereas the BER of the SIS EQ operating with strong LOS channels drops significantly. This phenomenon coincides with the analytical results we have obtained in Fig. 3-1 and Fig. 3-2. In other words, we have shown that, by both means of analyses and simulations, the SIS EQ indeed has the performance decay problem when the channel LOS is attenuated.

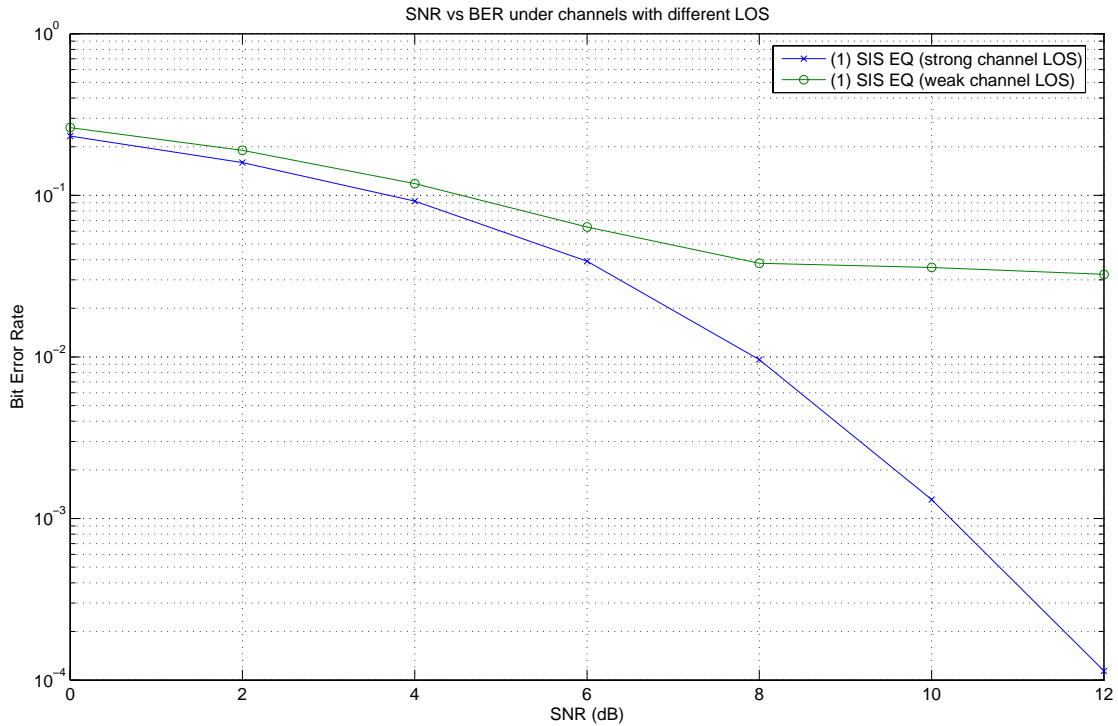


Fig. 5-2 BER versus SNR plots of the SIS EQ under channel with different LOS.

(Perfect CSI)

The next simulation is to compare the performance of the SIS EQ and the minimum-phase SIS DEF (both ZF and MMSE) solely in the case when channel LOS is weak. As shown in Fig. 5-3, compared with the SIS EQ, the minimum-phase SIS DFEs (either ZF or MMSE) successfully attain lower BERs, especially in the high SNR. We can see that the BER versus SNR curves of these two SIS DFEs act just like the one of the SIS EQ under a *good*⁸ channel (the one in Fig. 5-1). This implies the SIS DFE has successfully converted the *poor* channel into a *good* one by the minimum phase pre-filtering.

In addition, comparing the curves of the ZF and the MMSE SIS DFEs in Fig. 5-3, we can see that although they have little difference in the high SNR, the MMSE DFE indeed outperforms the ZF DFE in all SNR values. (The difference between the MMSE and the ZF in the low SNR looks very little. However, the difference is

⁸ In this chapter, we say that a channel is *good* if its LOS is strong (good for the SIS EQ), and, oppositely, a channel is said to be *poor* when its LOS is weak.

actually larger than what we can see in the figure because the BER is drawn in log scale.) This observation testifies the conclusions of 3.3.2, that is, the MMSE DFE would have a better performance than the ZF DFE in the low SNR, and their performance would come close to each other when the SNR increases (the noise variance approaches to zero).

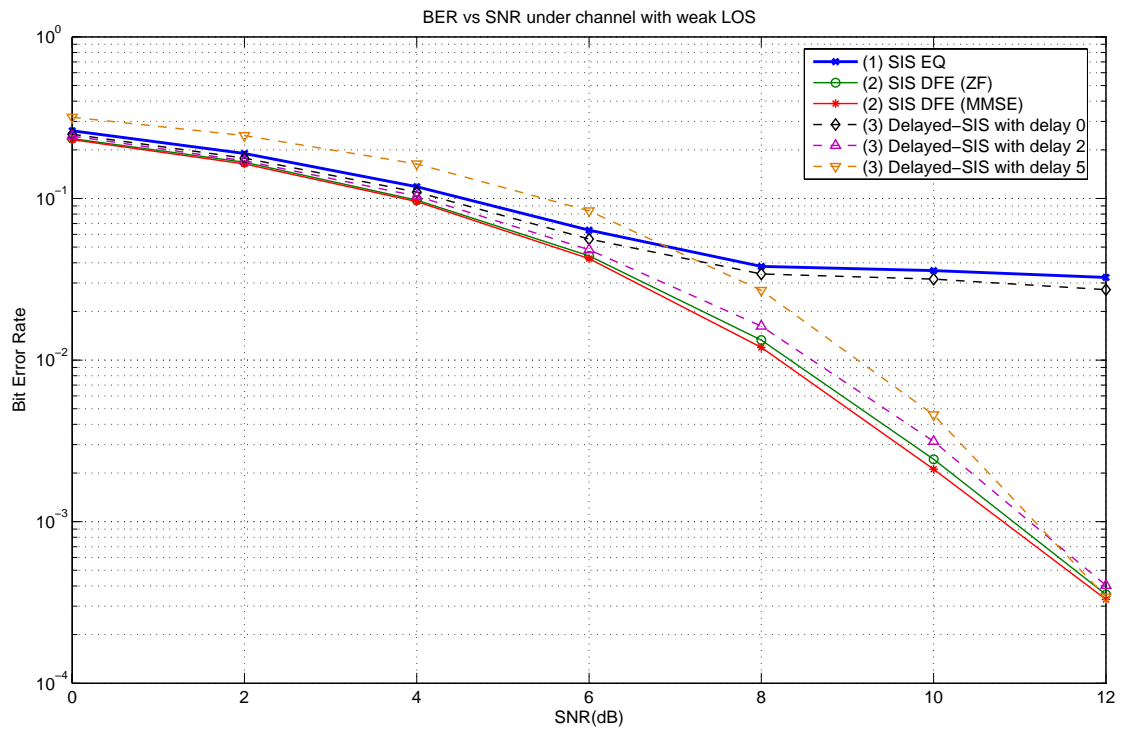


Fig. 5-3 BER versus SNR plots of the three equalizers under channel with weak LOS. (Perfect CSI)

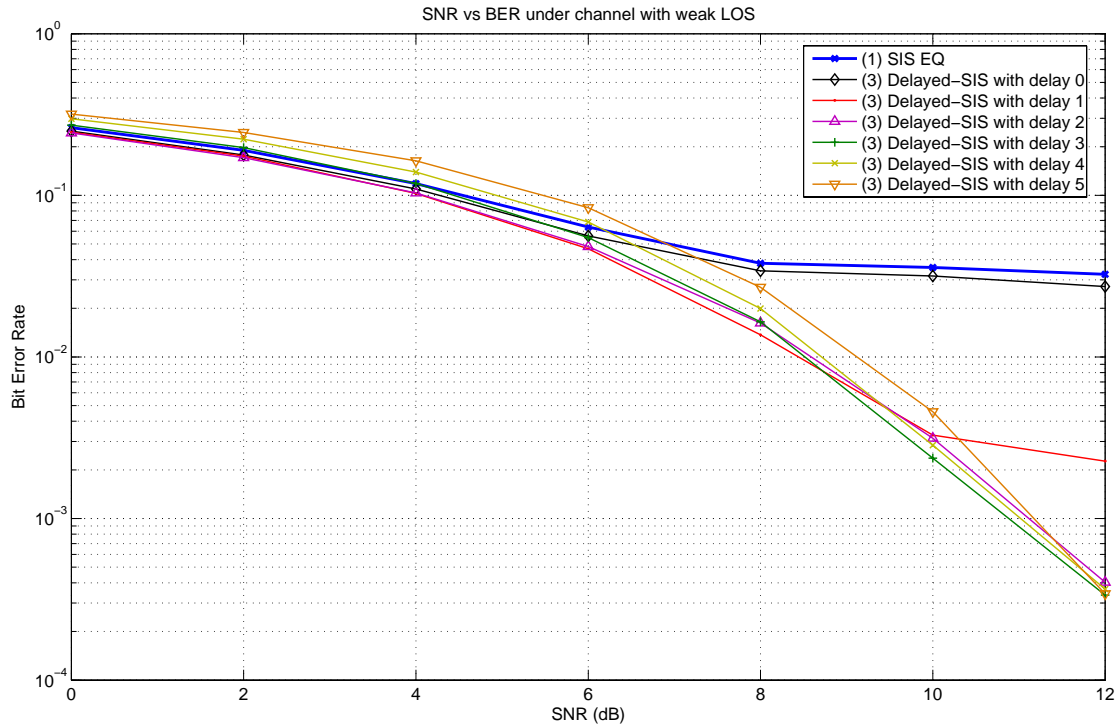


Fig. 5-4 Comparison of BER versus SNR plots of the D-SIS EQ with different delay d under channel with weak LOS. (Perfect CSI)

At last, let us take a close look at the effect of the delay d on the performance of the D-SIS EQ in Fig. 5-4. In this simulation, the channel has a poor LOS, and the highest channel gain appears at the 2nd impulse. That means the D-SIS EQ would not have good performance unless the delay d is large enough to cover the largest channel impulse (i.e. $d \geq 2$). If we observe the BER versus SNR curves of different values of d ($d = 0, 1, \dots, 5$) in Fig. 5-4, we can see that the D-SIS EQ has the best performance when $d = 2$ and 3 (around the real channel delay 2). However, it is interesting that the performance is worsen not only when the delay is not large enough ($d = 0, 1$), but also when the delay is too large ($d = 4, 5$). Therefore it is very important for the D-SIS EQ to have the delay d chosen properly, which is nearly impossible in the blind equalization scenarios. Furthermore we can observe that, despite of the high computation complexity, the D-SIS EQ, with all values of d , still does not have a better performance than the MMSE SIS DFE.

From the observations we have from the above simulation results, we can conclude that the minimum-phase SIS DFE scheme is definitely an attractive alternative of the particle filtering based equalization algorithm to overcome the channel impairment.

5.2 BLIND AND ADAPTIVE EQUALIZATIONS

In this section, we will provide the simulation results of the *blind SIS equalization* schemes as introduced in Chapter 4. The channel state information is assumed to be unknown to the receiver (equalizer) and is estimated/tracked using the recursive least-squares (RLS) algorithm in the simulation.

Same as the previous simulation, we assume that the CIR has unity power, i.e. $\sum_{l=0}^{M-1} |h_l|^2 = 1$. The number of the particles N_s is 100. The Max-Log SIS weighting parameter is set as $\gamma = 0.75$. The signals are transmitted equally likely with BPSK modulations. In this section, we consider four types of equalizers in the simulation:

- (1). *The blind SIS equalizer* (Blind SIS EQ): This is the original particle filter based equalization using the Max-Log SIS algorithm as proposed in 2.2.2 with the CIR estimated by a set of RLS adaptive filters for each particle.
- (2). *The blind SIS decision feedback equalizer* (Blind SIS DFE): This is the scheme we proposed in 4.2. The FFF and FBF filter coefficients are updated with the RLS algorithm introduced in 4.1.
- (3). *The blind delayed-SIS equalizer* (Blind D-SIS EQ): This is the delayed-SIS algorithm proposed in [3]. The CIR here is estimated in the same way as the blind SIS EQ in (1) through RLS adaptive filters.
- (4). *The Max-Weight blind SIS decision feedback equalizer* (MW blind SIS DFE): This is the complexity reduced version of (2), the blind SIS DFE, as

introduced in 4.3. Only one set RLS adaptive filter coefficients is used for 100 ($=N_s$) particles.

Again we consider two cases of CIR and see the performance of these blind equalizers.

5.2.1 Channel with a strong LOS

First in the case of strong LOS channels, the blind SIS EQ behaves as well as the blind SIS DFE does. This is because when the channel LOS is large, there is no much difference before and after the minimum-phase pre-filtering. These two algorithms result in the similar performance.

As far as the blind D-SIS algorithm is concerned in this case, the best choice of the delay d should be 0 because the channel has the largest impulse at h_0 . The blind D-SIS with delay $d = 0$ acts exactly as the blind SIS EQ. In addition, as we have testified in previous section (the perfect CSI scenario), choosing a larger d than the channel delay would not only rise the computation complexity but also cause the performance decay. Compared with the D-SIS EQ with $d = 2$ in the known CSI case in Fig. 5-1, the blind D-SIS EQ with $d = 2$ even has worse performance, especially in low SNR. This is because when the SNR is low, the particle filters are likely to draw the particles wrongly. The adaptation of the RLS filter according to the bad particles would be erroneous. The information exchange between the RLS filters and the SIS algorithm becomes a vicious circle, the error propagates through the iteration and hence raised the BER.

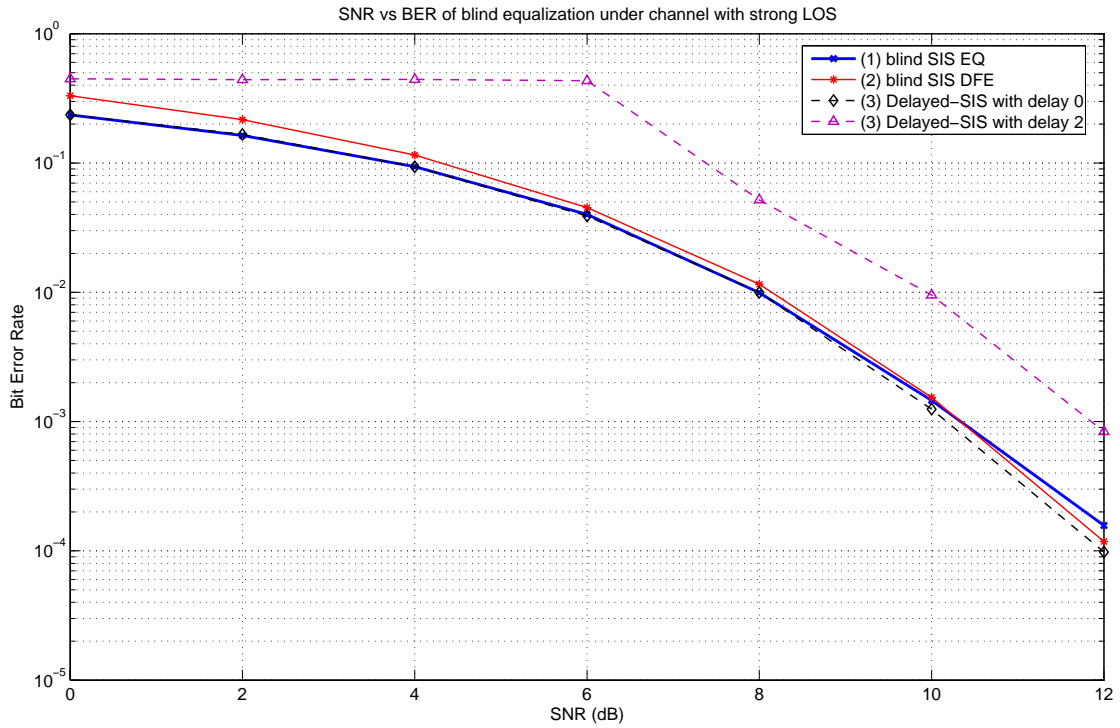


Fig. 5-5 BER versus SNR plots of the different blind equalizations under channel with strong LOS.

5.2.2 Channel with a weak LOS

Now we turn to the case when the channel LOS is seriously attenuated to see how the SIS DFE and D-SIS EQ improve the situation in blind equalization.

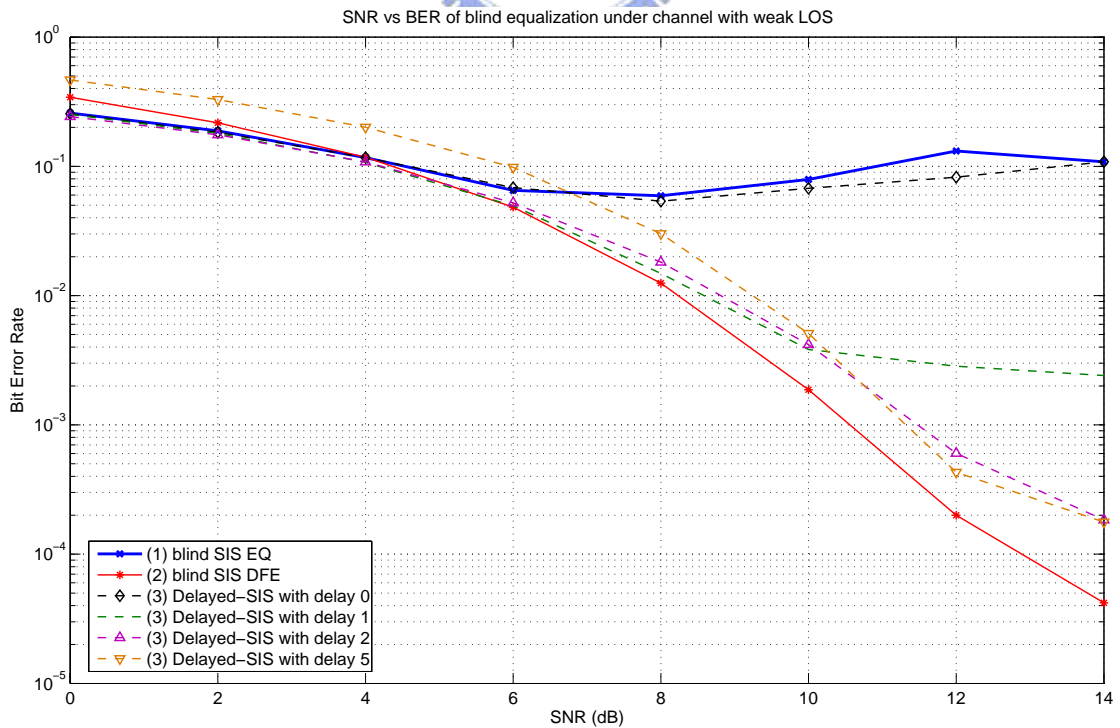


Fig. 5-6 BER versus SNR plots of the three blind equalizations under channel with weak LOS.

First we observe the curve of the SIS EQ: It is interesting that the curve climbs up with the SNR in the range of SNR = 8~12. This phenomenon has been found earlier in the analysis in Chapter 3, as drawn in Fig. 3-2, when the error propagation factor $\lambda_k^{(i)}$ is quite large. As expected, the SIS EQ does not attain a desirable low BER at the high SNR when it is used in the weak-LOS channel environment.

Second, the results of the D-SIS algorithm indicate again whether the performance is good enough or not is greatly dependent on the selection of d . Only when d is selected to be close to the channel delay (in this case, the channel delay is 2.) can the D-SIS EQ have the best performance. Similar to the case of the known channel with a weak LOS in Fig. 5-4, we can see that the D-SIS EQ with delay $d = 2$ has the best performance compared with that with other values of delay.

We end up with the simulation for the performance comparison of the blind SIS DFE and the Max-Weight blind SIS DFE, which utilizes the maximally-weighted particle to update *the only one* set of adaptive filter at each iteration (see 4.3). As we expected, this method would have a little worse performance than the blind SIS DFE because of the simplification. However, we have found (in Fig. 5-7) that the performance difference is small. This is a promising result because, we can use the Max-Weight blind SIS DFE, to save a lot of computations without sacrificing much performance. As shown in Fig 5-7, we can find that the Max-Weight blind SIS DFE even outperforms the D-SIS EQ (delay = 2) in the high SNR. Under the situations when the computation resources are relatively limited, the Max-Weight SIS DFE may become one of the appealing options for blind equalization.

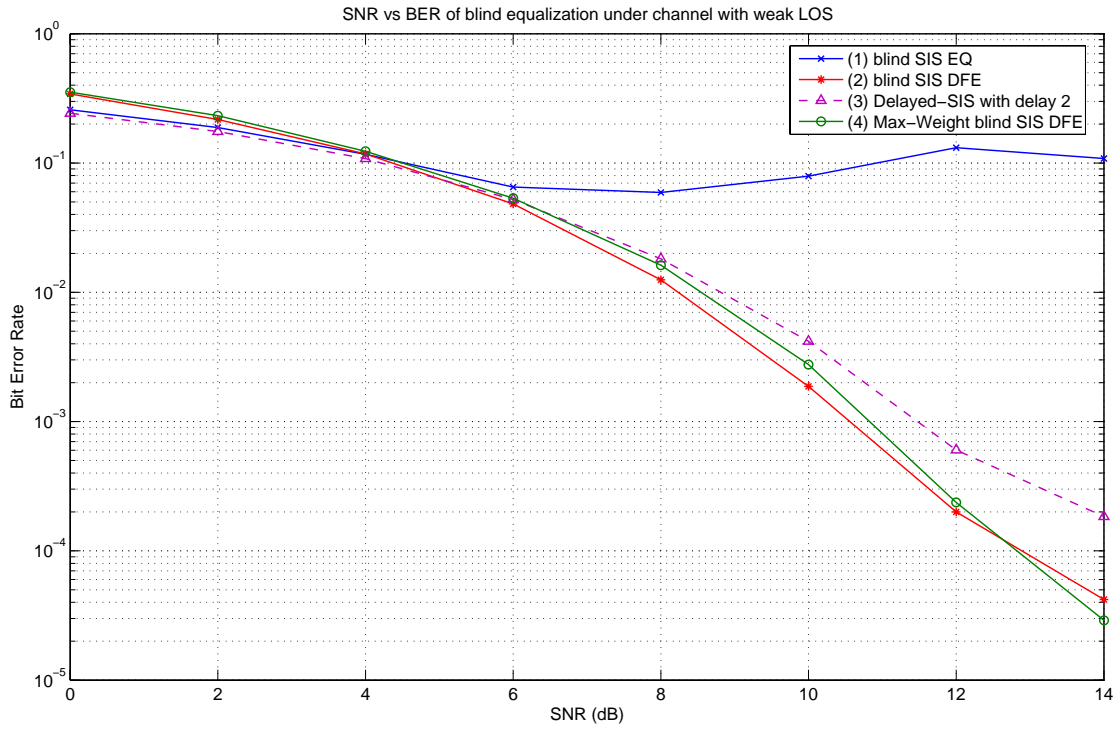


Fig. 5-7 BER versus SNR plots of the four different blind equalizations under channel with weak LOS.



5.3 CHAPTER SUMMARY

In this chapter, we have conducted the computer simulation of the particle filter based equalization algorithm described in the previous chapters. From the simulation results, we have testified that the weak LOS problem indeed affects the performance of the SIS EQ, as indicated in the BER analysis of the SIS EQ. We have performed the simulation on the ZF and the MMSE SIS DFE algorithms, showing their performance is not affected by the weak LOS channel, We also made simple comparison between the proposed SIS DFE and the D-SIS EQ.

In the second part, we simulated the blind SIS equalizers introduced in Chapter 4. Both the blind SIS DFE and the blind D-SIS EQ can improve the system performance under the weak LOS channel. However, the D-SIS EQ requires the knowledge of the channel delay in order to determine the best value of the delay d , and its computation complexity grows exponentially with its delay d . The blind SIS DFE can be further simplified to the Max-Weight blind SIS DFE without the significant loss of performance. The cost-effective Max-Weight blind SIS DFE can be considered a promising algorithm for the practical system implementation.

6 CONCLUSIONS AND PROSPECTIVE

6.1 CONCLUSION

Particle filter is a simulation-based algorithm utilizing the Monte Carlo methodologies. The Sequential Importance Sampling (SIS) algorithm is successfully applied to the channel equalization applications.

One of the major problems the SIS based equalization is the performance degeneracy caused by the weak-LOS in the channel response. In this thesis, we begin with the BER analysis of the SIS equalization and show how the energy of LOS affects the performance. To overcome this weak LOS problem, we employ the minimum-phase pre-filtering technique to transform the weak LOS channel into the equivalent strong LOS channel for the SIS equalization. This idea is realized by proposing SIS decision feedback equalizers. We mathematically prove that the MMSE DFE and the ZF PF DFE provide better performance than the original SIS EQ.

In addition, we compare the proposed algorithms with the Delay-SIS equalizer (D-SIS EQ), which was proposed to solve the weak LOS problem in [3] in the viewpoints of performance and computation complexity. We also indicate that the good performance of the D-SIS EQ is fatally determined by the proper selection of the delay d , which is hard to achieve in the real implementation.

The blind SIS based EQ algorithms are realized by using the techniques of adaptive filtering in the cases of unknown CSI and timing-varying channels. We propose a simplified blind SIS DFE algorithm, named the Max-Weight blind SIS DFE, which saves much computation complexity without causing obvious performance loss. From the computer simulation, we compare the performance of the particle filter based equalization algorithms and summarize the observations obtained from the

simulation results. We conclude that the Max-Weight blind SIS DFE is a promising algorithm, which has good performance, robustness (less limitations), and much lessened computation complexity.

6.2 PROSPECTIVE AND THE FUTURE WORKS

At the end of this thesis, we would like to propose some innovations or directions for the further research of the particle filter based equalizers.

The first work is the low-complexity blind SIS DFE algorithm. Although we have proposed the method of Max-Weight SIS DFE, there is still some room for improvement. For example, the Max-Weight method may not be stabilized in the beginning when the particle weights are very close to each other. Therefore, we may have to wait for a period of time T (after the convergence of the particle weights) and then pick up the particle with the largest weight. That is, we choose the maximally-weighted particle and update the adaptive filters for every $k = nT$. The operation would be similar to a certain kind of *windowing* method which has been frequently used in digital signal processing.

In addition, the most important information that the particle filters provide is the desired *probability distribution*. As some kind of *soft information*, the estimated distributions can be used to help the backend decoders. Therefore we may apply the technique of particle filtering to the *turbo equalizers*, which exchanges information between the equalizer and decoder iteratively to attain a desirable performance.

On the other hand, in sight of the great advance in the wireless communication, the multiple-input multiple-output (MIMO) wireless communication systems have drawn lots of interests in the recent years. The SIS-based algorithm can be further used for the blind equalization and detection in the MIMO system. I am convinced that the particle filter can provide some equilibrium between the implementation

complexity and the performance for the MIMO systems.

In sum, the importance sampling algorithm and the particle filtering can be applicable to any physical problems that acquire the knowledge of probability densities by estimating the them with the discrete random measures. We have introduced a probable application in this thesis. Whether the PF and the SIS algorithm may be applicable in other practical situations depends on their computation resources, the response time they require, and the performance they require, and so on.



REFERENCES

- [1]. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, “A tutorial on particle filters for on-line nonlinear/non-Gaussian Bayesian tracking,” *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 174-188, February 2002.
- [2]. P.M. Djuric, J.H. Kotecha, J. Zhang, Y. Huang, T. Ghirmai, M.F. Bugallo, and J. Miguez, “Particle filtering,” *IEEE Signal Processing Magazine*, vol. 20, no. 5, pp.19-38, September 2003.
- [3]. J. Míguez and P.M. Djurić, “Blind equalization of frequency-selective channels by sequential importance sampling,” *IEEE Transactions on Signal Processing*, vol. 52, no. 10, pp. 2738-2748, October 2004.
- [4]. Simon Haykin, *Modern Filters*. Macmillan, 1989.
- [5]. J.M. Cioffi, G.P. Dudevoir, M.V. Eyuboglu, and G.D. Forney, Jr, “MMSE decision-feedback equalizers and coding—part I: Equalization results,” *IEEE Transactions on Communications*, vol. 43, no. 10, pp. 2582-2594, October 1995.
- [6]. Athanasios Papoulis, *Probability, Random Variables, and Stochastic Processes*. 2nd Edition, McGraw-Hill, 1984.
- [7]. A.V. Oppenheim, R.W. Schaffer, and J.R. Buck, *Discrete-Time Signal Processing*. 2nd Edition, Prentice Hall, 1999.
- [8]. A. Doucet, *On sequential simulation-based methods for Bayesian filtering*. Technical report, CUED/F-INFENG/TR 310, Signal Processing Group, Department of Engineering, University of Cambridge, 1998.
- [9]. A. Doucet and X. Wang, “Monte Carlo methods for signal processing,” *IEEE Signal Processing Magazine*, vol. 22, no. 6, pp. 152-170, November 2005.
- [10]. Simon Haykin, *Adaptive Filter Theory*. 4th Edition, Prentice Hall, 2002.
- [11]. A. Doucet, S. Godsill, and C. Andrieu, “On sequential Monte Carlo sampling

- methods for Bayesian filtering,” *Statistics and Computing*, vol. 10, no. 3, pp. 197–208, 2000.
- [12]. A. Kong, J.S. Liu, and W.H. Wong, “Sequential imputations and Bayesian missing data problems,” *Journal of the American Statistical Association*, vol. 89, pp. 278–288, 1994.
- [13]. B. Dong, X. Wang, and A. Doucet, "A new class of soft MIMO demodulation algorithms," *IEEE Transactions on Signal Processing*, vol. 51, no. 11, pp. 2752-2763, November 2003.
- [14]. Y.T. Su, X.D. Zhang, and X.L. Zhu, "A low-complexity sequential Monte Carlo algorithm for blind detection in MIMO systems," *IEEE Transactions on Signal Processing*, vol. 54, no. 7, pp. 2485- 2496, July 2006.



APPENDIX

A. JACOBIAN APPROXIMATION

The Jacobian approximation is basically the following equation:

$$\log(\exp(a) + \exp(b)) \cong \max(a, b)$$

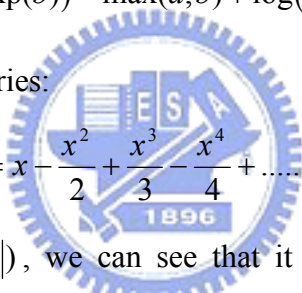
To obtain this equation, first we look at the case $a > b$, then

$$\begin{aligned} & \log(\exp(a) + \exp(b)) \\ &= \log(\exp(a) \cdot (1 + \exp(b) / \exp(a))) \\ &= \log(\exp(a) \cdot (1 + \exp(-(a - b)))) \quad \text{recall that } a > b \\ &= a + \log((1 + \exp(-(a - b)))) \end{aligned}$$

Considering jointly the case $b > a$, we can write the general form

$$\log(\exp(a) + \exp(b)) = \max(a, b) + \log(1 + \exp(-|a - b|))$$

Now recall from the Taylor Series:


$$\log(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots \quad \text{if } |x| < 1$$

By identifying x as $\exp(-|a - b|)$, we can see that it is always true that $|x| < 1$, and

$$\begin{aligned} \log(\exp(a) + \exp(b)) &= \max(a, b) + \log(1 + e^{-|a-b|}) \\ &= \max(a, b) + e^{-|a-b|} - \frac{e^{-2|a-b|}}{2} + \frac{e^{-3|a-b|}}{3} - \dots \end{aligned}$$

Note that the argument in the exponent only has to be a little bigger than zero to cause all of these extra terms to become small quickly. The *Max-Log algorithm* basically ignores the all the exponential terms and ends up with the following approximation:

$$\log(\exp(a) + \exp(b)) \cong \max(a, b)$$

B. DERIVATION OF THE EXPECTED VALUES IN THE BER DERIVATION (3-1)

To obtain the expected value $E[\psi_{k,+A}^{(i)} | b_k = 0]$ and $E[\psi_{k,+A}^{(i)} | b_k = 1]$, we consider the two cases that $b_k = 0$ and $b_k = 1$, respectively:

1. Case $b_k = 0$

To compute $E[\psi_{k,+A}^{(i)} | b_k = 0]$, we first consider the case when $b_k = 0$. Define

$\beta_{k,+A}^{(i)}$ as the error term $(y_k - A \cdot h_0 - \sum_{l=1}^{M-1} h_l \cdot x_{k-l}^{(i)})$ in the exponent of (2-13) and

$\beta_{k,-A}^{(i)}$ as $(y_k + A \cdot h_0 - \sum_{l=1}^{M-1} h_l \cdot x_{k-l}^{(i)})$ in (2-14) then:

$$\begin{aligned}\beta_{k,+A}^{(i)} \Big|_{b_k=0} &= (y_k - A \cdot h_0 - \sum_{l=1}^{M-1} h_l \cdot x_{k-l}^{(i)}) \Big|_{b_k=0} = \varepsilon_k^{(i)} \\ \beta_{k,-A}^{(i)} \Big|_{b_k=0} &= (y_k + A \cdot h_0 - \sum_{l=1}^{M-1} h_l \cdot x_{k-l}^{(i)}) \Big|_{b_k=0} = 2Ah_0 + \varepsilon_k^{(i)}\end{aligned}$$

where

$$\begin{aligned}\varepsilon_k^{(i)} &= (\sum_{l=1}^{M-1} h_l x_{k-l} + n_k) - \sum_{l=1}^{M-1} h_l x_{k-l}^{(i)} \\ &= \sum_{l=1}^{M-1} h_l d_{k-l}^{(i)} + n_k\end{aligned}$$

and

$$d_{k-l}^{(i)} = x_{k-l} - x_{k-l}^{(i)}$$

If M is sufficiently large, and the inter-symbol interference (ISI) is the sum of many similarly-sized components, we can approximate variables with Gaussian distribution through Central Limit Theorem:

$$\begin{aligned}\beta_{k,+A}^{(i)} \Big|_{b_k=0} = \varepsilon_k^{(i)} &\sim N(0, 4A^2 \lambda_k^{(i)} + \sigma_n^2) \\ \beta_{k,-A}^{(i)} \Big|_{b_k=0} = 2Ah_0 + \varepsilon_k^{(i)} &\sim N(2Ah_0, 4A^2 \lambda_k^{(i)} + \sigma_n^2)\end{aligned}$$

where $\lambda_k^{(i)} = \sum_{l=1}^{M-1} |h_l|^2 P_{err,k-l}^{(i)}$.

Under the assumption of equal priori probability, the condition $\rho_{k,+A}^{(i)} > \rho_{k,-A}^{(i)}$ in

(2-18) can be equivalent to the following conditions in terms of $\varepsilon_k^{(i)}$.

$$\begin{aligned}\rho_{k,+A}^{(i)} &\geq \rho_{k,-A}^{(i)} \\ \Leftrightarrow |\varepsilon_k^{(i)}|^2 &\leq |2Ah_0 + \varepsilon_k^{(i)}|^2 \\ \Leftrightarrow \varepsilon_k^{(i)} Ah_0 &\geq -A^2 |h_0|^2\end{aligned}$$

and the threshold $\psi_{k,+A}^{(i)}$ can also be written in terms of $\varepsilon_k^{(i)}$:

$$\psi_{k,+A}^{(i)} = \begin{cases} 1 - \frac{1}{2} \cdot \exp\left(-2\gamma(A^2|h_0|^2 + Ah_0\varepsilon_k^{(i)})/\sigma_n^2\right), \varepsilon_k^{(i)} Ah_0 \geq -A^2|h_0|^2 \\ \frac{1}{2} \cdot \exp\left(+2\gamma(A^2|h_0|^2 + Ah_0\varepsilon_k^{(i)})/\sigma_n^2\right), \varepsilon_k^{(i)} Ah_0 < -A^2|h_0|^2 \end{cases}$$

According to all the equations above in this section, we can finally obtain:

$$\begin{aligned} E[\psi_{k,+A}^{(i)} | b_k = 0] &= \frac{1}{\sqrt{2\pi(4A^2\lambda_k^{(i)} + \sigma_n^2)}} \left(\int_{\varepsilon_k^{(i)} Ah_0 \geq -A^2|h_0|^2} \left(1 - \frac{1}{2} e^{-2\gamma(A^2|h_0|^2 + Ah_0\varepsilon_k^{(i)})/\sigma_n^2}\right) \cdot e^{-|\varepsilon_k^{(i)}|^2/2(4A^2\lambda_k^{(i)} + \sigma_n^2)} d\varepsilon_k^{(i)} + \int_{\varepsilon_k^{(i)} Ah_0 < -A^2|h_0|^2} \frac{1}{2} e^{+2\gamma(A^2|h_0|^2 + Ah_0\varepsilon_k^{(i)})/\sigma_n^2} \cdot e^{-|\varepsilon_k^{(i)}|^2/2(4A^2\lambda_k^{(i)} + \sigma_n^2)} d\varepsilon_k^{(i)} \right) \\ &= \mathcal{Q}\left(-A|h_0| \sqrt{\frac{1}{4A^2\lambda_k^{(i)} + \sigma_n^2}}\right) + \frac{1}{\sqrt{2\pi(4A^2\lambda_k^{(i)} + \sigma_n^2)}} \left(-\frac{1}{2} \int_{-A|h_0|}^{\infty} e^{-2\gamma(A^2|h_0|^2 + Ah_0|\varepsilon_k^{(i)}|)/\sigma_n^2} \cdot e^{-|\varepsilon_k^{(i)}|^2/2(4A^2\lambda_k^{(i)} + \sigma_n^2)} d\varepsilon_k^{(i)} + \frac{1}{2} \int_{-\infty}^{-A|h_0|} e^{+2\gamma(A^2|h_0|^2 + Ah_0|\varepsilon_k^{(i)}|)/\sigma_n^2} \cdot e^{-|\varepsilon_k^{(i)}|^2/2(4A^2\lambda_k^{(i)} + \sigma_n^2)} d\varepsilon_k^{(i)} \right) \\ &= \mathcal{Q}\left(-\sqrt{\frac{h_0^2}{4\lambda_k^{(i)} + 1/\zeta}}\right) - \frac{1}{4} e^{2h_0^2\zeta^2(4\zeta\lambda_k^{(i)} + 1/\zeta)} \cdot \left(2 \cdot \mathcal{Q}\left(\gamma(8\zeta\lambda_k^{(i)} + 2 - 1/\gamma) \sqrt{\frac{h_0^2}{4\lambda_k^{(i)} + 1/\zeta}}\right) \right) + \frac{1}{4} e^{2h_0^2\zeta^2(4\zeta\lambda_k^{(i)} + 1/\zeta)} \cdot \left(2 \cdot \mathcal{Q}\left(\gamma(8\zeta\lambda_k^{(i)} + 2 + 1/\gamma) \sqrt{\frac{h_0^2}{4\lambda_k^{(i)} + 1/\zeta}}\right) \right) \end{aligned}$$

2. Case $b_k = 1$

Similar to the case $b_k = 0$, we can write

$$\begin{aligned} \beta_{k,+A}^{(i)} \Big|_{b_k=1} &= \left(y_k - A \cdot h_0 - \sum_{l=1}^{M-1} h_l \cdot x_{k-l}^{(i)} \right) \Big|_{b_k=0} = -2Ah_0 + \varepsilon_k^{(i)} \sim N(-2Ah_0, 4A^2\lambda_k^{(i)} + \sigma_n^2) \\ \beta_{k,-A}^{(i)} \Big|_{b_k=1} &= \left(y_k + A \cdot h_0 - \sum_{l=1}^{M-1} h_l \cdot x_{k-l}^{(i)} \right) \Big|_{b_k=0} = \varepsilon_k^{(i)} \sim N(0, 4A^2\lambda_k^{(i)} + \sigma_n^2) \end{aligned}$$

and the condition :

$$\begin{aligned} \rho_{k,+A}^{(i)} &\geq \rho_{k,-A}^{(i)} \\ \Leftrightarrow \left| -2Ah_0 + \varepsilon_k^{(i)} \right|^2 &\leq \left| \varepsilon_k^{(i)} \right|^2 \\ \Leftrightarrow \varepsilon_k^{(i)} Ah_0 &\geq A^2|h_0|^2 \end{aligned}$$

And thus $\psi_{k,-A}^{(i)}$ can be written as:

$$\psi_{k,-A}^{(i)} = \begin{cases} \frac{1}{2} \cdot \exp\left(+2\gamma(A^2|h_0|^2 + Ah_0\varepsilon_k^{(i)})/\sigma_n^2\right), \varepsilon_k^{(i)} Ah_0 \geq A^2|h_0|^2 \\ 1 - \frac{1}{2} \cdot \exp\left(-2\gamma(A^2|h_0|^2 + Ah_0\varepsilon_k^{(i)})/\sigma_n^2\right), \varepsilon_k^{(i)} Ah_0 < A^2|h_0|^2 \end{cases}$$

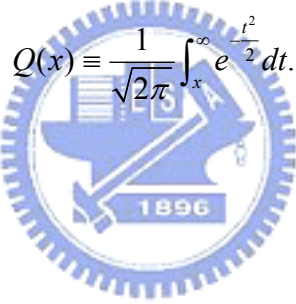
And finally the expected value:

$$\begin{aligned} E[\psi_{k,-A}^{(i)} | b_k = 1] &= \frac{1}{\sqrt{2\pi(4A^2\lambda_k^{(i)} + \sigma_n^2)}} \left(\int_{\varepsilon_k^{(i)} Ah_0 \geq A^2|h_0|^2} \frac{1}{2} e^{+2\gamma(A^2|h_0|^2 + Ah_0\varepsilon_k^{(i)})/\sigma_n^2} \cdot e^{-|\varepsilon_k^{(i)}|^2/2(4A^2\lambda_k^{(i)} + \sigma_n^2)} d\varepsilon_k^{(i)} + \int_{\varepsilon_k^{(i)} Ah_0 < A^2|h_0|^2} \left(1 - \frac{1}{2} e^{-2\gamma(A^2|h_0|^2 + Ah_0\varepsilon_k^{(i)})/\sigma_n^2}\right) \cdot e^{-|\varepsilon_k^{(i)}|^2/2(4A^2\lambda_k^{(i)} + \sigma_n^2)} d\varepsilon_k^{(i)} \right) \\ &= \mathcal{Q}\left(-A|h_0| \sqrt{\frac{1}{2(4A^2\lambda_k^{(i)} + \sigma_n^2)}}\right) + \frac{1}{\sqrt{2\pi(4A^2\lambda_k^{(i)} + \sigma_n^2)}} \left(\frac{1}{2} \int_{A|h_0|}^{\infty} e^{+2\gamma(A^2|h_0|^2 + Ah_0|\varepsilon_k^{(i)}|)/\sigma_n^2} \cdot e^{-|\varepsilon_k^{(i)}|^2/2(4A^2\lambda_k^{(i)} + \sigma_n^2)} d\varepsilon_k^{(i)} - \frac{1}{2} \int_{-\infty}^{-A|h_0|} e^{-2\gamma(A^2|h_0|^2 + Ah_0|\varepsilon_k^{(i)}|)/\sigma_n^2} \cdot e^{-|\varepsilon_k^{(i)}|^2/2(4A^2\lambda_k^{(i)} + \sigma_n^2)} d\varepsilon_k^{(i)} \right) \\ &= \mathcal{Q}\left(-\sqrt{\frac{h_0^2}{4\lambda_k^{(i)} + 1/\zeta}}\right) + \frac{1}{4} e^{2h_0^2\zeta^2(4\zeta\lambda_k^{(i)} + 1/\zeta)} \cdot \left(2 \cdot \mathcal{Q}\left(\gamma(8\zeta\lambda_k^{(i)} + 2 + 1/\gamma) \sqrt{\frac{h_0^2}{4\lambda_k^{(i)} + 1/\zeta}}\right) \right) - \frac{1}{4} e^{2h_0^2\zeta^2(4\zeta\lambda_k^{(i)} + 1/\zeta)} \cdot \left(2 \cdot \mathcal{Q}\left(\gamma(8\zeta\lambda_k^{(i)} + 2 - 1/\gamma) \sqrt{\frac{h_0^2}{4\lambda_k^{(i)} + 1/\zeta}}\right) \right) \end{aligned}$$

Substitute the results of these two cases into (3-1), we can obtain the averaged bit error probability as:

$$\begin{aligned}
P_{err,k}^{(i)} &= 1 - \frac{1}{2}E[\psi_{k,+A}^{(i)} | b_k = 0] - \frac{1}{2}E[\psi_{k,-A}^{(i)} | b_k = 1] \\
&= 1 - \frac{1}{2}(E[\psi_{k,+A}^{(i)} | b_k = 0] + E[\psi_{k,-A}^{(i)} | b_k = 1]) \\
&= Q\left(\sqrt{\frac{h_0^2}{4\lambda_k^{(i)} + 1/\zeta}}\right) + \frac{1}{2} \cdot \left[e^{2h_0^2\zeta\gamma^2(4\zeta\lambda_k^{(i)} + 1 - 1/\gamma)} \cdot Q\left(\gamma(8\zeta\lambda_k^{(i)} + 2 - 1/\gamma)\sqrt{\frac{h_0^2}{4\lambda_k^{(i)} + 1/\zeta}}\right) \right. \\
&\quad \left. - e^{2h_0^2\zeta\gamma^2(4\zeta\lambda_k^{(i)} + 1 + 1/\gamma)} \cdot Q\left(\gamma(8\zeta\lambda_k^{(i)} + 2 + 1/\gamma)\sqrt{\frac{h_0^2}{4\lambda_k^{(i)} + 1/\zeta}}\right) \right] \\
&\equiv F_0(\lambda_k^{(i)}, |h_0|^2, \zeta) + F_1(\lambda_k^{(i)}, |h_0|^2, \zeta, \gamma)
\end{aligned}$$

where $Q(x)$ is the "Q function" encountered in integrating the normal distribution (which is a normalized form of the Gaussian function), and is defined by:

$$Q(x) \equiv \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{t^2}{2}} dt.$$


C. PROOF OF THE BER AS A MONOTONICALLY DECREASING FUNCTION

In this section we intend to prove that the BER $P_{err,k}^{(i)}$ of the i -th particle, under certain assumptions, would be a monotonically decreasing function of $\eta \equiv |h_0|^2 \cdot \zeta$. This may seem to be trivial from the perspective of communication theory, however, here we still need to have some validated proofs to show that $P_{err,k}^{(i)}$ indeed decreases as η grows to simplify the proof of $P_{err,k,ZF}^{(i)} \leq P_{err,k}^{(i)}$ and $P_{err,k,MMSE}^{(i)} \leq P_{err,k}^{(i)}$ in Chapter 3.

First we assume that previous N_g particles before time k are all detected (or drawn) correctly, i.e. $P_{err,k-l}^{(i)} = 0$, for $l = 1, 2, \dots, N_g - 1$, and $\lambda_k^{(i)} = \sum_{l=1}^{M-1} |h_l|^2 P_{err,k-l}^{(i)} = 0$, which is the same assumption in all the analysis of BER in the DFE (ZF and MMSE-DFE) part of this thesis. Since $\lambda_k^{(i)} = 0$, we can write (from the result of Appendix B):

$$\begin{aligned} & P_{err,k}^{(i)} \Big|_{\lambda_k^{(i)}=0} \\ &= Q(\sqrt{\eta}) + \frac{1}{2} \cdot \left[e^{2\eta\gamma(\gamma-1)} \cdot Q\left((2\gamma-1)\sqrt{\eta}\right) - e^{2\eta\gamma(\gamma+1)} \cdot Q\left((2\gamma+1)\sqrt{\eta}\right) \right] \\ &= P_{err,k}^{(i)}(\eta, \gamma), \text{ where } \eta \equiv |h_0|^2 \cdot \zeta. \end{aligned}$$

To prove that $P_{err,k}^{(i)}(\eta, \gamma)$ is a monotonically decreasing function of η (when γ is fixed), we next take derivatives of $P_{err,k}^{(i)}(\eta, \gamma)$ w.r.t (with respect to) η as:

$$\begin{aligned}
\frac{\partial P_{err,k}^{(i)} \Big|_{\lambda_k^{(i)}=0}}{\partial \tau} &= -\frac{1}{2\sqrt{2\pi\eta}} \cdot \exp\left(-\frac{\eta}{2}\right) + \frac{1}{2} 2\gamma(\gamma-1) \cdot \exp(2\eta\gamma(\gamma-1)) \cdot Q\left((2\gamma-1)\sqrt{\eta}\right) \\
&\quad - \frac{1}{2} \cdot \exp(2\eta\gamma(\gamma-1)) \cdot \left[\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(2\gamma-1)^2 \cdot \eta}{2}\right) \cdot \frac{(2\gamma-1)}{2\sqrt{\eta}} \right] \\
&\quad - \frac{1}{2} 2\gamma(\gamma+1) \cdot \exp(2\eta\gamma(\gamma+1)) \cdot Q\left((2\gamma+1)\sqrt{\eta}\right) \\
&\quad + \frac{1}{2} \cdot \exp(2\eta\gamma(\gamma+1)) \cdot \left[\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(2\gamma+1)^2 \cdot \eta}{2}\right) \cdot \frac{(2\gamma+1)}{2\sqrt{\eta}} \right]
\end{aligned}$$

where the derivative of the Q -function is given by:

$$\frac{d}{dx} Q(x) = -\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

By combining the 3rd and the 5th terms above, we have

$$\begin{aligned}
&\frac{1}{2} \cdot \exp(2\eta\gamma(\gamma+1)) \cdot \left[\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(2\gamma+1)^2 \cdot \eta}{2}\right) \cdot \frac{(2\gamma+1)}{2\sqrt{\eta}} \right] \\
&- \frac{1}{2} \cdot \exp(2\eta\gamma(\gamma-1)) \cdot \left[\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(2\gamma-1)^2 \cdot \eta}{2}\right) \cdot \frac{(2\gamma-1)}{2\sqrt{\eta}} \right] \\
&= \frac{1}{4} \cdot \exp(2\eta\gamma^2) \cdot \frac{1}{\sqrt{2\pi\eta}} \exp\left(-\left(4\gamma^2+1\right) \cdot \frac{\eta}{2}\right) \left[e^{2\eta\gamma} \cdot e^{-2\eta\gamma} (2\gamma+1) - e^{-2\eta\gamma} \cdot e^{2\eta\gamma} (2\gamma-1) \right] \\
&= \frac{1}{4\sqrt{2\pi\eta}} \cdot \exp\left(2\eta\gamma^2 - 2\eta\gamma^2 - \frac{\eta}{2}\right) \cdot 2 \\
&= \frac{1}{2\sqrt{2\pi\eta}} \cdot \exp\left(-\frac{\eta}{2}\right)
\end{aligned}$$

We may soon discover that the sum of the 3rd term and 5th term is exactly the minus of the 1st term in the derivatives of $P_{err,k}^{(i)}$ so that the sum of these three terms would be zero. Therefore the derivative can be further simplified as:

$$\begin{aligned}
\frac{\partial P_{err,k}^{(i)} \Big|_{\lambda_k^{(i)}=0}}{\partial x} &= +\gamma(\gamma-1) \cdot \exp(2\eta\gamma(\gamma-1)) \cdot Q\left((2\gamma-1)\sqrt{\eta}\right) \\
&\quad - \gamma(\gamma+1) \cdot \exp(2\eta\gamma(\gamma+1)) \cdot Q\left((2\gamma+1)\sqrt{\eta}\right)
\end{aligned}$$

We can see that for $\gamma \in [0.5, 1]$, $\frac{\partial P_{err,k}^{(i)} \big|_{\lambda_k^{(i)}=0}}{\partial \eta}$ would be always less than 0, and

therefore we have proved that $P_{err,k}^{(i)}$ is a monotonically decreasing function of

$$\eta \equiv |h_0|^2 \cdot \zeta.$$



D. PRE-FILTER COEFFICIENTS

In this appendix, we give the procedure of how the coefficients of the feedforward filters (F_{ZF} and F_{MMSE}) are calculated when the CIR is known. The details of the related material can be referenced as to [5]. First, we take the ZF scheme as an example, and consider the following spectral factorization:

$$R_{hh}(z) \equiv H(z) \cdot H^*(1/z^*) = S_g G(z) \cdot G^*(1/z^*) \quad (3-4)$$

We can obtain the minimum phase response $G(z)$, (or equivalently, g_k) from the spectral factorization. And then we may calculate the coefficients of F_{ZF} by:

$$F_{ZF}(z) = \frac{z^{-d} H^*(1/z^*)}{S_g G^*(1/z^*)} \quad (3-5)$$

In order to make the feedforward filter stable, F_{ZF} is chosen to be an anti-causal filter. In practical, we make it a FIR filter by delay it with d slots and truncating the anti-causal component in (3-5).

Theoretically, the spectral factorization is expended with Laurent series and then using Taylor series expansion to obtain the coefficients g_k , $k = 0, 1, \dots, N_g$. (see reference [6]) In practical, however, the coefficients of $G(z)$ and $1/G^*(1/z^*)$, can be calculated as the following:

Step 1. Calculate the cepstrum⁹ of the autocorrelation function $R_{hh}(n)$:

$$\hat{R}_{hh}(n) = F^{-1} \left\{ \log \left(H(e^{j\omega}) H^*(1/e^{-j\omega}) \right) \right\}$$

Step 2. The coefficient of $G(z)$ can be calculated by

⁹ Let the Fourier Transform of $x(n)$ be $X(e^{j\omega})$, and $\hat{X}(e^{j\omega}) \equiv \log[X(e^{j\omega})]$, then the *complex cepstrum* of $x(n)$ is $\hat{x}(n) = F^{-1} \{ \hat{X}(e^{j\omega}) \}$, where F^{-1} denotes for the inverse Fourier Transform and $\hat{X}(e^{j\omega}) = \log[X(e^{j\omega})] = \log|X(e^{j\omega})| + j \arg[X(e^{j\omega})]$ (p788~p789, [7])

$$g_k = \sum_{i=1}^k \frac{i}{k} g_{k-i} \hat{R}_{hh}(i), \text{ with } g_0 = 1, \text{ for } 0 \leq k \leq N-1.$$

$$S_g = \exp(\hat{R}_{hh}(0))$$

Step 3. The coefficient of $1/G^*(1/z^*)$ can be calculated by:

$$\tilde{g}_k = \sum_{i=1}^{\min\{N-1, -k\}} g_i \cdot \tilde{g}_{-k+i}, \text{ with } \tilde{g}_0 = 1, \text{ and } k > 0$$

Step 4. Then we can obtain the coefficients of the feedforward filter:

$$F_{ZF}(z) = H^*(1/z^*) \cdot \frac{z^{-d}}{S_g G^*(1/z^*)}$$

$$\Rightarrow f_{ZF}(n) = \bar{h}(n) * (\tilde{g}(n)/S_g), \text{ where } \bar{h}(n) = Z^{-1}\{H^*(1/z^*)\}$$

The coefficients of the F_{MMSE} can be calculated in similar procedures by replacing the $\log(H(e^{j\omega})H^*(1/e^{-j\omega}))$ term in Step 1 with

$$\log(\sigma_x^2 H(e^{j\omega})H^*(1/e^{-j\omega}) + \sigma_n^2).$$

