


# 非線性回音消除之收斂性分析

學生：江冠諭

指導教授：謝世福

國立交通大學電信工程學系碩士班

## 摘要



爲了補償非線性失真在免持聽筒或是視訊會議系統，回音的消除，通常使用一個無記憶性的多項式結合 NLMS 演算法的適應性濾波器。在傳統上，多項式採用次方級數展開的形式；而在本篇論文中，爲了提昇非線性適應性濾波器的收斂速度我們使用正交多項式的非線性適應性濾波器。不論是級數型多項式或是正交多項式，我們都分析出它們的收斂理論值，並且從電腦模擬結果得知和我們的理論值是符合的，而模擬的結果也說明了我們提出的方式的確有較好的收斂性。

除了使用適應性的方式之外，在實際的語音傳送之前，我們使用了訓練序列信號的方式來評估非線性濾波器的係數。我們分析它們的收斂理論值，並且由電腦模擬得到驗證，而且在訊雜比不佳的情況下，它的收斂值會比適應性濾波器來得佳。

# **Convergence Analyses of Nonlinear Acoustic Echo Cancellation**

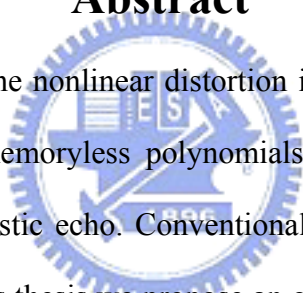
**Student : G. Y. Jiang**

**Advisor : S. F. Hsieh**

**Department of Communication Engineering**

**National Chiao Tung University**

## **Abstract**



In order to compensate the nonlinear distortion in the hands -free telephones or teleconferencing system, a memoryless polynomials NLMS adaptive filter can be used to cancel nonlinear acoustic echo. Conventional polynomials model employs a power-series expansion. In this thesis we propose an orthogonal polynomials adaptive filter and perform theoretical convergence analysis of residual echo power which proves its faster convergence rate owing to the reduced eigen spread of the input signal. Computer simulations justify our analysis and show the improved performance of the proposed nonlinear acoustic echo canceller.

In addition to the adaptive method, the training sequence (TS) can be used to estimate the coefficients of the nonlinear acoustic echo cancellation. The convergence rate of the training method is derived analytically. Computer simulations show that the TS method performs better than the NLMS method at low SNR.

# Acknowledgments

I am deeply grateful to my advisor Dr. S. F. Hsieh, for his support, patience, and encouragement throughout my master studies. My thanks also go to the members of my lab they give me many helpful comments.

Last but by no means least; I would like to thank my parents, my family, and all of people who have helped me in the past two years.



# Contents

中文摘要.....	i
English Abstract.....	ii
Acknowledgement.....	iii
Contents.....	iv
List of Figures.....	vii
List of Tables.....	ix
1 Introduction.....	1
2 Nonlinear Adaptive AEC.....	7
2.1 Adaptive nonlinear NLMS AEC.....	7
2.2 Convergence analysis of nonlinear NLMS algorithm.....	11
2.2.1 Variance of nonlinear coefficients error.....	11
2.2.2 Residual echo power analysis.....	14
2.3 Convergence analysis of linear NLMS algorithm.....	15
2.3.1 Variance of linear coefficients error.....	15
2.3.2 Residual echo power analysis.....	16
2.4 Nonlinear processor with orthogonal polynomials.....	17
2.4.1 Signal dependent orthogonal polynomials.....	17
2.4.2 Eigenvalue spread analysis.....	21
2.5 Application of dual loudspeakers system.....	28
2.6 Summary.....	31
3 Training Sequence and Coefficients Estimation.....	32
3.1 Correlation based nonlinear AEC.....	33
3.2 Training sequence estimation algorithm.....	35

3.2.1	Linear TS coefficients estimation algorithm.....	35
3.2.2	Nonlinear TS coefficients estimation algorithm.....	38
3.3	Recursive training sequence algorithm.....	41
3.4	Linear convergence analysis.....	44
3.4.1	Variance of linear coefficients error.....	44
3.4.2	Residual echo power analysis.....	47
3.5	Nonlinear convergence analysis.....	48
3.5.1	Variance of nonlinear coefficients error.....	48
3.5.2	Residual echo power analysis.....	49
3.6	TS for dual loudspeakers system.....	50
3.6	Summary.....	51
4	Computer Simulations.....	52
4.1	Parameters of simulations.....	52
4.2	Coefficients estimation based on NLMS adaptive algorithm.....	57
4.2.1	Individual coefficients and residual echo power convergence.....	57
4.2.2	Convergence rate using orthogonal bases.....	60
4.2.3	Joint convergence rate using orthogonal bases.....	62
4.2.4	Convergence rate of a speech input.....	65
4.2.5	Convergence rate of dual loudspeakers system.....	66
4.3	TS-based coefficients estimation.....	70
4.3.1	Convergences of linear TS coefficients and residual TS echo power	70
4.3.2	Joint TS convergence rate.....	73
4.3.3	Joint residual echo comparison of TS and NLMS algorithms.....	74
4.3.4	Training Sequences for dual loudspeakers system.....	77
4.4	Experiments with a real echo path.....	78
5	Conclusions.....	84

Appendix.....85  
Bibliography.....90



# List of Figures

1.1	The simplified diagram of hands-free telephone system.....	1
1.2	Nonlinear acoustic echo canceller.....	2
2.1	Polynomial nonlinear acoustic echo canceller.....	7
2.2	Nonlinear processor.....	17
2.3	Nonlinear AEC structure for dual loudspeakers system.....	28
3.1	Structure of nonlinear AEC based on training sequence.....	34
4.1	Room impulse response of a typical office.....	54
4.2	Room impulse response from the left loudspeaker.....	54
4.3	Room impulse response from the right loudspeaker.....	55
4.4	I/O mapping characteristic of the left nonlinear loudspeaker.....	55
4.5	I/O mapping characteristic of the right nonlinear loudspeaker.....	56
4.6	The speech signal.....	56
4.7	Comparison of nonlinear coefficients misalignments (perfect <b>h</b> ).....	58
4.8	Comparison of residual echo power (perfect <b>h</b> ).....	58
4.9	Comparison of linear coefficients misalignments (perfect <b>a</b> ).....	59
4.10	Comparison of residual echo power (perfect <b>a</b> ).....	59
4.11	Residual echo powers for uniform input (perfect <b>h</b> ).....	60
4.12	Residual echo powers for WGN input (perfect <b>h</b> ).....	61
4.13	Residual echo powers for Laplacian input (perfect <b>h</b> ).....	61
4.14	Joint-residual echo power for uniform input.....	63
4.15	Joint-residual echo power for Gaussian input.....	64
4.16	Joint-residual echo power for Laplacian input.....	64
4.17	ERLE for a true speech input signal with perfect linear coeff. ....	65
4.18	Histogram of input speech.....	66

4.19 Nonlinear coefficients misalignments for dual loudspeakers.....	68
4.20 Residual echo power for dual loudspeakers (perfect $\mathbf{h}$ 's) .....	68
4.21 Joint residual echo power for dual loudspeakers.....	69
4.22 Comparison of linear TS coeff. misal. (SNR=5 dB, perfect nonlinear coeff.)....	70
4.23 Comparison of linear residual TS echo power (SNR=5 dB).....	71
4.24 Comparison of nonlinear TS coeff. misal. (SNR=5 dB, perfect linear coeff.)....	72
4.25 Comparison of residual nonlinear TS echo power (SNR=5 dB).....	72
4.26 Joint linear coefficients misalignment for training method (SNR=5 dB).....	73
4.27 Joint residual echo power for training method (SNR=5 dB).....	74
4.28 Joint residual echo power comparison of NLMS and TS (SNR=10 dB).....	75
4.29 Joint residual echo power comparison of NLMS and TS (SNR=5 dB).....	76
4.30 Joint residual echo power comparison of NLMS and TS (SNR=0 dB).....	76
4.31 Comparison of residual echo powers for dual loudspeakers and training method (SNR=5 dB) .....	77
4.32 Comparison of joint residual echo powers for dual loudspeakers for TS and NLMS (SNR=5 dB) .....	78
4.33 The speech input and its microphone output signal.....	79
4.35 ERLE comparison between linear and nonlinear AEC for a true echo path.....	80
4.35 Average ERLE comparison of different nonlinear orders.....	82
4.36 Average ERLE comparison at different SNRs.....	83



# List of Tables

2.1	Comparison of computational cost, no. multiplication per iteration.....	10
2.2	Signal-dependent orthogonal polynomials.....	18
2.3	Orthogonal polynomials for a uniformly distributed signal.....	19
2.4	Orthogonal polynomials for a Gaussian-distributed input.....	20
2.5	Orthogonal polynomials for a Laplacian-distributed input.....	21
2.6	Eigenvalue spread comparison.....	27
3.1	Comparison of complexity computational.....	43
4.1	Average joint-ERLE(dB) comparison for a speech input signal.....	66
4.2	Average ERLE comparison between different linear filter lengths for true echo path.....	81
4.3	Nonlinear coefficients for each AEC with different linear orders.....	81
4.4	Average ERLE comparison of NLMS and TS methods for a true echo path.....	82

# Chapter 1

## Introduction

Hands-free telephone or teleconferencing usually suffers from the annoying acoustic echo problem [1], which is the far end speech transmitted back to the far end user as a result of the coupling of the loudspeaker and microphone at the near end. A simplified diagram of hands-free telephone system is shown in Fig. 1.1. The main object of acoustic echo cancellation (AEC) is to estimate the unknown echo path and subtract the estimated echo components from the microphone output. Since the echo path may be time-variant due to objects moving around the room, an adaptive filter is commonly used for tracking the echo path to provide satisfactory speech communication quality [1]-[6].

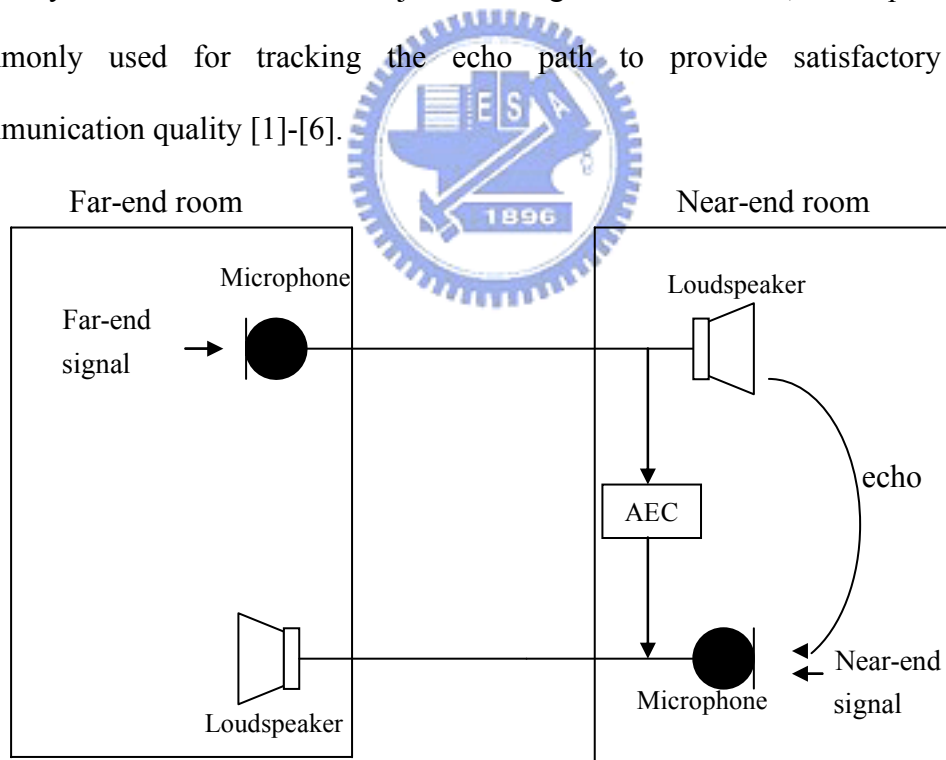


Fig. 1.1 The simplified diagram of hands-free telephone system

The performance of AEC relies on its tracking capability. There are many recursive algorithms that have been proposed [6]. The least-mean-square (LMS)

algorithm is famous for its low computational cost and the recursive least-squares (RLS) algorithm has its advantage of fast convergence rate. Beside the tracking capability, the performance of AEC is restricted by noise, finite precision, truncation effects, under-model of the echo path and loudspeaker nonlinearities [7].

In this thesis, the issue of nonlinearities is our main focus. In [8], the nonlinearities of loudspeaker can be classified as nonlinearities with and without memory. Dynamic loudspeakers cause nonlinearity with memory when the power amplifier is not driven into saturation. The other is nonlinearity without memory when the power amplifier is overdriven. To compensate these kinds of nonlinearities distortions, several nonlinear AECs have been proposed recently [9]-[20]. The nonlinear AEC system is shown in Fig.1.2. The signal from the far end is passing through the nonlinear loudspeaker and the room impulse response and then is picked up by the microphone. The nonlinear AEC is supposed to cancel the echo signal. The echo can be cancelled perfectly if the nonlinear AEC filter is identical to the nonlinear loudspeaker and room impulse response. Different nonlinear structure has its own computation complexity, convergence speed and robustness.

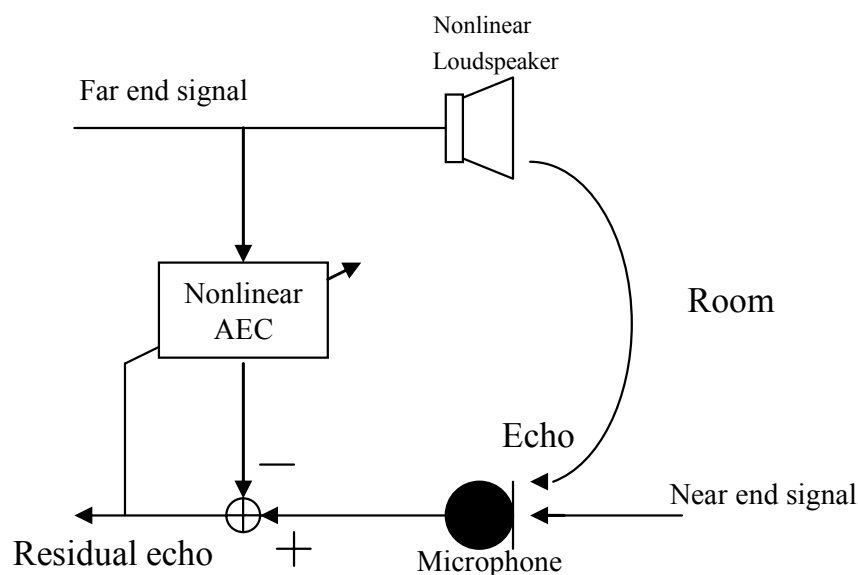


Fig.1.2 Nonlinear acoustic echo canceller

We summary some important nonlinear adaptive structures.

- Volterra model, which is attractive because it is a straightforward generalization of the linear system description and the behavior of many physical systems can be described with a Volterra filter. The Volterra series based filters have been proposed [10]-[12] etc. for line echo canceling. It can represent a large of class of nonlinear system. However, due to their high computational complexity they are limited use in practical systems.
- Bilinear model, which is a parametric model that contains cross terms; it corresponds to a subclass of the NARMAX structure [13]. The NARMAX structure is a general parametric model but needs a pre-identification procedure.
- Cascade model, which can be considered as a particular subclass of a Volterra series filter and its main advantage is to introduce fewer parameters for estimation. The neural network [14] with a cascade structure offers a new perspective but needs an extra reference microphone. Others cascade models include:
  - Hammerstein model [15], cascade of a memoryless polynomials filter and a FIR filter.
  - Wiener model [4], cascade of a FIR filter and a memoryless polynomials filter.
  - Wiener-Hammerstein model [26], cascade of a FIR filter, a memoryless polynomials filter and a FIR filter.

Among these nonlinear structures, the Hammerstein model will be used in this thesis for following reasons. First, the nonlinearity with memory only occurs in application of high quality loudspeakers [8]; so that for hands-free or power limited low cost application, the compensation for nonlinearity with memory is not necessary.

Second, considering that the far end signal passes through the loudspeaker first and then the room impulse response, the cascade model of a nonlinear processor and a FIR filter is a natural choice. Third, the Hammerstein model is widely used in nonlinear system identification such as nonlinear AEC, neural networks [15-17] etc. Its joint NLMS-type adaptation algorithm is well known [18].

Besides the polynomial function, a sigmoid function [6, 9] can also be used to model the nonlinear saturation. Similarly, a raised cosine function can also be used for nonlinear compensation [19]. The objective in [19] is to achieve a low computational complexity in implementing a nonlinear AEC. Both these two nonlinear models can use the NLMS adaptive algorithm to update its coefficients, except that its nonlinear component is generated by an exponential function. It only uses one parameter to control the nonlinearity therefore it has less freedom. By contrast, the polynomial type has more freedom in that each nonlinear order can be controlled with one coefficient.

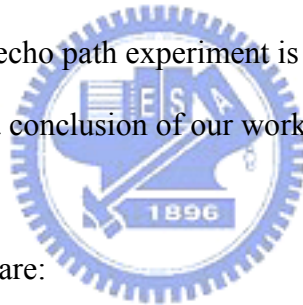
The power series polynomials are simple to implement but its high correlation among different polynomials orders leads to low convergence rate. To overcome this problem, recently some orthogonal structures have been developed. In [10], Mathews suggested to perform an orthogonalization procedure on the nonlinear bases outputs when the input signal is Gaussian distributed. Jenkins et al. in [25] proposed an orthogonal basis to represent the Volterra series thus orthogonalization procedure is not required but the input signal is also assumed to be have a Gaussian distribution and unity variance. Similarly, in [26] the Wiener-Hammerstein model is used and its nonlinearities is assumed to be expandable in a series of fixed orthogonal Hermite polynomials. The Hermite polynomials are a set of orthogonal polynomials on the infinite interval with respect to the  $e^{-x^2}$  weight function therefore the works in [25, 26] are limited to the unit-variance Gaussian input signal only. In [20] Kuech et al. proposed an adaptive orthogonalized power filter to improve the convergence rate for

the input signal with any distribution, stationary or time-variant. The orthogonal basis is updated online at each iteration and the Gram Schmidt procedure is employed to find out the orthogonalization coefficients, as a result, computational complexity is increased. In this thesis we use fixed orthogonal polynomials to produce the nonlinear components. Its low computational complexity and fast convergence rate makes the orthogonal polynomials filter very promising for nonlinear AEC. For practical applications, the influence of the probability distribution of the input signal is insignificant because the performance of the fixed orthogonal polynomials remains relatively well. Unlike the previous researches, we perform the convergence analysis of the nonlinear AEC, from which we examine the effects of orthogonality due to various input probability distributions and conclude the superiority of the orthogonal polynomials.

In addition to the orthogonal polynomial basis, sending a white sequence to train the coefficients of nonlinear AEC in advance of speech communication can also be used for speeding up the convergence rate. The training sequence, used in channel estimation, adaptive equalizer applications or echo path estimation, have been well studied in [21]-[24]. During the training mode, we can fast start up the adaptive filter, especially in noisy environments. The training sequence is generated by a training sequence generator. The estimation is done with the correlation method, where a portion of the training sequence is correlated with shifted versions of the received signal. Based on the difference between this known sequence and the received sequence, the coefficients of the unknown can be determined. Although the training method is to solve the Wiener-Hopf solution directly and often involves a matrix inverse, the solution is simple because the matrix is only a function of the known training sequence, and a pre-computed inverse of the matrix can be stored. For some special training sequences, an inverse matrix is not even required.

The other chapters of the thesis are organized as follows.

- Chapter 2, we perform convergence analyses of coefficients and residual echo power and we compare the convergence rates of orthogonal and non-orthogonal basis. Eigenvalue spread analysis is introduced for better illustration.
- Chapter 3, we perform the linear coefficients convergence analysis based on the training sequence method. The recursive analytical form will also be introduced. We also show the analysis of dual loudspeakers system.
- Chapter 4, we include many computer simulations that have been developed to illustrate the analyses in chapter 2 and 3. These simulations help us to compare the performances of different nonlinear AEC structure. Finally, for practical use, a true echo path experiment is performed.
- Chapter 5, we give a conclusion of our work.



The main efforts in this thesis are:

- (1) For NLMS nonlinear algorithm, we derive individual convergence analyses of the linear and nonlinear coefficients and its residual echo power.
- (2) An orthogonal basis for Gaussian and nonGaussian input signal is used to accelerate the nonlinear coefficients convergence rate.
- (3) A training sequence algorithm is proposed for a nonlinear Hammerstein model.
- (4) Convergence analyses of linear coefficients based on training sequence algorithm is derived.

# Chapter 2

## Nonlinear Adaptive filter

In Chapter 2 we will introduce the nonlinear AEC with cascade nonlinear processor and describe the joint NLMS-type adaptation algorithm to adaptive the both linear and nonlinear coefficients. In Section 2.2 we will analyses the nonlinear coefficients and residual echo power convergence rates under the assumption of linear coefficients have perfectly known. In Section 2.3 we will analyses the linear coefficients and its residual echo power when nonlinear coefficients have perfectly known. In Section 2.4, we discuss the eigenvalue spread of signal dependent orthogonal bases. Finally, in Section 2.5, we extend the analysis for dual loudspeakers system.

### 2.1 Adaptive nonlinear NLMS AEC

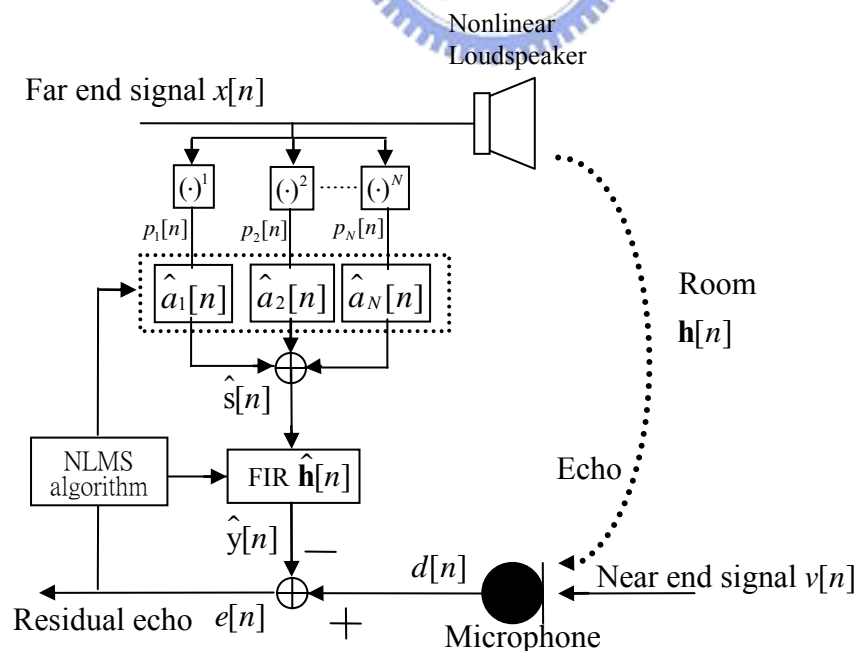


Fig.2.1 Polynomial nonlinear acoustic echo canceller



As shown in Fig 2.1, the signal  $x[n]$  from the far end is assumed to be nonlinearly distorted only in the power amplifier of loudspeaker. It is then passing through a room impulse response  $\mathbf{h}[n]$ . Hence, the nonlinear processor is modeled as the loudspeaker and linear filter is modeled as a room impulse. The cascade filter structure is the same as the loudspeaker and the room impulse response. Let  $d[n]$  denote the desired signal. The nonlinear AEC output signal  $\hat{y}[n]$  can be written as

$$\hat{y}[n] = \hat{\mathbf{h}}^T [n] \hat{\mathbf{s}}[n]$$

$$\hat{\mathbf{h}}[n] = [\hat{h}_0[n], \hat{h}_1[n], \dots, \hat{h}_{M-1}[n]]^T$$

where  $\hat{\mathbf{h}}[n]$  represents the estimated coefficients vector of the linear FIR filter,  $M$  denotes the length of the FIR filter.  $\hat{\mathbf{s}}[n]$  is the output vector of the nonlinear filter

$$\hat{\mathbf{s}}[n] = [\hat{s}[n], \hat{s}[n-1], \dots, \hat{s}[n-M+1]]^T.$$

For each  $\hat{s}[n]$  is given by

$$\hat{s}[n] = [x[n] \ x^2[n] \ \dots \ x^N[n]] [\hat{a}_1[n] \ \hat{a}_2[n] \ \dots \ \hat{a}_N[n]]^T$$

$$= \mathbf{x}[n]^T \hat{\mathbf{a}}[n]$$

therefore,  $\hat{\mathbf{s}}[n]$  is given by

$$\hat{\mathbf{s}}[n] = [ \mathbf{x}[n]^T \hat{\mathbf{a}}[n], \mathbf{x}[n-1]^T \hat{\mathbf{a}}[n-1], \dots, \mathbf{x}[n-M+1]^T \hat{\mathbf{a}}[n-M+1] ]^T.$$

$p_i$  is the polynomial basis of order  $i$ , for example  $p_1[n] = x[n]$  and  $p_2[n] = x^2[n]$  in case of a power series expansion basis.  $N$  is the order of the polynomials, and  $\hat{\mathbf{a}}[n]$  is the estimated coefficients vector of the nonlinear processor. The estimated error is

$$e[n] = d[n] - \hat{y}[n]$$

$$= d[n] - \hat{\mathbf{h}}^T [n] \hat{\mathbf{s}}[n]$$

The gradient of the error power  $e^2[n]$ , as derived for linear transversal filter in [18]

can be calculated according to:

$$\nabla_h = \frac{\partial e^2[n]}{\partial \hat{\mathbf{h}}[n]} = -2e[n]\hat{\mathbf{s}}[n]$$

$$\nabla_a = \frac{\partial e^2[n]}{\partial \hat{\mathbf{a}}[n]} = -2e[n]\mathbf{P}^T[n]\hat{\mathbf{h}}[n]$$

where  $\mathbf{P}[n]$  is nonlinear expanded matrix is defined by

$$\begin{bmatrix} p_1[n] & p_2[n] & \cdots & p_N[n] \\ p_1[n-1] & p_2[n-1] & \cdots & p_N[n-1] \\ \vdots & \vdots & \ddots & \vdots \\ p_1[n-M+1] & p_2[n-M+1] & \cdots & p_N[n-M+1] \end{bmatrix}$$

The definition of error signal is different from [18], here, we only calculate the scalar  $\hat{s}[n]$  of the vector  $\hat{\mathbf{s}}[n]$  at each iteration. If the coefficients vectors are updated with step size  $\mu_h$  and  $\mu_a$ , a joint NLMS-type adaptive algorithm is given by

$$\hat{\mathbf{h}}[n+1] = \hat{\mathbf{h}}[n] + \frac{\mu_h}{\|\hat{\mathbf{s}}[n]\|_2^2} \hat{\mathbf{s}}[n]e[n] \quad (2.1.1)$$

$$\hat{\mathbf{a}}[n+1] = \hat{\mathbf{a}}[n] + \frac{\mu_a}{\|\mathbf{P}^T[n]\hat{\mathbf{h}}[n]\|_2^2 + \delta} \mathbf{P}^T[n]\hat{\mathbf{h}}[n]e[n] \quad (2.1.2)$$

At each iteration, the echo signal  $e[n]$  is the same for coefficients update in both (2.1.1) and (2.1.2).

For computational complexity, we examine the number of multiplications required to make one complete iteration of the algorithm (2.1.1) and (2.1.2).  $\hat{\mathbf{s}}[n]$  in (2.1.1) and its 2-norm need  $N$  and  $M$  multiplications respectively thus the total requirement of (2.1.1) is about  $2M + N$ . For (2.1.2),  $\mathbf{P}^T[n]\hat{\mathbf{h}}[n]$  and its 2-norm need  $MN$  and  $N$  multiplications respectively thus the total requirement of (2.1.2) is about  $MN+2N$ . It needs  $MN+2M+3N$  multiplications to find out the coefficients at each iteration. In addition to the adaptation algorithm, the nonlinear components need  $N-1$

multiplications for power-series basis,  $x^2[n] = x[n]x[n]$ ,  $x^3[n] = x^2[n]x[n]$ , ...,  $x^N[n] = x^{N-1}[n]x[n]$ . When an orthogonal basis is used it needs more multiplications to produce the nonlinear components. According to Gram-Schmidt orthogonalization procedure it needs  $\frac{1}{2}(N^2+N-2)$ . In Table 2.1, we show the total requirement of multiplication at each iteration when Gram-Schmidt is used or not. The computational complexity only increases slightly when an orthogonal basis is used.

Table 2.1 Comparison of computational cost, no. multiplication per iteration

	Number of multiplication
Without Gram-Schmidt procedure	$MN+2M+4N-1$
With Gram-Schmidt procedure	$MN+2M+3.5N-1+0.5 N^2$

In the following sections, we assume that the nonlinear loudspeaker and room impulse response are time invariant. The near end signal  $v[n]$  only contains a white Gaussian noise (WGN) and double talk is not present.

## 2.2 Convergence analysis of nonlinear NLMS algorithm

### 2.2.1 Variance of nonlinear coefficients error

In this section we will derive the convergence rate of nonlinear coefficients under the assumption of perfect linear coefficients i.e.,  $\hat{\mathbf{h}}[n] = \mathbf{h}$ . The following analysis is similarly to [6]. First, the estimation error produced by the nonlinear AEC filter is expressed as

$$\begin{aligned} e[n] &= d[n] - \hat{\mathbf{a}}^T[n] \mathbf{P}^T[n] \hat{\mathbf{h}}[n] \\ &= \mathbf{a}^T \mathbf{P}^T[n] \mathbf{h} + v[n] - [\mathbf{a} + \boldsymbol{\varepsilon}_a[n]]^T \mathbf{P}^T[n] [\mathbf{h} + \boldsymbol{\varepsilon}_h[n]]. \end{aligned} \quad (2.2.1)$$

Because the cascade structure we have more estimated error terms than [6], the joint error term produced by linear and nonlinear is difficult to perform its convergence analysis. For this reason, we assume that the linear coefficients are perfectly known,  $\boldsymbol{\varepsilon}_h[n]$  is equal to zeros. We express the estimated error as follows.

$$e[n] = v[n] + \boldsymbol{\varepsilon}_a^T[n] \mathbf{P}^T[n] \mathbf{h}. \quad (2.2.2)$$

In (2.2.2)  $\mathbf{P}^T[n] \mathbf{h}$  contains not only linear but also nonlinear order of input signal, this is different from [6] but the analyses procedures in [6] can still be used for here.

We denote the nonlinear coefficients weight error by

$$\boldsymbol{\varepsilon}_a[n+1] = \mathbf{a} - \hat{\mathbf{a}}[n+1]$$

Using (2.1.2), (2.2.2) and let  $T$  denote  $\|\mathbf{P}^T[n] \mathbf{h}\|_2^2 + \delta$ , we may rewrite  $\boldsymbol{\varepsilon}_a[n+1]$  as

$$\begin{aligned} \boldsymbol{\varepsilon}_a[n+1] &= \mathbf{a} - \hat{\mathbf{a}}[n] - \frac{\mu_a}{T} \mathbf{P}^T[n] \mathbf{h}[n] e[n] \\ &= \boldsymbol{\varepsilon}_a[n] - \frac{\mu_a}{T} \mathbf{P}^T[n] \mathbf{h}[n] (v[n] + \boldsymbol{\varepsilon}_a^T[n] \mathbf{P}^T[n] \mathbf{h}) \\ &= \left[ I - \frac{\mu_a}{T} \mathbf{P}^T[n] \mathbf{h} \mathbf{h}^T \mathbf{P}[n] \right] \boldsymbol{\varepsilon}_a[n] - \frac{\mu_a}{T} \mathbf{P}^T[n] \mathbf{h} v[n]. \end{aligned}$$

According to the direct averaging method [6], when  $\mu_a \ll 1$ ,  $\boldsymbol{\varepsilon}_a[n+1]$  can be approximated as follows:

$$\boldsymbol{\varepsilon}_a[n+1] \approx \left[ I - \frac{\mu_a}{T} \mathbf{R}_{p^T[n]h} \right] \boldsymbol{\varepsilon}_a[n] + \mathbf{f}_a[n] \quad (2.2.3)$$

where

$$\mathbf{f}_a[n] = -\frac{\mu_a}{T} \mathbf{P}^T[n] \mathbf{h} v[n]$$

and  $\mathbf{R}_{p^T[n]h}$  is the correlation matrix of  $\mathbf{P}^T[n] \mathbf{h}$ . By applying the unitary similarity transformation,  $\mathbf{R}_{p^T[n]h}$  is transformed into a simpler form:

$$\mathbf{Q}_a^T \mathbf{R}_{p^T[n]h} \mathbf{Q}_a = \mathbf{D}_a$$

where  $\mathbf{Q}_a$  is a unitary matrix and  $\mathbf{D}_a$  is a diagonal matrix consisting of the eigenvalues  $\lambda_{ai}$ . Let  $\mathbf{K}_a[n] = \mathbf{Q}_a^T \boldsymbol{\varepsilon}_a[n]$  then we may transform (2.2.3) into the form

$$\mathbf{Q}_a^T \boldsymbol{\varepsilon}_a[n+1] = \mathbf{Q}_a^T \left[ I - \frac{\mu_a}{T} \mathbf{R}_{p^T[n]h} \right] \boldsymbol{\varepsilon}_a[n] + \mathbf{Q}_a^T \mathbf{f}_a[n]$$

$$\mathbf{K}_a[n+1] = \left[ I - \frac{\mu_a}{T} \mathbf{D}_a \right] \mathbf{K}_a[n] + \boldsymbol{\Phi}_a[n]$$

where  $\boldsymbol{\Phi}_a[n] = \mathbf{Q}_a^T \mathbf{f}_a[n]$ . The natural mode  $k_{ai}[n]$ ,  $i$ -th entry of  $\mathbf{K}_a[n]$ , is stochastic with a mean and mean square value of its own. Let  $k_{ai}[0]$  denote the initial value of  $k_{ai}[n]$  and  $\Phi_{ai}[n]$  is denoted  $i$ -th entry of  $\boldsymbol{\Phi}_a[n]$ . We may rewrite  $k_{ai}[n]$  as follows.

$$\begin{aligned} k_{ai}[n] &= \left(1 - \frac{\mu_a}{T} \lambda_{ai}\right) k_{ai}[n-1] + \Phi_{ai}[n-1] \\ &= \left(1 - \frac{\mu_a}{T} \lambda_{ai}\right)^n k_{ai}[0] + \sum_{j=0}^{n-1} \left[1 - \frac{\mu_a}{T} \lambda_{ai}\right]^{n-1-j} \Phi_{ai}[j] \end{aligned}$$

Hence, the first moment of  $k_{ai}[n]$  is given by

$$\begin{aligned} E[k_{ai}[n]] &= \left(1 - \frac{\mu_a}{T} \lambda_{ai}\right)^n E[k_{ai}[0]] + \sum_{j=0}^{n-1} \left[1 - \frac{\mu_a}{T} \lambda_{ai}\right]^{n-1-j} E[\Phi_{ai}[j]] \\ &= \left(1 - \frac{\mu_a}{T} \lambda_{ai}\right)^n k_{ai}[0] \end{aligned}$$

where

$$\begin{aligned} E[\boldsymbol{\Phi}_a[n]] &= \mathbf{Q}_a^T E[\mathbf{f}_a[n]] \\ &= -\frac{\mu_a}{T} \mathbf{Q}_a^T \mathbf{h} E[\mathbf{P}[n] v[n]] \\ &= 0. \end{aligned}$$

Since  $\mathbf{P}[n]$  only contains the input signal and is independent of noise, we further

assume that the initial value of  $k_{ai}[0]$  is independent of  $\Phi_{ai}$ , therefore the second moment of  $k_{ai}[n]$  is given by

$$E[|k_{ai}[n]|^2] = (1 - \frac{\mu_a}{T} \lambda_{ai})^{2n} |k_{ai}[0]|^2 + \sum_{g=0}^{n-1} \sum_{j=0}^{n-1} (1 - \frac{\mu_a}{T} \lambda_{ai})^{n-1-g} (1 - \frac{\mu_a}{T} \lambda_{ai})^{n-1-j} E[\Phi_{ai}[g] \Phi_{ai}[j]]. \quad (2.2.4)$$

The  $|k_{ai}[0]|^2$  in the right-hand side of (2.2.4) is equal to the nonlinear coefficients vector square 2-norm. The second term in right-hand side of (2.2.4) is zero when summation index  $g$  is not equal to  $j$  otherwise we can express  $E[\Phi_a[n] \Phi_a^T[n]]$  as

$$\begin{aligned} E[\Phi_a[n] \Phi_a^T[n]] &= (\frac{\mu_a}{T})^2 \mathbf{Q}_a^T \mathbf{h} E[\mathbf{P}[n] \mathbf{v}[n] \mathbf{v}[n] \mathbf{P}^T[n]] \mathbf{h}^T \mathbf{Q}_a \\ &= (\frac{\mu_a}{T})^2 \sigma_v^2 \mathbf{Q}_a^T \mathbf{R}_{P^T[n]h} \mathbf{Q}_a \\ &= (\frac{\mu_a}{T})^2 \sigma_v^2 \mathbf{D}_a. \end{aligned} \quad (2.2.5)$$

From (2.2.5), (2.2.4) can be written as

$$\begin{aligned} E[|k_{ai}[n]|^2] &= (1 - \frac{\mu_a}{T} \lambda_{ai})^{2n} \|\mathbf{a}\|_2^2 + \sigma_v^2 (\frac{\mu_a}{T})^2 \lambda_{ai} (1 - \frac{\mu_a}{T} \lambda_{ai})^{2n-2} \sum_{j=0}^{n-1} (1 - \frac{\mu_a}{T} \lambda_{ai})^{-2j} \\ &= \frac{\frac{\mu_a}{T} \sigma_v^2}{2 - \frac{\mu_a}{T} \lambda_{ai}} + (\|\mathbf{a}\|_2^2 - \frac{\frac{\mu_a}{T} \sigma_v^2}{2 - \frac{\mu_a}{T} \lambda_{ai}}) (1 - \frac{\mu_a}{T} \lambda_{ai})^{2n}. \end{aligned} \quad (2.2.6)$$

In (2.2.6) the error variance of nonlinear coefficients is given. Again, that (2.2.6) is under the assumption of perfect linear coefficients. The error variance of nonlinear coefficients can be determined with the knowledge of step size  $\mu_a$ , noise power  $\sigma_v^2$ , square 2-norm of nonlinear coefficients vector, eigenvalues  $\lambda_{ai}$  of correlation matrix of  $\mathbf{s}[n]$  and the sum of all eigenvalues. Because the step size and the eigenvalues are both positive, the second term of (2.2.6) will disappear when the iteration number approaches to infinity. The steady state of  $E[|k_{ai}[n]|^2]$  is given in the first term of (2.2.6).

## 2.2.2 Residual echo power analysis

From (2.2.2), the mean square error (i.e., residual echo) due to an estimated error of nonlinear coefficients is given by

$$\begin{aligned}
 J_a[n] &= E\left[|e[n]|^2\right] \\
 &= E[(v[n] + \boldsymbol{\varepsilon}_a^T[n]\mathbf{P}^T[n]\mathbf{h})(v[n] + \mathbf{h}^T\mathbf{P}[n]\boldsymbol{\varepsilon}_a[n])] \\
 &= \sigma_v^2 + E\left[\boldsymbol{\varepsilon}_a^T[n]\mathbf{P}^T[n]\mathbf{h}\mathbf{h}^T\mathbf{P}[n]\boldsymbol{\varepsilon}_a[n]\right]. \tag{2.2.7}
 \end{aligned}$$

Assume the variation of  $\boldsymbol{\varepsilon}_a[n]$  is slow compared with  $\mathbf{P}^T[n]\mathbf{h}$ , hence

$$\begin{aligned}
 &E\left[\boldsymbol{\varepsilon}_a^T[n]\mathbf{P}^T[n]\mathbf{h}\mathbf{h}^T\mathbf{P}[n]\boldsymbol{\varepsilon}_a[n]\right] \\
 &\approx E\left[\boldsymbol{\varepsilon}_a^T[n]E\left[\mathbf{P}^T[n]\mathbf{h}\mathbf{h}^T\mathbf{P}[n]\right]\boldsymbol{\varepsilon}_a[n]\right] \\
 &= E\left[\boldsymbol{\varepsilon}_a^T[n]\mathbf{R}_{\mathbf{P}^T[n]\mathbf{h}}\boldsymbol{\varepsilon}_a[n]\right] \\
 &= E[\mathbf{K}_a^T[n]\mathbf{Q}^T\mathbf{R}_{\mathbf{P}^T[n]\mathbf{h}}\mathbf{Q}\mathbf{K}_a[n]] \\
 &= E[\mathbf{K}_a^T[n]\mathbf{D}_a\mathbf{K}_a[n]] \\
 &= \sum_{i=1}^N \lambda_{ai} E\left[|k_{ai}[n]|^2\right]. \tag{2.2.8}
 \end{aligned}$$

From (2.2.6) and (2.2.8), the mean square error can be written as

$$J_a[n] = \sigma_v^2 + \frac{\mu_a}{T} \sigma_v^2 \sum_{i=1}^N \frac{\lambda_{ai}}{2 - \frac{\mu_a}{T} \lambda_{ai}} + \sum_{i=1}^N \lambda_{ai} \left[ \|\mathbf{a}\|_2^2 - \frac{\frac{\mu_a}{T} \sigma_v^2}{2 - \frac{\mu_a}{T} \lambda_{ai}} \right] \left(1 - \frac{\mu_a}{T} \lambda_{ai}\right)^{2n}. \tag{2.2.9}$$

From (2.2.8) and (2.2.9), when the nonlinear coefficients error variance has been known the residual echo power can also be obtained. The nonlinear convergence rate depends on the values  $\lambda_{ai}$  of  $\mathbf{P}^T[n]\mathbf{h}$ , we may change the basis of  $\mathbf{P}$  to have a faster convergence rate.

## 2.3 Convergence analysis of linear NLMS algorithm

### 2.3.1 Variance of linear coefficients error

In this section we assume that the nonlinear coefficients are perfectly known, i.e.  $\hat{\mathbf{a}}[n] = \mathbf{a}$  or  $\boldsymbol{\varepsilon}_a[n] = 0$ . The estimation error produced by the nonlinear AEC filter is similar to (2.2.1) but here  $\boldsymbol{\varepsilon}_a[n] = 0$  and  $\boldsymbol{\varepsilon}_h[n] \neq 0$

$$\begin{aligned} e[n] &= d[n] - \hat{\mathbf{a}}^T[n] \mathbf{P}^T[n] \hat{\mathbf{h}}[n] \\ &= \mathbf{a}^T \mathbf{P}^T[n] \mathbf{h} + v[n] - [\mathbf{a} + \boldsymbol{\varepsilon}_a[n]]^T \mathbf{P}^T[n] [\mathbf{h} + \boldsymbol{\varepsilon}_h[n]] \\ &= v[n] + \boldsymbol{\varepsilon}_h^T[n] \mathbf{s}[n] \end{aligned} \quad (2.3.1)$$

where  $\mathbf{s}[n] = \mathbf{P}[n] \mathbf{a}$ . We denote the linear coefficients weight error by

$$\boldsymbol{\varepsilon}_h[n+1] = \mathbf{h} - \hat{\mathbf{h}}[n+1].$$

The following analysis procedures are similar to Section 2.2 thus we will omit the details. Using (2.1.1) and (2.3.1), we may rewrite  $\boldsymbol{\varepsilon}_h[n+1]$  as

$$\boldsymbol{\varepsilon}_h[n+1] \approx \left[ I - \frac{\mu_h}{\|\mathbf{s}[n]\|_2^2} \mathbf{R}_{s[n]} \right] \boldsymbol{\varepsilon}_h[n] + \mathbf{f}_h[n] \quad (2.3.2)$$

where

$$\mathbf{f}_h[n] = - \frac{\mu_h}{\|\mathbf{s}[n]\|_2^2} \mathbf{s}[n] v[n].$$

$\mathbf{R}_{s[n]}$  is the correlation matrix of  $\mathbf{s}[n]$ . The convergence of linear coefficients error variance can be obtained by using the same procedures as before in Section 2.2.1, the linear coefficients error variance is given by

$$E[|k_{hi}[n]|^2] = \frac{\frac{\mu_h}{\|\mathbf{s}[n]\|_2^2} \sigma_v^2}{2 - \frac{\mu_h}{\|\mathbf{s}[n]\|_2^2} \lambda_{si}} + (\|\mathbf{h}\|_2^2 - \frac{\frac{\mu_h}{\|\mathbf{s}[n]\|_2^2} \sigma_v^2}{2 - \frac{\mu_h}{\|\mathbf{s}[n]\|_2^2} \lambda_{si}}) (1 - \frac{\mu_h}{\|\mathbf{s}[n]\|_2^2} \lambda_{si})^{2n} \quad (2.3.3)$$



where  $\mathbf{Q}_s^T \mathbf{R}_s \mathbf{Q}_s = \mathbf{D}_s$ ,  $\mathbf{Q}_s$  is a unitary matrix and  $\mathbf{D}_s$  is a diagonal matrix consisting of the eigenvalues  $\lambda_{si}$ ,  $\mathbf{K}_h[n] = \mathbf{Q}_s^T \boldsymbol{\varepsilon}_h[n]$  and  $k_{hi}[n]$  is  $i$ -th entry of  $\mathbf{K}_h[n]$ .

In (2.3.3) the error variance of linear coefficients is given under assumption of perfect nonlinear coefficients. It can be determined with the knowledge of step size  $\mu_h$ , noise power  $\sigma_v^2$ , square 2- norm of room impulse response, eigenvalues of correlation matrix of  $\mathbf{s}[n]$  and the sum of all eigenvalues. Similarly, the steady state of  $E[k_{hi}[n]]^2$  is given in the first term of (2.3.3).

### 2.3.2 Residual echo power analysis

From (2.3.1), the mean square error (i.e., residual echo) due to estimated error of linear coefficients is given by

$$\begin{aligned} J_h[n] &= E\left[|e[n]|^2\right] \\ &= \sigma_v^2 + E\left[\boldsymbol{\varepsilon}_h^T[n] \mathbf{s}[n] \mathbf{s}[n]^T \boldsymbol{\varepsilon}_h[n]\right]. \end{aligned} \quad (2.3.4)$$

Assume the variation of  $\boldsymbol{\varepsilon}_h[n]$  is slow compared with  $\mathbf{s}[n]$ , hence the residual echo power can be obtained similarly by

$$\begin{aligned} J_h[n] &= \sigma_v^2 + \frac{\mu_h}{\|\mathbf{s}[n]\|_2^2} \sigma_v^2 \sum_{i=1}^M \frac{\lambda_{si}}{2 - \frac{\mu_h}{\|\mathbf{s}[n]\|_2^2} \lambda_{si}} \\ &\quad + \sum_{i=1}^M \lambda_{si} \left[ \|\mathbf{h}\|_2^2 - \frac{\frac{\mu_h}{\|\mathbf{s}[n]\|_2^2} \sigma_v^2}{2 - \frac{\mu_h}{\|\mathbf{s}[n]\|_2^2} \lambda_{si}} \right] \left(1 - \frac{\mu_h}{\|\mathbf{s}[n]\|_2^2} \lambda_{si}\right)^{2n} \end{aligned} \quad (2.3.5)$$

Unlike Section 2.2, the nonlinear basis has no effect upon the eigenvalue in (2.3.5) when nonlinear coefficients are perfectly known. Therefore, in next section, we discuss the eigenvalues and nonlinear basis relationship only when linear coefficients are error free and nonlinear coefficients are not known.

## 2.4 Nonlinear processor with orthogonal polynomials

### 2.4.1 Signal dependent orthogonal polynomials

In this section, we discuss the orthogonal and non-orthogonal basis of nonlinear processor. The nonlinear processor is shown in Fig. 2.2. To accelerate the convergence rate, we define a set of orthogonal sets  $\{p_i[n], i=1, 2, \dots\}$  if their outputs are uncorrelated with each other

$$\begin{aligned} E[p_m[n]p_n[n]] &\triangleq \sum_x p_m[n]p_n[x]f_x[x] \\ &= q \delta[m-n] \end{aligned} \quad (2.4.1)$$

where  $f_x[x]$  is the probability density function (pdf) of the input signal  $x[n]$ . If  $q$  is equal to 1, then the polynomials are not only orthogonal but also orthonormal.

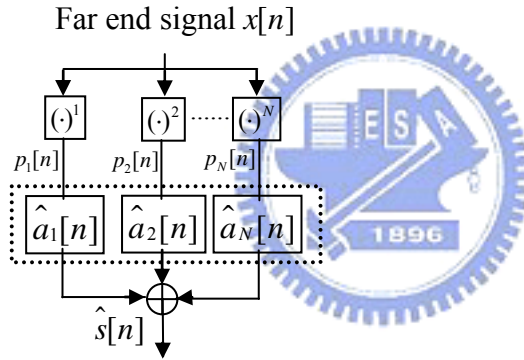


Fig.2.2 Nonlinear processor

The non-orthogonal polynomials set can be systematically modified to yield an orthogonal set by using Gram Schmidt orthonormalization in any interval with the weighting function  $f_x[x]$ . The weighting function depends on the probability distribution of the input signal, therefore we derive the general form first and then apply it to uniform, WGN, and Laplacian pdf's. The orthogonal bases are described as follow

$$\begin{aligned} p_1[n] &= x[n] \\ p_j[n] &= x^j[n] + \sum_{i=1}^{j-1} C_{j,i} x^i[n] \quad 1 < j \leq N. \end{aligned}$$

The orthogonalization coefficients  $C_{j,i}$  are chosen such that (2.4.1) is satisfactory,

i.e., the coefficients  $C_{j,i}$  for the  $j$ -th order polynomial is obtained by solving

$$\begin{bmatrix} m_{j+1} \\ \vdots \\ m_{2j-1} \end{bmatrix} = - \begin{bmatrix} m_2 & \cdots & m_j \\ \vdots & \ddots & \vdots \\ m_j & \cdots & m_{2j-2} \end{bmatrix} \begin{bmatrix} C_{j,1} \\ \vdots \\ C_{j,j-1} \end{bmatrix}$$

where  $m_i = E[x^i[n]]$  is  $i$ -th moment of the input signal  $x[n]$ . We can find out the orthogonalization coefficients when the input signal pdf is known *a priori*. The orthogonalization coefficients are constant for stationary input otherwise the coefficients are time dependent. Here, we assume that the input signal is stationary and its pdf is pre-known therefore the ensemble average is performed to find out the orthogonalization coefficients. The moment estimation is performed when the input signal pdf is not known, in such case, the orthogonalization coefficients are time dependent as described in [12]. After the orthogonalization procedures, the first six orders of orthogonal polynomials are provided in Table 2.2. Here we have assumed pdf's are symmetric with zero mean and all the odd-ordered moments are zero,  $m_i = 0$ , for  $i$ : odd.

Table 2.2 Signal-dependent orthogonal polynomials

Polynomials	Coefficients
$p_1[n] = x[n]$	
$p_2[x] = x^2[n] - C_{2,1}$	$C_{2,1} = m_2$
$p_3[x] = x^3[n] - C_{3,1}x[n]$	$C_{3,1} = \frac{m_4}{m_2}$
$p_4[x] = x^4[n] - C_{4,1}x^2[n] - C_{4,2}$	$C_{4,1} = \frac{m_6 - m_2m_4}{m_4 - m_2^2}$ $C_{4,2} = -C_{4,1}m_2 + m_4$

$p_5[x] = x^5[n] - C_{5,1}x^3[n] - C_{5,2}x[n]$	$C_{5,1} = \frac{m_8 - \frac{m_4}{m_2} m_6}{m_6 - 2\frac{m_4}{m_2} m_4 + \frac{m_4^2}{m_2}}$ $C_{5,2} = -C_{5,1} \frac{m_4}{m_2} + \frac{m_6}{m_2}$
$p_6[x] = x^6[n] - C_{6,1}x^4[n] - C_{6,2}x^2[n] - C_{6,3}$	$C_{6,1} = \frac{m_{10} - C_{4,1}m_8 - C_{4,2}m_6}{m_8 + C_{4,1}^2 m_4 + C_{4,2}^2 - 2(C_{4,1}m_6 + C_{4,2}m_4 + C_{4,1}C_{4,2}m_2)}$ $C_{6,2} = -C_{6,1}C_{4,1} + \frac{m_8 - m_2 m_6}{m_4 - m_2^2}$ $C_{6,3} = -C_{6,1}C_{4,2} - \frac{m_8 - m_2 m_6}{m_4 - m_2^2} m_2 + m_6$

Next, we discuss three kinds of signal distribution models, uniform, Gaussian and Laplacian.

#### A. Uniform input

For uniform signal in the interval  $[-1,1]$ , the moments are

$$m_2 = \frac{1}{3}, m_4 = \frac{1}{5}, m_6 = \frac{1}{7}, m_8 = \frac{1}{9}, m_{10} = \frac{1}{11}.$$

According to Table 2.2, the orthogonal polynomials for the uniformly distributed input are listed in Table 2.3.

Table 2.3 Orthogonal polynomials for a uniformly distributed signal

Order	polynomials	Order	polynomials
1	$p_1[n] = x[n]$	4	$p_4[x] = x^4[n] - \frac{6}{7}x^2[n] - \frac{3}{35}$
2	$p_2[x] = x^2[n] - \frac{1}{3}$	5	$p_5[x] = x^5[n] - 1.11x^3[n] + 0.24x[n]$
3	$p_3[x] = x^3[n] - \frac{3}{5}x[n]$	6	$p_6[x] = x^6[n] - 0.07x^4[n] - 0.65x^2[n] + 0.09$

In Table 2.3, these polynomials are not like Legendre polynomial form. Because we do not set  $p_n[1] = 1, n = 1 \sim 6$

### B. Gaussian input

For WGN input distribution signal we let  $m_2$  be equal to  $1/9=0.1111$ . The white Gaussian signal  $x[n]$  with unit variance  $\sigma_x^2=1$  is generated first and it is widely distributed over  $[-\infty, \infty]$ . In practice we let the signal within the interval  $\pm 3\sigma_x^2$  be normalized into  $[-1, 1]$ , the signal outside this interval can be ignored. Therefore, we use the normalized  $\frac{1}{3}x[n]$  as the Gaussian noise for comparison with other pdf's. Its 2<sup>nd</sup> moment is  $m_2 = 0.1111$ , the other moments are expressed as follows.

$$m_2 = 0.1111, m_4 = 3m_2^2, m_6 = 15m_2^3, m_8 = 105m_2^4, m_{10} = 945m_2^5.$$

Using these moments and Table 2.2, the orthogonal polynomials are listed in Table 2.4.

Table 2.4 Orthogonal polynomials for a Gaussian-distributed input

Order	polynomials	Order	polynomials
1	$p_1[n] = x[n]$	4	$p_4[x] = x^4[n] - 0.666x^2[n] + 0.030$
2	$p_2[x] = x^2[n] - 0.111$	5	$p_5[x] = x^5[n] - 1.11x^3[n] + 0.185x[n]$
3	$p_3[x] = x^3[n] - 0.333x[n]$	6	$p_6[x] = x^6[n] - 0.416x^4[n] - 0.277x^2[n] + 0.025$

### C. Laplacian input

For Laplacian input distribution signal, in practice, the same reason as WGN signal we let  $m_2=0.1111$  and the other moments are given by

$$m_2 = 0.1111, m_4 = 6m_2^2, m_6 = 90m_2^3, m_8 = 2520m_2^4, m_{10} = 113400m_2^5.$$

The orthogonal polynomials for Laplacian distribution signal are listed in Table 2.5.

Table 2.5 Orthogonal polynomials for a Laplacian-distributed input

Order	polynomials	Order	polynomials
1	$p_1[n] = x[n]$	4	$p_4[x] = x^4[n] - 1.8667x^2[n] + 0.1333$
2	$p_2[x] = x^2[n] - 0.1111$	5	$p_5[x] = x^5[n] - 4.074x^3[n] + 1.605x[n]$
3	$p_3[x] = x^3[n] - 0.6666x[n]$	6	$p_6[x] = x^6[n] - 4.45x^4[n] + 2.307x^2[n] + 0.050$

### 2.4.2 Eigenvalue spread analysis

In Section 2.2 we show that the convergence rate depends on the normalized eigenvalues. Equivalently, the eigenvalue spread is commonly used for convergence analysis, instead of normalized eigenvalues [6]. Here, we will examine the eigenvalue spreads for different input pdf's versus different orthogonal polynomials.

For simplicity we normalize  $\|\mathbf{h}\|_2$  to be 1 and nonlinear distortion only contains odd order harmonics and even orders are excluded, in case of a symmetric input/output nonlinear characteristic curve.

#### A. Power-series nonorthogonal polynomials basis

##### I. Uniformly distributed input

The correlation matrix,  $\mathbf{R}_{p^T[n]h}$ , is expressed as follows and the orthogonal polynomials can be found in Table 2.3.

$$\mathbf{R}_{p^T[n]h} = E[\mathbf{P}^T[n]\mathbf{h}\mathbf{h}^T\mathbf{P}[n]]$$

$$= \begin{bmatrix} r_{11} & r_{13} & r_{15} \\ r_{31} & r_{33} & r_{35} \\ r_{51} & r_{53} & r_{55} \end{bmatrix} \quad (2.4.2)$$

with  $r_{11} = E[(p_1[n]h_0 + p_1[n-1]h_1 + \dots + p_1[n-M+1]h_{M-1})^2]$

$$= m_2 \|\mathbf{h}\|_2^2.$$

$$r_{33} = E[(p_3[n]h_0 + p_3[n-1]h_1 + \dots + p_3[n-M+1]h_{M-1})^2]$$

$$= m_6 \|\mathbf{h}\|_2^2.$$

$$r_{55} = E[(p_5[n]h_0 + p_5[n-1]h_1 + \dots + p_5[n-M+1]h_{M-1})^2]$$

$$= m_{10} \|\mathbf{h}\|_2^2.$$

Unlike the orthogonal basis, the correlation term of  $\mathbf{R}_{p^T[n]h}$  is not equal to zero as a

result of the basis is not orthogonal

$$\begin{aligned} r_{13} &= r_{31} \\ &= E[(p_1[n]h_0 + p_1[n-1]h_1 + \dots + p_1[n-M+1]h_{M-1}) \\ &\quad (p_3[n]h_0 + p_3[n-1]h_1 + \dots + p_3[n-M+1]h_{M-1})] \\ &= E\left[\sum_{i=0}^{M-1} p_1[n-i]p_3[n-i]h_i^2\right] + E\left[\sum_{i=0}^{M-1} \sum_{j \neq i}^{M-1} p_1[n-i]p_3[n-j]h_i h_j\right] \\ &= m_4 \|\mathbf{h}\|_2^2. \end{aligned}$$

$$\begin{aligned} r_{15} &= r_{51} \\ &= E[(p_1[n]h_0 + p_1[n-1]h_1 + \dots + p_1[n-M+1]h_{M-1}) \\ &\quad (p_5[n]h_0 + p_5[n-1]h_1 + \dots + p_5[n-M+1]h_{M-1})] \\ &= m_6 \|\mathbf{h}\|_2^2. \end{aligned}$$

$$\begin{aligned} r_{35} &= r_{53} \\ &= E[(p_3[n]h_0 + p_3[n-1]h_1 + \dots + p_3[n-M+1]h_{M-1}) \\ &\quad (p_5[n]h_0 + p_5[n-1]h_1 + \dots + p_5[n-M+1]h_{M-1})] \end{aligned}$$

$$\begin{aligned}
& (p_5[n]h_0 + p_5[n-1]h_1 + \dots + p_5[n-M+1]h_{M-1}) \\
& = m_8 \|\mathbf{h}\|_2^2.
\end{aligned}$$

Therefore, the correlation matrix becomes

$$\mathbf{R}_{p^T[n]h} = \begin{bmatrix} 0.3333 & 0.2 & 0.1428 \\ 0.2 & 0.1428 & 0.1111 \\ 0.1428 & 0.1111 & 0.0909 \end{bmatrix},$$

which is no more diagonal. Its three eigenvalues are 0.0006, 0.0334, 0.5331 and its eigenvalue spread is as large as 888.5.

## II. Gaussian input

The non-diagonal correlation matrix is

$$\mathbf{R}_{p^T[n]h} = \begin{bmatrix} 0.1111 & 0.0370 & 0.0206 \\ 0.0370 & 0.0206 & 0.0160 \\ 0.0206 & 0.0160 & 0.0160 \end{bmatrix}.$$

Three eigenvalues are 0.0008, 0.0170, 0.1297, thus the large eigenvalue spread is 162.1.

## III. Laplacian input

The correlation matrix for the Laplacian input is

$$\mathbf{R}_{p^T[n]h} = \begin{bmatrix} 0.1111 & 0.0741 & 0.1235 \\ 0.0741 & 0.2344 & 0.3841 \\ 0.1235 & 0.3841 & 1.9204 \end{bmatrix}.$$

Three eigenvalues are 0.0177, 0.1285, 2.0088, and its eigenvalue spread is 113.5.

## B. Uniform orthogonal polynomials basis

### I. Uniform input

In the first discussion, we use uniformly orthogonal polynomials which are found under uniform input distribution. When the input signal is uniformly



distributed (matched to the polynomial bases), the nonlinear components between different orders have perfect orthogonality. The correlation matrix,  $\mathbf{R}_{p^T[n]h}$  is expressed the same as (2.4.2) with

$$\begin{aligned} r_{11} &= E \left[ (p_1[n]h_0 + p_1[n-1]h_1 + \dots + p_1[n-M+1]h_{M-1})^2 \right] \\ &= E \left[ \sum_{i=0}^{M-1} p_1^2[n-i]h_i^2 \right] + E \left[ \sum_{i=0}^{M-1} \sum_{j \neq i}^{M-1} p_1[n-i]p_1[n-j]h_i h_j \right]. \end{aligned}$$

Because the Input signal is zero mean and uncorrelated between different time index, (2.4.2) can be expressed as

$$r_{11} = m_2 \|\mathbf{h}\|_2^2.$$

Similarly,  $r_{33}$  and  $r_{55}$  can be written as

$$\begin{aligned} r_{33} &= E \left[ (p_3[n]h_0 + p_3[n-1]h_1 + \dots + p_3[n-M+1]h_{M-1})^2 \right] \\ &= (m_6 - \frac{6}{5}m_4 + \frac{9}{25}m_2) \|\mathbf{h}\|_2^2. \\ r_{55} &= E \left[ (p_5[n]h_0 + p_5[n-1]h_1 + \dots + p_5[n-M+1]h_{M-1})^2 \right] \\ &= (m_{10} - 2.22m_8 + 1.71m_6 - 0.53m_4 + 0.057m_2) \|\mathbf{h}\|_2^2. \end{aligned}$$

The other terms of the correlation matrix are zeros as a result of orthogonal property.

Therefore, we can write the correlation matrix as

$$\mathbf{R}_{p^T[n]h} = \begin{bmatrix} 0.3333 & 0 & 0 \\ 0 & 0.0029 & 0 \\ 0 & 0 & 0.0015 \end{bmatrix},$$

whose eigenvalue spread is 222.22.

## II. Gaussian input

The correlation matrix becomes

$$\mathbf{R}_{p^T[n]h} = \begin{bmatrix} 0.1111 & -0.0296 & 0.0061 \\ -0.0296 & 0.0161 & -0.0016 \\ 0.0061 & -0.0016 & 0.0024 \end{bmatrix}$$

with an eigenvalue spread of 59.1.

### III. Laplacian input

The correlation matrix becomes

$$\mathbf{R}_{p^T[n]h} = \begin{bmatrix} 0.1111 & 0.0074 & 0.0679 \\ 0.0074 & 0.0746 & 0.2241 \\ 0.0679 & 0.2241 & 1.2461 \end{bmatrix}$$

with an eigenvalue spread of 39.3.

## C. Gaussian orthogonal polynomials basis

### I. Gaussian input

For its matched input signal with Gaussian distribution, the correlation matrix can be found by the same procedure with the orthogonal polynomials given in Table 2.4.

$$\mathbf{R}_{p^T[n]h} = \begin{bmatrix} 0.0020 & 0 & 0 \\ 0 & 0.0082 & 0 \\ 0 & 0 & 0.1111 \end{bmatrix},$$

Three eigenvalues are 0.0020, 0.0082, 0.1111, thus the eigenvalue spread is 55.6.

### II. Uniformly distributed input

The correlation matrix becomes

$$\mathbf{R}_{p^T[n]h} = \begin{bmatrix} 0.3333 & 0.0889 & -0.0175 \\ 0.0889 & 0.0466 & -0.0046 \\ -0.0175 & -0.0046 & 0.0024 \end{bmatrix}$$

with an eigenvalue spread of 246.

### III. Laplacian input

The correlation matrix becomes

$$\mathbf{R}_{p^r_{[n]h}} = \begin{bmatrix} 0.1111 & 0.0370 & 0.0618 \\ 0.0370 & 0.0864 & 0.2402 \\ 0.0618 & 0.2402 & 1.2389 \end{bmatrix}$$

with an eigenvalue spread of 42.

#### D. Laplacian orthogonal polynomials basis

##### I. Uniformly distributed input

The correlation matrix becomes

$$\mathbf{R}_{p^r_{[n]h}} = \begin{bmatrix} 0.3333 & -0.0222 & -0.1369 \\ -0.0222 & 0.0243 & -0.0586 \\ -0.1369 & -0.0586 & 0.2584 \end{bmatrix}$$

with three eigenvalues of 0.0001, 0.1773, and 0.4386, and an eigenvalue spread of 3000.

##### II. Gaussian input

The correlation matrix becomes

$$\mathbf{R}_{p^r_{[n]h}} = \begin{bmatrix} 0.1111 & -0.0370 & 0.0480 \\ -0.0370 & 0.0206 & -0.0404 \\ 0.0480 & -0.0404 & 0.0950 \end{bmatrix}$$

with three eigenvalues of 0.0002, 0.05512, and 0.1714, and an eigenvalue spread of 871.4.

##### III. Laplacian input.

The correlation matrix becomes

$$\mathbf{R}_{p^r_{[n]h}} = \begin{bmatrix} 0.1111 & 0 & 0 \\ 0 & 0.0741 & 0 \\ 0 & 0 & 0.5538 \end{bmatrix}$$

which has a small eigenvalue spread of 7.5.

### E. General comparison

Finally, we discuss the eigenvalue spread under non-perfect orthogonality, a case of mismatch between the pdf of the input signal and the nonlinear polynomial basis. Table 2.6 compares the eigenvalue spreads for different input pdf's versus different orthogonal polynomial bases. If the nonlinear components have perfect orthogonality, the minimum eigenvalue spread can be achieved, as can be seen from the diagonal entries in Table 2.6. When the input signal's pdf is unknown, according to Table 2.6 the uniformly orthogonal polynomial basis is recommended.

Table 2.6 Eigenvalue spread comparison

PDF \ Basis	Uniform	Gaussian	Laplacian	Power-series
Uniform	222.2	246	3000 <sup>★</sup>	888.5
Gaussian	59.1	55.5	871.4 <sup>◆</sup>	162
Laplacian	39.3	42	7.473	113.5

<sup>★</sup> eigenvalues are 0.0001, 0.17728, 0.43863

<sup>◆</sup> eigenvalues are 0.00019671, 0.055117, 0.17141

We recall the residual echo power in (2.2.9)

$$J_a[n] = \sigma_v^2 + \frac{\mu_a}{T} \sigma_v^2 \sum_{i=0}^N \frac{\lambda_i}{2 - \frac{\mu_a}{T} \lambda_i} + \sum_{i=1}^N \lambda_i [\|a\|_2^2 - \frac{\mu_a \sigma_v^2}{2 - \frac{\mu_a}{T} \lambda_i}] (1 - \frac{\mu_a}{T} \lambda_i)^{2n}.$$

The convergence rate is dependent on the eigenvalue spread; a smaller eigenvalue has faster convergence rate. The Laplacian polynomial basis has larger eigenvalue spread than the others because it has an eigenvalue closely to zero. But, it is worthy to note that a smaller eigenvalue has less contribution to the residual echo power. Therefore, we may calculate the eigenvalue spread excluding zero eigenvalue in such case the Laplacian type basis has a smaller eigenvalue than the power series basis.

## 2.5 Application of dual loudspeakers system

In this section we consider the hands-free system with dual loudspeakers. 3D sound is an essential element of the new services (mobile, multimedia, etc.). It enriches sound playback more vividly. In fact the 3D sound effect can be generated by dual loudspeakers. The AEC is also essential to achieve satisfactory speech quality in such system. We use two nonlinear AEC with a cascade structure to cancel the left and right channel echo signal. The nonlinear AEC structure for dual loudspeakers system is shown in Fig. 2.3. Let  $x_L[n]$ ,  $x_R[n]$  denote the left-channel and right-channel signal and uncorrelated with each other. Nonlinear AEC1 and AEC2 are used to cancel the echo components from the left and right loudspeaker, respectively.

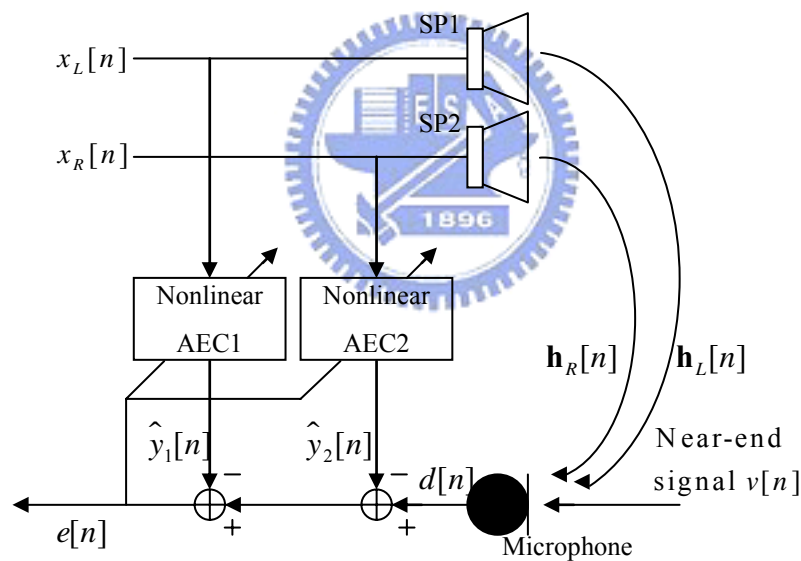


Fig. 2.3 Nonlinear AEC structure for dual loudspeakers system

The signals, passing through the echo path  $\mathbf{h}_L[n]$  and  $\mathbf{h}_R[n]$ , respectively, are picked up by the microphone. The microphone output signal is expressed as

$$d[n] = [a_1 a_3] \begin{bmatrix} x_{L,1}[n] & x_{L,1}[n-1] & \cdots & x_{L,1}[n-M+1] \\ x_{L,3}[n] & x_{L,3}[n-1] & \cdots & x_{L,3}[n-M+1] \end{bmatrix} \begin{bmatrix} h_{L0} \\ h_{L1} \\ \vdots \\ h_{LM-1} \end{bmatrix} \\ + [b_1 b_3] \begin{bmatrix} x_{R,1}[n] & x_{R,1}[n-1] & \cdots & x_{R,1}[n-M+1] \\ x_{R,3}[n] & x_{R,3}[n-1] & \cdots & x_{R,3}[n-M+1] \end{bmatrix} \begin{bmatrix} h_{R0} \\ h_{R1} \\ \vdots \\ h_{RM-1} \end{bmatrix} + v[n].$$

We rewrite  $d[n]$  as vector inner products:

$$d[n] = v[n] + \mathbf{a}^T \mathbf{P}_{x_L}^T[n] \mathbf{h}_L + \mathbf{b}^T \mathbf{P}_{x_R}^T[n] \mathbf{h}_R. \quad (2.5.1)$$

In this section we derive the residual echo power under the assumption of perfect linear coefficients. We derive the nonlinear coefficients convergence rate first. It is easy to see that (2.5.1) is not the same with the case of a single loudspeaker, here, we have two kinds of nonlinear coefficients in the right-hand side of (2.5.1). We can cascade up the nonlinear coefficients  $\mathbf{a}$  and  $\mathbf{b}$  together into a new vector  $\mathbf{c}$ . By the same procedure, we merge  $\mathbf{P}_{x_L}^T[n]$  and  $\mathbf{P}_{x_R}^T[n]$  into a block-diagonal matrix  $\mathbf{P}_s^T[n]$ .

The desired signal can be rewritten as

$$d[n] = v[n] + [\mathbf{a}^T \mathbf{b}^T] \begin{bmatrix} \mathbf{P}_{x_L}^T[n] \mathbf{h}_L \\ \mathbf{P}_{x_R}^T[n] \mathbf{h}_R \end{bmatrix} \\ = v[n] + [a_1[n] a_3[n] b_1[n] b_3[n]] \begin{bmatrix} \mathbf{P}_{x_L}^T[n] & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_{x_R}^T[n] \end{bmatrix}^T \begin{bmatrix} \mathbf{h}_L \\ \mathbf{h}_R \end{bmatrix} \\ = v[n] + \mathbf{c}^T[n] \mathbf{P}_s^T[n] \mathbf{h}_s \quad (2.5.2)$$

where  $\mathbf{h}_s$  denotes  $[\mathbf{h}_L \ \mathbf{h}_R]^T$

(2.5.2) is almost the same as the single loudspeaker case. The nonlinear AEC output and residual echo become

$$\hat{y}[n] = \hat{\mathbf{c}}^T [n] \mathbf{P}_s^T [n] \mathbf{h}_s$$

$$\begin{aligned} e[n] &= d[n] - \hat{y}[n] \\ &= d[n] - \hat{\mathbf{c}}^T [n] \mathbf{P}_s^T [n] \mathbf{h}_s. \end{aligned}$$

The NLMS adaptation algorithm becomes

$$\begin{aligned} \nabla_c &= \frac{\partial e^2[n]}{\partial \hat{\mathbf{c}}[n]} = -2e[n] \mathbf{P}_s^T [n] \mathbf{h}_s \\ \mathbf{c}[n+1] &= \mathbf{c}[n] + \mu_c \frac{\mathbf{P}_s^T [n] \mathbf{h}_s}{\|\mathbf{P}_s^T [n] \mathbf{h}_s\|_2^2 + \delta} e[n]. \end{aligned} \quad (2.5.3)$$

The residual echo convergence rate can be found following the same procedures in Section 2.2.2. The nonlinear coefficients error variance and residual echo power are described in (2.5.4) and (2.5.5).

$$E[k_{ai}[n]^2] = \frac{\frac{\mu_a}{T} \sigma_v^2}{2 - \frac{\mu_a}{T} \lambda_i} + \left( \|\mathbf{c}\|_2^2 - \frac{\frac{\mu_a}{T} \sigma_v^2}{2 - \frac{\mu_a}{T} \lambda_i} \right) \left( 1 - \frac{\mu_a}{T} \lambda_i \right)^{2n} \quad (2.5.4)$$

$$J_a[n] = \sigma_v^2 + \frac{\mu_a}{T} \sigma_v^2 \sum_{i=0}^N \frac{\lambda_i}{2 - \frac{\mu_a}{T} \lambda_i} + \sum_{i=1}^N \lambda_i \left[ \|\mathbf{c}\|_2^2 - \frac{\frac{\mu_a}{T} \sigma_v^2}{2 - \frac{\mu_a}{T} \lambda_i} \right] \left( 1 - \frac{\mu_a}{T} \lambda_i \right)^{2n} \quad (2.5.5)$$

where

$$T = \|\mathbf{P}_s^T [n] \mathbf{h}_s\|_2^2 + \delta$$

$$\mathbf{P}_s^T [n] \mathbf{h}_s = \begin{bmatrix} \mathbf{P}_{x_L}^T [n] & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_{x_R}^T [n] \end{bmatrix}^T \begin{bmatrix} \mathbf{h}_L \\ \mathbf{h}_R \end{bmatrix}$$

$$\|\mathbf{c}\|_2^2 = \|\mathbf{a} \mathbf{b}\|_2^2.$$

## 2.6 Summary

In this chapter we show a nonlinear AEC with a signal dependent orthogonal basis to decrease the correlation among different polynomial orders and increase the convergence rate. We have presented convergence analyses of linear and nonlinear NLMS algorithm, under the assumption of the other part of coefficients are perfectly known. Because the linear and nonlinear coefficients errors affect each other in the cascade structure, it is difficult to perform the joint error analysis theoretically. For an input with unknown pdf, the orthogonality of the polynomial bases may not be perfect, the eigenvalue spread analyses in Section 2.4 shows that we also have faster convergence rate than conventional power series basis. For dual loudspeakers case, we cascade the coefficients together therefore the analyses in single loudspeaker case can be easily extended to the dual loudspeakers case.





# Chapter 3

## Training Sequence and Coefficients Estimation

The training sequence (TS) can also be used in adaptive equalizer or echo path estimation. Here we will apply it to nonlinear AEC with a cascade structure. There are applications where it is necessary to compare one reference signal with one or more signals to determine the similarity between these two and to determine additional information based on the similarity. For example, in digital communications, a set of data symbols are represented by a set of unique discrete time sequences. If one of these sequences is transmitted, the receiver has to determine which particular sequence has been received by comparing the received signal with every member of possible sequences from the set. Similarly, in radar and sonar applications, the received signal reflected from the target is the delayed version of the transmitted signal and by measuring the delay; one can determine the location of the target.

In Section 3.2, we will derive the TS algorithm under expectation operator. In practice, we replace expectation with sample mean. We will show the recursive form of TS algorithm in Section 3.3. In Section 3.4, we will perform the convergence analysis of linear TS coefficients and nonlinear convergence analysis in Section 3.5. In Section 3.6, we extend applications to dual loudspeakers system where the linear convergence will be given.

### 3.1 Correlation based nonlinear AEC

In Chapter 2 we have known that the NLMS adaptive algorithm is based on a simple stochastic gradient. The kind of adaptive algorithm has good performance when the background noise or double talk is not present. In Fig. 2.1 the error signal  $e[n]$  is used to update both the linear filter and the nonlinear processor. However, when the background noise is present and/or larger than the echo signal  $y[n]$ , the desired signal contains not only the echo signal but also the background noise. The coefficients may diverge when the error signal includes a significant near-end speech signal.

In this chapter we use a white sequence (non polar signal) to train both linear and nonlinear coefficients to overcome the problem of adaptive filter under low SNR condition. Because of the cascade structure we have to modify the Wiener-Hopf equation to fulfill our requirement. In the case of our structure, the TS runs through a multipath nonlinear processor and weighted by nonlinear coefficients, the sum of multipath signals passes through a linear filter. Therefore, the vector direction of the estimated coefficients is parallel to the vector of the room impulse response and its vector length is composed of nonlinear coefficients. The other problem is how to separate each nonlinear coefficient from the length of estimated coefficients. Here we only have one cross correlation length between the input and the microphone signal, in order to get more information we can use the nonlinear order of input signal and the microphone output signal to generate the other information about the nonlinear loudspeaker. The nonlinear coefficients can be found by solving these equation sets. According to this concept, the structure of a nonlinear AEC based on TS is shown in Fig. 3.1.

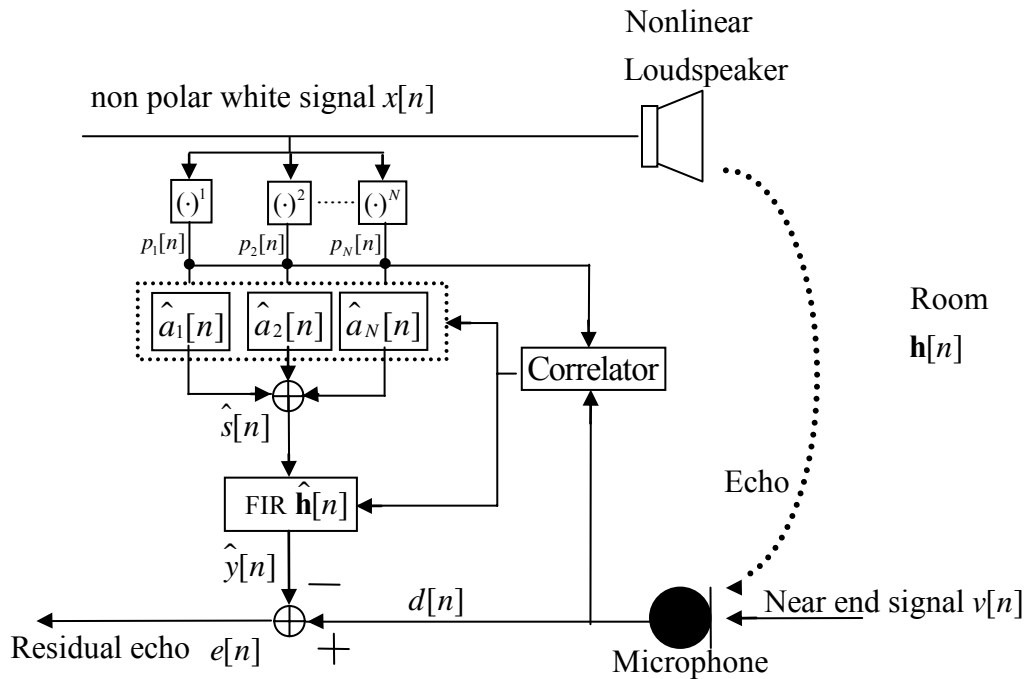


Fig.3.1 Structure of a nonlinear AEC based on training sequence

In Fig. 3.1, the correlator produces the correlation between  $p_1[n]$  and  $d[n]$ . To find the nonlinear coefficients, the correlator will also create the correlation between  $p_j[n]$  and  $d[n]$ , for  $2 \leq j \leq N$

## 3.2 Training sequence estimation algorithm

In this section we will derive the linear TS coefficients algorithm first in Section 3.2.1. In Section 3.2.2 we find the nonlinear TS coefficients algorithm. We also give the TS algorithm for nonlinear coefficients when an orthogonal basis is used.

### 3.2.1 Linear TS coefficients estimation algorithm

In the section we will derive how to use the TS to find out the linear coefficients. First, we generate a white and zero mean sequence  $x[n]$  by TS generator at the near end. This sequence cannot be a polar signal in order to have nontrivial higher order moments, which will be explained later. The sequence is injected to the loudspeaker and through the near end room it is picked up by the microphone. The desired signal is expressed as follows:

$$\begin{aligned}
 d[n] &= \sum_{t=1}^N a_t \sum_{i=0}^{M-1} h_i x^t[n-i] + v[n] \\
 &= a_1 \sum_{i=0}^{M-1} h_i x[n-i] + a_2 \sum_{i=0}^{M-1} h_i x^2[n-i] + \dots + a_N \sum_{i=0}^{M-1} h_i x^N[n-i] + v[n]. \quad (3.2.1)
 \end{aligned}$$

$M$  is the length of the linear filter,  $N$  is the nonlinear order. We multiply (3.2.1) by  $x[n-k]$  and take the expectation value to get

$$\begin{aligned}
 E[x[n-k]d[n]] &= \sum_{t=1}^N a_t \sum_{i=0}^{M-1} h_i E[x[n-k]x^t[n-i]] \\
 &= a_1 \sum_{i=0}^{M-1} h_i E[x[n-k]x[n-i]] + a_2 \sum_{i=0}^{M-1} h_i E[x[n-k]x^2[n-i]] + \dots \\
 &\quad + a_N \sum_{i=0}^{M-1} h_i E[x[n-k]x^N[n-i]] + E[x[n-k]v[n]] \quad , \quad k = 0, 1, \dots \quad (3.2.2)
 \end{aligned}$$

The two expectations in (3.2.2) may be written as

$$\begin{aligned}
 r_{xd}[-k] &= \sum_{t=1}^N a_t \sum_{i=0}^{M-1} h_i r_{xx^t}[i-k] \\
 &= a_1 \sum_{i=0}^{M-1} h_i r_{xx}[i-k] + a_2 \sum_{i=0}^{M-1} h_i r_{xx^2}[i-k] + \dots + a_N \sum_{i=0}^{M-1} h_i r_{xx^N}[i-k] \quad (3.2.3)
 \end{aligned}$$

where

$$r_{xd}[-k] = E[x[n-k]d[n]]$$

is defined as the cross correlation function between the input signal  $x[n]$  and the microphone output signal  $d[n]$  for a lag of  $-k$ ; and

$$r_{xx'}[i-k] = E[x[n-k]x'[n-i]],$$

$r_{xx'}[i-k]$  is defined as the autocorrelation function of the input signal for a lag of  $i-k$ . The last term of (3.2.2),  $E[x[n-k]v[n]]$ , is equal to zero due to the input signal is independent of the background noise. We define

$$\mathbf{r}_{xx} = [r_{xx}[0-k], r_{xx}[1-k], \dots, r_{xx}[M-1-k]]$$

and

$$\mathbf{h} = [h_0, h_1, \dots, h_{M-1}]^T,$$

then we rewrite (3.2.3) into the form of inner products:

$$\begin{aligned} r_{xd}[-k] &= \sum_{t=1}^N a_t \mathbf{r}_{xx^t} \mathbf{h} \\ &= a_1 \mathbf{r}_{xx^1} \mathbf{h} + a_2 \mathbf{r}_{xx^2} \mathbf{h} + \dots + a_N \mathbf{r}_{xx^N} \mathbf{h}. \end{aligned} \quad (3.2.4)$$

According to (3.2.4), it is similar to the Wiener-Hopf equation. The left-hand side of (3.2.4) contains the cross correlation, but its right-hand side contains not only linear but also nonlinear coefficients. Therefore, we need to take some procedures to estimate the linear and nonlinear coefficients. Before the procedure, we extend (3.2.4) into matrix form. The impulse response is defined by the finite set of tap weights, i.e.,  $\mathbf{h} = [h_0, h_1, \dots, h_{M-1}]^T$ , hence we let the lag index  $k$  go from 0 to  $M-1$ . We extend (3.2.4) in matrix form as follows:

$$\mathbf{R}_{xd} = E[\mathbf{x}[n]d[n]] = \begin{bmatrix} r_{xd}[0] \\ r_{xd}[-1] \\ \vdots \\ r_{xd}[1-M] \end{bmatrix}$$

$$\begin{aligned}
&= \sum_{t=1}^N a_t \mathbf{R}_{xx^t} \mathbf{h} \\
&= a_1 \mathbf{R}_{xx} \mathbf{h} + a_2 \mathbf{R}_{xx^2} \mathbf{h} + \cdots + a_N \mathbf{R}_{xx^N} \mathbf{h}
\end{aligned} \tag{3.2.5}$$

$\mathbf{x}[n]$  denotes the  $M$ -by-1 vector of the tap inputs  $x[n], x[n-1], \dots, x[n-M+1]$ .  $\mathbf{R}_{xd}$  denotes the  $M$ -by-1 cross correlation vector between the input signal and the microphone output signal.  $\mathbf{R}_{xx}$  denotes the  $M$ -by- $M$  correlation matrix of  $\mathbf{x}[n]$

$$\begin{aligned}
\mathbf{R}_{xx} &= E \begin{bmatrix} x[n]x[n] & x[n]x[n-1] & \cdots & x[n]x[n-M+1] \\ x[n-1]x[n] & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ x[n-M+1]x[n] & \cdots & \cdots & x[n-M+1]x[n-M+1] \end{bmatrix} \\
&= m_2 \mathbf{I}_{M \times M} .
\end{aligned}$$

$$\mathbf{R}_{xx^3} = m_4 \mathbf{I}_{M \times M}$$

$\vdots$

$$\mathbf{R}_{xx^t} = m_{t+1} \mathbf{I}_{M \times M} \text{ for } 1 \leq t \leq N \text{ and } t \text{ is odd}$$

where the moment is defined as  $m_t = E[x^t[n]]$ . We note that  $\mathbf{R}_{xx^2}, \mathbf{R}_{xx^4}, \dots, \mathbf{R}_{xx^t}$  are zero matrices for  $2 \leq t \leq N$ ,  $t$  is even, since the input signal is white zero mean, a symmetric pdf, and uncorrelated between different orders,  $E[x[n]x^2[n]] = 0$ . Here we let  $N$  be an odd integer and rewrite (3.2.5) as

$$\begin{aligned}
\mathbf{R}_{xd} &= \sum_{\substack{t=1 \\ t \text{ is odd}}}^N a_t m_{t+1} \mathbf{h} \\
&= a_1 m_2 \mathbf{h} + a_3 m_4 \mathbf{h} + \cdots + a_N m_{N+1} \mathbf{h} .
\end{aligned} \tag{3.2.6}$$

We may then pre-multiply both sides of  $(a_1 m_2 + a_3 m_4 + \cdots + a_N m_{N+1})^{-1} \mathbf{I}_{M \times M}$  and solve (3.2.6) for  $\mathbf{h}$ , but the nonlinear coefficients are unknown yet. We assume that  $\|\mathbf{h}\|_2^2$  equals to 1 thus the direction of  $\mathbf{R}_{xd}$  is equal to  $\mathbf{h}$ , then the room impulse can be found by

$$\mathbf{h} = \frac{\mathbf{R}_{xd}}{\|\mathbf{R}_{xd}\|_2}. \quad (3.2.7)$$

### 3.2.2 Nonlinear TS coefficients estimation algorithm

We have to solve the nonlinear coefficients but we only have one equation in (3.2.6). Because the input signal pdf is symmetric, the odd-order moments are zeros. We divide the nonlinear TS algorithm into two groups, odd and even-ordered. First, we solve the odd-ordered nonlinear coefficients. Again, we multiply (3.2.1) by  $x^3[n-k], \dots, x^t[n-k]$  for  $1 \leq t \leq N$ ,  $t$  is odd, respectively and take the expectation value. Similar to (3.2.6), we can get the other equations:

$$\mathbf{R}_{x^3d} = (a_1 m_4 + a_3 m_6 + \dots + a_N m_{N+3}) \mathbf{h} \quad (3.2.8)$$

⋮

$$\mathbf{R}_{x^N d} = (a_1 m_{N+1} + a_3 m_{N+3} + \dots + a_N m_{2N}) \mathbf{h}. \quad (3.2.9)$$

We multiply (3.2.6), (3.2.8), and (3.2.9) by  $\mathbf{h}^T$  respectively and then we can solve the following equations to find the odd-ordered nonlinear coefficients  $a_1, a_3, \dots, a_N$

$$\begin{bmatrix} \mathbf{h}^T \mathbf{R}_{xd} \\ \mathbf{h}^T \mathbf{R}_{x^3d} \\ \vdots \\ \mathbf{h}^T \mathbf{R}_{x^N d} \end{bmatrix} = \underbrace{\begin{bmatrix} m_2 & m_4 & \dots & m_{N+1} \\ m_4 & m_6 & \dots & m_{N+3} \\ \vdots & \vdots & \ddots & \vdots \\ m_{N+1} & m_{N+3} & \dots & m_{2N} \end{bmatrix}}_{\mathbf{G}_{odd}} \begin{bmatrix} a_1 \\ a_3 \\ \vdots \\ a_N \end{bmatrix}.$$

Hence, the odd-ordered nonlinear coefficients can be found as

$$\begin{bmatrix} a_1 \\ a_3 \\ \vdots \\ a_N \end{bmatrix} = \underbrace{\begin{bmatrix} m_2 & m_4 & \dots & m_{N+1} \\ m_4 & m_6 & \dots & m_{N+3} \\ \vdots & \vdots & \ddots & \vdots \\ m_{N+1} & m_{N+3} & \dots & m_{2N} \end{bmatrix}}_{\mathbf{G}_{odd}}^{-1} \begin{bmatrix} \mathbf{h}^T \mathbf{R}_{xd} \\ \mathbf{h}^T \mathbf{R}_{x^3d} \\ \vdots \\ \mathbf{h}^T \mathbf{R}_{x^N d} \end{bmatrix}. \quad (3.2.10)$$

We note that the matrix  $\mathbf{G}_{odd}$  in (3.2.10) has to be nonsingular in order to have a unique solution for the nonlinear coefficients  $a_1, a_2, \dots, a_N$ . As mentioned earlier, the input training sequence  $x[n]$  cannot be a polar  $\pm 1$  signal. Up to now, we have found

the linear coefficients  $\mathbf{h}$  and odd-ordered nonlinear coefficients  $a_t$  for  $1 \leq t \leq N$ ,  $t$  is odd. Next we derive the even-ordered nonlinear coefficients. Similarly, we start with (3.2.2) but there are some differences from previous procedure. Although the input signal is white and zero mean but the even-ordered moments are not equal to zero. In order to get a diagonal correlation matrix we multiply (3.2.1) by  $x^{2[n-k]-m_2}$  and take expectation to get

$$\mathbf{R}_{(x^{2-m_2})d} = a_2(m_4 - m_2^2)\mathbf{h} + a_4(m_6 - m_2m_4)\mathbf{h} + \cdots + a_4(m_{N+2} - m_2m_N)\mathbf{h} \quad (3.2.11)$$

Again, we multiply (3.2.1) by  $x^4[n-k]-m_4, \dots, x^{N-1}[n-k]-m_{N-1}$  and take expectation value

$$\begin{aligned} & \vdots \\ & \mathbf{R}_{(x^{N-1-m_{N-1}})d} = a_2(m_{N+1} - m_2m_{N-1})\mathbf{h} + a_4(m_{N+3} - m_4m_{N-1})\mathbf{h} + \cdots + a_{N-1}(m_{2N-2} - m_{N-1}^2)\mathbf{h}. \end{aligned} \quad (3.2.12)$$

By multiplying (3.2.11) and (3.2.12) by  $\mathbf{h}^T$  respectively, we can find out the nonlinear coefficients  $a_t$  for  $1 \leq t \leq N$ ,  $t$  is even, by solving the following equations

$$\begin{bmatrix} \mathbf{h}^T \mathbf{R}_{(x^{2-m_2})d} \\ \mathbf{h}^T \mathbf{R}_{(x^4-m_4)d} \\ \vdots \\ \mathbf{h}^T \mathbf{R}_{(x^{N-1-m_{N-1}})d} \end{bmatrix} = \underbrace{\begin{bmatrix} m_4 - m_2^2 & m_6 - m_2m_4 & \cdots & m_{N+2} - m_2m_N \\ m_6 - m_2m_4 & m_8 - m_4^2 & \cdots & m_{N+3} - m_4m_{N-1} \\ \vdots & \vdots & \ddots & \vdots \\ m_{N+1} - m_2m_{N-1} & m_{N+3} - m_4m_{N-1} & \cdots & m_{2N-2} - m_{N-1}^2 \end{bmatrix}}_{\mathbf{G}_{even}} \begin{bmatrix} a_2 \\ a_4 \\ \vdots \\ a_{N-1} \end{bmatrix}.$$

Hence, the even-ordered nonlinear coefficients can be found as

$$\begin{bmatrix} a_2 \\ a_4 \\ \vdots \\ a_{N-1} \end{bmatrix} = \underbrace{\begin{bmatrix} m_4 - m_2^2 & m_6 - m_2m_4 & \cdots & m_{N+2} - m_2m_N \\ m_6 - m_2m_4 & m_8 - m_4^2 & \cdots & m_{N+3} - m_4m_{N-1} \\ \vdots & \vdots & \ddots & \vdots \\ m_{N+1} - m_2m_{N-1} & m_{N+3} - m_4m_{N-1} & \cdots & m_{2N-2} - m_{N-1}^2 \end{bmatrix}}_{\mathbf{G}_{even}}^{-1} \begin{bmatrix} \mathbf{h}^T \mathbf{R}_{(x^{2-m_2})d} \\ \mathbf{h}^T \mathbf{R}_{(x^4-m_4)d} \\ \vdots \\ \mathbf{h}^T \mathbf{R}_{(x^{N-1-m_{N-1}})d} \end{bmatrix}. \quad (3.2.13)$$



Again, we have assumed that the training white sequence  $x[n]$  is non polar signal to avoid the singularity of the matrix  $\mathbf{G}_{even}$ . Finally, in Chapter 2 we mentioned a nonlinear processor with orthogonal basis. When the orthogonal basis is used the same procedure can also be used to find out the coefficients. The desired signal can be written as

$$\begin{aligned} d[n] &= \sum_{t=1}^N a_t \sum_{i=0}^{M-1} h_i p_t[n-i] \\ &= a_1 \sum_{i=0}^{M-1} h_i p_1[n-i] + a_2 \sum_{i=0}^{M-1} h_i p_2[n-i] + \dots + a_N \sum_{i=0}^{M-1} h_i p_N[n-i] + v[n] \end{aligned}$$

where  $p_i$  becomes the nonlinear orthogonal polynomial. Following the same procedure from (3.2.8) to (3.2.13), we can find the linear and nonlinear coefficients

$$\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{bmatrix} = \underbrace{\begin{bmatrix} m_{p_1} & 0 & \dots & 0 \\ 0 & m_{p_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & m_{p_N} \end{bmatrix}}_{\mathbf{G}_{orthogonal}}^{-1} \begin{bmatrix} \mathbf{h}^T \mathbf{R}_{p_1 d} \\ \mathbf{h}^T \mathbf{R}_{p_2 d} \\ \vdots \\ \mathbf{h}^T \mathbf{R}_{p_N d} \end{bmatrix}. \quad (3.2.14)$$

(3.2.14) shows the nonlinear coefficients, where  $m_{p_i}$  is equal to the second moment of the basis,  $p_i$ . The matrix inversion in (3.2.14) is simpler than that of non-orthogonal basis because it is a diagonal matrix. Because of the orthogonality (3.2.16) can represent the odd and even-ordered coefficients simultaneously.

### 3.3 Recursive Training Sequence algorithm

In order to compute the linear and nonlinear coefficients in (3.2.7), (3.2.10), and (3.2.13), we need to compute the autocorrelation function and cross correlation function. In practical system, we replace the expectation operation with sample mean. The sample mean of a set  $x[1], x[2], \dots, x[n]$  of  $n$  observations is defined by

$$\text{sample mean} = \frac{1}{n} \sum_{i=1}^n x[i]$$

In Section 3.2, the expectation operator is performed to compute coefficients thus the true coefficients notation  $\mathbf{h}$  is used. In this section the sample mean method is used, therefore we use the estimation's notation as that in Chapter 2. At first, we use the sample mean as follows

$$\begin{aligned} \hat{\mathbf{R}}_{xd}[n] &= \frac{1}{n} \sum_{i=1}^n \mathbf{x}[i]d[i] \\ &= \frac{1}{n} \left[ \frac{n-1}{n-1} \sum_{i=1}^{n-1} \mathbf{x}[i]y[i] + \mathbf{x}[n]y[n] \right] \\ &= \frac{n-1}{n} \hat{\mathbf{R}}_{xd}[n-1] + \frac{1}{n} \mathbf{x}[n]d[n]. \end{aligned} \quad (3.3.1)$$

According to (3.3.1), we can rewrite (3.2.7) as follows

$$\begin{aligned} \hat{\mathbf{h}}[n] &= \frac{\hat{\mathbf{R}}_{xd}[n]}{\|\hat{\mathbf{R}}_{xd}[n]\|_2} \\ &= \frac{1}{\|\hat{\mathbf{R}}_{xd}[n]\|_2} \left( \frac{n-1}{n} \hat{\mathbf{R}}_{xd}[n-1] + \frac{1}{n} \mathbf{x}[n]d[n] \right) \\ &= \frac{n-1}{n} \hat{\mathbf{h}}[n-1] + \frac{1}{n \|\hat{\mathbf{R}}_{xd}[n]\|_2} \mathbf{x}[n]d[n]. \end{aligned} \quad (3.3.2)$$

In (3.3.2),  $\|\hat{\mathbf{R}}_{xd}[n]\|_2$  is equal to  $(\hat{a}_1[n-1]m_2 + \hat{a}_3[n-1]m_4 + \dots + \hat{a}_N[n-1]m_{N+1})$ , the nonlinear coefficients can be obtained from the last iteration of (3.3.4). (3.3.2) is similar to (2.1.1), the second term on the right hand side of (3.3.2) represents the

adjustment that is applied to the current estimate of linear coefficients. But the error signal does not appear in (3.3.2), it is different from the NLMS-type adaptation algorithm with an error feedback structure.

The TS estimation of nonlinear coefficients in (3.2.10) can be written as a recursive form too. First, we have to replace the correlation matrix of (3.2.10) with the sample mean. We can rewrite the cross correlation as

$$\begin{bmatrix} \widehat{\mathbf{R}}_{x^d}[n] \\ \widehat{\mathbf{R}}_{x^3d}[n] \\ \vdots \\ \widehat{\mathbf{R}}_{x^Nd}[n] \end{bmatrix} = \frac{n-1}{n} \begin{bmatrix} \widehat{\mathbf{R}}_{x^d}[n-1] \\ \widehat{\mathbf{R}}_{x^3d}[n-1] \\ \vdots \\ \widehat{\mathbf{R}}_{x^Nd}[n-1] \end{bmatrix} + \frac{1}{n} \begin{bmatrix} \mathbf{x}[n]d[n] \\ \mathbf{x}^3[n]d[n] \\ \vdots \\ \mathbf{x}^N[n]d[n] \end{bmatrix}. \quad (3.3.3)$$

According to (3.3.3), (3.2.10) can be written into a recursive form

$$\begin{bmatrix} \widehat{a}_1[n] \\ \widehat{a}_3[n] \\ \vdots \\ \widehat{a}_N[n] \end{bmatrix} = \mathbf{G}_{odd}^{-1} \begin{bmatrix} \frac{n-1}{n} \begin{bmatrix} \widehat{\mathbf{h}}^T[n] \widehat{\mathbf{R}}_{x^d}[n-1] \\ \widehat{\mathbf{h}}^T[n] \widehat{\mathbf{R}}_{x^3d}[n-1] \\ \vdots \\ \widehat{\mathbf{h}}^T[n] \widehat{\mathbf{R}}_{x^Nd}[n-1] \end{bmatrix} + \frac{1}{n} d[n] \begin{bmatrix} \widehat{\mathbf{h}}^T[n] \mathbf{x}[n] \\ \widehat{\mathbf{h}}^T[n] \mathbf{x}^3[n] \\ \vdots \\ \widehat{\mathbf{h}}^T[n] \mathbf{x}^N[n] \end{bmatrix} \end{bmatrix} \\ = \frac{n-1}{n} \begin{bmatrix} \widehat{a}_1[n-1] \\ \widehat{a}_3[n-1] \\ \vdots \\ \widehat{a}_N[n-1] \end{bmatrix} + \frac{1}{n} d[n] \mathbf{G}_{odd}^{-1} \begin{bmatrix} \widehat{\mathbf{h}}^T[n] \mathbf{x}[n] \\ \widehat{\mathbf{h}}^T[n] \mathbf{x}^3[n] \\ \vdots \\ \widehat{\mathbf{h}}^T[n] \mathbf{x}^N[n] \end{bmatrix}. \quad (3.3.4)$$

The matrix inverse  $\mathbf{G}_{odd}^{-1}$  in the right-hand side of (3.3.4) can be pre-computed so long as each moment of input signal in (3.2.10) are known. The same procedure can be applied to (3.2.13); the recursive form of (3.2.13) becomes

$$\begin{bmatrix} \widehat{a}_2[n] \\ \widehat{a}_4[n] \\ \vdots \\ \widehat{a}_{N-1}[n] \end{bmatrix} = \frac{n-1}{n} \begin{bmatrix} \widehat{a}_2[n-1] \\ \widehat{a}_4[n-1] \\ \vdots \\ \widehat{a}_{N-1}[n-1] \end{bmatrix} + \frac{1}{n} d[n] \mathbf{G}_{even}^{-1} \begin{bmatrix} \widehat{\mathbf{h}}^T[n](\mathbf{x}^2[n] - m_2) \\ \widehat{\mathbf{h}}^T[n](\mathbf{x}^4[n] - m_4) \\ \vdots \\ \widehat{\mathbf{h}}^T[n](\mathbf{x}^{N-1}[n] - m_{N-1}) \end{bmatrix} \quad (3.3.5)$$

where the moment matrix  $\mathbf{G}_{even}$  are given in (3.2.13).

The steady state of linear and nonlinear TS coefficients will achieve the optimum Wiener solution, when the iteration numbers approaches to infinity. By contrast, the coefficients computed by the NLMS-type algorithm might not achieve the optimum solution. It has excess mean square error resulting from gradient noise at small step size.

In view of computational complexity, it needs about  $2M+N$  and  $MN+2N+ N^2$  multiplications per iteration to find the linear and nonlinear TS coefficients, respectively. Here, we do not take into account the computational cost of the matrix inverse  $\mathbf{G}^{-1}$ , since it can be computed a priori. To generate the nonlinear components, it needs about  $N-1$  multiplications. Table 3.1 compares the computational complexities between the TS method and NLMS adaptive filter. The computational complexity of the TS method is almost the same as the NLMS adaptive filter. Both methods can be used to find the AEC coefficients. In case of a noisy environment, the TS method is more attractive.

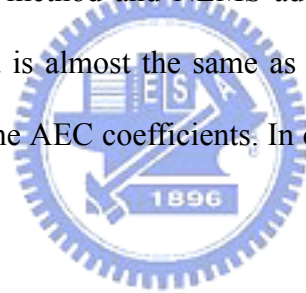


Table 3.1 Comparison of complexity computational

	Number of multiplication
Training method	$MN+2M+4N-1+ N^2$
NLMS-adaptive filter	$MN+2M+4N-1$

### 3.4 Linear convergence analysis

The linear filter dominates the overall system convergence rate because it has more coefficients than the nonlinear processor order. Through analysis, the linear coefficients convergence rate can be used to estimate the whole system performance. In this section we will perform the convergence analysis of linear TS coefficients and residual echo power under the assumptions of perfect nonlinear coefficients with 5<sup>th</sup> nonlinear order. The nonlinear TS coefficients estimation is performed on a system of equation sets. Theoretical analysis is difficult because of the moment matrices inverse in (3.3.4) and (3.3.5).

#### 3.4.1 Variance of linear coefficients error

Because the input signal is independent of the background noise we divide the analysis into two parts by the law of superposition. In the first part, we discuss the effect of the background noise under the assumption of perfect linear coefficients. The other part deals with the error due to sample mean estimation method under the assumption of no background noise. In the first discussion, the estimated error vector due to noise can be written as follows

$$\begin{aligned}\boldsymbol{\varepsilon}_h[n] &= E\left[\mathbf{h} - \frac{1}{n(a_1m_2 + a_3m_4 + a_5m_6)} \sum_{i=1}^n \mathbf{x}[i]d[i]\right] \\ &= E\left[\mathbf{h} - \frac{1}{(a_1m_2 + a_3m_4 + a_5m_6)} \mathbf{R}_{xd} - \frac{1}{n(a_1m_2 + a_3m_4 + a_5m_6)} \sum_{i=1}^n \mathbf{x}[i]v[i]\right] \\ &= \frac{-1}{n(a_1m_2 + a_3m_4 + a_5m_6)} \sum_{i=1}^n \mathbf{x}[i]v[i].\end{aligned}$$

The mean square estimation error is given by

$$\begin{aligned}E[\boldsymbol{\varepsilon}_h^T[n]\boldsymbol{\varepsilon}_h[n]] &= E\left[\left(\frac{1}{n(a_1m_2 + a_3m_4 + a_5m_6)} \sum_{i=1}^n \mathbf{x}[i]v[i]\right)^2\right] \\ &= \frac{1}{n^2(a_1m_2 + a_3m_4 + a_5m_6)^2} (Mm_2\sigma_v^2 + Mm_2\sigma_v^2 + \dots + Mm_2\sigma_v^2)\end{aligned}$$

$$= \frac{1}{n(a_1 m_2 + a_3 m_4 + a_5 m_6)^2} M m_2 \sigma_v^2. \quad (3.4.1)$$

The steady state of (3.4.1) is equal to zero as  $n \rightarrow \infty$ . It means the estimation has no excess mean square error resulting from the background noise. Next, we assume that the noise is not present and we focus on the sample mean estimation error, which is equal to the variance of  $\hat{\mathbf{R}}_{xd}[n]$ . When the variance of  $\hat{\mathbf{R}}_{xd}[n]$  is equal to zero an unbiased estimation of the  $\hat{\mathbf{h}}[n]$  can be obtained. Hence, the mean square estimation error of  $\hat{\mathbf{h}}[n]$  can be written as follows

$$\begin{aligned} E[\boldsymbol{\varepsilon}_h^T[n]\boldsymbol{\varepsilon}_h[n]] &= E\left[\left(\frac{1}{n(a_1 m_2 + a_3 m_4 + a_5 m_6)} \sum_{i=1}^n \mathbf{x}[i]d[i]\right)^2\right] \\ &= \frac{(E[\mathbf{x}[1]d[1]]^2 + E[\mathbf{x}[2]d[2]]^2 + \dots + E[\mathbf{x}[n]d[n]]^2)}{n^2(a_1 m_2 + a_3 m_4 + a_5 m_6)^2} \\ &= \frac{1}{n^2(a_1 m_2 + a_3 m_4 + a_5 m_6)^2} n \sigma_{xy}^2 \|\mathbf{h}\|_2^2 \end{aligned} \quad (3.4.2)$$

where

$$\begin{aligned} E[(\mathbf{x}[n]d[n])^2] &= E[(a_1 \mathbf{x}[n]\mathbf{x}^T[n] + a_2 \mathbf{x}[n]\mathbf{x}^{2T}[n] + a_3 \mathbf{x}[n]\mathbf{x}^{3T}[n] \\ &\quad + a_4 \mathbf{x}[n]\mathbf{x}^{4T}[n] + a_5 \mathbf{x}[n]\mathbf{x}^{5T}[n])\mathbf{h})^2]. \end{aligned} \quad (3.4.3)$$

Usually, the characteristic of nonlinear distortion of a power amplifier can be modeled as a saturation curve. Therefore, we assume that the nonlinear coefficients  $a_2$  and  $a_4$  are both much less than  $a_1$ . For this reason, we rewrite (3.4.3) as

$$\begin{aligned} E[(\mathbf{x}[n]d[n])^2] &= \mathbf{h}^T \{ E[a_1^2 \mathbf{x}[n]\mathbf{x}^T[n]\mathbf{x}[n]\mathbf{x}^T[n]] + E[a_3^2 \mathbf{x}[n]\mathbf{x}^{3T}[n]\mathbf{x}[n]\mathbf{x}^{3T}[n]] \\ &\quad + E[a_5^2 \mathbf{x}[n]\mathbf{x}^{5T}[n]\mathbf{x}[n]\mathbf{x}^{5T}[n]] + 2E[a_1 a_3 \mathbf{x}[n]\mathbf{x}^T[n]\mathbf{x}[n]\mathbf{x}^{3T}[n]] \\ &\quad + 2E[a_1 a_5 \mathbf{x}[n]\mathbf{x}^T[n]\mathbf{x}[n]\mathbf{x}^{5T}[n]] + 2E[a_3 a_5 \mathbf{x}[n]\mathbf{x}^{3T}[n]\mathbf{x}[n]\mathbf{x}^{5T}[n]] \} \mathbf{h} \\ &= \mathbf{h}^T \{ a_1^2 (m_4 + (M-1)m_2^2) \mathbf{I} + 2a_1 a_3 (m_6 + (M-1)m_2 m_4) \mathbf{I} \\ &\quad + a_3^2 (m_8 + (M-1)m_4^2) \mathbf{I} + 2a_1 a_5 (m_8 + (M-1)m_2 m_6) \mathbf{I} \\ &\quad + a_5^2 (m_{12} + (M-1)m_6^2) \mathbf{I} + 2a_3 a_5 (m_{10} + (M-1)m_4 m_6) \mathbf{I} \} \mathbf{h} \end{aligned}$$

$$= \sigma_{xy}'^2 \|\mathbf{h}\|_2^2$$

where

$$\begin{aligned} \sigma_{xy}'^2 &= (M-1)(a_1 m_2 + a_3 m_4 + a_5 m_6)^2 + a_1^2 m_4 + a_3^2 m_8 + a_5^2 m_{12} \\ &\quad + 2a_1 a_3 m_6 + 2a_1 a_5 m_8 + 2a_3 a_5 m_{10} \\ &\simeq M(a_1 m_2 + a_3 m_4 + a_5 m_6)^2 \\ &\triangleq \sigma_{xy}^2. \end{aligned} \tag{3.4.4}$$

According to (3.4.4), (3.4.2) becomes

$$\begin{aligned} E[\boldsymbol{\varepsilon}_h^T[n] \boldsymbol{\varepsilon}_h[n]] &= \frac{1}{n(a_1 m_2 + a_3 m_4 + a_5 m_6)^2} \sigma_{xy}^2 \|\mathbf{h}\|_2^2 \\ &\simeq \frac{M}{n} \|\mathbf{h}\|_2^2. \end{aligned} \tag{3.4.5}$$

By the law of superposition, we can sum up (3.4.1) and (3.4.5) to get the mean square estimation error of linear coefficients

$$\begin{aligned} E[\boldsymbol{\varepsilon}_h^T[n] \boldsymbol{\varepsilon}_h[n]] &= \frac{M}{n} \|\mathbf{h}\|_2^2 + \frac{M m_2 \sigma_v^2}{n(a_1 m_2 + a_3 m_4 + a_5 m_6)^2} \\ &= \frac{M}{n} \left[ \|\mathbf{h}\|_2^2 + \frac{m_2 \sigma_v^2}{(a_1 m_2 + a_3 m_4 + a_5 m_6)^2} \right] \\ &= \frac{M}{n} \left[ 1 + \frac{m_2 \sigma_v^2}{(a_1 m_2 + a_3 m_4 + a_5 m_6)^2} \right]. \end{aligned} \tag{3.4.6}$$

Once the noise power, the nonlinear coefficients, and the length of the linear filter are known, we can get the convergence rate of linear coefficients. It is different from the NLMS-type adaptation, the effect of the noise will disappear asymptotically as  $n \rightarrow \infty$ . According to (3.4.6), there is no excess mean square estimation error because the coefficients is found directly by the Wiener-Hopf equations, which is the optimum solution for such systems. For a NLMS algorithm, (2.3.3) will never achieve

the optimum solution even when  $n \rightarrow \infty$  unless there is no background noise. But in the general, a background noise always exists in hands-free environment.

### 3.4.2 Residual echo power analysis

In this section we focus on the residual echo power, with similar analysis to that in Section 2.3.2. Therefore, the residual echo power can be written as

$$\begin{aligned}
J_h[n] &= E\left[|e[n]|^2\right] \\
&= E[(v[n] + \boldsymbol{\varepsilon}_h^T[n]\mathbf{s}[n])(v[n] + \mathbf{s}[n]^T \boldsymbol{\varepsilon}_h[n])] \\
&= \sigma_v^2 + E\left[\boldsymbol{\varepsilon}_h^T[n]\mathbf{s}[n]\mathbf{s}[n]^T \boldsymbol{\varepsilon}_h[n]\right] \\
&\cong \sigma_v^2 + E\left[\boldsymbol{\varepsilon}_h^T[n]E[\mathbf{s}[n]\mathbf{s}[n]^T] \boldsymbol{\varepsilon}_h[n]\right] \\
&= \sigma_v^2 + E\left[\boldsymbol{\varepsilon}_h^T[n]\mathbf{R}_{s[n]} \boldsymbol{\varepsilon}_h[n]\right]. \tag{3.4.7}
\end{aligned}$$

The non-diagonal term of  $\mathbf{R}_{s[n]}$  can be ignored under the assumption that even-ordered nonlinear coefficients are much less than those of the odd-ordered. The diagonal terms of  $\mathbf{R}_{s[n]}$  are given by

$$\begin{aligned}
&E[(a_1x[n] + a_2x^2[n] + a_3x^3[n] + a_4x^4[n] + a_5x^5[n])^2] \\
&\cong a_1^2m_2 + a_3^2m_6 + a_5^2m_{10} + 2a_1a_3m_4 + 2a_1a_5m_6 + 2a_3a_5m_8 \\
&= \sigma_{s[n]}^2.
\end{aligned}$$

Thus  $\mathbf{R}_{s[n]}$  is approximate to a diagonal matrix  $\sigma_{s[n]}^2 \mathbf{I}_{M \times M}$ . Using (3.4.6) and  $\mathbf{R}_{s[n]}$ , (3.4.7) can be written as

$$J_h[n] = \sigma_v^2 + \sigma_{s[n]}^2 \frac{M}{n} \left[ 1 + \frac{m_2 \sigma_v^2}{(a_1 m_2 + a_3 m_4 + a_5 m_6)^2} \right]. \tag{3.4.8}$$



### 3.5 Nonlinear convergence analysis

Up to now, we have introduced the analyses of linear coefficients convergence, which shows that the convergence rate depends on the estimated error of  $\widehat{\mathbf{R}}_{xd}[n]$ . Similarly, the nonlinear coefficients convergence depends not only on  $\widehat{\mathbf{R}}_{xd}[n]$  but also  $\widehat{\mathbf{R}}_{x^3d}[n]$  and  $\widehat{\mathbf{R}}_{x^5d}[n]$  as shown in (3.2.10). For each nonlinear coefficient, its convergence rate depends on all the cross correlation vectors ( $\widehat{\mathbf{R}}_{xd}[n], \widehat{\mathbf{R}}_{x^3d}[n], \widehat{\mathbf{R}}_{x^5d}[n]$ ) because a moment matrix is performed on the nonlinear coefficients vector, therefore it has more complex mathematical procedure than the analysis of linear coefficients.

#### 3.5.1 Variance of nonlinear coefficients error

First, we define the covariance matrix of  $\boldsymbol{\varepsilon}_a[n]$  as

$$\begin{aligned}
 E[\boldsymbol{\varepsilon}_a[n]\boldsymbol{\varepsilon}_a^T[n]] &= \begin{bmatrix} E[\varepsilon_{a,1}^2[n]] & E[\varepsilon_{a,1}[n]\varepsilon_{a,3}[n]] & E[\varepsilon_{a,1}[n]\varepsilon_{a,5}[n]] \\ E[\varepsilon_{a,1}[n]\varepsilon_{a,3}[n]] & E[\varepsilon_{a,3}^2[n]] & E[\varepsilon_{a,3}[n]\varepsilon_{a,5}[n]] \\ E[\varepsilon_{a,1}[n]\varepsilon_{a,5}[n]] & E[\varepsilon_{a,3}[n]\varepsilon_{a,5}[n]] & E[\varepsilon_{a,5}^2[n]] \end{bmatrix} \\
 &= \mathbf{G}_{odd}^{-1} \begin{bmatrix} \mathbf{h}^T \widehat{\mathbf{R}}_{xd} \mathbf{h}^T \widehat{\mathbf{R}}_{xd} & \mathbf{h}^T \widehat{\mathbf{R}}_{xd} \mathbf{h}^T \widehat{\mathbf{R}}_{x^3d} & \mathbf{h}^T \widehat{\mathbf{R}}_{xd} \mathbf{h}^T \widehat{\mathbf{R}}_{x^5d} \\ \mathbf{h}^T \widehat{\mathbf{R}}_{x^3d} \mathbf{h}^T \widehat{\mathbf{R}}_{xd} & \mathbf{h}^T \widehat{\mathbf{R}}_{x^3d} \mathbf{h}^T \widehat{\mathbf{R}}_{x^3d} & \mathbf{h}^T \widehat{\mathbf{R}}_{x^3d} \mathbf{h}^T \widehat{\mathbf{R}}_{x^5d} \\ \mathbf{h}^T \widehat{\mathbf{R}}_{x^5d} \mathbf{h}^T \widehat{\mathbf{R}}_{xd} & \mathbf{h}^T \widehat{\mathbf{R}}_{x^5d} \mathbf{h}^T \widehat{\mathbf{R}}_{x^3d} & \mathbf{h}^T \widehat{\mathbf{R}}_{x^5d} \mathbf{h}^T \widehat{\mathbf{R}}_{x^5d} \end{bmatrix} \mathbf{G}_{odd}^{-T}
 \end{aligned}$$

Again, for the same reason in Section 3.4, we divide the analysis into two parts by the law of superposition, one is *sample mean estimation* (SME) error and the other is the *noise effect* (NE). Therefore, we can rewrite  $E[\boldsymbol{\varepsilon}_a[n]\boldsymbol{\varepsilon}_a^T[n]]$  as

$$E[\boldsymbol{\varepsilon}_a[n]\boldsymbol{\varepsilon}_a^T[n]] = \mathbf{G}_{odd}^{-1} [\mathbf{SME} + \mathbf{NE}] \mathbf{G}_{odd}^{-T} \quad (3.5.1)$$

In order to keep the main object of this section, the detailed mathematical procedures about (3.5.1) will be given in Appendix. Thus, the variance of nonlinear coefficients

error can be found as

$$tr(E[\boldsymbol{\varepsilon}_a[n]\boldsymbol{\varepsilon}_a^T[n]]) = tr(\mathbf{G}_{odd}^{-1}[\mathbf{SME} + \mathbf{NE}]\mathbf{G}_{odd}^{-T})$$

where  $tr(\cdot)$  denotes the trace operator. The same result is found that the nonlinear coefficients error variance can also achieve zero, the optimum nonlinear coefficients can be found as  $n \rightarrow \infty$ .

### 3.5.2 Residual echo power analysis

Once the nonlinear coefficients convergence is given, its residual echo power can also be found as

$$\begin{aligned} J_a[n] &= E\left[|e[n]|^2\right] \\ &= E[(v[n] + \boldsymbol{\varepsilon}_a^T[n]\mathbf{P}^T[n]\mathbf{h})(v[n] + \mathbf{h}^T\mathbf{P}[n]\boldsymbol{\varepsilon}_a[n])] \\ &\cong \sigma_v^2 + E[\boldsymbol{\varepsilon}_a^T E[\mathbf{P}^T[n]\mathbf{h}\mathbf{h}^T\mathbf{P}[n]]\boldsymbol{\varepsilon}_a[n]] \\ &= \sigma_v^2 + E\left[\boldsymbol{\varepsilon}_a^T \begin{bmatrix} m_2 & m_4 & m_6 \\ m_4 & m_6 & m_8 \\ m_6 & m_8 & m_{10} \end{bmatrix} \boldsymbol{\varepsilon}_a[n]\right] \\ &= \sigma_v^2 + m_2 E[\varepsilon_{a,1}^2[n]] + m_6 E[\varepsilon_{a,3}^2[n]] + m_{10} E[\varepsilon_{a,5}^2[n]] \\ &\quad + 2(m_4 E[\varepsilon_{a,1}[n]\varepsilon_{a,3}[n]] + m_6 E[\varepsilon_{a,1}[n]\varepsilon_{a,5}[n]] \\ &\quad + m_8 E[\varepsilon_{a,3}[n]\varepsilon_{a,5}[n]]) \end{aligned} \quad (3.5.2)$$

The covariance terms of nonlinear coefficients in (3.5.2) can be obtained from (3.5.1). In addition to the background noise, there is no residual echo power when the nonlinear coefficients error is equal to zero.

### 3.6 TS for dual loudspeakers system

We have introduced the dual loudspeakers system in Section 2.5 as shown in Fig. 2.3. In this section we will use the white TS to train the linear and nonlinear coefficients of two nonlinear AEC's. If we use two uncorrelated white noises to train the coefficients, (3.4.8) can be used here with some modifications. From the nonlinear AEC1 (as showed in Fig. 2.3) viewpoint, the desired signal is

$$d[n] = \mathbf{a}^T \mathbf{P}_{x_L}^T[n] \mathbf{h}_L + (v[n] + \mathbf{b}^T \mathbf{P}_{x_R}^T[n] \mathbf{h}_R). \quad (3.6.1)$$

Both  $v[n]$  and  $\mathbf{b}^T \mathbf{P}_{x_R}^T[n] \mathbf{h}_R$  in the right-hand side of (3.6.1) are viewed as noises.

Therefore, the power of noise can be expressed as

$$\sigma_{noise}^2 = \sigma_v^2 + \sigma_{b^T P_{x_R}^T[n] h_R}^2.$$

From (3.4.8), the residual echo power of AEC1 can be expressed as

$$J_{h-AEC1}[n] = (\sigma_v^2 + \sigma_{b^T P_{x_R}^T[n] h_R}^2) + \sigma_{s[n]}^2 \frac{M}{n} \left[ 1 + \frac{m_2 (\sigma_v^2 + \sigma_{b^T P_{x_R}^T[n] h_R}^2)}{(a_1 m_2 + a_3 m_4 + a_5 m_6)^2} \right]. \quad (3.6.2)$$

Similarly, the residual echo power of AEC2 is

$$J_{h-AEC2}[n] = (\sigma_v^2 + \sigma_{a^T P_{x_L}^T[n] h_L}^2) + \sigma_{s[n]}^2 \frac{M}{n} \left[ 1 + \frac{m_2 (\sigma_v^2 + \sigma_{a^T P_{x_L}^T[n] h_L}^2)}{(a_1 m_2 + a_3 m_4 + a_5 m_6)^2} \right]. \quad (3.6.3)$$

The total residual echo power is given by

$$\begin{aligned} J_h[n] &= E \left[ |e[n]|^2 \right] \\ &= J_{h-AEC1}[n] + J_{h-AEC2}[n] - (\sigma_v^2 + \sigma_{a^T P_{x_L}^T[n] h_L}^2 + \sigma_{b^T P_{x_R}^T[n] h_R}^2) \\ &= \sigma_v^2 + \sigma_{s[n]}^2 \frac{M}{n} \left[ 1 + \frac{m_2 (\sigma_v^2 + \sigma_{b^T P_{x_R}^T[n] h_R}^2)}{(a_1 m_2 + a_3 m_4 + a_5 m_6)^2} \right] \\ &\quad + \sigma_{s[n]}^2 \frac{M}{n} \left[ 1 + \frac{m_2 (\sigma_v^2 + \sigma_{a^T P_{x_L}^T[n] h_L}^2)}{(a_1 m_2 + a_3 m_4 + a_5 m_6)^2} \right]. \end{aligned} \quad (3.6.4)$$

When  $n \rightarrow \infty$ , the residual echo power only contains the background noise. Compared

with the single loudspeaker case, AEC with dual loudspeakers has slower convergence rate. The convergence rate for dual loudspeakers is similar to single loudspeaker with 0 dB SNR, if we use two white sequences of equal powers to train the coefficients.

### 3.6 Summary

In this chapter we proposed the coefficients estimation based on a TS. We showed that the nonlinear AEC with a cascade structure can also use this TS to find out both linear and nonlinear coefficients. Although the TS algorithm only performs under a *white* sequence input and can not use a step size to control its convergence rate but its non-error-feedback structure makes the coefficients achieve the optimum solution and more robust to the background noise. Compared with a popular and computational efficient NLMS algorithm, the training method has similar computational cost. If we modify the noise variance, the analysis for a single loudspeaker system can be readily extended to dual loudspeakers system.

# Chapter 4

## Computer Simulations

In this chapter we will show the simulation results of the nonlinear AEC to verify previous analyses in Chapter 2 and Chapter 3. In Section 4.1, we will define some parameters used in following simulations. In Section 4.2 and 4.3, we will compare simulations results and theoretical analyses. In Section 4.4, an experiment with real echo path is made.

### 4.1 Parameters of simulations

The room impulse response shown in Fig. 4.1 is measured in a typical office. In Fig. 4.1, the length of the room impulse is 1000, which is 125 msec long, at a 8 KHz sampling rate. For simplicity, we use a 128-tap room impulse response as shown in Fig.4.2. It is generated by a random number generator with an exponential damping factor. Fig. 4.2 is also used as the left room impulse response for dual loudspeakers simulation, and the right room impulse response is shown in Fig. 4.3.

The sigmoid function, commonly used in neural network community [6], is shown in Fig. 4.4. In this thesis a polynomial type nonlinear processor is used. The sigmoid function is

$$\varphi[x] = \left( \frac{2}{1 + \exp(-\alpha x)} - 1 \right) \beta .$$

The nonlinear coefficients are found by using power-series polynomials to fit the sigmoid curve. When  $\alpha = 6$  and  $\beta = 1$ , the nonlinear coefficients  $a_1 \sim a_5$  corresponding to  $p_1[n] \sim p_5[n]$  are given by

$$a_1 = 2.5967 \quad a_3 = -3.3283 \quad a_5 = 1.7833 \quad a_2 = a_4 = 0 .$$

Again, for dual loudspeakers system (right-loudspeaker) we find the other nonlinear

coefficients set by using the power series polynomials to fit the other sigmoid curve as show in Fig. 4.5. When  $\alpha = 3$  and  $\beta = 1$  the nonlinear coefficients for the right nonlinear loudspeaker are

$$a_1 = 1.4775 \quad a_3 = -0.8928 \quad a_5 = 0.3259 \quad a_2 = a_4 = 0$$

We also use the speech signal as the input signal to examine the performance. The speech signal in Fig 4.6 is sampled with 8 KHz sampling rate.

In simulation, we add a background noise. The signal to noise ration is defined as

$$SNR = 10 \log_{10} \frac{P_{echo}}{P_v}$$

where  $P_{echo}$  is the power of the input signal to the microphone output signal and  $P_v$  is the power of the background noise. Setting the SNR we can observe the AEC performance under different environment conditions. The linear coefficients misalignment is defined as the normalized norm of the coefficients error

$$\frac{\epsilon_h[n]}{\|\mathbf{h}\|_2} \triangleq \frac{\|\mathbf{h} - \hat{\mathbf{h}}[n]\|_2}{\|\mathbf{h}\|_2}$$

Finally, in the following simulations, unless otherwise stated, we let the step size  $\mu_h = 0.1$   $\mu_a = 0.25$ ,  $\delta = 1$ ; the length of the room impulse response is set to be 128, which is identical to the number of taps of the room impulse response; the highest nonlinear order is equal to 5<sup>th</sup> (excluding the even order) and SNR = 20 dB.

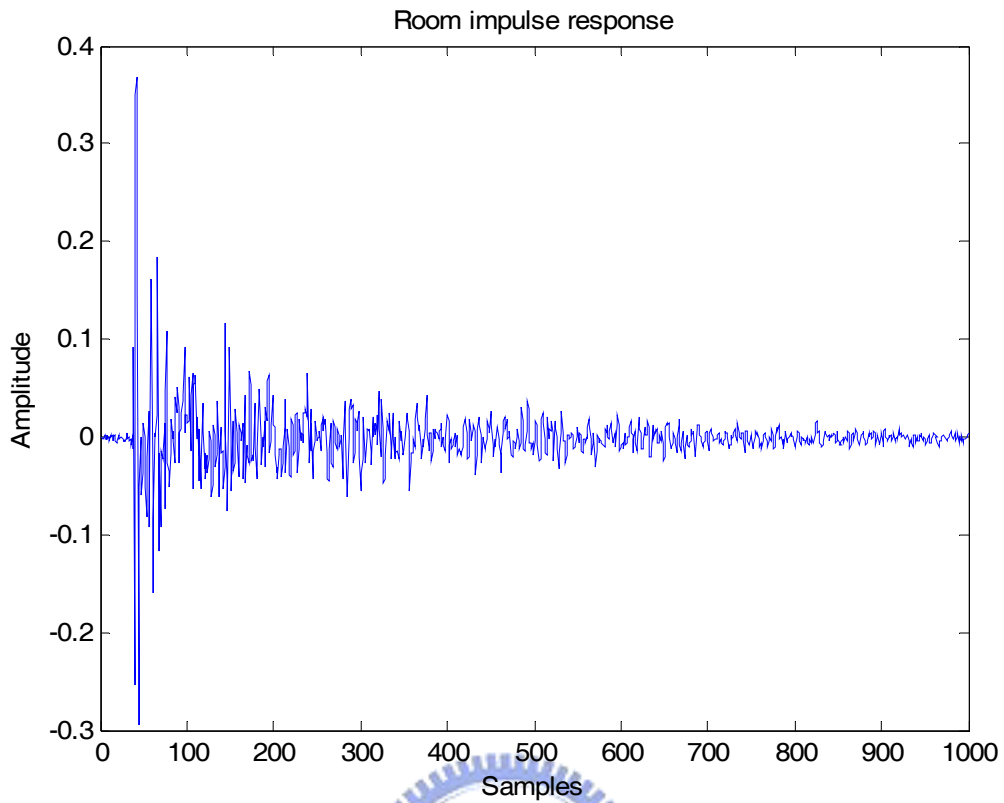


Fig. 4.1 Room impulse response of a typical office

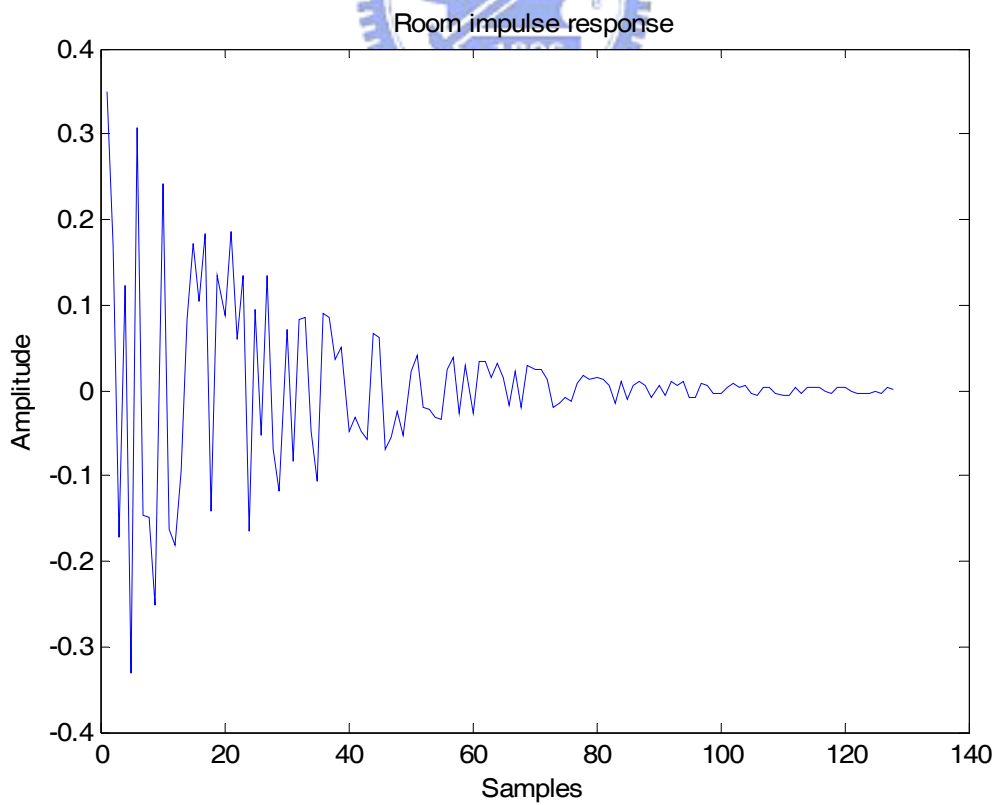


Fig. 4.2 Room impulse response from the left loudspeaker

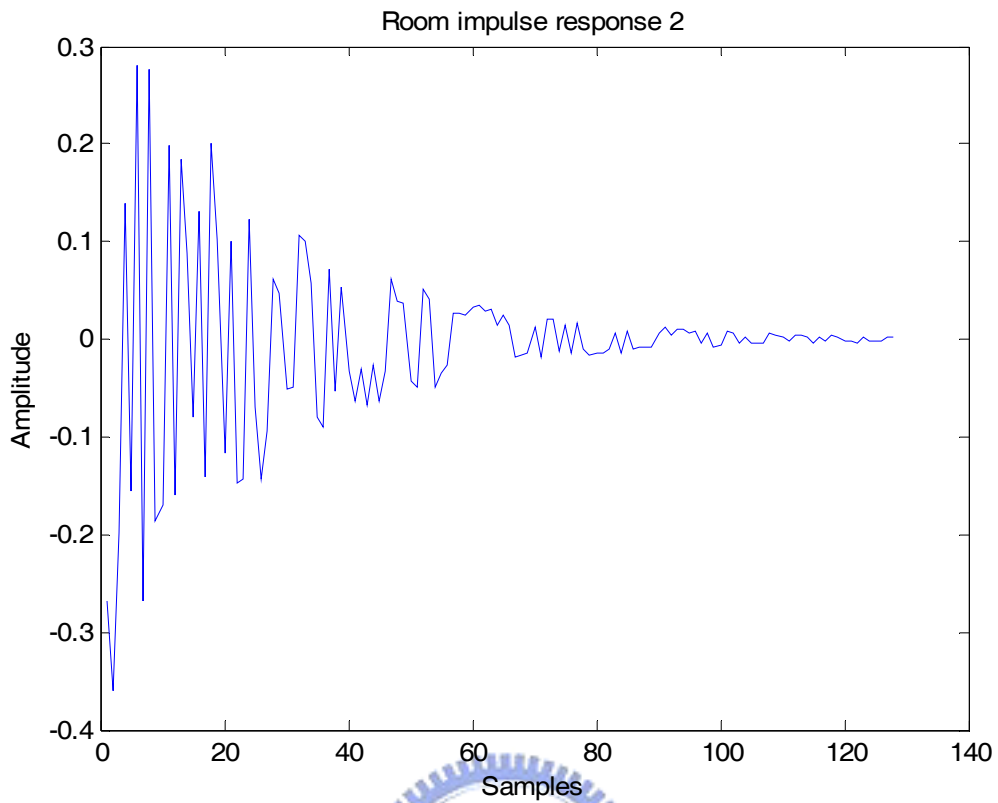


Fig. 4.3 Room impulse response from the right loudspeaker

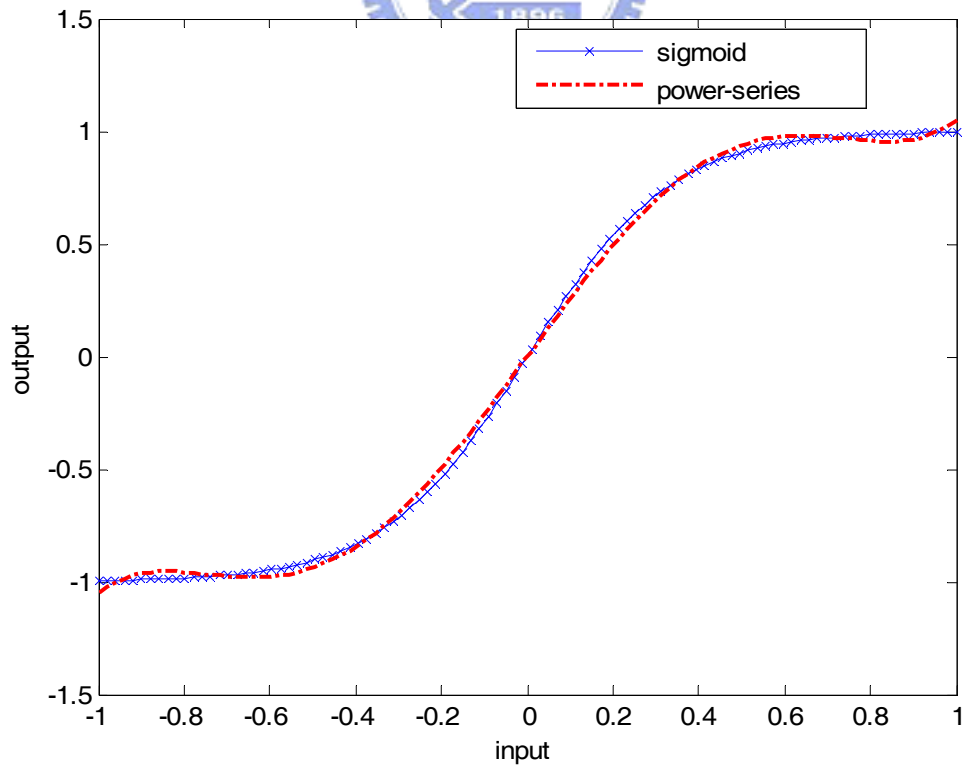


Fig. 4.4 I/O mapping characteristic of the left nonlinear loudspeaker



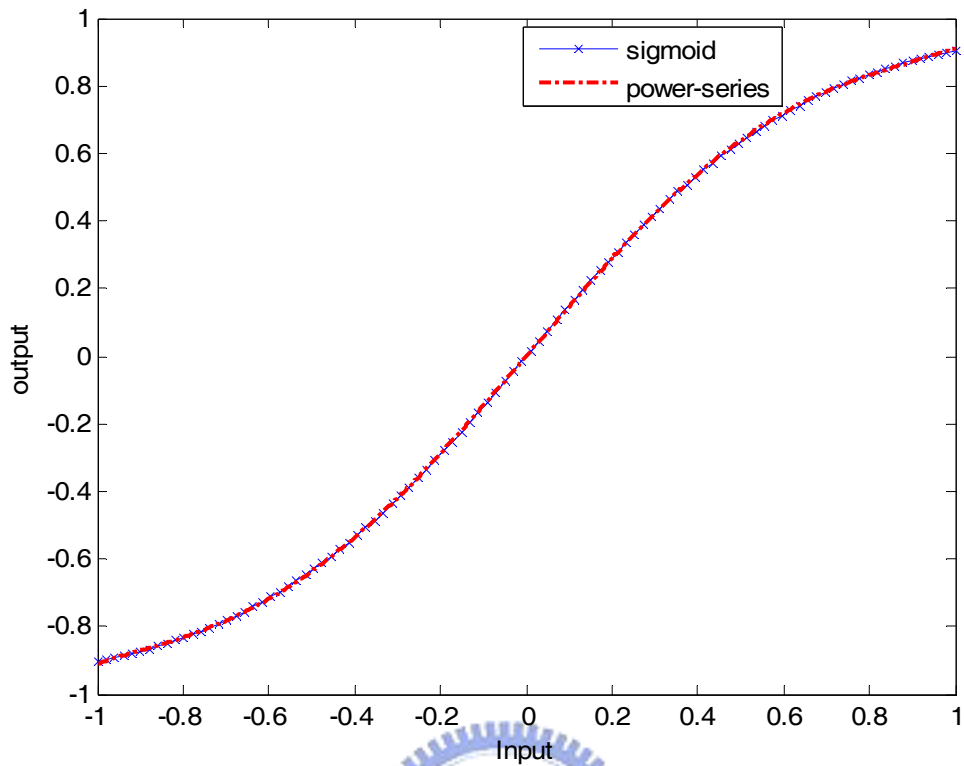


Fig. 4.5 I/O mapping characteristic of the right nonlinear loudspeaker

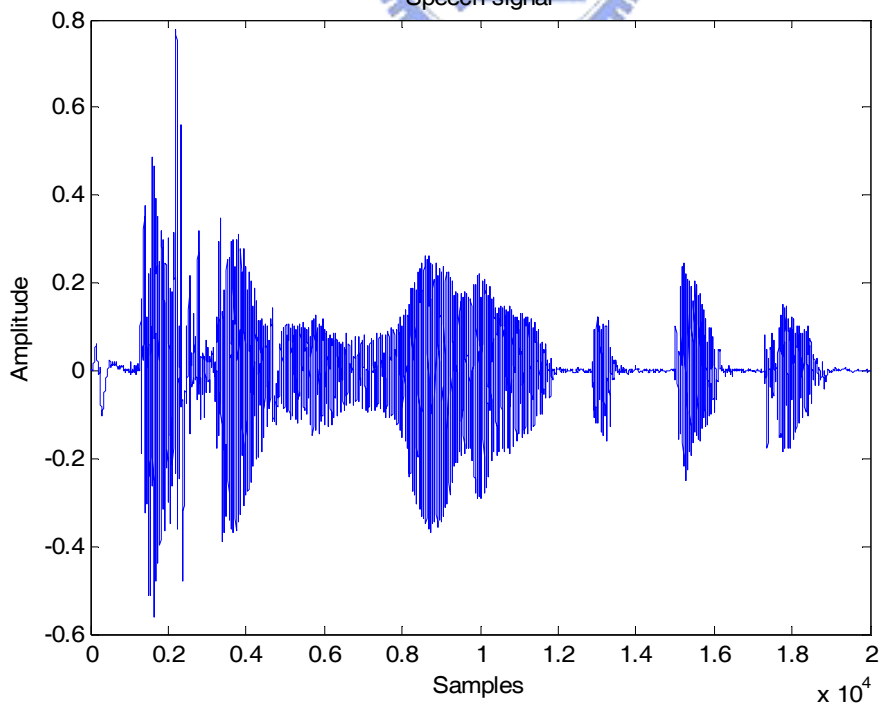


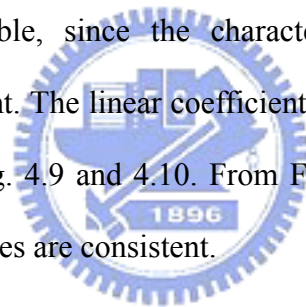
Fig. 4.6 The speech signal

## 4.2 Coefficients estimation based on NLMS adaptive algorithm

### 4.2.1 Individual coefficients and residual echo power convergence

This section corresponds to Section 2.2 and 2.3. We use the simulations to verify our convergent analyses in (2.2.6), (2.2.9), (2.3.3), and (2.3.5). We let the input signal be uniformly distributed over  $\pm 1$  and the basis is given in Table 2.3. First, we examine (2.2.6) and (2.2.9) by assuming that the linear coefficients are perfectly known. The assumption is reasonable, since linear coefficients can be found first when transmit the low power signal, nonlinearity of the power amplifier has not been excited yet. The nonlinear coefficients misalignments and residual echo powers are shown in Fig. 4.7 and 4.8. The simulation results match very well with the theoretical curve.

Next, we verify (2.3.3) and (2.3.5) with perfectly known nonlinear coefficients. This assumption is reasonable, since the characteristic of loudspeaker can be considered to be time invariant. The linear coefficients misalignments and its residual echo power are plotted in Fig. 4.9 and 4.10. From Fig. 4.7 to 4.10, we can see that simulated and theoretical curves are consistent.



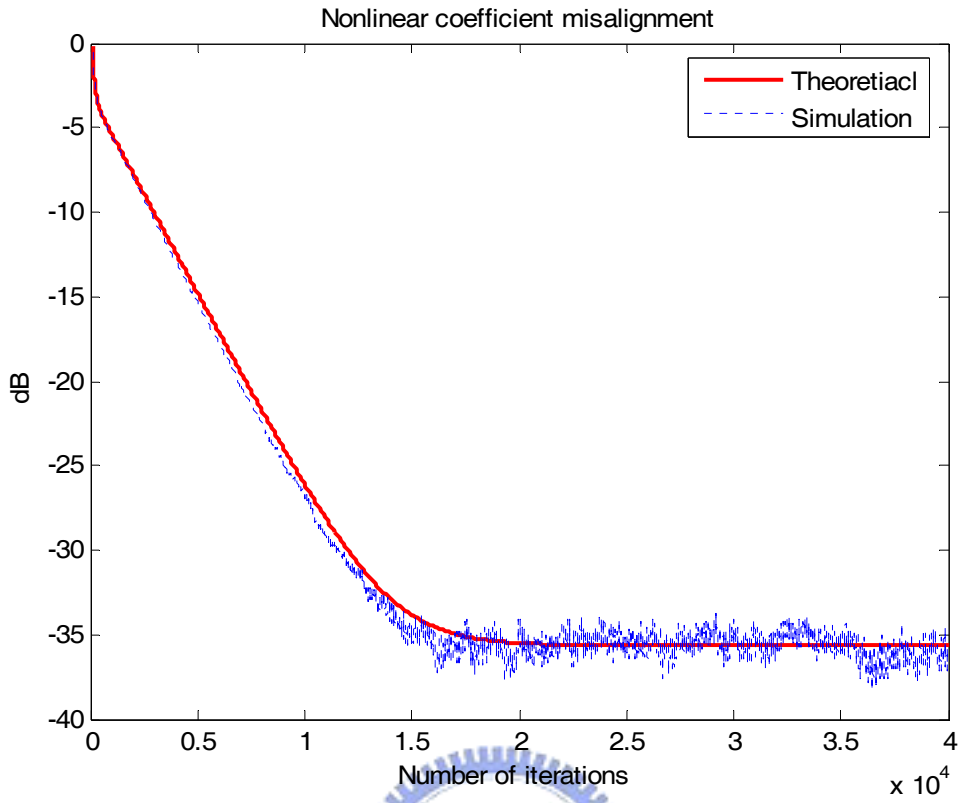


Fig. 4.7 Comparison of nonlinear coefficients misalignments (perfect  $\mathbf{h}$ )

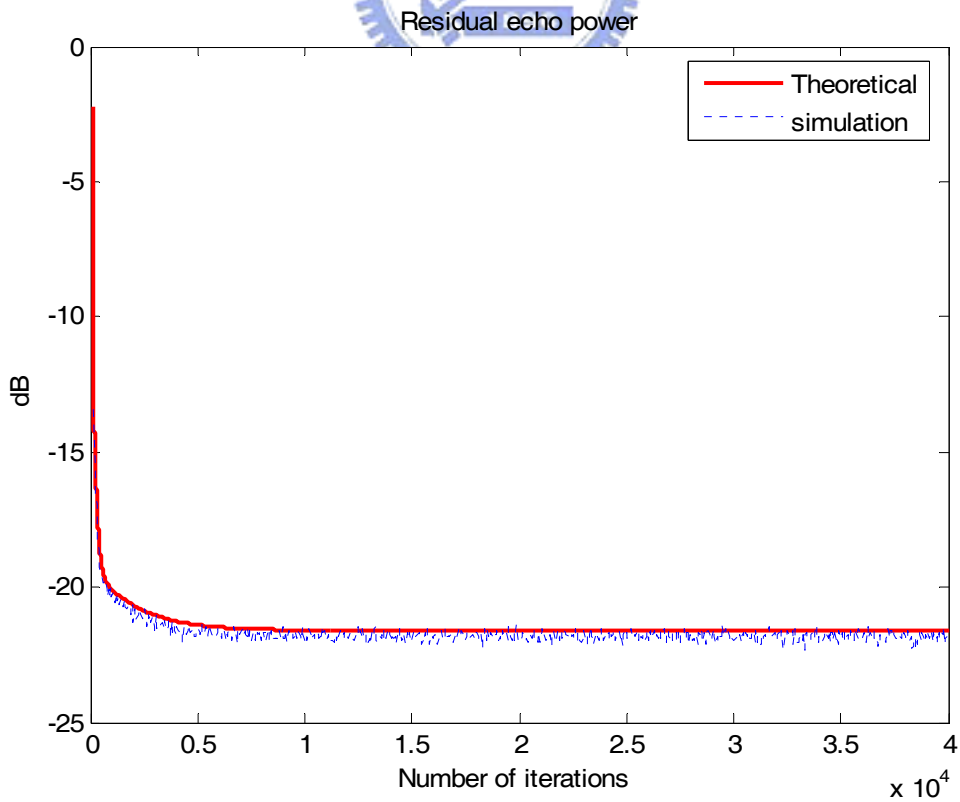


Fig. 4.8 Comparison of residual echo power (perfect  $\mathbf{h}$ )

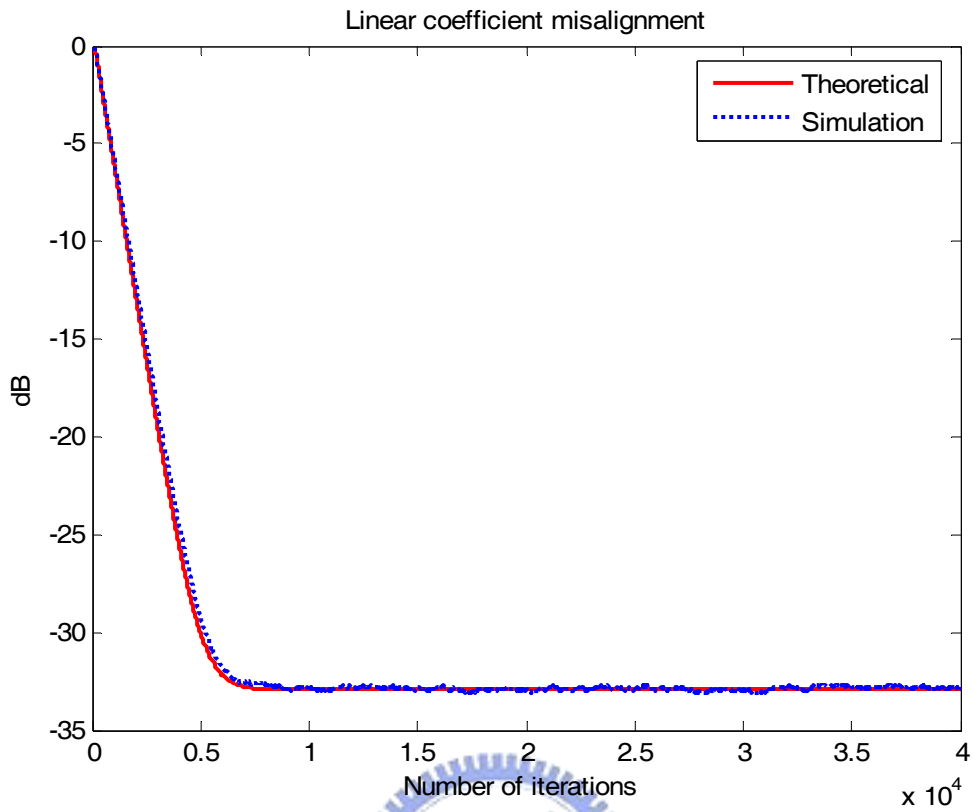


Fig. 4.9 Comparison of linear coefficients misalignments (perfect a)

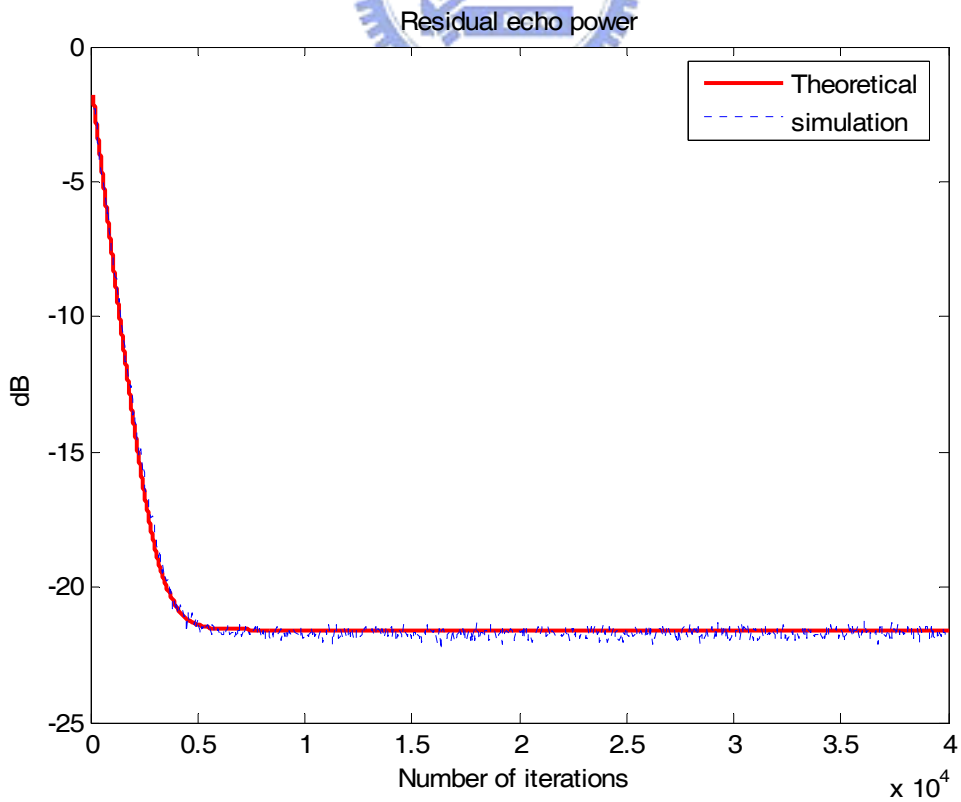


Fig. 4.10 Comparison of residual echo power (perfect a)

## 4.2.2 Convergence rate using orthogonal bases

We compare the residual echo convergence rates of the nonlinear processor using orthogonal and non-orthogonal basis as given in Section 2.4. Convergence rate improvement using the orthogonal basis is considered, and simulations are performed under perfect linear coefficients. The residual echo powers for a uniformly distributed input to 4 different polynomial bases nonlinear processors are shown in Fig. 4.11.

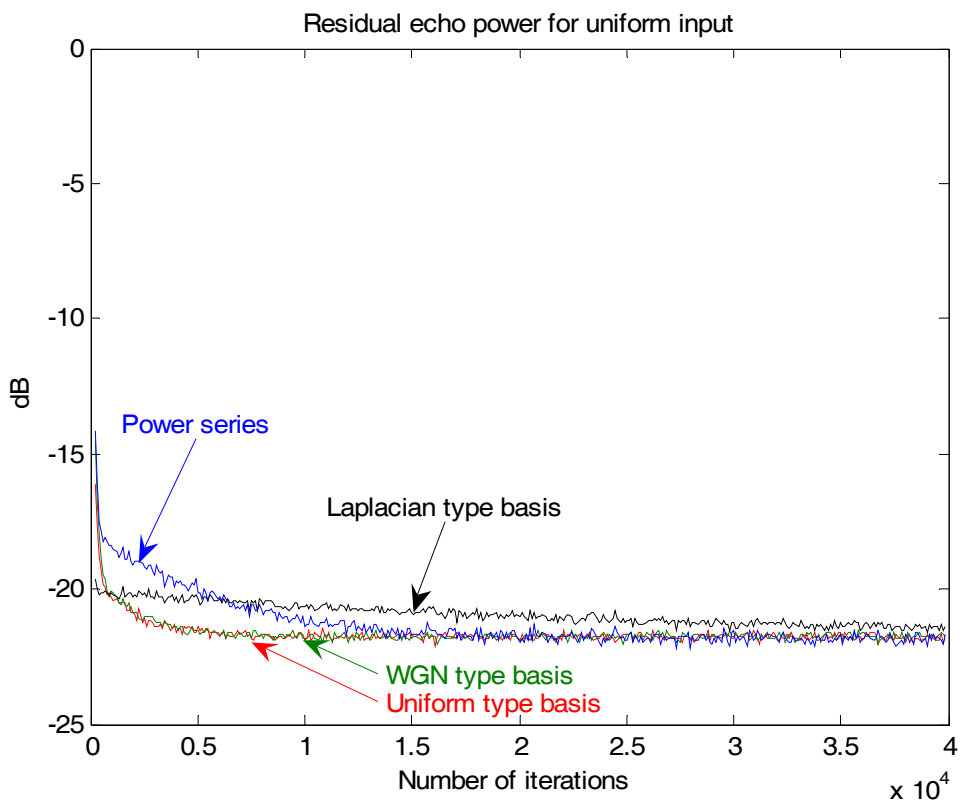


Fig. 4.11 Residual echo powers for uniform input (perfect  $\mathbf{h}$ )

As for a Gaussian and Laplacian input signal, the simulations are shown in Fig.4.12 and Fig. 4.13, respectively.

According to Fig. 4.11, 4.12 and 4.13, the AEC with orthogonal polynomials indeed converges faster due to its smaller eigenvalue spread. The simulation results agree well with the eigenvalue spread analyses in Section 2.4.

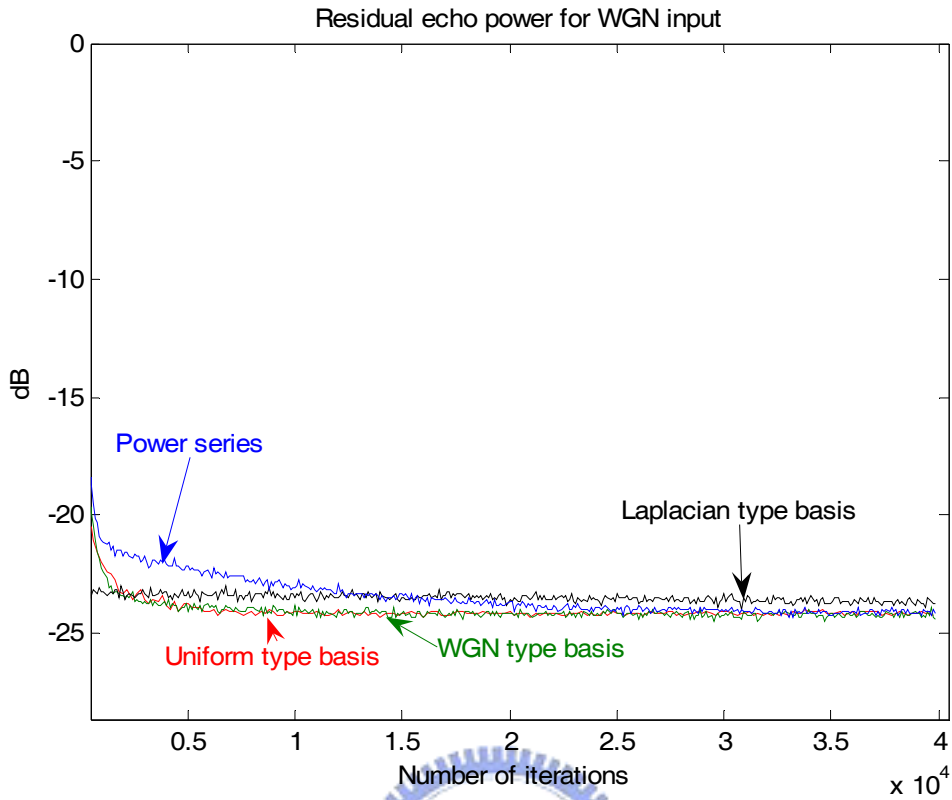


Fig. 4.12 Residual echo powers for WGN input (perfect  $\mathbf{h}$ )

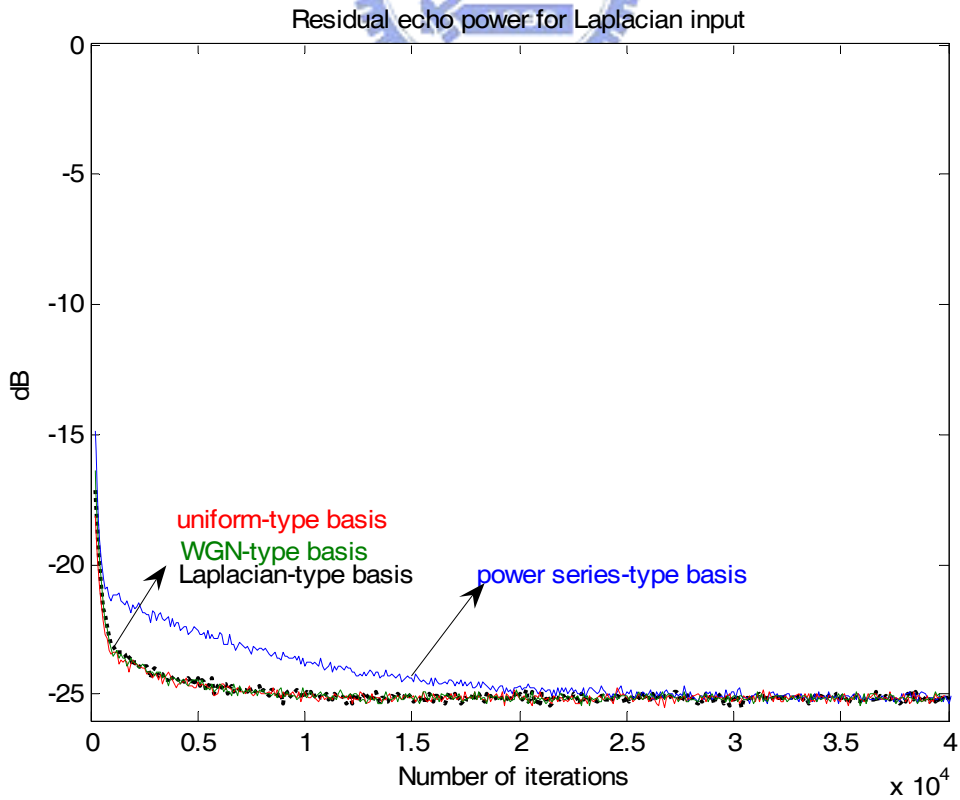
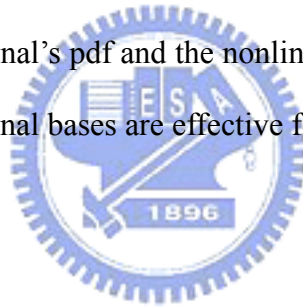


Fig. 4.13 Residual echo powers for Laplacian input (perfect  $\mathbf{h}$ )

At last, we observe the convergence rate when the nonlinear components do not have perfect orthogonality as discussed in Section 2.4.2. When the input signal is uniformly distributed, the convergence rates between four types of bases are shown in Fig. 4.11. In Table 2.6, the eigenvalue spreads are almost the same for uniform and Gaussian nonlinear polynomial bases; and the Laplacian one has the slowest convergence rate.

Fig. 4.12 and 4.13 show the convergence rates using four types of polynomial bases when the input signal is Gaussian and Laplacian, respectively. From Fig. 4.11 to 4.13, we can see that the orthogonal bases have faster convergence rate than non-orthogonal bases because of its reduced eigenvalue spread. The simulation results validate the theoretical analyses in Table 2.6. When the orthogonality is not perfectly matched between the input signal's pdf and the nonlinear orthogonal polynomial basis, uniform and Gaussian orthogonal bases are effective for either uniform, Gaussian, and Laplacian input signals.



### **4.2.3 Joint convergence rate using orthogonal bases**

The word joint means that the NLMS algorithm updates both linear and nonlinear coefficients without either assumption of perfect linear or nonlinear coefficients. The joint convergence rates of residual echo powers are plotted from Fig. 4.14 to Fig. 4.16.

The joint residual echo power for uniformly distributed input signal is shown in Fig. 4.14. The nonlinear AEC with orthogonal basis has faster convergence rate. For a Gaussian and Laplacian input signal, the simulation results are shown in Fig. 4.15 and 4.16, respectively.

Comparing Fig. 4.11 and Fig. 4.14, we can see the joint residual echo power has slower convergence because its linear coefficients are unknown. Although joint

NLMS update algorithm has a slower convergence rate, the proposed nonlinear orthogonal AEC improves the convergence among these simulations without increased computation cost.

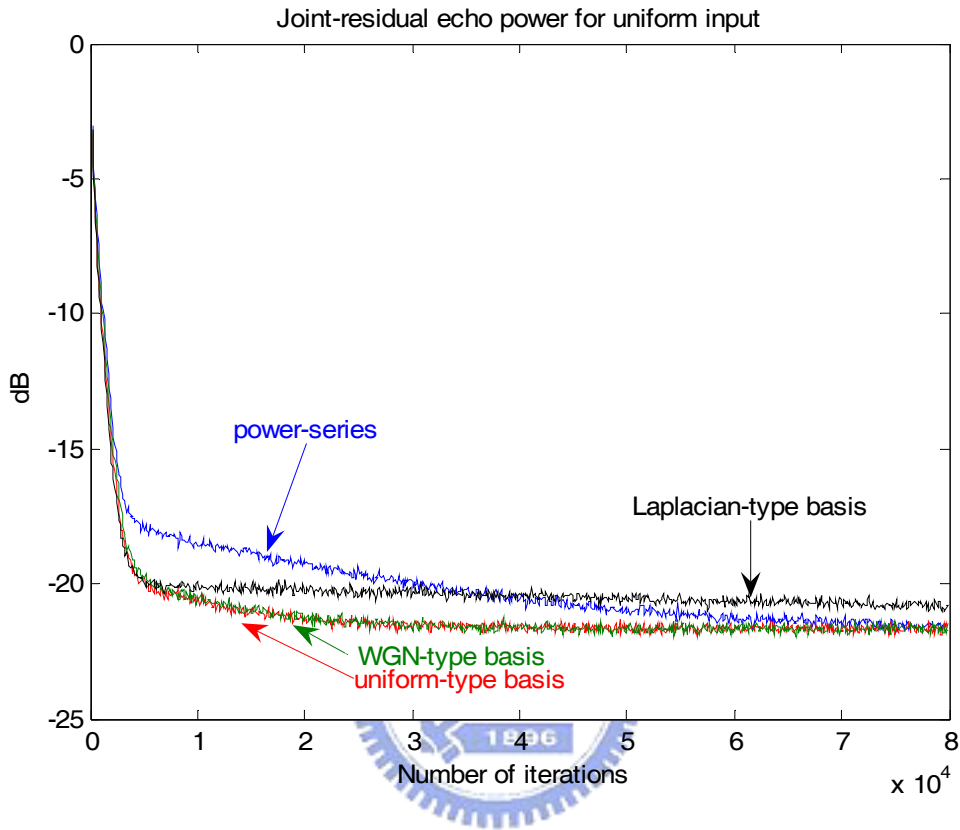


Fig. 4.14 Joint-residual echo power for uniform input



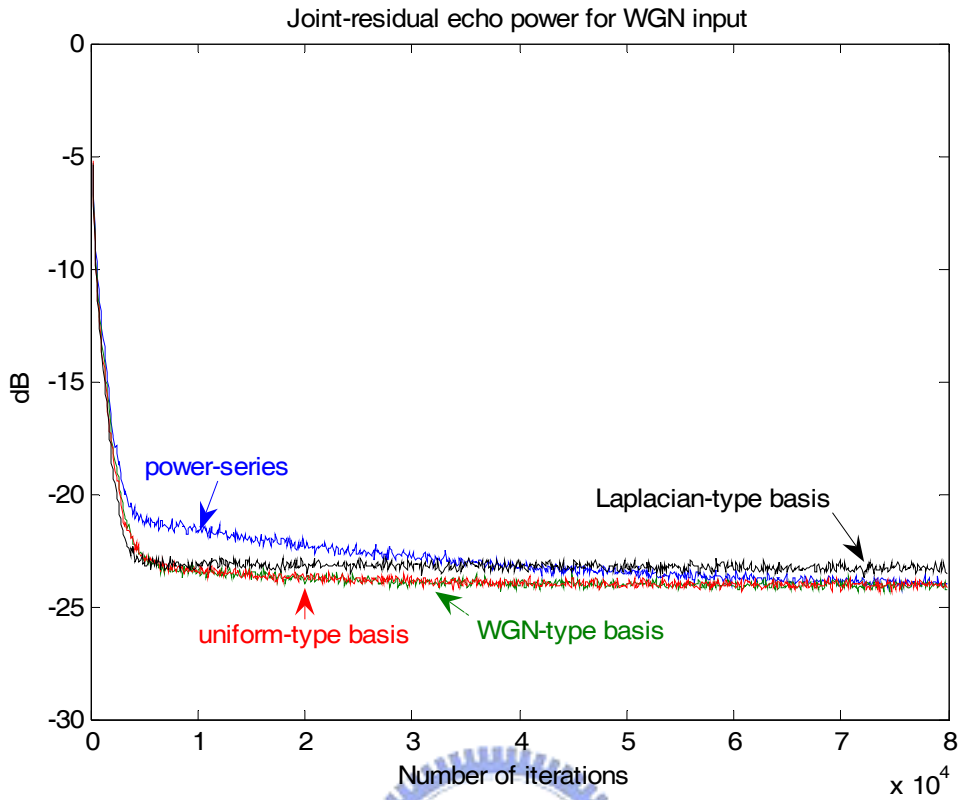


Fig. 4.15 Joint-residual echo power for Gaussian input

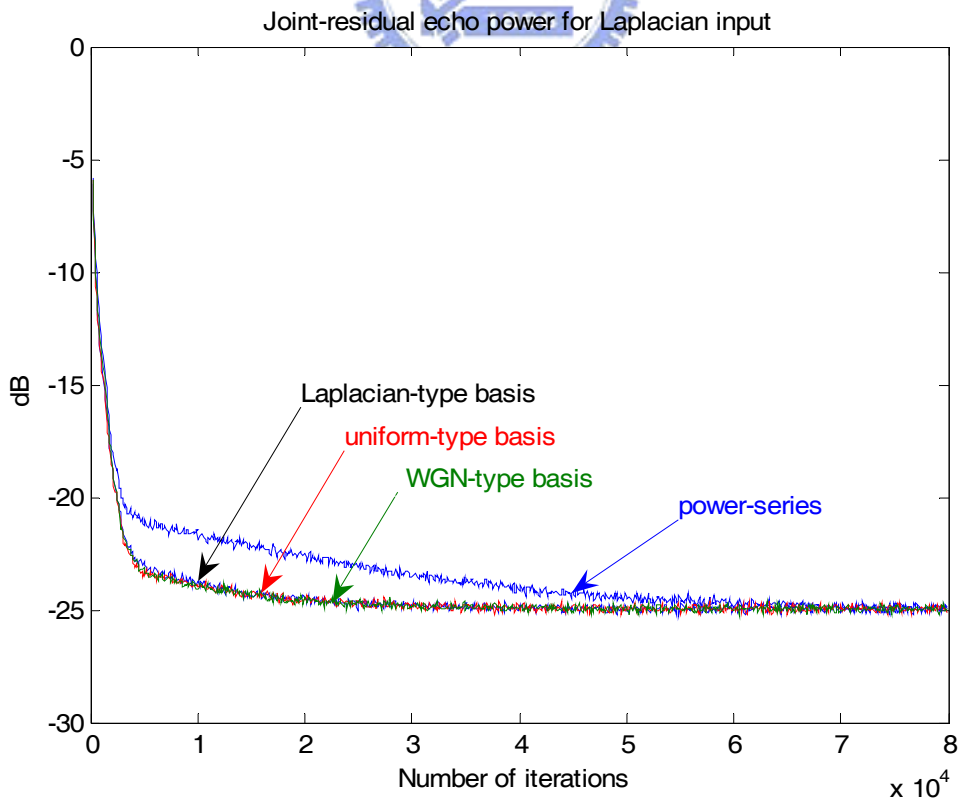


Fig. 4.16 Joint-residual echo power for Laplacian input

#### 4.2.4 Convergence rate of a speech input

In this section we compare the performance between orthogonal bases and power series polynomials when the input signal is a real speech. The performance is measured by echo return loss enhancement (ERLE), defined as

$$ERLE(dB) = 10 \log_{10} \frac{E[y^2[n]]}{E[e^2[n]]}$$

The comparisons of ERLE are shown in Fig. 4.17 when linear coefficients are error free. In Fig. 4.17, the Laplacian type orthogonal basis is used and it has better performance than power series even the input is a real speech signal.

In the case of joint adaptation, the average joint-ERLE using different orthogonal bases are given in Table 4.1. We can see that the Laplacian type orthogonal basis has slightly better performance than the others. This is well known fact a speech signal has a Laplacian probability distribution as shown in Fig. 4.18.

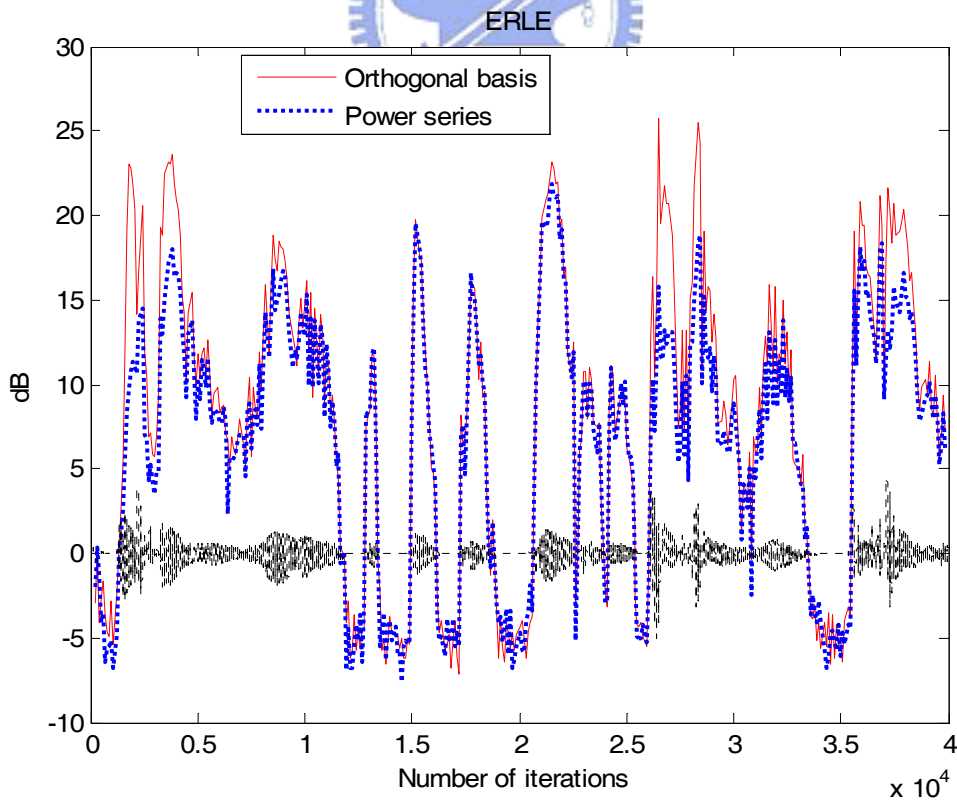


Fig. 4.17 ERLE for a true speech input signal with perfect linear coeff.

Table 4.1 Average joint-ERLE(dB) comparison for a speech input signal

AEC type	Linear	Nonlinear			
Nonlinear basis		Non orthogonal	Orthogonal		
		Power series	Uniform	WGN	Laplacian
Average ERLE	10.2	10.5	10.7	10.6	10.9

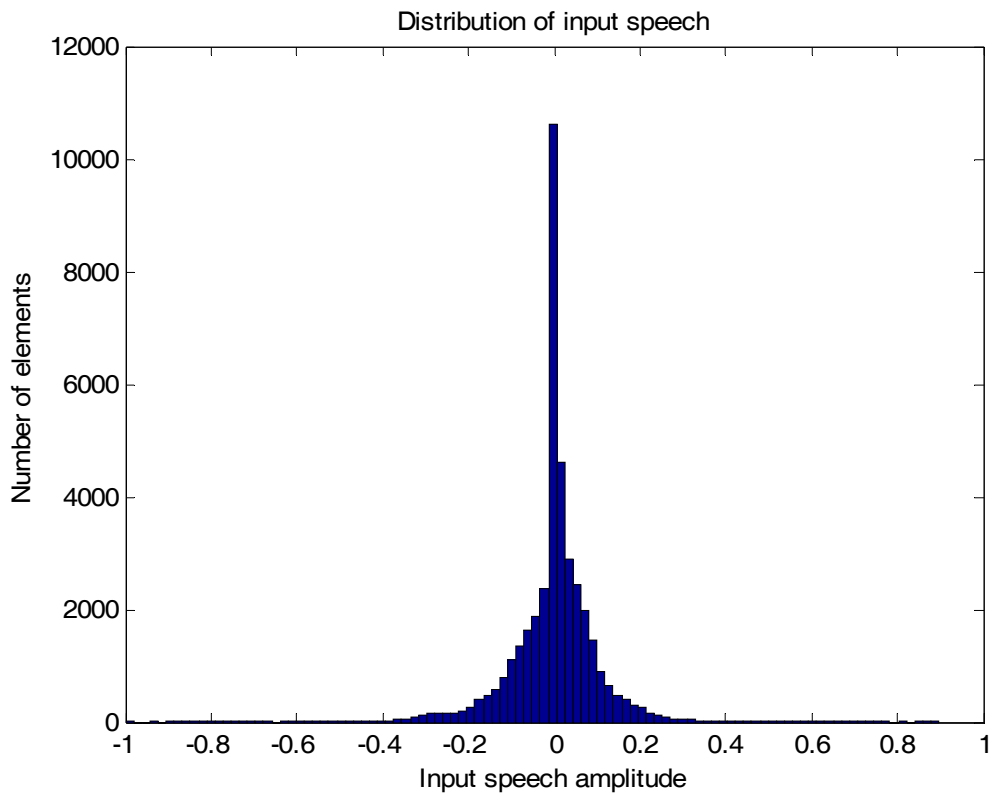


Fig. 4.18 Histogram of input speech

#### 4.2.5 Convergence rate of dual loudspeakers system

In this section we examine the theoretical analyses in Section 2.5 and compare its convergence rate with that of one loudspeaker system. Two echo paths corresponding to loudspeaker 1 and 2 are plotted in Fig. 4.2 and 4.3 respectively. The I/O nonlinear characteristics of two loudspeakers are given in Fig 4.4 and 4.5. The parameters of adaptive filter are the same as earlier.

First, we examine (2.5.4) in Section 2.5 with uniform input and perfect linear coefficients. The uniform type orthogonal basis is used for nonlinear processor. The simulation results of nonlinear coefficients and its residual echo power convergence rates are shown in Fig. 4.19 and 4.20. According to the simulations, we confirm that the convergence analysis in one loudspeaker can be easily extended to dual loudspeakers case.

Compared with Fig. 4.7 and 4.8, the system with dual loudspeakers has slower convergence rate and worse steady state. From the nonlinear AEC1 viewpoint, the microphone signal picked up from the 2<sup>nd</sup> echo path can be viewed as a background noise source. Therefore, each nonlinear AEC in dual loudspeakers system has more noise sources than single loudspeaker system; consequently, the convergence rate becomes slower.

Next, we examine the joint residual echo for orthogonal and non-orthogonal bases. As show in Fig. 4.21, the nonlinear processor with uniform orthogonal basis has faster convergence rate. According to Section 4.2, the proposed orthogonal basis can be used in either single or dual loudspeaker to enhance the convergence rate.

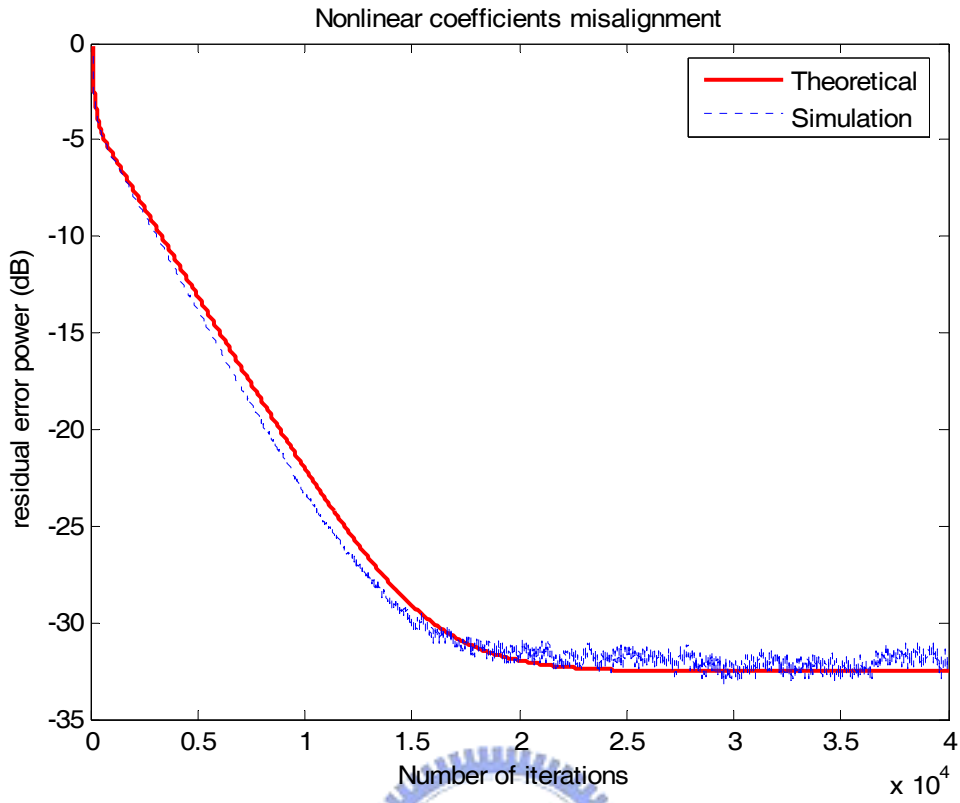


Fig. 4.19 Nonlinear coefficients misalignments for dual loudspeakers

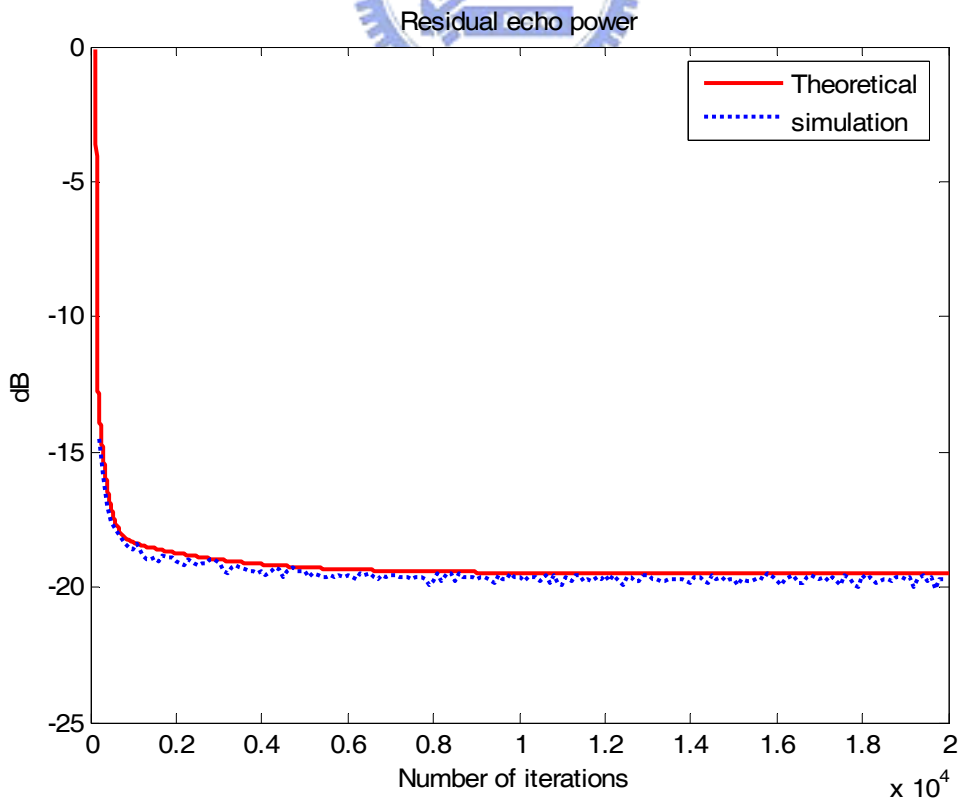


Fig. 4.20 Residual echo power for dual loudspeakers (perfect  $\mathbf{h}$ 's)

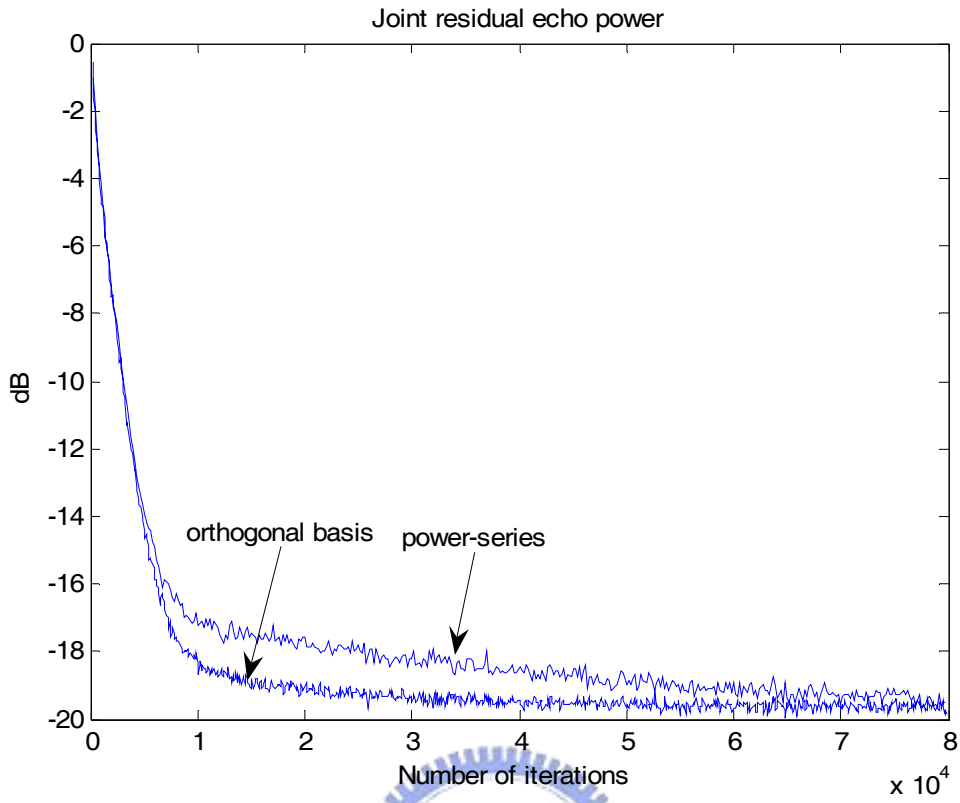


Fig. 4.21 Joint residual echo power for dual loudspeakers



### 4.3 TS-based coefficients estimation

In this section the coefficients are estimated by using training sequences. In Section 4.3.1, we examine the theoretical analyses. In Section 4.3.2, we show the simulation results in case of joint estimation. In Section 4.3.3, we compare the training method with NLMS adaptive filter. In Section 4.3.4, we use two TS's to train the coefficients for dual loudspeakers. In the following simulation the power series basis is used in the nonlinear processor. The orthogonal basis has similar result.

#### 4.3.1 Convergences of linear TS coefficients and residual TS echo power

In this section we use the white uniformly distributed sequence to train the linear coefficients. With 5 dB SNR and the assumption of perfectly known nonlinear coefficients, the linear coefficients convergence is shown in Fig. 4.22 where the theoretical curve is plotted from (3.4.6).

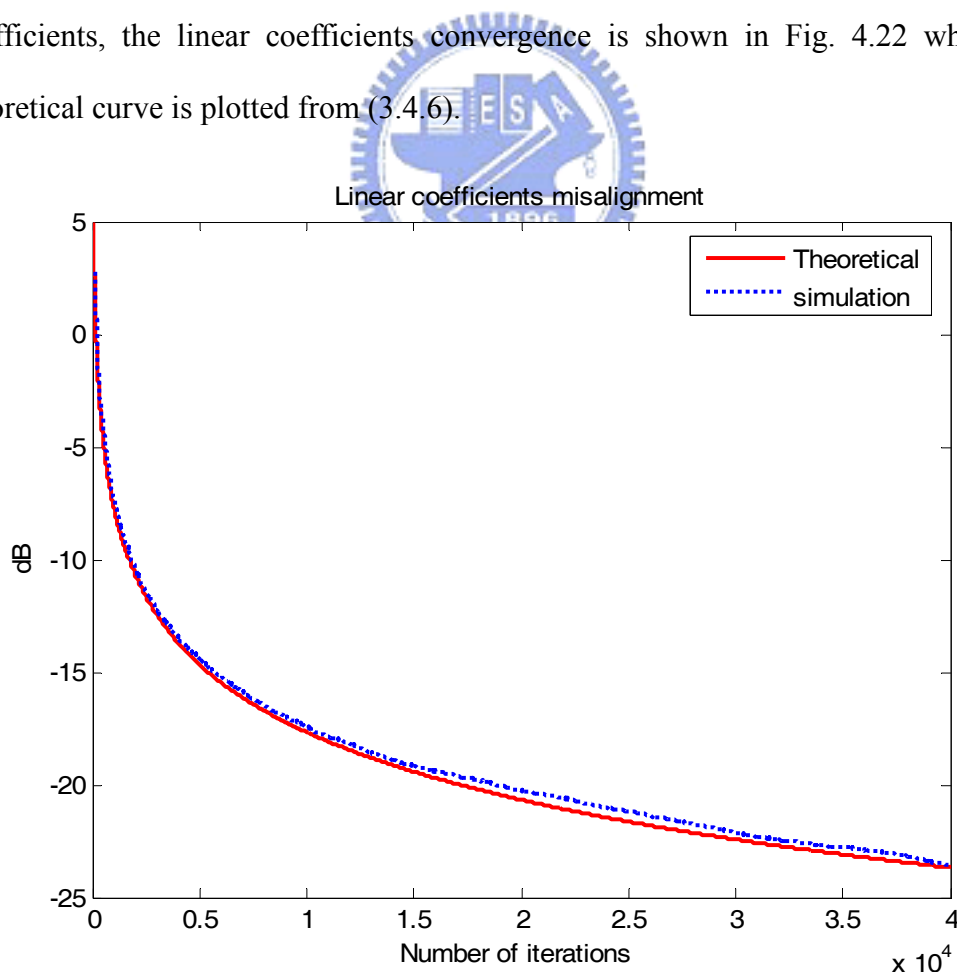


Fig. 4.22 Comparison of linear TS coeff. misal. (SNR=5 dB, perfect nonlinear coeff.)

In Fig. 4.22, we do not show the steady states which can be achieved only when  $n \rightarrow \infty$ . The convergence curve is inversely proportional of the iteration number. The residual echo power is only a scaled version of Fig. 4.22 as plotted in Fig. 4.23, we only show that *residual* echo term in error signal  $e[n]$  therefore the performance is not bounded by the background noise. Both curves show that our theoretical analyses are valid.

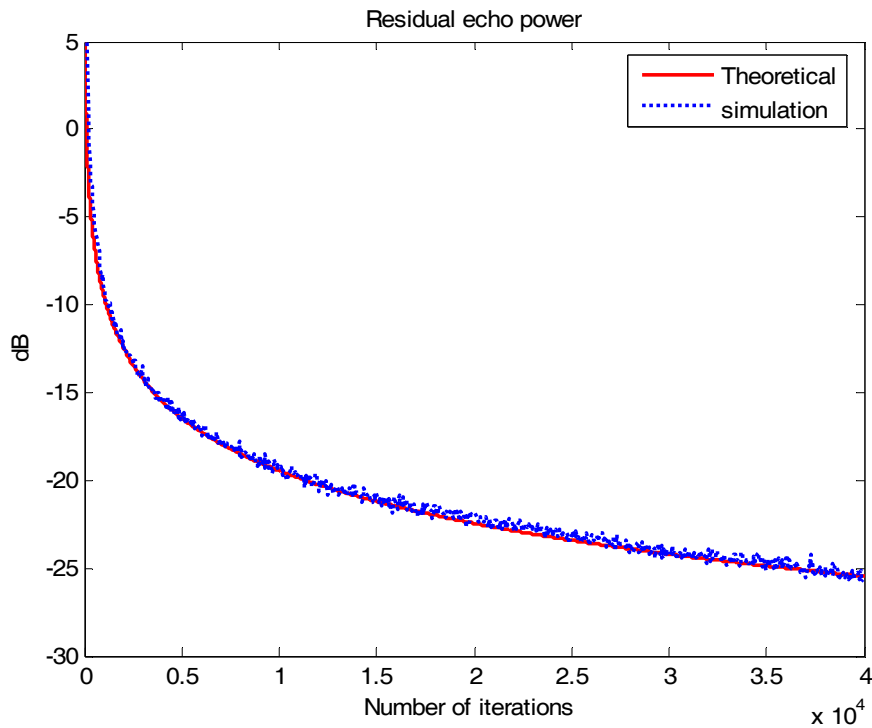


Fig. 4.23 Comparison of residual linear TS echo power (SNR=5 dB)

Next, we show the simulation result of nonlinear coefficients and its residual echo power in Fig. 4.24 and Fig. 4.25, respectively. In Fig. 4.25, we have faster convergence rate than Fig. 4.23 due to the assumption of perfect linear coefficients. Both show that our analyses are correct and optimum solution can be achieved when  $n \rightarrow \infty$ .



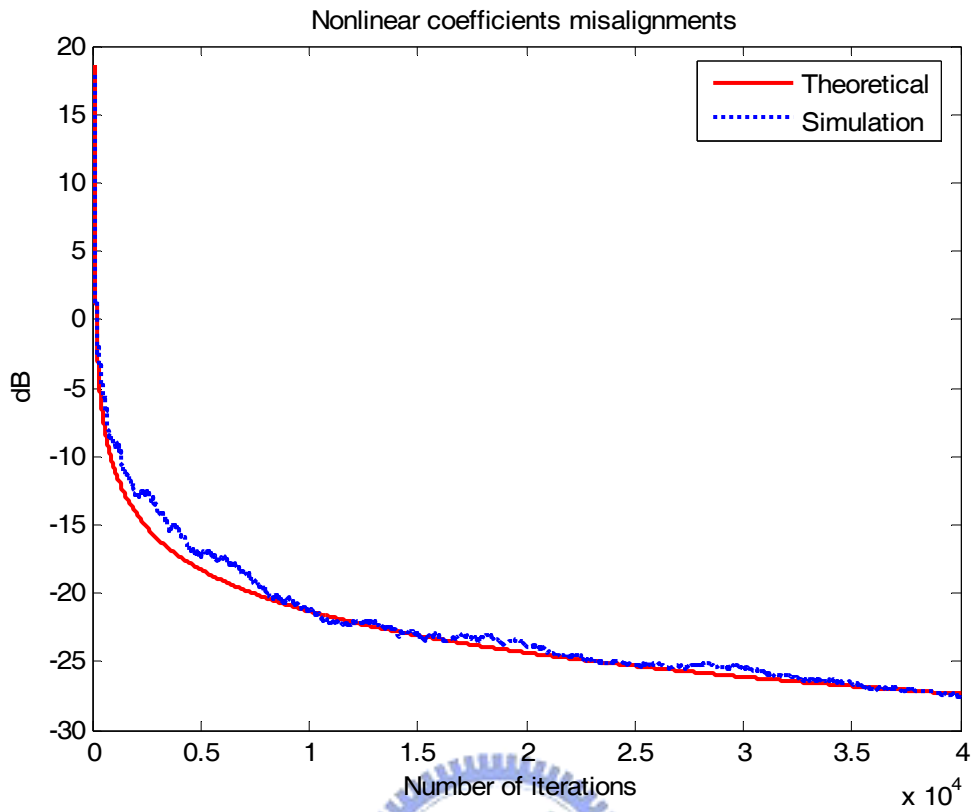


Fig. 4.24 Comparison of nonlinear TS coeff. misal. (SNR=5 dB, perfect linear coeff.)

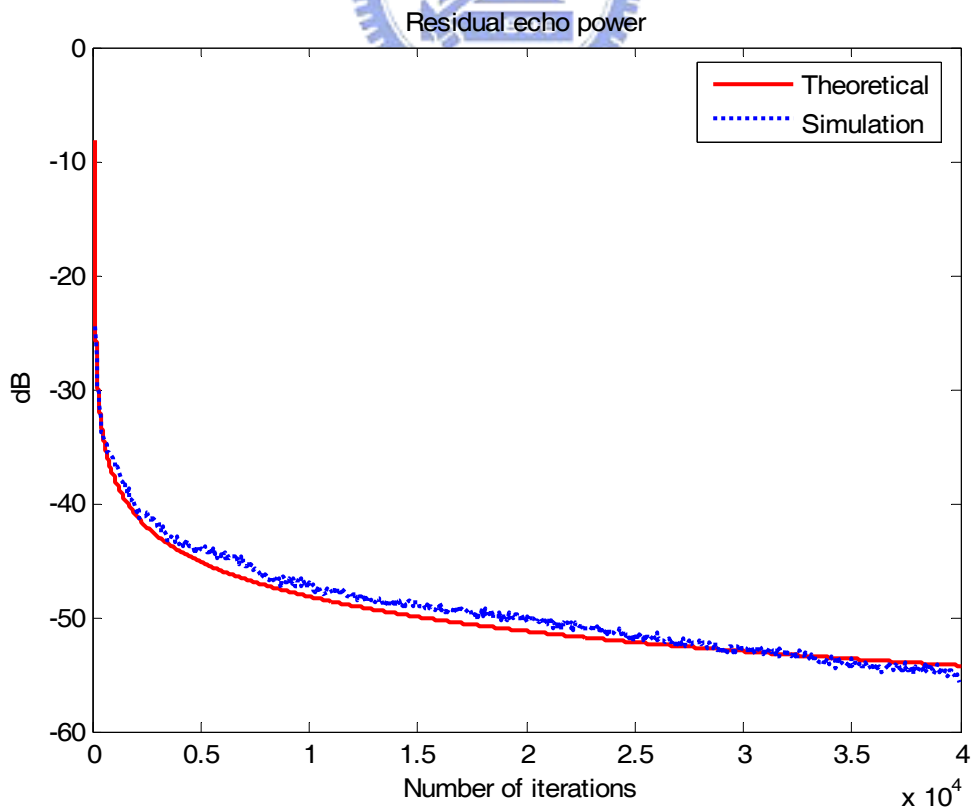


Fig. 4.25 Comparison of residual nonlinear TS echo power (SNR=5 dB)

### 4.3.2 Joint TS convergence rate

Now the assumption of perfectly known nonlinear coefficients is discarded. The joint linear coefficient and residual echo power convergences are given in Fig. 4.26 and 4.27. By comparing Fig. 4.26 and Fig. 4.22; the simulation results are almost identical. In Fig. 4.27, the joint residual echo power is slightly inferior to the case of perfect nonlinear coefficients. According to both simulations, the nonlinear coefficients estimation error has less effect on linear coefficients and residual echo power.

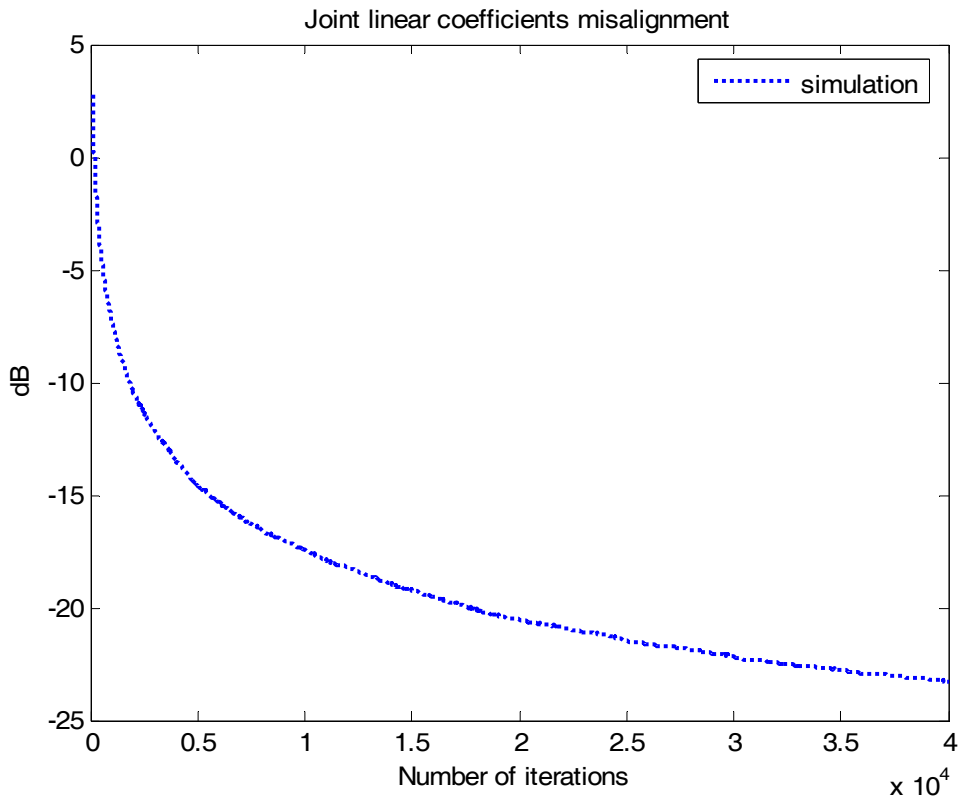


Fig. 4.26 Joint linear coefficients misalignment for training method (SNR=5 dB)

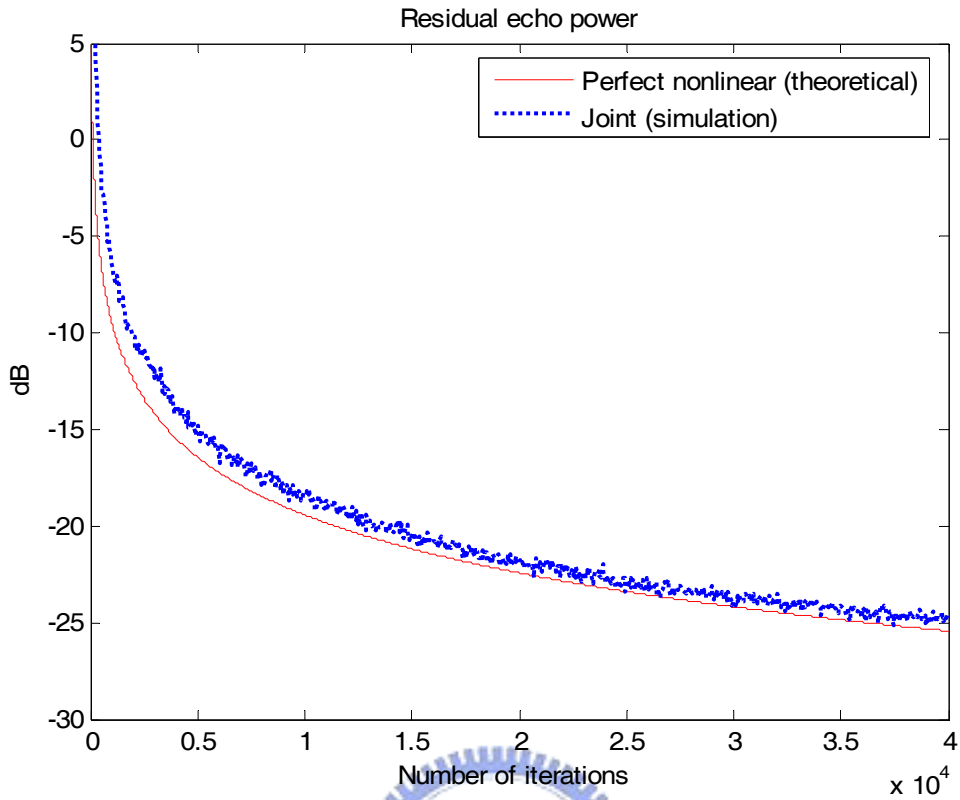


Fig. 4.27 Joint residual echo power for training method (SNR=5 dB)

### 4.3.3 Joint residual echo comparison of TS and NLMS algorithms

In this section we compare the training method with NLMS-adaptive filter. The joint residual echo is shown in Fig. 4.28 when SNR is equal to 10 dB. The convergence rate of the NLMS adaptive filter is dependent on the step size, thus we use two sets of step size,  $\mu_a = 0.05$ ,  $\mu_h = 0.05$ , and  $\mu_a = 0.25$ ,  $\mu_h = 0.1$ .

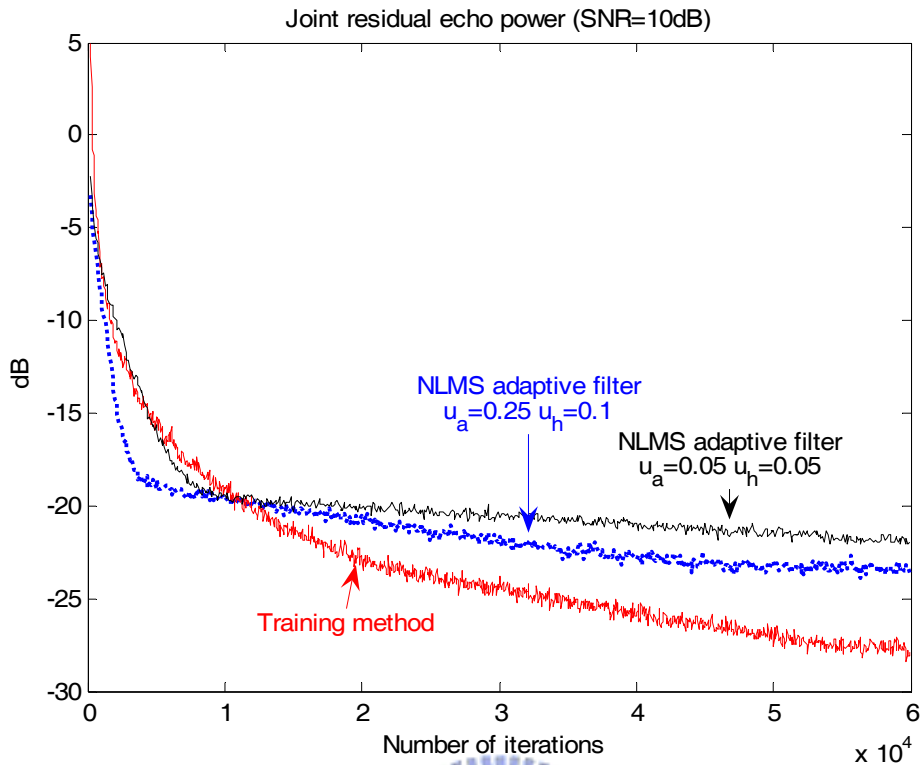


Fig. 4.28 Joint residual echo power comparison of NLMS and TS (SNR=10 dB)

The training method has better steady state performance than the NLMS method since its convergence rate is inversely proportional to the iteration number as indicated in (3.4.8). As for NLMS, its performance is bounded by the background noise. Although the adaptive filter has faster convergence rate at the beginning of iteration, but its steady state performance is poorer as a tradeoff.

Next in Fig. 4.29 and 4.30, comparisons are made when SNR is equal to 5 dB and 0 dB, respectively. TS convergence rate is relatively insensitive to SNR variations, except for a bias due to the noise. This is the main advantage of the training method. Again, the training is preferable than NLMS adaptive filter at low SNRs.

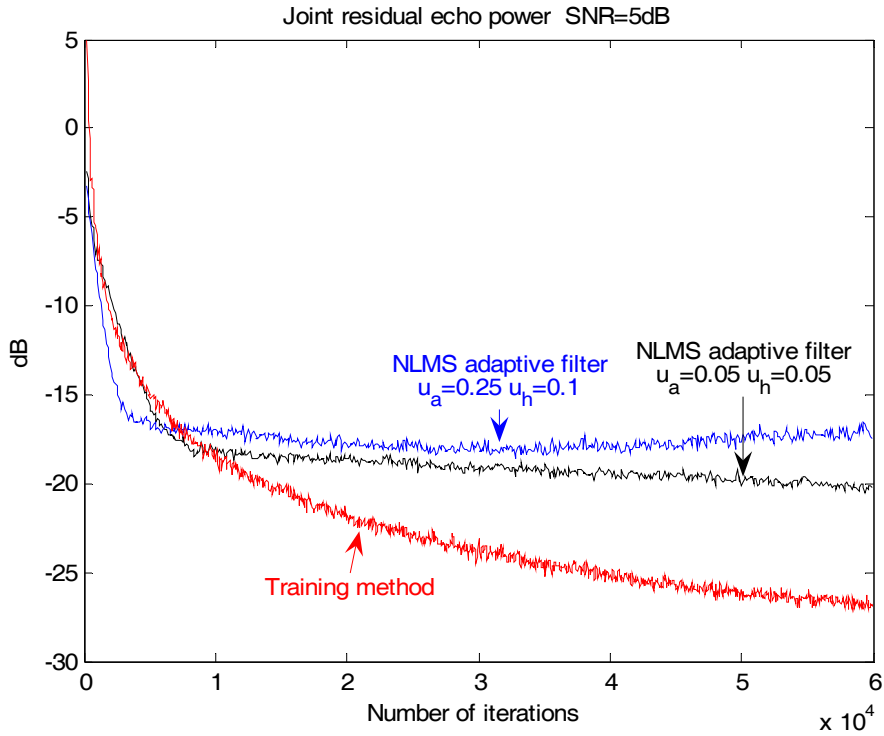


Fig. 4.29 Joint residual echo power comparison of NLMS and TS (SNR=5 dB)

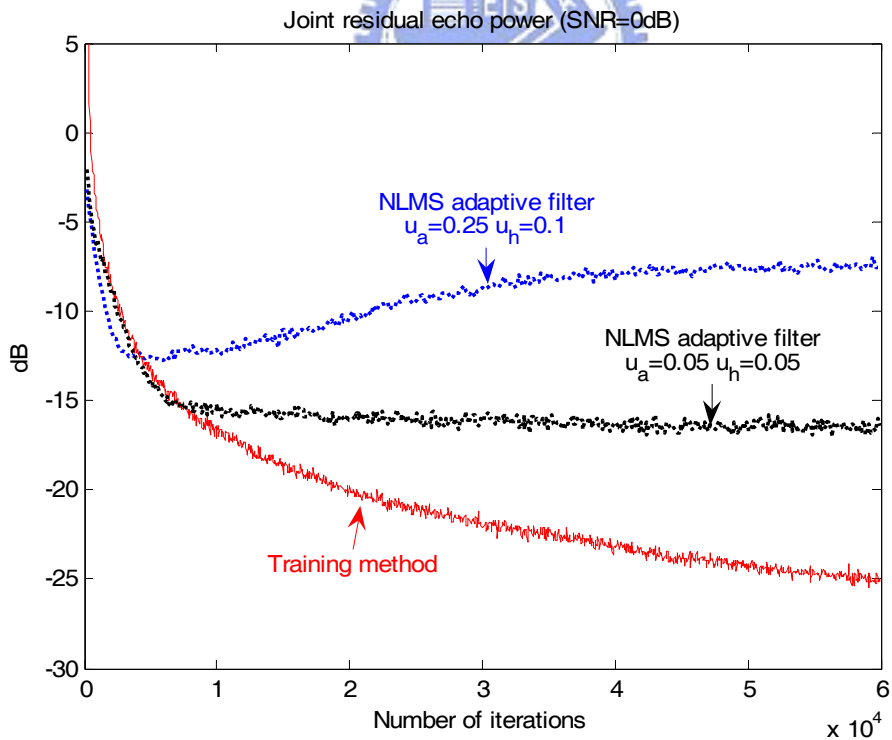


Fig. 4.30 Joint residual echo power comparison of NLMS and TS (SNR=0 dB)

#### 4.3.4 Training Sequences for dual loudspeakers system

We will examine the analyses in Section 3.6 and compare the joint residual echo powers of TS and NLMS algorithms. First, in Fig. 4.31, the simulated residual echo power is plotted and compared with the theoretical curve from (3.6.4). It shows that the modification of one loudspeaker theoretical curve is correct, which can be extended to dual loudspeakers case.

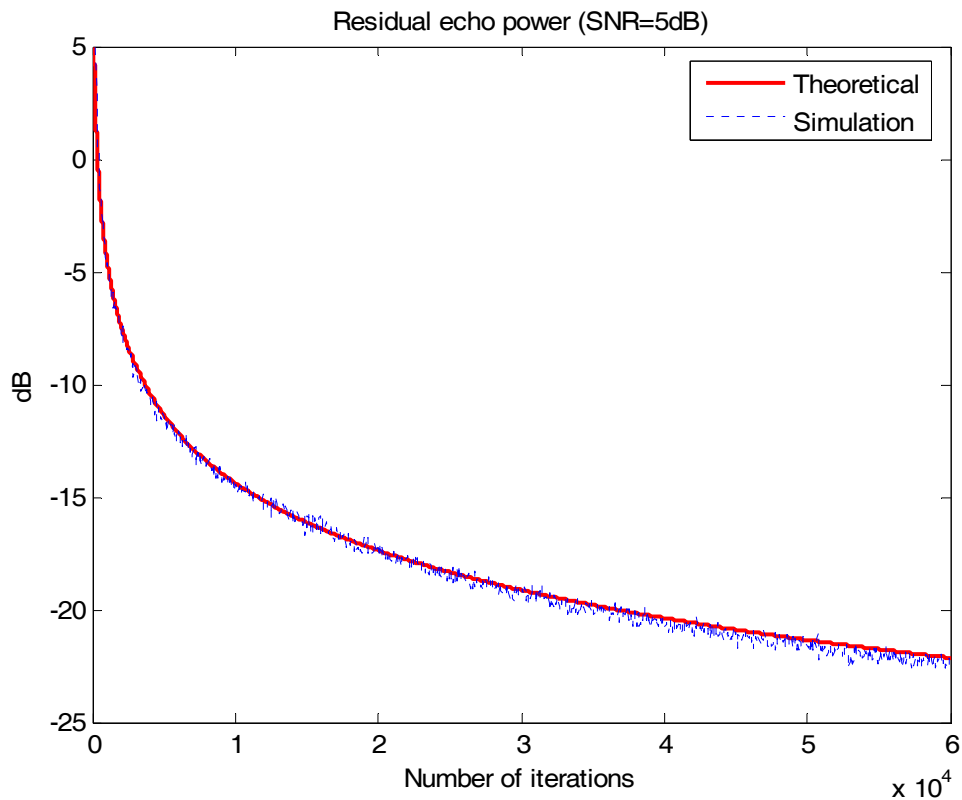


Fig. 4.31 Comparison of residual echo powers for dual loudspeakers and training method (SNR=5 dB)

Next, the joint residual echo power comparison for TS and NLMS are shown in Fig. 4.32. Compared with one loudspeaker case, the performance is worse at same SNR condition since two echoes signals can affect each other as noise sources.

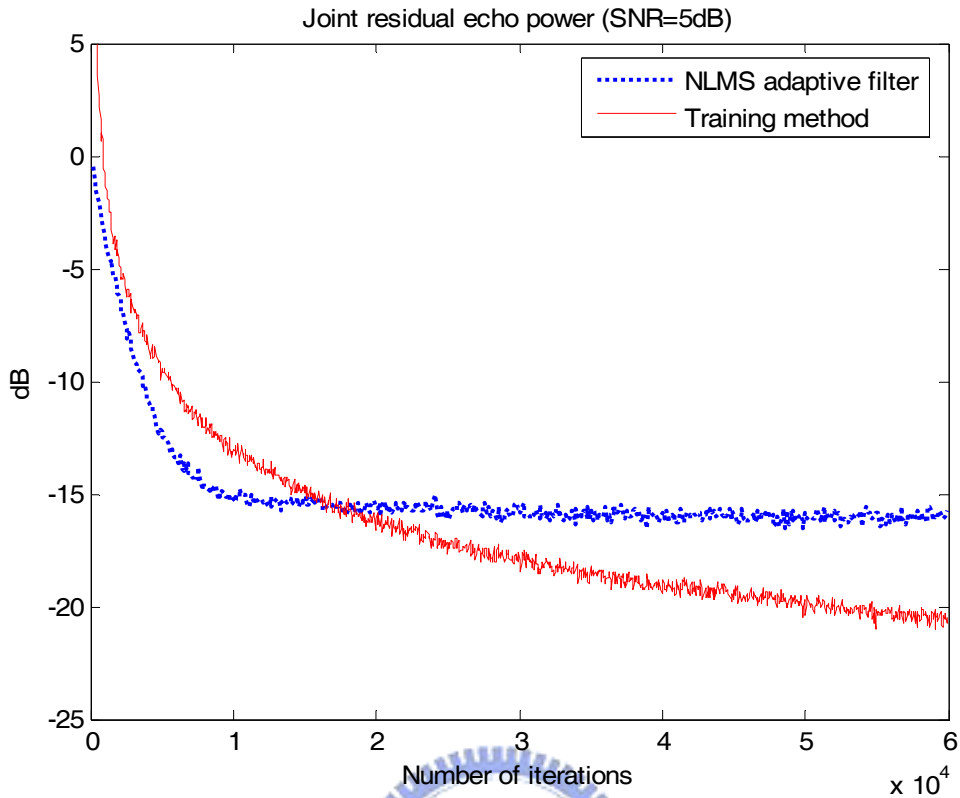


Fig. 4.32 Comparison of joint residual echo powers for dual loudspeakers for TS and NLMS (SNR=5 dB)



#### 4.4 Experiments with a real echo path

So far, for simplicity, we only consider that the nonlinear AEC with an artificial room impulse response and nonlinear loudspeaker. For practice use, we pass the signal through a real nonlinear loudspeaker, a room impulse response, and then pick up the signal from a microphone. The equipment includes a personal computer with a low-cost 2.5 inch diameter desktop loudspeaker and a Creative-MC1000 microphone. We put the loudspeaker above the microphone about 4 inches, to emulate the set-up of a cell phone. Both the loudspeaker and microphone are positioned toward the wall 3.3 feet away. We show the performances of different AEC structures, which are linear AEC, power series type nonlinear AEC, Laplacian type nonlinear AEC, and the

sigmoid type nonlinear AEC. The far-end speech input signal and the microphone output signal are shown in Fig. 4.33.

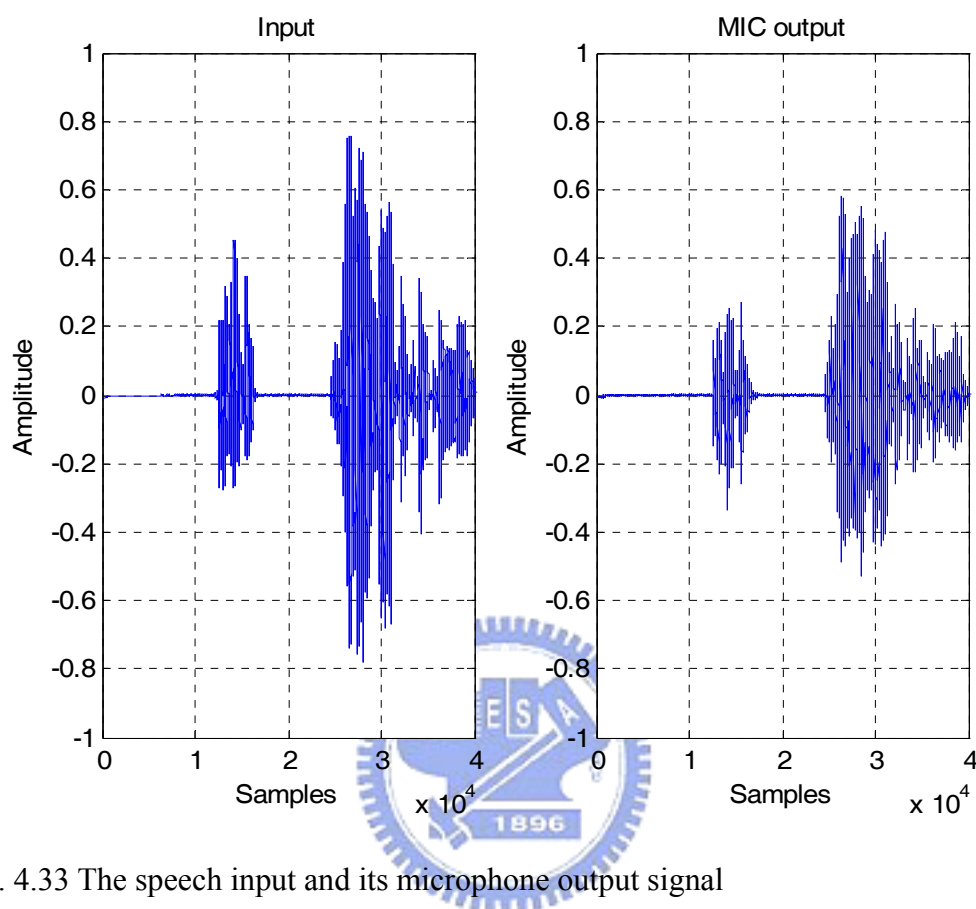


Fig. 4.33 The speech input and its microphone output signal

The speech signal is sampled with 8 KHz sampling rate and 16 bits/samples. Here, we choose 3 different lengths of linear filter, which are 1024, 512 and 256, and the nonlinear filter order is equal to 5, with the even order excluded. The step sizes  $\mu_h$  and  $\mu_a$  of linear and nonlinear filter are equal to 0.5. The sigmoid function is given by

$$\varphi[x] = \left( \frac{2}{1 + \exp(-\alpha x)} - 1 \right) \beta .$$

We use the NLMS algorithm to update  $\alpha$  and  $\beta$  in the sigmoid function which has been introduced in [6, Ch 17]. The step sizes are 0.5 which are chosen to keep the system stable.

As seen from the input signal, the nonlinear distortion arises between 25K and



30K samples. Thus, the nonlinear AEC has better performance than the linear AEC during this period. We show the ERLE between linear and nonlinear AECs in Fig.4.34 when linear filter order is equal to 1024 and the Laplacian type basis is used in the nonlinear AEC. The nonlinear AEC has about 3 to 5 dB ERLE gain over the linear AEC. The ERLE using different linear filter lengths are shown in Table 4.2 and the nonlinear coefficients for each AEC are listed in Table 4.3.

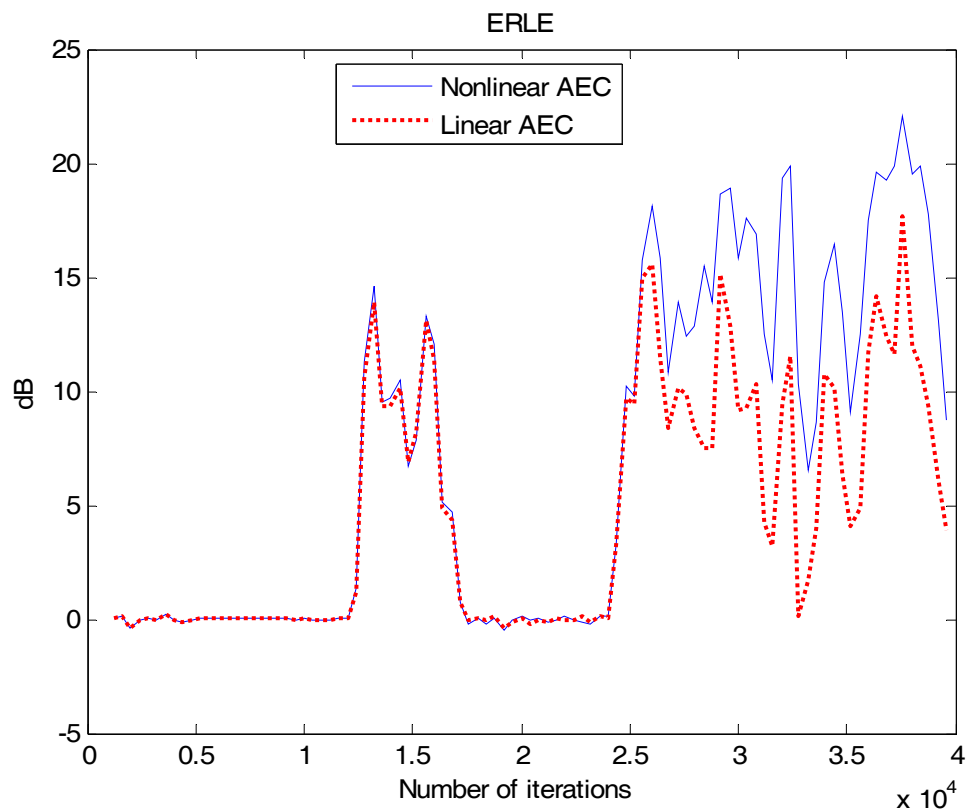
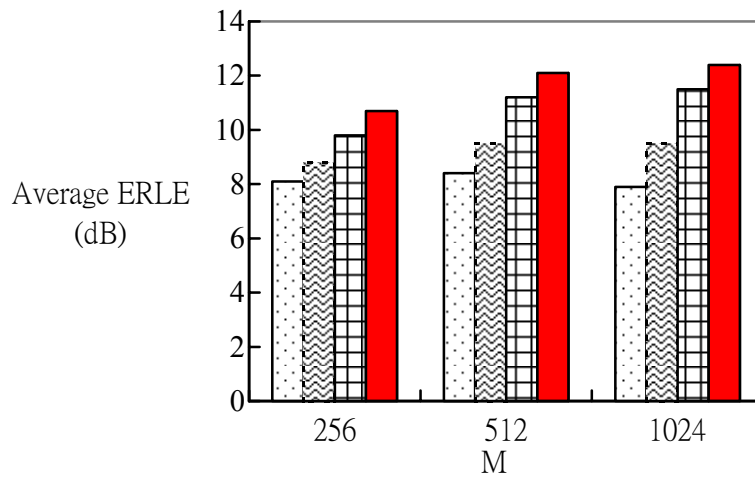


Fig. 4.34 ERLE comparison between linear and nonlinear AEC for a true echo path

Table 4.2 Average ERLE comparison between different linear filter lengths for true echo path



	256	512	1024
□ Linear	8.1	8.4	7.9
▨ Nonlinear_Sigmoid	8.8	9.5	9.5
▤ Nonlinear_Power series	9.8	11.2	11.5
■ Nonlinear_Laplacian	10.7	12.1	12.4

Table 4.3 Nonlinear coefficients for each AEC with different linear orders

M \ AEC	Linear [a1]	Nonlinear (Sigmoid) [α β]	Nonlinear (Power series) [a1 a3 a5]	Nonlinear (Laplacian type) [a1 a3 a5]
1024	[0.9]	[2.53 -1.57]	[1.23 -1.33 -0.5]	[1.3 -2.19 0.52]
512	[0.96]	[2.53 -1.52]	[1.24 -1.57 -0.59]	[1.3 -2.25 0.54]
256	[0.9]	[2.57 -1.54]	[1.06 -1.76 -0.63]	[1.14 -2.21 0.55]

In Fig. 4.35 we examine the performance of different nonlinear orders for a 1024<sup>th</sup> order linear FIR and power series basis nonlinear processor.

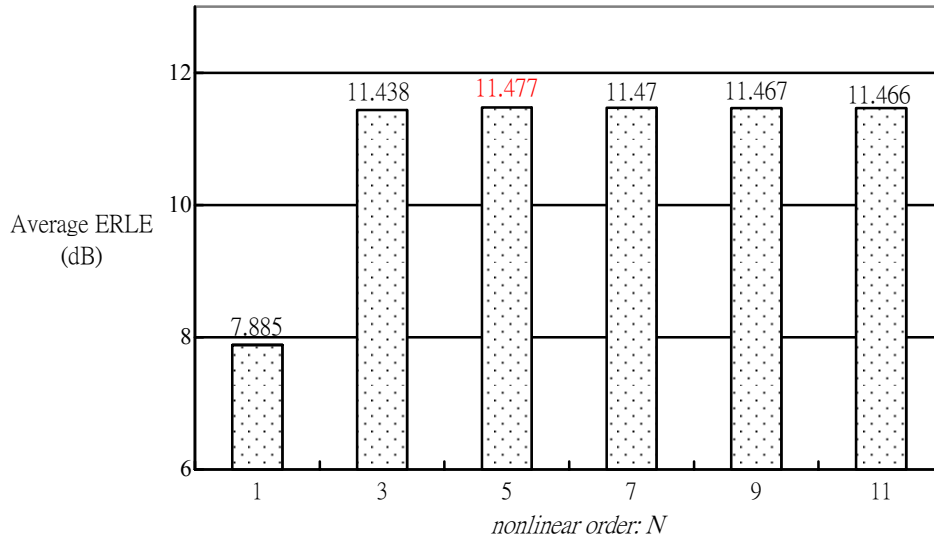


Fig. 4.35 Average ERLE comparison of different nonlinear orders

Next, we compare the performances of TS and NLMS algorithms for a real echo path. In this experiment, 5<sup>th</sup> power series basis is used and the number of FIR is equal to 256. We show the average ERLE in Table 4.3. The steady state of training method is about 7dB, which is different from the theoretical analysis. One possible explanation is as follows. The optimum coefficients can be achieved for training method only for a perfect white signal. In this simulation, the loudspeaker may not be memoryless thus the optimum coefficients can not be achieved.

Table 4.4 Average ERLE comparison of NLMS and TS methods for a true echo path

	NLMS adaptive filter	Training method
Average ERLE	14.9dB	7.7dB

Finally, we check the performance in a noisy environment such as the car cabin. Now we add a background interference speech (i.e., double talk) to the microphone output signal. The performances at different SNRs (signal to interference ration) are shown in Fig. 4.36. The NLMS algorithm heavily depends on the SNR conditions; its

adaptive filter will diverge when SNR is below 3 dB. By contrast, the training method is more robust to noise and/or double talk.

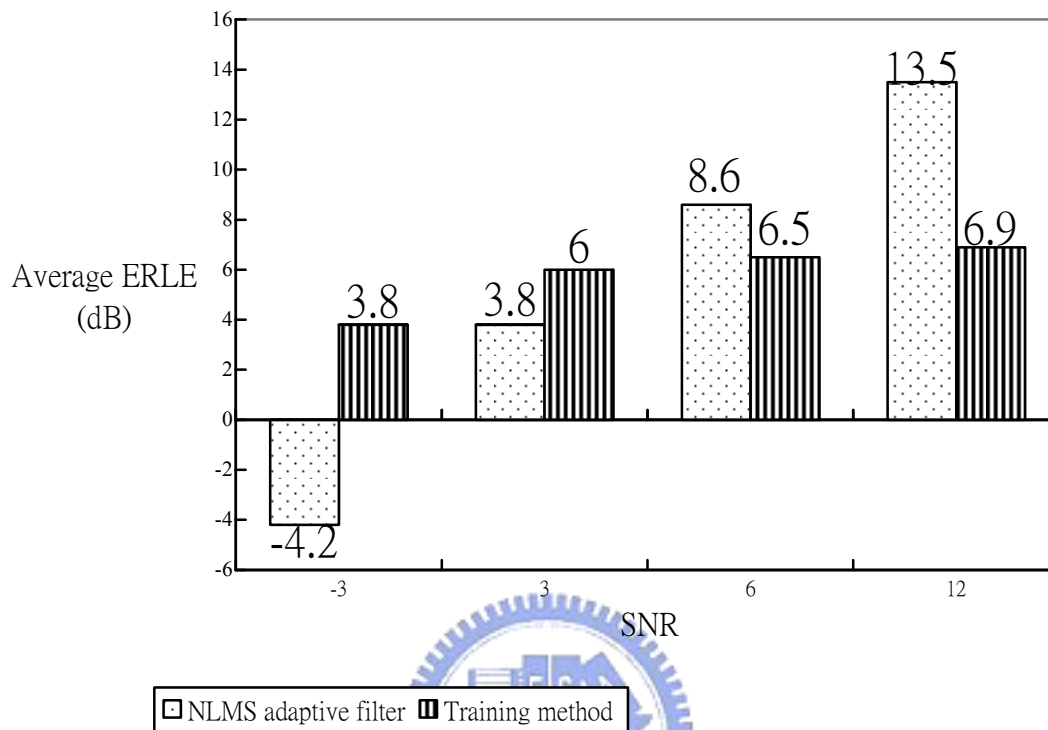


Fig. 4.36 Average ERLE comparison at different SNRs

# Chapter 5

## Conclusions

In this thesis we have proposed the orthogonal basis polynomials for nonlinear AEC. For either uniform, Gaussian and Laplacian inputs, the simulation results agree well with theoretical curves. Convergence rate analysis indicates that a smaller eigenvalue spread is closely related to the correlation between polynomials bases. The proposed orthogonal basis does not incur more computation cost and has better performance (smaller eigenvalue spread) than conventional power-series basis, even in case of imperfect orthogonality. The proposed basis can also improve the convergence rate for dual loudspeakers system.

We also proposed and analyzed the coefficients estimation based on a training sequence. It can achieve the optimum solution thus it has better steady state performance than the NLMS adaptation algorithm under the same computational complexity. Although the TS algorithm can not use the step size to control its convergence rate, but its non-error feedback structure has more resistance against noise, especially at low SNR. Finally, the training method can also be used in dual loudspeakers case.

From the experiment of a real speech into a real echo path, we know that the proposed nonlinear processor structure has better performance than a linear AEC. In case of a large noise or double talk, the NLMS adaptive filter diverges; the training method has its advantage in such noisy environment.

The future work includes study on a computation-effective AEC structure for dual loudspeakers, one AEC to cancel two echo path echoes signals and its convergence rate.

# Appendix

In this appendix, we will give the detailed the mathematical derivation in (3.5.1)

The covariance matrix of  $\boldsymbol{\varepsilon}_a[n]$  can be written as

$$\begin{aligned}
 & E[\boldsymbol{\varepsilon}_a[n]\boldsymbol{\varepsilon}_a^T[n]] \\
 &= \begin{bmatrix} E[\varepsilon_{a,1}^2[n]] & E[\varepsilon_{a,1}[n]\varepsilon_{a,3}[n]] & E[\varepsilon_{a,1}[n]\varepsilon_{a,5}[n]] \\ E[\varepsilon_{a,1}[n]\varepsilon_{a,3}[n]] & E[\varepsilon_{a,3}^2[n]] & E[\varepsilon_{a,3}[n]\varepsilon_{a,5}[n]] \\ E[\varepsilon_{a,1}[n]\varepsilon_{a,5}[n]] & E[\varepsilon_{a,3}[n]\varepsilon_{a,5}[n]] & E[\varepsilon_{a,5}^2[n]] \end{bmatrix} \\
 &= \mathbf{G}_{odd}^{-1} \begin{bmatrix} \mathbf{h}^T \widehat{\mathbf{R}}_{xd} \\ \mathbf{h}^T \widehat{\mathbf{R}}_{x^3d} \\ \mathbf{h}^T \widehat{\mathbf{R}}_{x^5d} \end{bmatrix} \begin{bmatrix} \mathbf{h}^T \widehat{\mathbf{R}}_{xd} & \mathbf{h}^T \widehat{\mathbf{R}}_{x^3d} & \mathbf{h}^T \widehat{\mathbf{R}}_{x^5d} \end{bmatrix} \mathbf{G}_{odd}^{-T} \\
 &= \mathbf{G}_{odd}^{-1} \begin{bmatrix} \mathbf{h}^T \widehat{\mathbf{R}}_{xd} \mathbf{h}^T \widehat{\mathbf{R}}_{xd} & \mathbf{h}^T \widehat{\mathbf{R}}_{xd} \mathbf{h}^T \widehat{\mathbf{R}}_{x^3d} & \mathbf{h}^T \widehat{\mathbf{R}}_{xd} \mathbf{h}^T \widehat{\mathbf{R}}_{x^5d} \\ \mathbf{h}^T \widehat{\mathbf{R}}_{x^3d} \mathbf{h}^T \widehat{\mathbf{R}}_{xd} & \mathbf{h}^T \widehat{\mathbf{R}}_{x^3d} \mathbf{h}^T \widehat{\mathbf{R}}_{x^3d} & \mathbf{h}^T \widehat{\mathbf{R}}_{x^3d} \mathbf{h}^T \widehat{\mathbf{R}}_{x^5d} \\ \mathbf{h}^T \widehat{\mathbf{R}}_{x^5d} \mathbf{h}^T \widehat{\mathbf{R}}_{xd} & \mathbf{h}^T \widehat{\mathbf{R}}_{x^5d} \mathbf{h}^T \widehat{\mathbf{R}}_{x^3d} & \mathbf{h}^T \widehat{\mathbf{R}}_{x^5d} \mathbf{h}^T \widehat{\mathbf{R}}_{x^5d} \end{bmatrix} \mathbf{G}_{odd}^{-T} \quad (\text{A.1})
 \end{aligned}$$

$$= \mathbf{G}_{odd}^{-1} [\text{SME} + \text{NE}] \mathbf{G}_{odd}^{-T} \quad (\text{A.2})$$

where

$$\begin{aligned}
 \widehat{\mathbf{R}}_{xd} &= \frac{1}{n} \sum_{i=1}^n \mathbf{h}^T \mathbf{x}[i] d[i], \\
 d[i] &= a_1 \mathbf{h}^T \mathbf{x}[i] + a_3 \mathbf{h}^T \mathbf{x}^3[i] + a_5 \mathbf{h}^T \mathbf{x}^5[i] + v[i].
 \end{aligned}$$

First, by the law of superposition we let the noise term  $v[i] = 0$ , noise effect **NE** in (A.2), is equal to a zero matrix. First, the *sample mean estimation* error term, **SME**.

$\mathbf{h}^T \widehat{\mathbf{R}}_{xd} \mathbf{h}^T \widehat{\mathbf{R}}_{xd}$  in **SME** is given by

$$\begin{aligned}
 \mathbf{h}^T \widehat{\mathbf{R}}_{xd} \mathbf{h}^T \widehat{\mathbf{R}}_{xd} &= \frac{1}{n^2} E \left[ \left( \sum_{i=1}^n a_1 \mathbf{h}^T \mathbf{x}[i] \mathbf{x}^T[i] \mathbf{h} + a_3 \mathbf{h}^T \mathbf{x}[i] \mathbf{x}^{3T}[i] \mathbf{h} + a_5 \mathbf{h}^T \mathbf{x}[i] \mathbf{x}^{5T}[i] \mathbf{h} \right)^2 \right] \\
 &\simeq \frac{1}{n^2} a_1^2 E \left[ \sum_{i=1}^n (\mathbf{h}^T \mathbf{x}[i] \mathbf{x}^T[i] \mathbf{h})^2 \right] + \frac{1}{n^2} a_3^2 E \left[ \sum_{i=1}^n (\mathbf{h}^T \mathbf{x}[i] \mathbf{x}^{3T}[i] \mathbf{h})^2 \right]
 \end{aligned}$$

$$\begin{aligned}
& + \frac{1}{n^2} a_3^2 E \left[ \sum_{i=1}^n (\mathbf{h}^T \mathbf{x}[i] \mathbf{x}^{5^T} [i] \mathbf{h})^2 \right] \\
& + \frac{2}{n^2} E \left[ \sum_{i=1}^n (\mathbf{h}^T \mathbf{x}[i] \mathbf{x}^T [i] \mathbf{h}) (\mathbf{h}^T \mathbf{x}[i] \mathbf{x}^{3^T} [i] \mathbf{h}) \right] \\
& + \frac{2}{n^2} E \left[ \sum_{i=1}^n (\mathbf{h}^T \mathbf{x}[i] \mathbf{x}^T [i] \mathbf{h}) (\mathbf{h}^T \mathbf{x}[i] \mathbf{x}^{5^T} [i] \mathbf{h}) \right] \\
& + \frac{2}{n^2} E \left[ \sum_{i=1}^n (\mathbf{h}^T \mathbf{x}[j] \mathbf{x}^{3^T} [j] \mathbf{h}) (\mathbf{h}^T \mathbf{x}[i] \mathbf{x}^{5^T} [i] \mathbf{h}) \right]
\end{aligned} \tag{A.3}$$

where the first term in (A.3) is

$$\begin{aligned}
& \frac{1}{n^2} a_1^2 E \left[ \sum_{i=1}^n (\mathbf{h}^T \mathbf{x}[i] \mathbf{x}^T [i] \mathbf{h})^2 \right] \\
& = \frac{a_1^2}{n^2} \left\{ E \left[ \sum_{i=1}^n (\mathbf{h}^T \mathbf{x}[i])^4 \right] + E \left[ 2 \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n (\mathbf{h}^T \mathbf{x}[i])^2 (\mathbf{h}^T \mathbf{x}[j])^2 \right] \right\}
\end{aligned} \tag{A.4}$$

where  $E \left[ \sum_{i=1}^n (\mathbf{h}^T \mathbf{x}[i])^4 \right]$  in (A.4) is

$$\begin{aligned}
& E \left[ \left( h_0 x[i] + h_1 x[i-1] + \dots + h_{M-1} x[i-M+1] \right)^4 \right] \\
& = n \left[ m_4 \sum_{i=0}^{M-1} h_i^4 + m_2 \sum_{i=0}^{M-1} \sum_{\substack{j=0 \\ j \neq i}}^{M-1} h_i^2 h_j^2 \right] \\
& \simeq n m_4 \sum_{i=0}^{M-1} h_i^4
\end{aligned} \tag{A.5}$$

where  $E \left[ 2 \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n (\mathbf{h}^T \mathbf{x}[i])^2 (\mathbf{h}^T \mathbf{x}[j])^2 \right]$  in (A.4) is approximately as

$$n m_4 \sum_{i=0}^{M-1} \sum_{\substack{j=0 \\ j \neq i}}^{M-1} h_i^2 h_j^2 \tag{A.6}$$

Therefore (A.4) is found as sum of (A.5) and (A.6),

$$\simeq \frac{a_1^2 m_4}{n} \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} h_i^2 h_j^2 \tag{A.7}$$

Similarly, the other terms in (A.3) is

$$\frac{1}{n^2} a_3^2 E \left[ \sum_{i=1}^n (\mathbf{h}^T \mathbf{x}[i] \mathbf{x}^{3T} [i] \mathbf{h})^2 \right] \simeq a_3^2 \frac{m_8}{n} \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} h_i^2 h_j^2 \quad (\text{A.8})$$

$$\frac{1}{n^2} a_5^2 E \left[ \sum_{i=1}^n (\mathbf{h}^T \mathbf{x}[i] \mathbf{x}^{5T} [i] \mathbf{h})^2 \right] \simeq a_5^2 \frac{m_{12}}{n} \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} h_i^2 h_j^2 \quad (\text{A.9})$$

$$\frac{2}{n^2} E \left[ \sum_{i=1}^n (\mathbf{h}^T \mathbf{x}[i] \mathbf{x}^T [i] \mathbf{h})(\mathbf{h}^T \mathbf{x}[i] \mathbf{x}^{3T} [i] \mathbf{h}) \right] \simeq 2a_1 a_3 \frac{m_6}{n} \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} h_i^2 h_j^2 \quad (\text{A.10})$$

$$\frac{2}{n^2} E \left[ \sum_{i=1}^n (\mathbf{h}^T \mathbf{x}[i] \mathbf{x}^T [i] \mathbf{h})(\mathbf{h}^T \mathbf{x}[i] \mathbf{x}^{5T} [i] \mathbf{h}) \right] \simeq 2a_1 a_5 \frac{m_8}{n} \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} h_i^2 h_j^2 \quad (\text{A.11})$$

$$\frac{2}{n^2} E \left[ \sum_{i=1}^n (\mathbf{h}^T \mathbf{x}[i] \mathbf{x}^{3T} [i] \mathbf{h})(\mathbf{h}^T \mathbf{x}[i] \mathbf{x}^{5T} [i] \mathbf{h}) \right] \simeq 2a_3 a_5 \frac{m_{10}}{n} \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} h_i^2 h_j^2 \quad (\text{A.12})$$

According to the equations from (A.4) to (A.12),

$$\begin{aligned} & \mathbf{h}^T \widehat{\mathbf{R}}_{xd} \mathbf{h}^T \widehat{\mathbf{R}}_{xd} \\ & \simeq \frac{1}{n} \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} h_i^2 h_j^2 (a_1^2 m_4 + a_3^2 m_8 + a_5^2 m_{12} + 2(a_1 a_3 m_6 + a_1 a_5 m_8 + a_3 a_5 m_{10})) \end{aligned} \quad (\text{A.13})$$

Similarly, the other term in (A.1) can be found by the same procedures are

$$\mathbf{h}^T \widehat{\mathbf{R}}_{x^3d} \mathbf{h}^T \widehat{\mathbf{R}}_{x^3d} \simeq \frac{1}{n} \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} h_i^2 h_j^2 (a_1^2 m_8 + a_3^2 m_{12} + a_5^2 m_{16} + 2(a_1 a_3 m_{10} + a_1 a_5 m_{12} + a_3 a_5 m_{14}))$$

$$\mathbf{h}^T \widehat{\mathbf{R}}_{x^5d} \mathbf{h}^T \widehat{\mathbf{R}}_{x^5d} \simeq \frac{1}{n} \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} h_i^2 h_j^2 (a_1^2 m_{12} + a_3^2 m_{16} + a_5^2 m_{20} + 2(a_1 a_3 m_{14} + a_1 a_5 m_{16} + a_3 a_5 m_{18}))$$

$$\mathbf{h}^T \widehat{\mathbf{R}}_{xd} \mathbf{h}^T \widehat{\mathbf{R}}_{x^3d} \simeq \frac{1}{n} \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} h_i^2 h_j^2 (a_1^2 m_6 + a_3^2 m_{10} + a_5^2 m_{14} + 2(a_1 a_3 m_8 + a_1 a_5 m_{10} + a_3 a_5 m_{12}))$$

$$\mathbf{h}^T \widehat{\mathbf{R}}_{xd} \mathbf{h}^T \widehat{\mathbf{R}}_{x^5d} \simeq \frac{1}{n} \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} h_i^2 h_j^2 (a_1^2 m_8 + a_3^2 m_{12} + a_5^2 m_{16} + 2(a_1 a_3 m_{10} + a_1 a_5 m_{12} + a_3 a_5 m_{14}))$$

$$\mathbf{h}^T \widehat{\mathbf{R}}_{x^3d} \mathbf{h}^T \widehat{\mathbf{R}}_{x^5d} \simeq \frac{1}{n} \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} h_i^2 h_j^2 (a_1^2 m_{10} + a_3^2 m_{14} + a_5^2 m_{18} + 2(a_1 a_3 m_{12} + a_1 a_5 m_{14} + a_3 a_5 m_{16}))$$

Thus, **SME** in (A.2) is given as



$$\begin{bmatrix}
\frac{a_1^2 m_4 + a_3^2 m_8 + a_5^2 m_{12}}{n} H & \frac{a_1^2 m_6 + a_3^2 m_{10} + a_5^2 m_{14}}{n} H & \frac{a_1^2 m_8 + a_3^2 m_{12} + a_5^2 m_{16}}{n} H \\
\frac{a_1^2 m_6 + a_3^2 m_{10} + a_5^2 m_{12}}{n} H & \frac{a_1^2 m_8 + a_3^2 m_{12} + a_5^2 m_{16}}{n} H & \frac{a_1^2 m_{10} + a_3^2 m_{14} + a_5^2 m_{18}}{n} H \\
\frac{a_1^2 m_8 + a_3^2 m_{12} + a_5^2 m_{16}}{n} H & \frac{a_1^2 m_{10} + a_3^2 m_{14} + a_5^2 m_{18}}{n} H & \frac{a_1^2 m_{12} + a_3^2 m_{16} + a_5^2 m_{20}}{n} H
\end{bmatrix}
\tag{A.14}$$

where  $H = \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} h_i^2 h_j^2$ .

Next we discuss the effect of the noise term, NE. In this case the desired signal  $d[i]$  only contains the background noise  $v[i]$ , thus,  $\mathbf{h}^T \widehat{\mathbf{R}}_{xd} \mathbf{h}^T \widehat{\mathbf{R}}_{xd}$  in (A.1) can be found by

$$\begin{aligned}
\mathbf{h}^T \widehat{\mathbf{R}}_{xd} \mathbf{h}^T \widehat{\mathbf{R}}_{xd} &= \frac{1}{n^2} E \left[ \left( \frac{1}{n} \sum_{i=1}^n \mathbf{h}^T \mathbf{x}[i] v[i] \right) \left( \frac{1}{n} \sum_{j=1}^n \mathbf{h}^T \mathbf{x}[j] d[j] \right) \right] \\
&= \frac{1}{n^2} E \left[ \left( \mathbf{h}^T \mathbf{x}[1] v[1] + \mathbf{h}^T \mathbf{x}[2] v[2] + \dots + \mathbf{h}^T \mathbf{x}[n] v[n] \right)^2 \right] \\
&= \frac{1}{n^2} n m_2 \sigma_v^2 \sum_{i=0}^{M-1} h_i^2.
\end{aligned}$$

Similarly, the other terms in (A.1) are

$$\begin{aligned}
\mathbf{h}^T \widehat{\mathbf{R}}_{x^3 d} \mathbf{h}^T \widehat{\mathbf{R}}_{x^3 d} &= \frac{1}{n} m_6 \sigma_v^2 \sum_{i=0}^{M-1} h_i^2 \\
\mathbf{h}^T \widehat{\mathbf{R}}_{x^5 d} \mathbf{h}^T \widehat{\mathbf{R}}_{x^5 d} &= \frac{1}{n} m_{10} \sigma_v^2 \sum_{i=0}^{M-1} h_i^2 \\
\mathbf{h}^T \widehat{\mathbf{R}}_x \mathbf{h}^T \widehat{\mathbf{R}}_{x^3 d} &= \frac{1}{n} m_4 \sigma_v^2 \sum_{i=0}^{M-1} h_i^2 \\
\mathbf{h}^T \widehat{\mathbf{R}}_x \mathbf{h}^T \widehat{\mathbf{R}}_{x^5 d} &= \frac{1}{n} m_6 \sigma_v^2 \sum_{i=0}^{M-1} h_i^2 \\
\mathbf{h}^T \widehat{\mathbf{R}}_{x^3 d} \mathbf{h}^T \widehat{\mathbf{R}}_{x^5 d} &= \frac{1}{n} m_8 \sigma_v^2 \sum_{i=0}^{M-1} h_i^2
\end{aligned}$$

therefore, in case of  $d[i] = v[i]$ , the noise term NE is

$$\mathbf{NE} = \begin{bmatrix} \frac{m_2 \sigma_{noise}^2 \sum_{i=0}^{M-1} h_i^2}{n} & \frac{m_4 \sigma_{noise}^2 \sum_{i=0}^{M-1} h_i^2}{n} & \frac{m_6 \sigma_{noise}^2 \sum_{i=0}^{M-1} h_i^2}{n} \\ \frac{m_4 \sigma_{noise}^2 \sum_{i=0}^{M-1} h_i^2}{n} & \frac{m_6 \sigma_{noise}^2 \sum_{i=0}^{M-1} h_i^2}{n} & \frac{m_8 \sigma_{noise}^2 \sum_{i=0}^{M-1} h_i^2}{n} \\ \frac{m_6 \sigma_{noise}^2 \sum_{i=0}^{M-1} h_i^2}{n} & \frac{m_8 \sigma_{noise}^2 \sum_{i=0}^{M-1} h_i^2}{n} & \frac{m_{10} \sigma_{noise}^2 \sum_{i=0}^{M-1} h_i^2}{n} \end{bmatrix} \quad (\text{A.15})$$

(A.15) is simpler than (A.14), therefore, when  $\mathbf{SME} = 0$  (A.2) can be written as

$$\begin{aligned} E[\boldsymbol{\varepsilon}_a[n] \boldsymbol{\varepsilon}_a^T[n]] &= \mathbf{G}_{odd}^{-1} [\mathbf{NE}] \mathbf{G}_{odd}^{-T} \\ &= \frac{\sigma_{noise}^2 \|\mathbf{h}\|_2^2}{n} \mathbf{G}_{odd}^{-T} \end{aligned} \quad (\text{A.16})$$

thus, the coefficients error due to background noise is

$$tr(E[\boldsymbol{\varepsilon}_a[n] \boldsymbol{\varepsilon}_a^T[n]]) = \frac{\sigma_{noise}^2 \|\mathbf{h}\|_2^2 tr(\mathbf{G}_{odd}^{-T})}{n}$$



# Bibliography

- [1] E. Hansler, "The hands-free telephone problem; an annotated bibliography," *Signal Processing*, vol. 27 pp. 259-271, 1992.
- [2] S. M. Kuo, Y. C. Huang, and Z. Pan, "Acoustic noise and echo cancellation microphone system for videoconferencing," *IEEE Trans. Consumer Electronics* vol. 41, no. 4, pp. 1150-1158, Nov. 1995.
- [3] W. Hsu, F. Chui, and D. A. Hodges, "An acoustic echo canceler," *IEEE Journal Solid State Circuits* vol. 24, no. 6, pp. 1639-1646, Dec. 1989.
- [4] J. P. Costa, A. Lagrange, and A. Arliaud, "Acoustic echo cancellation using nonlinear cascade filters," *ICASSP*, pp.v389-v392, 2003.
- [5] S. Gudvangen, and S.J. Flockton, "Modelling of acoustic transfer functions for echo cancellers," *IEE Proceedings-Vision, Image and Signal Processing*, vol. 42 no. 1 pp.47-51, Feb 1995.
- [6] S. Haykin: *Adaptive filter theory*, 4<sup>th</sup> ed., Prentice-Hall, 2002.
- [7] A. N. Birkett, and R. A. Goubran, "Limitations of handsfree acoustic echo cancellers due to nonlinear loudspeaker distortion and enclosure vibration effects," *ICASSP*, pp. 103-106, Oct. 1995.

- [8] F. X. Y. Gao, and W. M. Senlgrave, "Adaptive linearization of loudspeaker," ICASSP, vol. 5, pp. 3589-3592 Apr. 1991.
- [9] A. N. Birkett, and R. A. Goubran, "Nonlinear loudspeaker compensation for hands free acoustic echo cancellation," Electronics Letters vol. 32, issue 12, pp. 1063-1064, Jun. 1996.
- [10] V. J. Mathews, "Adaptive Volterra filters using orthogonal structures," IEEE Signal Processing Letter, vol. 3, issue 12, pp. 307-309, Dec. 1996.
- [11] A. Stenger, L. Trautmann, and R. Rabenstein, "Nonlinear acoustic echo cancellation with 2<sup>nd</sup> order adaptive Volterra filters," ICASSP proceeding, vol. 2, pp. 877-880, Nov. 1999.
- [12] A. Guerin, G. Faucon, and Le Bouquin-Jeannes, R., "Nonlinear acoustic echo cancellation based on Volterra filters," IEEE Trans. Speech and Audio Processing vol. 11, no. 6, pp. 672-683, Nov. 2003.
- [13] J. P. Costa, T. Pitarque, and E. Thierry, "Using orthogonal least square identification for adaptive nonlinear filtering of GSM signals," ICASSP, vol. 3 pp.2397-2400, 1997.
- [14] A. N. Birkett and R. A. Goubran, "Acoustic echo cancellation using NLMS-neural network structures," ICASSP, pp.3035-3038 vol. 5, May 1995.

- [15] K. Narendra and P. Gallman, "An iterative method for the identification of nonlinear systems using a Hammerstein model," IEEE Trans. Automatic Control, vol. 11, issue 13, pp. 546-550. Jul. 1966.
- [16] K. S. H. Ngia and J. Sjöbert, , "Nonlinear acoustic echo cancellation using a Hammerstein model," ICASSP, vol. 2 pp.1229-1232 May. 1998.
- [17] J. Schultheiss and L. del Re, "Neural networks identification of muscular response using extended Hammerstein models," Engineering in Medicine and Biology Society, vol. 5 pp. 2566-2569. Nov. 1998.
- [18] A. Stenger and W. Kellermann, "Nonlinear acoustic echo cancellation with fast converging memoryless preprocessor," ICASSP, vol. 2, pp. 805–808, Jun. 2000.
- [19] H. Dai and W. P. Zhu, "An acoustic echo cancellation scheme using raised-cosine function for nonlinear compensation," TENCOM, vol. A, pp.131-134, Nov. 2004.
- [20] F. Kuech, A. Mitnacht, and W. Kellermann, "Nonlinear acoustic echo cancellation using adaptive orthogonalized power filters," ICASSP, vol.3, pp. 105-108, March 2005.
- [21] A. Tchamkerten, and I. E. Telatar, " On the use of training sequences for channel estimation," IEEE Trans. information theory, vol. 52, no. 3, Mar. 2006.

- [22] P. B. Rapajic and B. Vucetic, "Adaptive receiver structures for asynchronous CDMA systems," *IEEE J. Select. Areas in Commun.*, vol. 12, pp. 685-697, May 1994.
- [23] G. Long, and F. Ling, "Fast initialization of data-driven Nyquist in-band echo cancellers," *IEEE Trans. Commun.* vol.41, no. 6, pp.893-904, June 1993
- [24] S. Qureshi, "Fast start-up equalization with periodic training sequences," *IEEE Trans. Information Theory*, vol. IT-23, pp. 553-563, Sept. 1977.
- [25] W. K. Jenkins, C. W. Therrien, and X. Li, "Orthogonal polynomial-based nonlinear adaptive filters," *IEEE Southeast Con*, pp. 444-446, Apr. 1996.
- [26] N. J. Bershad, P. Celka, and S. McLaughlin, "Analysis of stochastic gradient identification of Wiener–Hammerstein systems for nonlinearities with Hermite polynomial expansions," *IEEE Trans. Signal Process*, vol. 49, pp. 1060–1071, May 2001.