SPONTANEOUS AND EXPLICIT ESTIMATION OF TIME DELAYS IN THE ABSENCE/PRESENCE OF MULTIPATH PROPAGATION

BY

HING-CHEUNG SO

(蘇慶祥)

A THESIS

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DIVISION OF ELECTRONIC ENGINEERING

THE CHINESE UNIVERSITY OF HONG KONG

MAY 1995

To my parents,

and Wing-Chun.

Acknowledgement

First of all, I would like to express my earnest appreciation to my supervisor, Dr. Pak-Chung Ching, for the interesting problem he introduced to me and for providing always a stimulating feedback during the time of this research. His insightful advice, guidance and patience were invaluable. I also wish to thank Prof. Yiu-Tong Chan and Dr. King-Choi Ho for their helpful discussions and support.

I would like to remember all the colleagues and friends that shared with me in my graduate study. To name only some of them: Dr. Alan Kai-Hau Yeung, Quentin Kwok-Wai Wong, Man-Hong Ko, Sam Wing-Cheong Chan, Wai-Kuen Lai, Tan Lee, Simon Kim-Ming Chiu, Brian Kan-Wing Mak, Terry Kin-Chung Lo, William Chi-Min Pang, Francis Cho-Yiu Chik, Timothy Kai-Cheung Chung, Wai-Kit Lo and many others.

I wish to express my deepest gratitude to my parents for their disciplines and encouragement throughout my whole education. Special thanks are due to my dearest Wing-Chun for her sincere endurance and support and for staying with me in the long busy days that I needed to complete this work.

Finally and above all, I am grateful to my true and faithful God who has provided me peace and wisdom and let me experience "For nothing is impossible with God" throughout my life, particularly during the research.

This work was supported in part by Earmarked Grant for Research #CUHK 63/91 awarded by the Hong Kong Research Grant Council.

Abstract

In this dissertation, novel adaptive filtering techniques are proposed for the problem of time delay estimation (TDE) which has various applications such as source localization and speed measurement. A new TDE algorithm, called the explicit time delay estimator (ETDE), is first developed to find the differential delay of a signal received at two spatially separated sensors with the assumption that the signal arrives at each receiver through only one propagation path. Basically, the ETDE is an adaptive FIR filter whose coefficients are expressed as a function of the delay estimate which is adjusted using an LMS-type algorithm and thus it can provide direct delay measurements on a sample-by-sample basis. The ETDE performance surface is multimodal and an initialization method is suggested to ensure global convergence. Convergence dynamics and delay variance for both static and nonstationary delays are also given. Although the ETDE and the constrained adaptive time delay estimator (CATDE) impose the same constraint on the filter coefficients, the ETDE gives more accurate delay estimates under noisy environment. The ETDE has an advantage over another direct TDE method, the adaptive digital delay lock discriminator (ADDLD), which restricts the estimated delay to be an integral multiple of the sampling period, because it can model any real-valued delay. Furthermore, the ETDE is more computational efficient than the least mean square time delay estimator (LMSTDE) whose delay estimate is obtained by interpolating its filter weights and it is more reliable because wrong estimates due to false peak coefficients will not occur. The mean square delay error of the ETDE is shown to be comparable to the Cramér-Rao lower bound (CRLB) for a wide range of signal-to-noise ratio (SNR). Based on Wiener solution, a variable gain control is added to the ETDE to decouple

the effect of time-varying signal and/or noise power. As a result, this improved ETDE, namely, the explicit time delay and gain estimator (ETDGE), provides a smaller delay variance at low SNR and an unbiased delay estimate for any finite filter order. Experimental results are included to validate the theoretical analysis and evaluate the performance of these two algorithms.

With the use of the ETDE and the ETDGE, two adaptive systems, called the multipath cancellation time delay estimator (MCTDE) and the multipath equalization time delay estimator (METDE) respectively, are developed to generalize the TDE problem when multipath propagation is taken into account. Prior to using the ETDE to model the time difference between the two sensor outputs, the MCTDE employs two adaptive IIR filters to eliminate the multipath component at each transmission channel. When the multipath delays are integral multiples of the sampling interval, the MCTDE achieves the optimum performance. Otherwise, due to inherent truncation error in the multipath cancellers, the accuracy of the system parameter estimates decreases as the multipath gain increases or the interpath delay decreases. On the other hand, the METDE, which comprises the ETDE and the ETDGE only, allows exact modeling of the time delay as well as the multipath parameters. The idea of the METDE is to equalize the two received signals through minimization of a cost function whose global minimum contains the actual value of the time delay and multipath parameters, and it can be applied when each sensor sees more than one multipath signal. Computer simulations are presented which show that these two methods can accurately estimate and track nonstationary delay parameters under high SNR conditions.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Time Delay Estimation (TDE) and its Applications

The problem of estimating and tracking the time delay between two noisy versions of the same signal received at two spatially separated sensors has attracted much attention in many literatures [1], [2]. Time delay estimation (TDE) has been widely used in sonar and radar to locate the position and to detect the speed of a signal source [3], [4]. A typical passive sonar example is shown in Figure 1.1. The submarine radiates a noise-like signal due to machinery on board or propeller action. This signal propagates through the water and is received by an array of two omnidirectional sensors. Because of the extra propagation distance, the wavefront at sensor A lags that at sensor B by a delay, say, D. When the array is located far from the target, the bearing of the submarine, denoted by θ, can be approximated by [5]

$$\theta \approx \cos^{-1}\left(\frac{cD}{L}\right) \tag{1.1}$$

where L is the distance between the sensors and c is the speed of sound in water. Therefore, if we can estimate the time difference of the two received signals, then a bearing estimate will also be acquired. Notice that in practice, the received waveforms are embedded in corrupting noise and multipath propagation often arises from the sea surface reflection or bottom bounce [5]. These make accurate determination of D a very difficult task to achieve, particularly when the number of multipath signals is large.

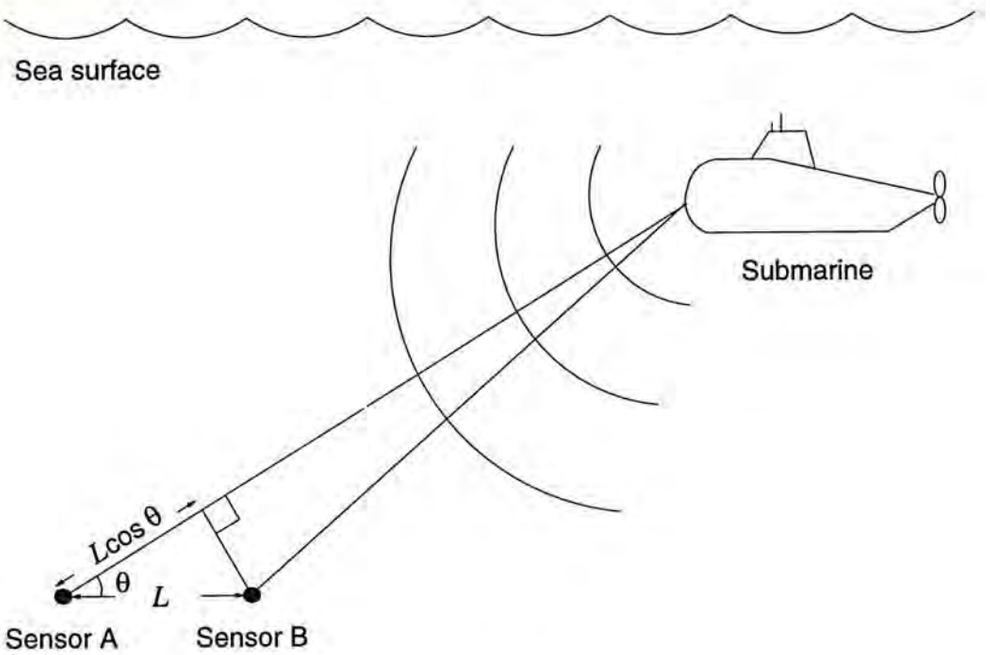

Figure 1.1: Source localization by passive sonar

If there are three sensors instead of two, two bearing estimates can be provided. The point at which these two bearings intersect will yield a position estimate of the radiated source [2]. Furthermore, if the source is moving, its velocity can be obtained from the time-varying delay measurements [4].

A similar but more recent application of TDE is in global positioning system (GPS) [6],[7]. In contrast with other geolocation approaches [8],[9], this method does not require disruption of the normal satellite operation since delay estimation is accomplished passively. Another important advantage of the method is that additional spaceborne hardware is not needed which makes the TDE approach more cost-effective. By using the differential delay measurements of an uplink signal received by three or four geostationary satellites, the location of the corresponding target transmitter can be determined [7]. Three satellites are needed when the emitter is on the earth surface and four if its altitude is not known *a priori*. In actual circumstances, the transmitter may be an interfering source to a domestic geosynchronous satellite communication system.

TDE has also found many useful applications in biomedical engineering [10]-[14]. As an example, Figure 1.2 illustrates a monopolar electrode configuration for sensory nerve conduction velocity measurement which is of importance to the diagnosis of neuromuscular disorders [10]. Excited by the stimulating electrodes at the finger, two evoked response signals, which differ by a delay, are obtained from the monopolar electrodes. The time displacement of one electrode signal with respect to the second gives the conduction delay over the distance d and hence the nerve conduction velocity at the wrist.

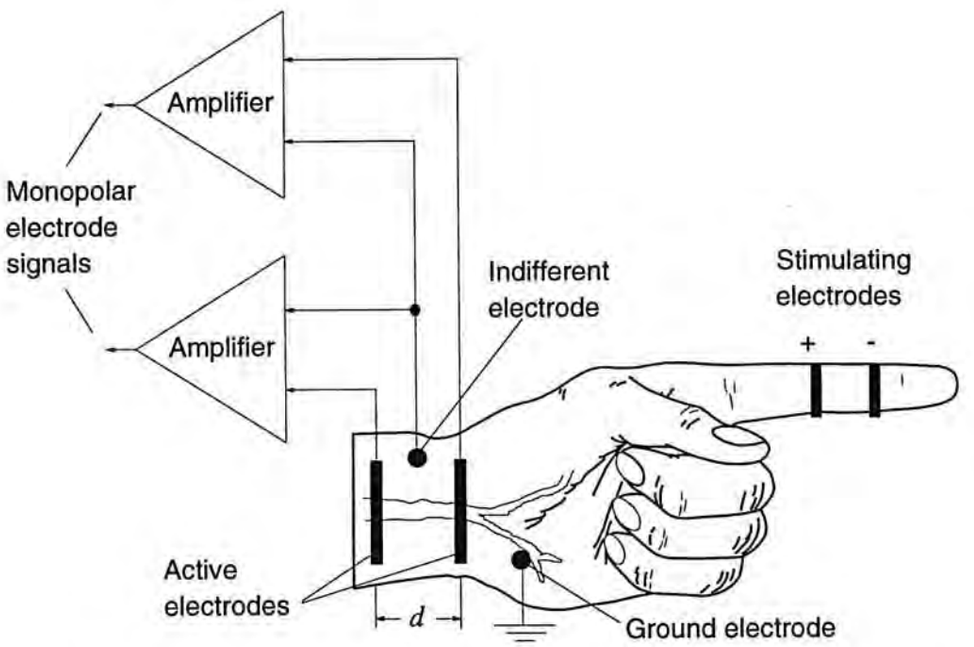

Figure 1.2: Sensory nerve conduction velocity measurement

Other areas that make use of TDE techniques include seismology [15], noncontact speed measurement [16]-[18], digital communication [19]-[22], image coding [23] and tomography [24]-[25].

Let the waveforms received at two spatially separated sensors be $x(t)$ and $y(t)$. A widely used mathematical model for time delay estimation can be expressed as follows [5],

$$x(t) = s(t) + n_1(t) \tag{1.2a}$$

and

$$y(t) = s(t - D) + n_2(t) \tag{1.2b}$$

where the unknown source signal $s(t)$ and the corrupting noises $n_1(t)$ and $n_2(t)$ are assumed to be Gaussian, jointly stationary and mutually uncorrelated with each other while the parameter D is the time difference of arrival (TDOA) between the two received signals. Although the TDE model described here is for two sensors, it can be easily extended to cases when there are more than two sensors.

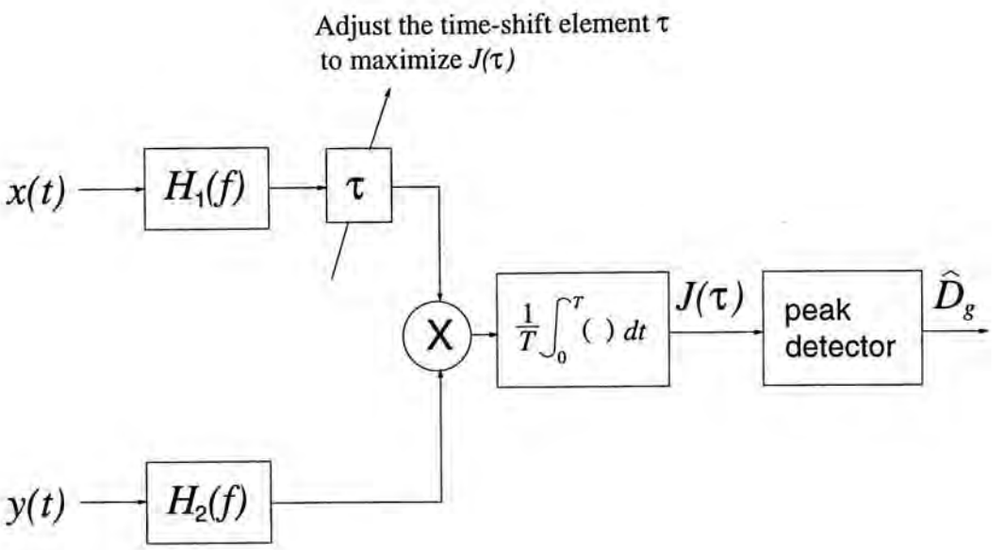

Figure 1.3: A generalized cross correlator configuration

One of the conventional approaches to estimate the delay D is based on the generalized cross correlation (GCC) methods [26]-[28]. The system block diagram of a generalized cross correlator is shown in Figure 1.3. It can be realized as a pair of receiver prefilters, $H_1(f)$ and $H_2(f)$, followed by a cross correlator. In general, the role of the prefilters is to enhance the frequency bands where the signal is strong and to attenuate the bands where the noise is strong [2]. When the observation time T is much longer than the correlation length of $x(t)$ and $y(t)$, the output of the correlator, $J(\tau)$, is given by [26]

$$J(\tau) = \int_{-\infty}^{\infty} H_1(f) H_2^*(f) G_{xy}(f) e^{j2\pi f\tau} df \tag{1.3}$$

where

$$G_{xy}(f) = \int_{-\infty}^{\infty} E\{x(t)y(t-\tau)\}e^{-j2\pi f\tau}d\tau \tag{1.4}$$

is the cross-power spectrum between $x(t)$ and $y(t)$. The symbol $*$ stands for the complex conjugate while E denotes the expectation operator. The delay estimate $\hat{D}_g$ is equal to the time argument at which $J(\tau)$ achieves its maximum value, that is,

$$\hat{D}_g = \arg\max_{\tau} J(\tau) \tag{1.5}$$

With proper choice of the weighting functions $H_1(f)$ and $H_2(f)$, minimum delay variance can be achieved [26]. However, the weights of the prefilters are dependent on the signal and noise spectra which are generally unknown and hence have to be estimated in practice.

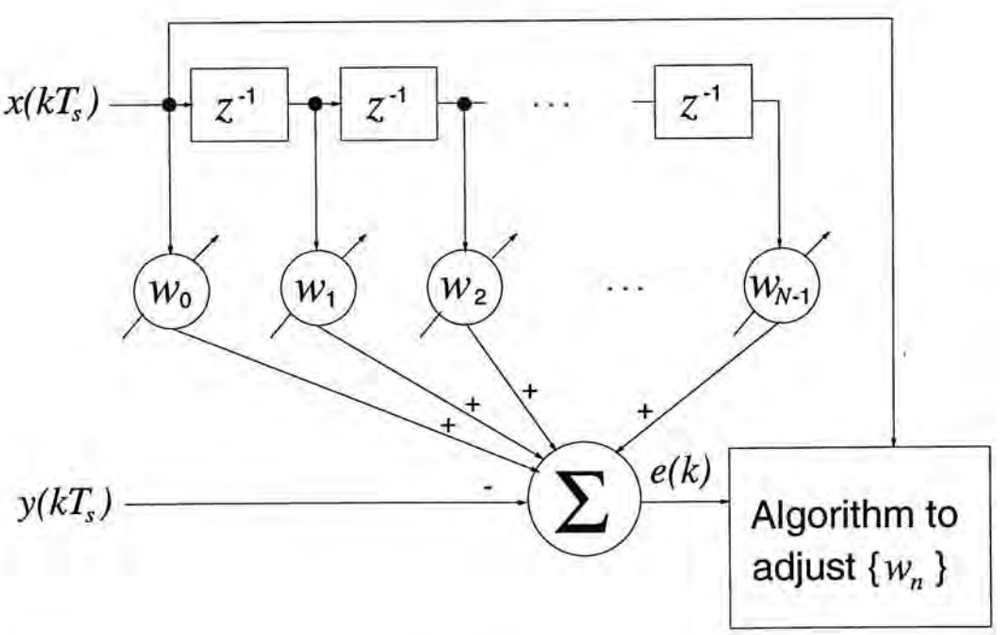

Figure 1.4: An adaptive filter system for time delay estimation

On the other hand, adaptive filtering techniques [29]-[43] have also been applied to the TDE problem with great success. Basically, an adaptive filter is a time-variant filter whose coefficients are adjusted in a way to satisfy some predetermined optimization criterion such as to achieve the minimum mean square error. These methods can operate in circumstances which are not known *a priori* and some of

them may even work under nonstationary environment. The performance of most adaptive TDE techniques is mainly governed by the filter structure as well as the updating algorithm. A typical adaptive filter system for time delay estimation is shown in Figure 1.4 where the received signals of (1.2) are sampled with a period of T_s and are of the form,

$$x(kT_s) = s(kT_s) + n_1(kT_s) \tag{1.6a}$$

and

$$y(kT_s) = s(kT_s - D) + n_2(kT_s) \tag{1.6b}$$

The transversal filter is used to provide a proper time shift to the input $x(kT_s)$ and its filter coefficients, $\{w_n\}$, for $0 \leq n \leq N-1$, are adjusted to minimize the mean square value of the output error $e(k)$. One can employ Widrow's LMS algorithm [44],[45] to adjust the filter weights and the resultant system is known as the least mean square time delay estimator (LMSTDE) [29]-[30]. Upon convergence, the delay estimate, denoted by $\hat{D}_a$, is determined by interpolating the filter coefficients and is given by [29]

$$\hat{D}_a = \arg\max_t \hat{w}(t) \tag{1.7}$$

where

$$\hat{w}(t) = \sum_{n=0}^{N-1} w_n \frac{\sin(2\pi\beta(t - nT_s))}{2\pi\beta(t - nT_s)} \tag{1.8}$$

and β is the source signal bandwidth. Apparently, $\hat{w}(t)$ is a truncated expansion of the continuous-time filter weights and the accuracy of $\hat{D}_a$ increases with the filter length [35]. Since only the peak of the filter coefficients is of interest, numerical methods such as bisection can be used to obtain the delay estimate.

1.2 Goal of the Work

Although many GCC based methods as well as adaptive filtering approaches have been proposed for time delay estimation in the past decades, the issue of multipath propagation is seldom taken into account in most of these studies. When the signal source arrives at each receiver through more than one path, which is very common

in many applications such as sonar [5] and radar [46], the propagation structure will become more complicated than the simple model of (1.2) or (1.6). In this case, many well known methods such as the LMSTDE will fail to provide accurate delay estimates unless the multipath effect is properly compensated.

In the presence of multipath transmissions, the received waveforms of two spatially separated sensors can be generalized as

$$x(kT_s) = s(kT_s) + \sum_{p=1}^{M_1} \alpha_{1p} s(kT_s - \Delta_{1p}) + n_1(kT_s) \tag{1.9a}$$

$$y(kT_s) = s(kT_s - D) + \sum_{q=1}^{M_2} \alpha_{2q} s(kT_s - \Delta_{2q}) + n_2(kT_s) \tag{1.9b}$$

The gain factors α_{1p} and α_{2q} as well as the interpath delays Δ_{1p} and Δ_{2q}, for $M_1 \geq p \geq 1$ and $M_2 \geq q \geq 1$, characterize the multipath propagation. While M_1 and M_2 are the numbers of multipaths contained in $x(kT_s)$ and $y(kT_s)$ respectively.

Some of the previous work on multipath time delay estimation only aims at finding the multipath parameters within a single sensor. Examples include cepstral filtering [47], autocorrelation, maximum likelihood estimation (MLE) [48]-[49] and adaptive inverse filtering [50]. The estimates of the multipath parameters can be used to improve the source localization performance [51]. Furthermore, the source position can be acquired with a single sensor in the presence of multipath transmissions [51],[52]. In order to determine the differential delay between the two sensors, they may combine with other standard TDE methods. Ching *et al* [53] appear to be the only authors that have suggested techniques for estimating the delay D as well as the multipath parameters in the multipath TDE problem when the source $s(kT_s)$ is an unknown signal. However, the algorithms proposed are designed for use when there is exactly one multipath in either one of the two receivers and thus their application is limited.

The main objective of this dissertation is to develop novel adaptive algorithms for accurate estimation and tracking of the differential delay between the two sensor outputs, D, and also the multipath parameters, α_{1p}, α_{2q}, Δ_{1p} and Δ_{2q}, if each sensor sees at least one multipath signal. Our study is focused on adaptive TDE techniques

because it has three advantages over the conventional GCC methods. Firstly, they do not require spectral estimation of the transmitted source nor the corrupting noises. Secondly, they are in general simple to implement and involve less computation. Last but not least, this method is capable of tracking nonstationary delays due to either relative source/receiver motion or time-varying characteristics of the transmission medium. We limit our investigation to the use of LMS-type stochastic gradient techniques although there exist more powerful methods for adaptive filtering [45],[54]-[56] such as the recursive least squares and the quasi-Newton algorithms. It is mainly because LMS-type algorithms do not require measurements of the pertinent correlation functions, nor do they need matrix inversion, and hence real-time implementation is allowed. Moreover, they have good tracking ability in time-varying systems and are often useful in nonstationary environment.

In the absence of multipath transmission, two new adaptive algorithms have been derived to estimate the differential delay between signals received at two spatially separated sensors. It is assumed that there is no multipath signal at each sensor and the received signals are given by (1.6). The aim of these two algorithms is to achieve better performance, both in computational complexity and estimation accuracy, over other adaptive TDE methods. By imposing constraints on the filter coefficients, the proposed methods provide explicit delay measurements and no interpolation using (1.7) as in [29]-[35] is necessary. As a result, less computations are involved and the delay estimates are also free from interpolation error.

For time delay estimation in the presence of multipath propagation, we attempt to remove any restriction on the number of multipaths. Using the idea of multipath cancellation and multipath equalizaton respectively, two adaptive systems are developed to determine the TDOA as well as the multipath parameters when the sensor outputs are given by (1.9). It has been demonstrated that the first method can give accurate delay and multipath parameter estimates when there is a multipath in each sensor while the second method provides exact delay modeling irrespective of the number of existing multipaths.

1.3 Thesis Outline

In Chapter 2, two different classic adaptive TDE approaches, namely, the LMSTDE and the adaptive digital delay-lock discriminator (ADDLD) [40], are first discussed briefly. Motivated by these two methods and using the property that a delayed version of a bandlimited signal can be represented by convolving a *sinc* function with the signal itself [35], the explicit time delay estimator (ETDE) is derived in Chapter 3. Generally speaking, the ETDE is an adaptive FIR filter which provides a direct measure of time delay between signals received at two spatially separated sensors on a sample-by-sample basis. Performance analysis of the ETDE for static and nonstationary delays, as well as comparison with other conventional TDE methods and the Cramér-Rao lower bound (CRLB), are included. It is shown that the delay error of the ETDE will increase when either the signal-to-noise ratio or the number of filter taps is decreased. An improvement is then made to the delay estimator to alleviate the problem and this is described in Chapter 4. This new time delay estimator is analyzed in detail and it is proved that smaller delay variances and unbiased delay estimates for any finite filter length can be attained if a variable gain control is added to the ETDE.

In the presence of multipath propagation, accurate delay estimation can still be achieved using the ETDE or the ETDGE if the multipath effect in the received signals is removed. Based on the concept of multipath cancellation and multipath equalization, and applying the explicit algorithms developed in Chapter 3 and Chapter 4, two adaptive systems are proposed in Chapter 5 to extract the TDOA as well as the multipath parameters when each sensor receives the source signal plus its attenuated and delayed replicas. Theoretical analysis of these two methods, namely, the multipath cancellation time delay estimator (MCTDE) and the multipath equalization time delay estimator (METDE) are given. Finally, concluding remarks and further research topics are presented in Chapter 6.

Chapter 2

Adaptive Methods for TDE

Many adaptive time delay estimation (TDE) methods [29]-[43] have been proposed in the past two decades, we shall briefly review the basic principle of operation of the least mean square time delay estimator (LMSTDE) [29]-[30] and the adaptive digital delay-lock discriminator (ADDLD) [40] in this chapter. These two different approaches have provided the groundworks and motivation for part of this research and as a result a novel TDE scheme has been designed. The so-called explicit time delay estimator (ETDE) is developed in the early stage of this work and then an improved version, namely, the explicit time delay and gain estimator (ETDGE), in the second phase of the project. In Section 2.1, the problem of TDE in the absence of multipath transmission is first formulated. Based on Widrow's least mean square (LMS) algorithm [44],[45], the LMSTDE uses an adaptive FIR filter to model the time difference of arrival (TDOA) between signals received at two spatially separated sensors and the delay estimate is obtained by interpolating the filter coefficients. This method together with some of the modifications for enhancement [31]-[32],[36] are discussed in Section 2.2. Section 2.3, on the other hand, describes another LMS-type TDE algorithm, the ADDLD, which provides direct measurement of time delay and no interpolation is necessary.

2.1 Problem Description

Typically, the problem of time delay estimation can be formulated as follows. Given the discrete-time outputs of two sensors,

$$x(k) = s(k) + n_1(k) \tag{2.1a}$$

and

$$y(k) = s(k - D) + n_2(k) \tag{2.1b}$$

where the source signal $s(k)$ and the corrupting noises $n_1(k)$ and $n_2(k)$ are assumed to be mutually uncorrelated, stationary and white Gaussian processes. Without loss of generality, the sampling period T_s is taken to be unity. Thus the spectra of the signal and additive noises are supposed to be flat and band-limited between $-\pi$ and π. For ease of analysis, the signal arrives at each sensor is assumed to go through only one propagation path in the same plane with the receivers and source. The objective is to estimate the time difference, D, from the received signals $x(k)$ and $y(k)$.

2.2 The Least Mean Square Time Delay Estimator (LMSTDE)

Figure 2.1 shows the schematic diagram of the conventional least mean square time delay estimator (LMSTDE) [29]-[30]. The two channel inputs, $x(k)$ and $y(k)$, are sensor outputs as described by (2.1). Basically, it is an adaptive FIR filter which is used to provide a time shift to the input signal $x(k)$. The transfer function of the LMSTDE is given by

$$W(z) = \sum_{i=-P}^{P} w_i z^{-i} \tag{2.2}$$

The filter in this case is noncausal and has an order of $2P$, but, of course, it can be easily realized in causal form. The output of the filter is subtracted from the other received signal $y(k)$ to give an error function $e(k)$. In noise-free condition, if $e(k)$

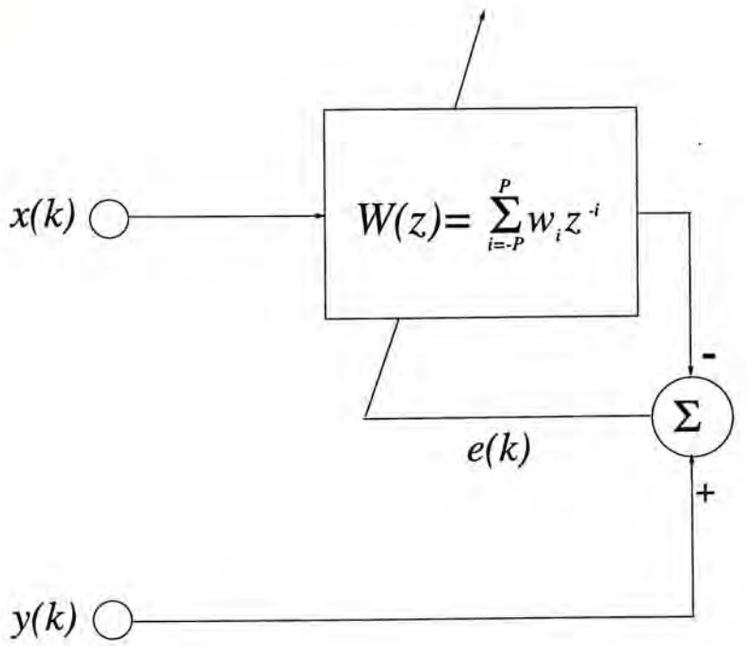

Figure 2.1: System block diagram of the LMSTDE

goes to zero, then the time shift introduced by the adaptive filter will be equal to the actual delay between the two signals. The error signal, $e(k)$, is given by

$$e(k) = y(k) - \sum_{i=-P}^{P} w_i(k)x(k-i) \tag{2.3}$$

Each filter coefficient of $W(z)$, viz., $w_i(k)$, is adjusted by minimizing the mean square error, $E\{e^2(k)\}$, according to Widrow's LMS algorithm [44],[45],

$$\begin{aligned} w_i(k+1) &= w_i(k) - \mu_w \frac{\partial e^2(k)}{\partial w_i(k)}, \qquad P \geq i \geq -P \\ &= w_i(k) + 2\mu_w e(k)x(k-i) \end{aligned} \tag{2.4}$$

The parameter μ_w is a positive scalar that controls the convergence rate and stability of the algorithm.

From the convolution theorem, it has been shown that a delayed version of $s(k)$, say, $s(k-D)$, can be expressed as [35]

$$s(k-D) = \sum_{i=-\infty}^{\infty} sinc(i-D)s(k-i) \tag{2.5}$$

where

$$sinc(v) = \frac{\sin(\pi v)}{\pi v} \tag{2.6}$$

From (2.3) and (2.5), if P tends to infinity and if the filter coefficients of $W(z)$ are samples of a *sinc* function, then the LMSTDE will introduce a proper delay to the input signal.

Assume that the channel inputs and the filter weights at time k are uncorrelated [57] and denote σ_s^2 and σ_n^2 as the signal power and the variance of $n_1(k)$ and $n_2(k)$ respectively. Substituting (2.1), (2.3) and (2.5) into (2.4) and taking expectation yields

$$\begin{aligned}
E\{w_i(k+1)\} &= E\{w_i(k)\} + 2\mu_w E\{(s(k-D) + n_2(k) - \sum_{n=-P}^{P} w_n(k)s(k-n) \\
&\quad - \sum_{n=-P}^{P} w_n(k)n_1(k-n)) \cdot (s(k-i) + n_1(k-i))\} \\
&= E\{w_i(k)\} + 2\mu_w E\{(s(k-D) - \sum_{n=-P}^{P} w_n(k)s(k-n))s(k-i)\} \\
&\quad -2\mu_w E\{\sum_{n=-P}^{P} w_n(k)n_1(k-n)n_1(k-i)\} \\
&= E\{w_i(k)\} + 2\mu_w E\{(\sum_{n=-\infty}^{\infty} sinc(n-D)s(k-n) \\
&\quad - \sum_{n=-P}^{P} w_n(k)s(k-n)) \cdot s(k-i)\} - 2\mu_w E\{w_i(k)n_1^2(k-i)\} \\
&= E\{w_i(k)\} + 2\mu_w \sigma_s^2 E\{sinc(i-D) - w_i(k)\} - 2\mu_w \sigma_n^2 E\{w_i(k)\} \\
&= E\{w_i(k)\}(1 - 2\mu_w(\sigma_s^2 + \sigma_n^2)) + 2\mu_w sinc(i-D) \qquad (2.7)
\end{aligned}$$

Solving (2.7) gives

$$\begin{aligned}
E\{w_i(k+1)\} &= \frac{\sigma_s^2}{\sigma_s^2 + \sigma_n^2} sinc(i-D)(1 - (1 - 2\mu_w(\sigma_s^2 + \sigma_n^2))^k) \\
&\quad + w_i(0)(1 - 2\mu_w(\sigma_s^2 + \sigma_n^2))^k, \qquad P \geq i \geq -P \qquad (2.8)
\end{aligned}$$

provided that the following is being satisfied,

$$0 < \mu_w < \frac{1}{\sigma_s^2 + \sigma_n^2} \qquad (2.9)$$

The initial values of the filter weights, $\{w_i(0)\}$, can be selected arbitrarily since the performance surface to be minimized is unimodal [45]. The filter coefficients will eventually converge to the Wiener solution, $\{w_i^o\}$, which is equal to the steady state

value of (2.8) and has the form

$$w_i^o = \frac{\text{SNR}}{1+\text{SNR}} sinc(i-D), \qquad P \geq i \geq -P \tag{2.10}$$

where $\text{SNR} = \sigma_s^2/\sigma_n^2$ is the signal-to-noise ratio. Upon convergence, the delay estimate of the LMSTDE, $\hat{D}_w(k)$, is obtained indirectly from $\{w_i(k)\}$ by interpolation and it is given by [34],[35]

$$\hat{D}_w(k) = \arg\max_t \{\psi(t)\} \tag{2.11}$$

where

$$\psi(t) = \sum_{i=-P}^{P} w_i(k) sinc(t-i) \tag{2.12}$$

Notice that $\psi(t)$ represents the continuous-time version of $\{w_i(k)\}$ since it is the convolution of $\{w_i(k)\}$ and $sinc(t)$ where $sinc(t)$ is the impulse response of an ideal low-pass filter with a bandwidth of 0.5.

2.2.1 Bias and Variance

Due to the finite number of filter taps being used, the expected value of $\psi(t)$ may not necessarily have a peak at exactly $t = D$. Using (2.10) and (2.11), the delay estimate can in fact be found by solving the following equation,

$$\frac{d}{dt} \sum_{i=-P}^{P} w_i^o sinc(t-i) = 0$$

$$\Rightarrow \sum_{i=-P}^{P} sinc(i-D) f(t-i) = 0 \tag{2.13}$$

where

$$f(v) = \frac{\cos(\pi v) - sinc(v)}{v} \tag{2.14}$$

This delay modeling error has been examined in [35] for different values of D and P. It is shown that the deviation from D is negligible if a sufficiently long filter length is employed. For example, given $D \in (0, 0.5)$, the largest possible error is about

8.2% when $P = 5$ but this error will drop to 4% if the filter length is increased to 21.

When steady state is reached, the variance of the delay estimate, denoted by $var(\hat{D}_w)$, can be expressed as [39]

$$\begin{aligned} var(\hat{D}_w) &\triangleq \lim_{k\to\infty} E\{(\hat{D}_w(k) - D)^2\} \\ &= \left. \frac{(1+\mathrm{SNR}^2)\sum_{i=-P}^{P} var(w_i) sinc'^2(i-t)}{\mathrm{SNR}^2\left(\sum_{i=-P}^{P} sinc(i-D)sinc''(i-t)\right)^2} \right|_{t=D} \end{aligned} \tag{2.15}$$

where the symbol $'$ and $''$ represent the first and second derivative respectively and $var(w_i)$ is the variance of the filter weight $w_i(k)$. Define

$$\eta = \sum_{i=-\infty}^{\infty} sinc(i-D)sinc''(i-t)|_{t=D} \tag{2.16}$$

By using L'Hôspital's rule, η can be evaluated as

$$\begin{aligned} \eta &= \sum_{i=-\infty}^{\infty} sinc(i-D)\left(\frac{2sinc(i-D) - 2\cos(\pi(i-D)) - \pi(i-D)\sin(\pi(i-D))}{(i-D)^2}\right) \\ &= \lim_{\Delta\to 0} \frac{2sinc(\Delta) - 2\cos(\pi\Delta) - \pi\Delta\sin(\pi\Delta)}{\Delta^2} \\ &= \lim_{\Delta\to 0}\left(\frac{2\sin(\pi\Delta)}{\pi\Delta^3} - \frac{2\cos(\pi\Delta) + \pi\Delta\sin(\pi\Delta)}{\Delta^2}\right) \\ &= -\frac{\pi^2}{3} \end{aligned} \tag{2.17}$$

Moreover, from [58] and [45], we have

$$\sum_{i=-\infty}^{\infty} sinc'^2(i-t)|_{t=D} = \sum_{i=-\infty}^{\infty} f^2(i-D) = \frac{\pi^2}{3} \tag{2.18}$$

for integral values of D and

$$var(w_i) \approx \mu_w \sigma_n^2 \left(1 + \frac{\mathrm{SNR}}{1+\mathrm{SNR}}\right) \tag{2.19}$$

Hence, (2.15) can be simplified to

$$var(\hat{D}_w) \approx \frac{3\mu_w\sigma_n^2(1+2\mathrm{SNR})(1+\mathrm{SNR})}{\pi^2\mathrm{SNR}^2} \tag{2.20}$$

For high and low SNR conditions, we can write

$$var(\hat{D}_w) = \begin{cases} \dfrac{6\mu_w\sigma_n^2}{\pi^2} & , \ \text{SNR} \gg 1 \\ \\ \dfrac{3\mu_w\sigma_n^2}{\pi^2\text{SNR}^2} & , \ \text{SNR} \ll 1 \end{cases} \tag{2.21}$$

It can be seen that the value of $var(\hat{D}_w)$ is proportional to μ_w and σ_n^2. Furthermore, the delay variance is inversely proportional to SNR^2 under very noisy environments.

2.2.2 Probability of Occurrence of False Peak Weight

From (2.11), it is observed that the value of $\hat{D}_w(k)$ is mainly determined by the location of the filter weight that has the largest magnitude. Let $w_L(k)$, $L \in [-P, P]$, be the largest filter tap, then L essentially gives the round-off integral value of the delay estimate. However, due to the noisy gradients in (2.4), the desired peak weight might not be located accurately [32],[59] at each iteration. When D is an integer, the probability of occurrence of a false peak, $P(e)$, is given by [59]

$$P(e) = 1 - \frac{1}{\sqrt{2\pi\mu_w\sigma_n^2}} \int_{-\infty}^{\infty} \prod_{i \neq D} Q\left(\frac{E\{w_D(k)\} - E\{w_i(k)\} + x}{\sqrt{\mu_w\sigma_n^2}}\right) e^{-\frac{x^2}{2\mu_w\sigma_n^2}} dx \tag{2.22}$$

where

$$Q(v) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{v} exp\left(-\frac{x^2}{2}\right) dx \tag{2.23}$$

Using (2.8) and setting all $\{w_i(0)\}$ to zero, it has been found [59] that $P(e)$ decreases with increasing I where I is defined as

$$I = \frac{\text{SNR}}{1 + \text{SNR}} \cdot \frac{1 - (1 - 2\mu_w\sigma_n^2(1 + \text{SNR}))^k}{\sqrt{\mu_w\sigma_n^2}} \tag{2.24}$$

It is noticed that when the SNR is small, the value of I increases with k and the SNR and decreases with μ_w.

2.2.3 Some Modifications of the LMSTDE

Youn *et al* [31]-[32] have modified (2.4) by replacing μ_w by $\mu_w(k)$, which is a time-varying convergence parameter of the form

$$\mu_w(k) = \frac{1-\gamma}{2\hat{\sigma}_x^2(k)} \tag{2.25}$$

where γ is a smoothing factor and has a value between 0 and 1. The quantity $\hat{\sigma}_x^2(k)$ is an estimate of the variance of $x(k)$ and is updated via the relation

$$\hat{\sigma}_x^2(k) = \gamma\hat{\sigma}_x^2(k-1) + (1-\gamma)x^2(k) \tag{2.26}$$

As $x(k)$ is stationary, $\hat{\sigma}_x^2(k)$ will approach its desired value of $(\sigma_s^2 + \sigma_n^2)$ in steady state. Using the optimal value for $\hat{\sigma}_x^2(k)$ and substituting (2.25) into (2.8), the learning trajectory of the filter coefficients becomes

$$E\{w_i(k)\} = \frac{\sigma_s^2}{\sigma_s^2+\sigma_n^2} sinc(i-D)(1-\gamma^k) + w_i(0)\gamma^k, \quad |i| \leq P \tag{2.27}$$

The advantage of this method is that the convergence rate of $\{w_i(k)\}$ is now controlled by the predefined parameter γ only and is not affected by the power of $x(k)$ which is often an unknown value.

From (2.5), we see that a delayed version of a bandlimited signal can be represented by convolving a *sinc* function with the signal itself. Using this property, Ching and Chan [36] have proposed a new constrained adaptive time delay estimation (CATDE) algorithm, which makes use of the constraint that the FIR filter coefficients must be samples of a *sinc* function, to model the estimated delay. It simplifies the LMSTDE algorithm considerably because in this case only $w_L(k)$ needs to be adapted according to (2.4) in each iteration. The delay estimate of this method, denoted by $\hat{D}_c(k)$, is related to $w_L(k)$ by the following unique mapping

$$w_L(k) = sinc(L - \hat{D}_c(k)) \tag{2.28}$$

The values of the remaining filter weights are determined by $w_i(k) = sinc(i - \hat{D}_c(k))$ and they can be easily found by a table lookup operation. As a result, the computation is reduced substantially and also no interpolation of filter weights

is required. Furthermore, it has been demonstrated [36] that the TDE process can converge much faster even under a high SNR environment by imposing this condition on the coefficients.

2.3 The Adaptive Digital Delay-Lock Discriminator (ADDLD)

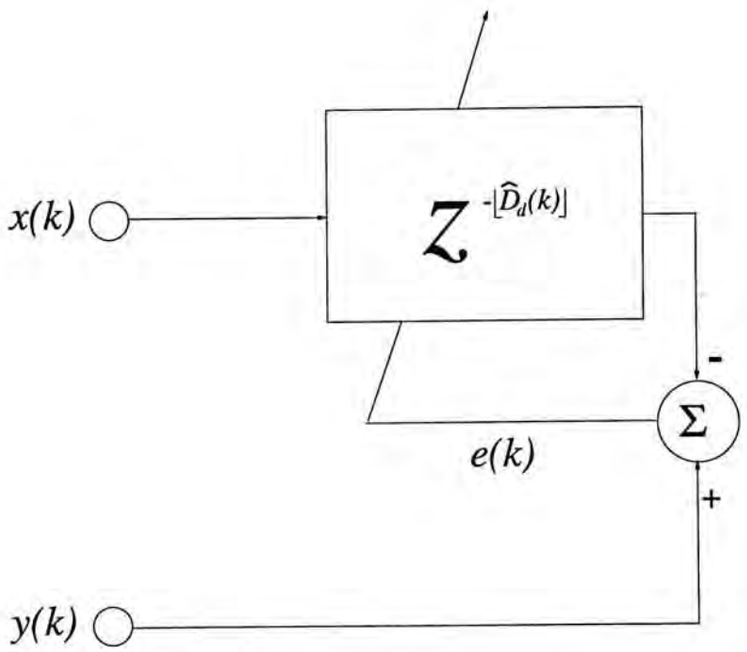

Figure 2.2: System block diagram of the ADDLD

The system block diagram of the ADDLD [40] for TDE is depicted in Figure 2.2. The delay element, $z^{-\lfloor \hat{D}_d(k) \rfloor}$ where $\lfloor \hat{D}_d(k) \rfloor$ is an integer that represents the estimate of D at time instant k, is used to provide a time shift of $\lfloor \hat{D}_d(k) \rfloor$ to $x(k)$. The error function of this system, $e(k)$, is given by

$$e(k) = y(k) - x(k - \lfloor \hat{D}_d(k) \rfloor) \tag{2.29}$$

The basis idea is to minimize the mean square value $E\{e^2(k)\}$ through the adjustment of the variable delay estimate.

The optimization algorithm of the ADDLD is of the form

$$\hat{D}_d(k+1) = \hat{D}_d(k) - \mu_d \nabla_k \tag{2.30}$$

where $\hat{D}_d(k)$, whose round-off value equals $\lfloor \hat{D}_d(k) \rfloor$, is a continuous variable and μ_d is the corresponding convergence factor. Similar to the adaptive algorithm developed by Widrow [44],[45], ∇_k is defined as the stochastic gradient of $E\{e^2(k)\}$ with respect to the delay estimate $\hat{D}_d(k)$, that is,

$$\begin{aligned} \nabla_k &= \frac{\partial e^2(k)}{\partial \hat{D}_d(k)} \\ &= -2e(k)\frac{\partial x(k-\hat{D}_d(k))}{\partial \hat{D}_d(k)} \end{aligned} \tag{2.31}$$

In [40], the symmetric difference is used as an approximation to the derivative of $x(k-\hat{D}_d(k))$ with respect to the delay estimate. Hence, (2.30) can be simplified to

$$\hat{D}_d(k+1) = \hat{D}_d(k) - \mu_d e(k)(x(k-\lfloor \hat{D}_d(k) \rfloor - 1) - x(k-\lfloor \hat{D}_d(k) \rfloor + 1)) \tag{2.32}$$

When the sampling interval goes to zero, it has been shown [40] that the gradient estimate of the algorithm is unbiased. In other words, its expected value will be equal to $E\{\nabla_k\}$. A stability bound for μ_d is given by [40]

$$0 < \mu_d < \frac{1}{10\sigma_s^2} \tag{2.33}$$

Although the computational complexity of the ADDLD is much simpler than that of the LMSTDE, it has two restrictions. Firstly, this algorithm is unable to give an accurate estimate of D if it is not an integral multiple of the sampling period. Secondly, the performance surface $E\{e^2(k)\}$ is multimodal. Using (2.1) and (2.29), the continuous form of $E\{e^2(k)\}$ can be shown to be [60]

$$E\{e^2(k)\} = 2\sigma_s^2(1 - sinc(\hat{D}_d - D)) + 2\sigma_n^2 \tag{2.34}$$

The time index k is dropped from the delay estimate for convenience. A sketch of (2.34) is illustrated in Figure 2.3. It can be seen that the global minimum occurs when $\hat{D}_d = D$ and the two local maxima that are closest to it appear at $\hat{D}_d = D + 1.45$ and $\hat{D}_d = D - 1.45$. Therefore, in order to guarantee global convergence of the ADDLD algorithm, $\hat{D}_d$ should satisfy

$$D + 1.45 \geq \hat{D}_d \geq D - 1.45 \tag{2.35}$$

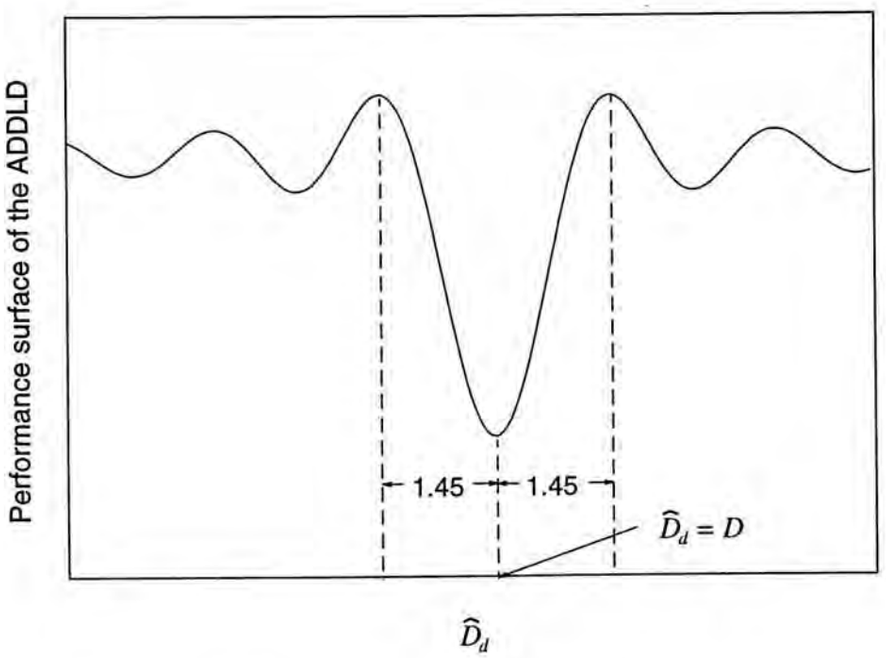

Figure 2.3: Performance surface of the ADDLD

This means that $\hat{D}_d(k)$ should be close enough to D for accurate delay estimation. To achieve this, we may need to initialize the value of the delay estimate by methods such as cross-correlation [61], [62].

2.4 Summary

To summarize, two adaptive approaches for estimating the time delay between two sensor outputs have been reviewed. The LMSTDE uses an FIR filter, whose coefficients are adjusted according to Widrow's LMS algorithm, to model the differential delay. The delay estimate of this method is obtained indirectly from interpolation of the filter weights. However, false peak weight might occur which will degrade the delay estimation accuracy. Time-varying convergence factor and constrained adaptation have been proposed to improve the performance of the LMSTDE. While, on the other hand, the ADDLD provides direct delay measurements and employs the stochastic gradient of the error function with respect to the estimated delay. The ADDLD algorithm is very efficient computationally but its delay estimates are

restricted to be integral multiples of the sampling period. In addition, it has a non-quadratic performance surface and may limit the operation range of the estimated delay.

Chapter 3

The Explicit Time Delay Estimator (ETDE)

In this chapter, a novel adaptive delay estimation algorithm for extracting the time difference of arrival (TDOA) between signals received at two spatially separated sensors, called the explicit time delay estimator (ETDE) [63],[64], is introduced. In order to illustrate the potential of the ETDE, we compare and contrast its performance with other adaptive TDE methods [29],[30],[36] and with the Cramér-Rao lower bound (CRLB) [65]. Similar to the LMSTDE [29],[30], the ETDE models the time delay by using an FIR transversal filter. But it has the merit of providing a direct measure of the TDOA on a sample-by-sample basis and does not involve any interpolation. Although the ADDLD [40] also updates the estimated delay directly, it cannot deal with delays that are non-integral multiples of the sampling period, while the ETDE has no such restriction. The ETDE can be considered as an alternative realization of the constrained adaptive time delay estimator (CATDE) proposed in [36] which constrains all coefficients of the FIR filter to be samples of a *sinc* function. However, we shall show that the ETDE algorithm provides a more accurate and robust delay measurement than the CATDE under noisy environment.

In Section 3.1, the ETDE is developed and analyzed. The performance surface of the ETDE algorithm is found to be multimodal and thus the operation range of the estimated delay is limited. To release this restriction, an initialization method is suggested to ensure global convergence. Both static and dynamic behaviours of the ETDE are also given. Section 3.2 compares the ETDE with the LMSTDE and the

CATDE in terms of estimation accuracy and computational complexity. A simple relationship between the mean square delay error of the ETDE and the CRLB is established. Computer simulations are presented in Section 3.3 to corroborate the theoretical analysis and to evaluate the performance of the ETDE. Finally, conclusion is drawn in Section 3.4.

3.1 Derivation and Analysis of the ETDE

3.1.1 The ETDE system

For ease of reference, the problem of time delay estimation is reiterated here. As given by (2.1), the received signals at two spatially separated sensors are expressed as

$$x(k) = s(k) + n_1(k)$$

and

$$y(k) = s(k - D) + n_2(k)$$

where the source signal $s(k)$ and the corrupting noises $n_1(k)$ and $n_2(k)$ are mutually uncorrelated, stationary and white Gaussian processes. Our task is to estimate the differential delay D from $x(k)$ and $y(k)$.

It has been observed from (2.5) that a delay can be modeled by an FIR filter whose coefficients are samples of a *sinc* function. Using this idea, the explicit time delay estimator (ETDE) is derived and its system block diagram is depicted in Figure 3.1. The transfer function of the ETDE, $\hat{D}(z)$, is given by

$$\hat{D}(z) = \sum_{i=-P}^{P} sinc(i - \hat{D}) z^{-i} \tag{3.1}$$

The ETDE has the same filter structure with the LMSTDE but the filter coefficients $\{w_i(k)\}$ are now replaced by $\{sinc(i - \hat{D}(k))\}$ for $P \geq i \geq -P$, where $\hat{D}(k)$ is the instantaneous estimated delay. Since the value of $sinc(i - D)$ decreases with increasing i, the finite-order ETDE can still provide an accurate time shift of $\hat{D}(k)$

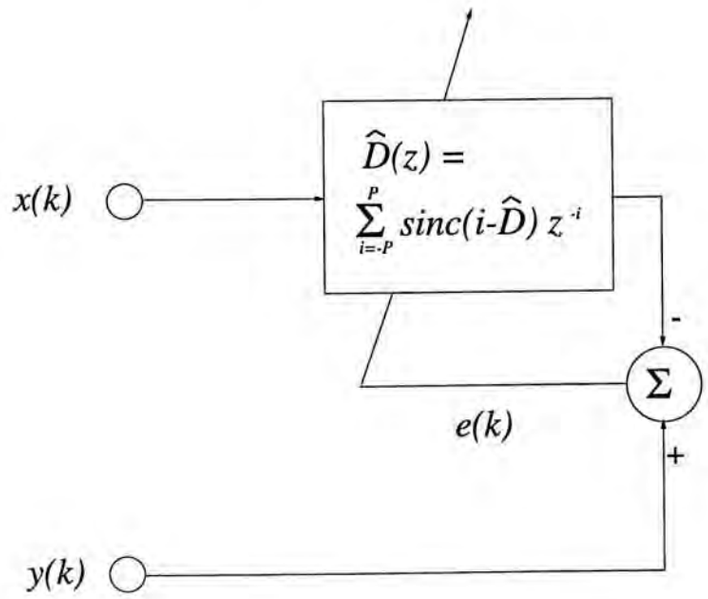

Figure 3.1: System block diagram of the ETDE

to $x(k)$ when P is chosen large enough to minimize the truncation error [35]. The output error $e(k)$ can be computed from

$$\begin{aligned} e(k) &= y(k) - \sum_{i=-P}^{P} sinc(i - \hat{D}(k))x(k-i) \\ &= y(k) - \tilde{x}(k - \hat{D}(k)) \end{aligned} \tag{3.2}$$

where $\tilde{x}(k-v)$ denotes an approximate version of $x(k)$ with time lag v and the approximation becomes exact when P goes to infinity. Similar to Widrow's LMS algorithm [45], the ETDE uses a stochastic gradient estimate which is obtained by differentiating the instantaneous error square, $e^2(k)$, with respect to $\hat{D}(k)$. The parameter updating equation is thus given by

$$\begin{aligned} \hat{D}(k+1) &= \hat{D}(k) - \mu_D \frac{\partial e^2(k)}{\partial \hat{D}(k)} \\ &= \hat{D}(k) - \mu_D \cdot 2e(k) \cdot \frac{\partial\{y(k) - \sum_{i=-P}^{P} sinc(i - \hat{D}(k))x(k-i)\}}{\partial \hat{D}(k)} \\ &= \hat{D}(k) + 2\mu_D e(k) \sum_{i=-P}^{P} x(k-i) \frac{\partial sinc(i - \hat{D}(k))}{\partial \hat{D}(k)} \\ &= \hat{D}(k) - 2\mu_D e(k) \sum_{i=-P}^{P} x(k-i) f(i - \hat{D}(k)) \end{aligned} \tag{3.3}$$

where μ_D is the convergence parameter that controls the convergence rate and stability of the ETDE algorithm. The function f here has the same definition as (2.14), i.e.,

$$f(v) = \frac{\cos(\pi v) - sinc(v)}{v}$$

Since

$$\cos(\pi v) = (-1)^{v_i} \cos(\pi v_d) \tag{3.4}$$

where v_i and v_d represent the integral and decimal part of v respectively, only one cosine operation is needed in (3.3). In order to reduce the computational complexity further, lookup tables of the cosine and *sinc* function are constructed. The former is a vector $\vec{M}$ of length L and has the following elements

$$\vec{M} = \left[\begin{array}{cccc} 1 & \cos\left(\frac{\pi}{2(L-1)}\right) & \cos\left(\frac{2\pi}{2(L-1)}\right) & \cdots \quad \cos(0.5\pi) \end{array} \right] \tag{3.5}$$

While the *sinc* table is a matrix $\mathcal{N}$ of dimension $K \times (2P+1)$ with entry

$$n_{ij} = sinc\left(\frac{0.5i}{K-1} - j\right) \tag{3.6}$$

for $K-1 \geq i \geq 0$ and $P \geq j \geq -P$. Thus, $\mathcal{N}$ is of the form

$$\mathcal{N} = \left[\begin{array}{cccc} sinc(-P) & sinc(-P+1) & \cdots & sinc(P) \\ sinc\left(\frac{0.5}{K-1} - P\right) & sinc\left(\frac{0.5}{K-1} - P + 1\right) & \cdots & sinc\left(\frac{0.5}{K-1} + P\right) \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ sinc(0.5-P) & sinc(0.5-P+1) & \cdots & sinc(0.5+P) \end{array} \right] \tag{3.7}$$

It is clear that the resolution of the delay estimate increases with the dimensions of $\vec{M}$ and $\mathcal{N}$. Note that $\mathcal{N}$ only accounts for delay ranging from 0 to 0.5 and thus the resolution of $\hat{D}(k)$ is $1/(2(K-1))$. Similarly, the cosine function has a resolution of $1/(2(L-1))$ where L is chosen to be comparable with K so that the accuracies of the f and *sinc* function are similar. Implementation details concerning the *sinc* table

have been discussed by Ching and Chan [36]. Denote $\hat{D}^r(k)$ and $\hat{D}^d(k)$ as the round-off integer and decimal part of $\hat{D}(k)$ respectively, then $\hat{D}^d(k) = |\hat{D}(k) - \hat{D}^r(k)|$, which has a value between 0 and 0.5. The variable $\hat{D}^d(k)$ is used to locate the closest element in $\vec{M}$ and the central column of $\mathcal{N}$. Whereas $\hat{D}^r(k)$ gives the sign of the cosine functions as well as the required vector in the *sinc* table. This information is then used to calculate $\{sinc(i - \hat{D}(k))\}$ as well as $\{f(i - \hat{D}(k))\}$. As a result, the computation involved in the ETDE algorithm is minimal. At each sampling interval, a total of $(8P + 4)$ additions and $(6P + 6)$ multiplications are needed to calculate $\hat{D}(k)$.

3.1.2 Performance Surface

For ease of analysis, we assume that the sensor outputs and the delay estimate are uncorrelated [57]. Using (3.2), the performance surface of the ETDE, $E\{e^2(k)\}$, can be found to be

$$\begin{aligned} E\{e^2(k)\} &= E\{(s(k-D) + n_2(k) - \tilde{s}(k-\hat{D}) - \tilde{n}_1(k-\hat{D}))^2\} \\ &= E\{(s(k-D) - \tilde{s}(k-\hat{D}))^2\} + E\{(n_2(k) - \tilde{n}_1(k-\hat{D}))^2\} \\ &= \sigma_s^2 \left(1 + \sum_{i=-P}^{P} sinc^2(i-\hat{D}) - 2\sum_{i=-P}^{P} sinc(i-\hat{D})sinc(i-D)\right) \\ &\quad + \sigma_n^2 \left(1 + \sum_{i=-P}^{P} sinc^2(i-\hat{D})\right) \end{aligned} \tag{3.8}$$

Since

$$\sum_{i=-\infty}^{\infty} sinc(i-\hat{D})sinc(i-D) = sinc(D-\hat{D}) \tag{3.9}$$

Equation (3.8) can be approximated by

$$E\{e^2(k)\} \approx 2\sigma_s^2(1 - sinc(D-\hat{D})) + 2\sigma_n^2 \tag{3.10}$$

if P is chosen sufficiently large. It is obvious that $E\{e^2(k)\}$ is a multimodal function of $\hat{D}$ which is in fact identical to that of the ADDLD as shown in Figure 2.3. Thus the ETDE algorithm is capable of searching for the global minimum if the time-varying delay estimate $\hat{D}$ is close enough to the actual delay and lies within the

bound,

$$D + 1.45 \geq \hat{D} \geq D - 1.45 \tag{3.11}$$

To satisfy (3.11), it is necessary that the initial delay estimate, $\hat{D}(0)$, lies within the above admissible range.

An efficient initialization scheme for $\hat{D}(k)$ has been considered for high SNR condition. In our studies, the initial delay estimate is obtained by using the LMSTDE algorithm at the very beginning of the adaptation and is given by

$$\hat{D}(0) = \arg\max_i \{w_i(k)\} \tag{3.12}$$

Note that the ETDE and the LMSTDE have a similar filter structure so only the adaptive algorithm is changed during initialization. Although other methods such as cross correlation techniques [61],[62] can also provide a reasonable initial estimate of $\hat{D}(0)$, the LMSTDE is capable of achieving proper initialization within a small number of iterations if the filter weights are all assigned to have zero values at the early stage of the adaptation. Putting $w_i(k) = 0$ for $P \geq i \geq -P$ into (2.8), the learning behaviour of the filter weights becomes

$$E\{w_i(k)\} = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_n^2} sinc(i - D)(1 - (1 - 2\mu_w(\sigma_s^2 + \sigma_n^2))^k), \qquad P \geq i \geq -P \tag{3.13}$$

In this case, we can obtain an accurate delay estimate by interpolating the expected filter weights of (3.13) at any time instant k since the vector $\{E\{w_i(k)\}\}$ is actually a scaled version of the Wiener solution, $\{w_i^o\}$, which is given by (2.10),

$$w_i^o = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_n^2} sinc(i - D), \qquad P \geq i \geq -P$$

This implies that accurate delay estimation can be attained after a few iterations even when the filter coefficients have not yet converged. In practice, the round-off integral value of D can be found by adapting the LMSTDE just for a few hundred samples.

3.1.3 Static Behaviour

In this section, the convergence dynamics of the ETDE for a static delay is investigated. For analysis purpose, it is assumed that the delay estimate $\hat{D}(k)$ is reasonably close to the actual time difference D. Taking the expected value of (3.3) yields

$$\begin{aligned}
&E\{\hat{D}(k+1)\} \\
&= E\{\hat{D}(k)\} - 2\mu_D E\left\{(y(k) - \tilde{x}(k-\hat{D}(k))) \cdot \sum_{i=-P}^{P} x(k-i) f(i-\hat{D}(k))\right\} \\
&= E\{\hat{D}(k)\} - 2\mu_D E\{(s(k-D) - \tilde{s}(k-\hat{D}(k)) + n_2(k) - \tilde{n}_1(k-\hat{D}(k))) \\
&\quad \cdot \sum_{i=-P}^{P} (s(k-i) + n_1(k-i)) f(i-\hat{D}(k))\} \\
&= E\{\hat{D}(k)\} - 2\mu_D E\left\{\sum_{j=-\infty}^{\infty} s(k-j)sinc(j-D) \sum_{i=-P}^{P} s(k-i) f(i-\hat{D}(k))\right\} \\
&\quad +2\mu_D E\left\{\sum_{j=-P}^{P} s(k-j)sinc(j-\hat{D}(k)) \sum_{i=-P}^{P} s(k-i) f(i-\hat{D}(k))\right\} \\
&\quad +2\mu_D E\left\{\sum_{j=-P}^{P} n_1(k-j)sinc(j-\hat{D}(k)) \sum_{i=-P}^{P} n_1(k-i) f(i-\hat{D}(k))\right\} \\
&= E\{\hat{D}(k)\} - 2\mu_D \sigma_s^2 E\left\{\sum_{i=-P}^{P} sinc(i-D) f(i-\hat{D}(k))\right\} \\
&\quad +2\mu_D(\sigma_s^2 + \sigma_n^2) E\left\{\sum_{i=-P}^{P} sinc(i-\hat{D}(k)) f(i-\hat{D}(k))\right\} \qquad (3.14)
\end{aligned}$$

Since

$$\sum_{i=-\infty}^{\infty} sinc(i-D) f(i-\hat{D}(k)) = f(D-\hat{D}(k)) \qquad (3.15)$$

and from L'Hôspital's rule, $f(0) = 0$, thus (3.14) can be approximated by, for sufficiently large P,

$$E\{\hat{D}(k+1)\} \approx E\{\hat{D}(k)\} - 2\mu_D \sigma_s^2 E\{f(D-\hat{D}(k))\} \qquad (3.16)$$

Expanding $f(D-\hat{D}(k))$ about $\hat{D}(k) = D$ using Taylor's series up to the first order term, (3.16) becomes

$$\begin{aligned}
E\{\hat{D}(k+1)\} &\approx E\{\hat{D}(k)\} - 2\mu_D \sigma_s^2 (E\{f(0) + f'(0)(D-\hat{D}(k))\}) \\
&= E\{\hat{D}(k)\} + \frac{2}{3}\mu_D \sigma_s^2 \pi^2 (D - E\{\hat{D}(k)\}) \qquad (3.17)
\end{aligned}$$

Solving (3.17) gives

$$E\{\hat{D}(k)\} = D + (\hat{D}(0) - D)\left(1 - \frac{2}{3}\mu_D\sigma_s^2\pi^2\right)^k \tag{3.18}$$

provided that μ_D lies within the range,

$$0 < \mu_D < \frac{3}{\pi^2\sigma_s^2} \tag{3.19}$$

Therefore, the ETDE algorithm is asymptotically unbiased since $E\{\hat{D}(k)\} \to D$ as $k \to \infty$.

The mean square delay error at time k, denoted by $\epsilon(k)$, has been shown to be (See Appendix A)

$$\begin{aligned} \epsilon(k) &\triangleq E\{(\hat{D}(k) - D)^2\} \\ &\approx \alpha^k(\hat{D}(0) - D)^2 + \beta\left(\frac{1-\alpha^k}{1-\alpha}\right) \end{aligned} \tag{3.20}$$

where

$$\alpha = 1 - \frac{4}{3}\mu_D\sigma_s^2\pi^2\left(1 - \mu_D\sigma_s^2\pi^2 - \frac{1}{3}\mu_D\sigma_n^2\pi^2\right) \tag{3.21}$$

and

$$\beta = \frac{8}{3}\mu_D^2\pi^2\sigma_n^2(\sigma_s^2 + \sigma_n^2) \tag{3.22}$$

The validity of (3.20) is subject to the following being satisfied,

$$|\alpha| < 1$$

$$\Rightarrow 0 < \mu_D < \frac{1}{\pi^2(\sigma_s^2 + \sigma_n^2/3)} \tag{3.23}$$

Notice that (3.23) is a more stringent bound for μ_D since the mean square delay error is confined to have a finite value [66].

For slow adaptation, that is, when μ_D is selected much smaller than $1/(\pi^2(\sigma_s^2 + \sigma_n^2/3))$, it is noted that $\epsilon(k)$ consists of three components as follows,

$$\begin{aligned} \epsilon(k) &= \left(1 - \frac{4}{3}\mu_D\sigma_s^2\pi^2\right)^k(\hat{D}(0) - D)^2 - \frac{2\mu_D\sigma_n^2(\sigma_s^2 + \sigma_n^2)}{\sigma_s^2}\left(1 - \frac{4}{3}\mu_D\sigma_s^2\pi^2\right)^k \\ &\quad + \frac{2\mu_D\sigma_n^2(\sigma_s^2 + \sigma_n^2)}{\sigma_s^2} \end{aligned} \tag{3.24}$$

The first and second term are transient errors due to gradient noises in the absence and presence of corrupting noise respectively, while the last term is the steady state error due to noisy gradients. Using (3.24), the delay variance, $var(\hat{D})$, which is equal to the steady state mean square delay error, can be expressed as

$$\begin{aligned} var(\hat{D}) &\triangleq \lim_{k\to\infty} \epsilon(k) \\ &= 2\mu_D\sigma_n^2\left(1+\frac{1}{\text{SNR}}\right) \end{aligned} \tag{3.25}$$

The delay variance of the ETDE increases with the step size μ_D and the noise power σ_n^2. Moreover, it is inversely proportional to SNR when under very noisy environment. In the absence of noise, $var(\hat{D})$ has zero value. It is intriguing to note that (3.25) is indeed identical to the variance of the closed-loop least mean square time delay estimator reported by H.Messer and Y.Bar-Ness [42].

In the extreme cases for high SNR and low SNR, $var(\hat{D})$ can be simplified to

$$var(\hat{D}) = \begin{cases} 2\mu_D\sigma_n^2 & , \ \text{SNR} \gg 1 \\ \\ \dfrac{2\mu_D\sigma_n^2}{\text{SNR}} & , \ \text{SNR} \ll 1 \end{cases} \tag{3.26}$$

3.1.4 Dynamic Behaviour

In this section, the tracking performance of the ETDE is examined for a linearly moving delay [30],[67]. In general, the trajectory of a nonstationary delay may be considered as a combination of many piecewise linear functions if it changes very slowly. The actual time-varying delay is denoted by $D(k)$ and has the form

$$D(k) = D_0 + \lambda k \tag{3.27}$$

where D_0 is the delay at $k = 0$ and λ is the Doppler time compression. By modifying (3.17) and assuming that the actual delay is changing slowly such that $\hat{D}(k)$ more or less follows $D(k)$ at each iteration, the expected value of the delay estimate in tracking a linearly varying delay can then be approximated by

$$E\{\hat{D}(k+1)\} \approx E\{\hat{D}(k)\} + \frac{2}{3}\mu_D\sigma_s^2\pi^2(D_0 + \lambda k - E\{\hat{D}(k)\}) \tag{3.28}$$

Solving (3.28) yields (See Appendix B)

$$E\{\hat{D}(k)\} \approx D_0 + \lambda k + \left(\hat{D}(0) - D_0 + \frac{3\lambda}{2\mu_D\sigma_s^2\pi^2}\right)\left(1 - \frac{2}{3}\mu_D\sigma_s^2\pi^2\right)^k - \frac{3\lambda}{2\mu_D\sigma_s^2\pi^2} \tag{3.29}$$

Equation (3.29) is similar to (3.18) since the time constant in both cases has the same value of $3/(2\mu_D\sigma_s^2\pi^2)$ which is independent of the noise power. In fact, substituting $\lambda = 0$ and $D_0 = D$ in (3.29) gives (3.18). Furthermore, the third term of (3.29) is a transient term which converges to zero as k goes to infinity while the last term, $3\lambda/(2\mu_D\sigma_s^2\pi^2)$, contributes to the time lag in steady state. It can be seen that this time lag is directly proportional to λ and is inversely proportional to μ_D and σ_s^2. It is worthy to note that the lower bound of μ_D in (3.23) should be increased to cope with nonstationary delays. This means that the value of μ_D cannot be chosen too small such that the assumption in deriving (3.28) becomes void. Let the maximum allowable time lag be t_{max} whose value is less than 0.5 in practice. Using (3.29), a time lag of t_{max} will correspond to a convergence step size of $3\lambda/(2\pi^2\sigma_s^2 t_{max})$. As a result, a modified bound for μ_D is given by

$$\frac{3\lambda}{(2\pi^2\sigma_s^2 t_{max})} < \mu_D < \frac{1}{\pi^2(\sigma_s^2 + \sigma_n^2/3)} \tag{3.30}$$

In this case, the steady state mean square delay error, $\xi(\hat{D})$, is equal to the delay variance plus the square of the time lag [68], that is,

$$\xi(\hat{D}) = \frac{2\mu_D\sigma_s^2(1 + \text{SNR})}{\text{SNR}^2} + \frac{9\lambda^2}{4\mu_D^2\sigma_s^4\pi^4} \tag{3.31}$$

We observe that the first term of (3.31) increases with μ_D whereas the second term decreases with μ_D^2. Thus, μ_D must be selected appropriately in order to achieve the optimal performance. As a rule of thumb, when noise dominates, a small value of μ_D should be used. While, on the other hand, a larger value of μ_D should be chosen if the delay is changing rapidly with time.

3.2 Performance Comparisons

3.2.1 With the LMSTDE

We shall contrast the convergence speed of the LMSTDE and the ETDE when a step change of time difference from D_0 to D is introduced between the two received signals. Without loss of generality, it is assumed that all filter weights have attained their desired values initially, that is, $\{sinc(i-D_0)\sigma_s^2/(\sigma_s^2+\sigma_n^2)\}$ for the LMSTDE and $\{sinc(i-D_0)\}$ for the ETDE.

In order to make a fair comparison, the relationship between μ_D and μ_w such that identical delay variance is kept in both systems is established by equating (2.21) and (3.26), and it is given by

$$\mu_D = \begin{cases} \dfrac{3\mu_w}{\pi^2} & , \ \text{SNR} \gg 1 \\ \\ \dfrac{3\mu_w}{2\pi^2 \text{SNR}} & , \ \text{SNR} \ll 1 \end{cases} \tag{3.32}$$

It has been shown from Section 3.1.2 that the LMSTDE provides a very fast response to a static delay when the initial filter weights are all set to zero. However, when the delay is time-varying, this is no longer applicable. From (2.8), the learning characteristics of $w_i(k)$ becomes

$$\begin{aligned} E\{w_i(k)\} &= \frac{\sigma_s^2}{\sigma_s^2+\sigma_n^2} sinc(i-D)(1-(1-2\mu_w(\sigma_s^2+\sigma_n^2))^k) \\ &\quad + \frac{\sigma_s^2}{\sigma_s^2+\sigma_n^2} sinc(i-D_0)(1-2\mu_w(\sigma_s^2+\sigma_n^2))^k \end{aligned} \tag{3.33}$$

Due to different initial conditions, (3.33) contains an additional term as compared with (3.13). In this case, accurate delay estimates can be obtained only if all the filter coefficients have reached their steady state values. Substituting (3.33) into (2.11), the mean value of $\hat{D}_w(k)$, $E\{\hat{D}_w(k)\}$, is given by

$$\begin{aligned} E\{\hat{D}_w(k)\} &= \arg\max_t \left\{ \sum_{i=-P}^{P} E\{w_i(k)\} sinc(t-i) \right\} \\ &\approx \arg\max_t \{(1-(1-2\mu_w(\sigma_s^2+\sigma_n^2))^k) sinc(t-D) \\ &\quad + (1-2\mu_w(\sigma_s^2+\sigma_n^2))^k sinc(t-D_0)\} \end{aligned} \tag{3.34}$$

Since the first term of (3.34) peaks at $t = D$ with a value of $(1-(1-2\mu_w(\sigma_s^2+\sigma_n^2))^k)$ and the second term has a maximum value of $(1-2\mu_w(\sigma_s^2+\sigma_n^2))^k$ when $t = D_0$, we may approximate $E\{\hat{D}_w(k)\}$ by the following weighted sum

$$\begin{aligned} E\{\hat{D}_w(k)\} &\approx (1-(1-2\mu_w(\sigma_s^2+\sigma_n^2))^k)D + (1-2\mu_w(\sigma_s^2+\sigma_n^2))^k D_0 \\ &= D + (D_0 - D)(1-2\mu_w(\sigma_s^2+\sigma_n^2))^k \end{aligned} \tag{3.35}$$

Note that the second term represents the mean delay error of the LMSTDE at the kth iteration. The mean delay error of the ETDE, $E\{\hat{D}(k) - D\}$, can be easily obtained from (3.18)

$$E\{\hat{D}(k)\} - D = (D_0 - D)\left(1 - \frac{2}{3}\mu_D\sigma_s^2\pi^2\right)^k \tag{3.36}$$

The ratio of the mean delay errors at the kth iteration is thus equal to

$$\frac{E\{\hat{D}(k)\} - D}{E\{\hat{D}_w(k)\} - D} = \left(\frac{1 - 2\mu_D\sigma_s^2\pi^2/3}{1-2\mu_w(\sigma_s^2+\sigma_n^2)}\right)^k \tag{3.37}$$

If the ETDE converges faster than the LMSTDE, then the numerator of (3.37) will be smaller than the denominator which implies

$$\mu_D > \begin{cases} \dfrac{3\mu_w}{\pi^2} & , \ \text{SNR} \gg 1 \\ \\ \dfrac{3\mu_w}{\pi^2 \text{SNR}} & , \ \text{SNR} \ll 1 \end{cases} \tag{3.38}$$

Comparing (3.32) with (3.38), we see that the performance of both methods are identical when the SNR is high. When SNR $\ll$ 1, (3.38) is not satisfied which means that the convergence speed of the ETDE is slower than that of the LMSTDE when the delay variance of both systems are kept identical. However, (2.15) has assumed that the delay estimate $\hat{D}_w(k)$ is at the neighborhood of D and this is not true when false peaks of $w_i(k)$ are located [59]. As discussed in Section 2.2.2, the probability of occurrence of false peak weight increases with μ_w and decreases with the SNR under noisy environments. When a wrong peak is located, the value of $\hat{D}_w(k)$ may deviate significantly from the original delay. On the other hand, the ETDE has no such problem and therefore it is a more reliable delay estimation scheme.

The computational load of updating $w_i(k)$ requires $(4P+2)$ additions and $(4P+4)$ multiplications. It seems that the LMSTDE algorithm involves less operation counts than that of the ETDE. But if interpolation of $w_i(k)$ is also considered, a further $20P$ additions and $(20P + 10)$ multiplications will be needed assuming that the delay estimate has a resolution of 0.001. Whilst the ETDE requires no interpolation but a larger memory size is required to store the elements of the lookup tables.

3.2.2 With the CATDE

From Section 2.2.3, we mention that the CATDE simplifies the LMSTDE algorithm by restricting the filter parameters to be samples of a *sinc* function and the delay estimate $\hat{D}_c(k)$ is related to the peak weight $w_L(k)$ by (2.28),

$$w_L(k) = sinc(L - \hat{D}_c(k))$$

In this constrained algorithm, only the largest filter coefficient is being adapted, and the values of other filter weights are simply retrieved from the *sinc* lookup table as given by (3.7). As a result, this algorithm is computationally more efficient since only $(2P+2)$ additions and $(2P+4)$ multiplications are required to update $w_L(k)$.

However, we shall show the drawbacks of the CATDE shortly. By using (2.4) and (2.28), the adaptation rule of $w_L(k)$ can be written as

$$\begin{aligned} sinc(L - \hat{D}_c(k+1)) &= sinc(L - \hat{D}_c(k)) + 2\mu_w x(k-L) \\ &\quad \cdot(y(k) - \tilde{x}(k - \hat{D}_c(k)) \end{aligned} \tag{3.39}$$

Taking expectation of (3.39) gives

$$\begin{aligned} E\{sinc(L - \hat{D}_c(k+1))\} &= E\{sinc(L - \hat{D}_c(k))\}(1 - 2\mu_w(\sigma_s^2 + \sigma_n^2)) \\ &\quad +2\mu_w\sigma_s^2 sinc(L - D) \end{aligned} \tag{3.40}$$

Examining the steady state condition of (3.40), we get

$$\begin{aligned}\lim_{k\to\infty} E\{sinc(L-\hat{D}_c(k))\} &= \lim_{k\to\infty} E\{w_L(k)\} \\ &= \frac{\sigma_s^2}{\sigma_s^2+\sigma_n^2} sinc(L-D) \qquad (3.41)\end{aligned}$$

From (3.41), it is noted that at very high SNR environment or noise-free condition, accurate delay estimation will be acquired since $\sigma_s^2/(\sigma_s^2+\sigma_n^2)$ is close to 1 and this has been demonstrated in [36]. However, large delay errors are introduced through inaccurate mapping of (2.28) at lower SNRs. Furthermore, unlike the ETDE, this algorithm is not capable of tracking a nonstationary delay. It is because in this situation the position of the peak weight may need to change from time to time and hence it has to be relocated by freely adapting all filter coefficients again and again. This in turn degrades the convergence speed of the algorithm and thus its performance is inferior to that of the ETDE.

3.2.3 With the CRLB

It is useful to compare (3.24) with the Cramér-Rao lower bound (CRLB) [65] which provides a lower bound on the variance attainable by any unbiased estimator using the same input data. In order to establish a simple relationship between the mean square delay error of the ETDE and the CRLB, we assume that D lies between -1 and 1 with an uniform distribution. With this *a priori* information about the parameter D, the CRLB for time delay estimation [3] is given by

$$\text{CRLB} = \min\left\{\frac{3(1+2\text{SNR})}{\pi^2 T \text{SNR}^2}, \frac{1}{3}\right\} \qquad (3.42)$$

where T is the observation time. Suppose that the initial estimated delay of the ETDE is not close to D, say, $\hat{D}(0)=0$ while $D=1$. In this case, the second term of (3.24) can be neglected and the continuous-time version of $\epsilon(k)$, $\epsilon(T)$, can be approximated by

$$\epsilon(T) \approx \exp\left(-\frac{4}{3}\mu_D\sigma_s^2\pi^2 T\right) + \frac{2\mu_D\sigma_s^2(1+\text{SNR})}{\text{SNR}^2} \qquad (3.43)$$

The term $\mu_D\sigma_s^2$ can be removed from (3.43) if an optimum step size is used. The optimal value, denoted by μ_D^o, is obtained by differentiating $\epsilon(T)$ with respect to μ_D

and equating the result to zero, and is given by

$$\mu_D^o = \frac{3}{4\mu_D \sigma_s^2 \pi^2 T} \ln\left(\frac{2\pi^2 T \mathrm{SNR}^2}{3(1+\mathrm{SNR})}\right) \tag{3.44}$$

When T is too small such that $\ln(2\pi^2 T\mathrm{SNR}^2/(3(1+\mathrm{SNR}))) < 0$, the step size μ_D^0 will be set to zero. In this case, the smallest mean square delay error is unity. Substituting μ_D^o into (3.43), we obtain a minimum value for $\epsilon(T)$ which is represented by $\epsilon^o(T)$,

$$\epsilon^o(T) = \begin{cases} \dfrac{3(1+\mathrm{SNR})}{2\pi^2 T\mathrm{SNR}^2}\left(1+\ln\left(\dfrac{2\pi^2 T\mathrm{SNR}^2}{3(1+\mathrm{SNR})}\right)\right) & , \ \ln\left(\dfrac{2\pi^2 T\mathrm{SNR}^2}{3(1+\mathrm{SNR})}\right) > 0 \\ 1 & , \ \text{otherwise} \end{cases} \tag{3.45}$$

Defining the performance index R as

$$R = \frac{\epsilon^o(T)}{\mathrm{CRLB}} \tag{3.46}$$

then the closer the value of R to unity, the better the delay estimation performance of the ETDE. Using (3.42) and (3.45), the variations of R against T for $\mathrm{SNR} \gg 1$ and $\mathrm{SNR} \ll 1$ are shown in Figure 3.2 and Figure 3.3 respectively. From Figure 3.2, it is found that R will increase by 0.5 if T is increased by 10 times or the SNR is increased by 10 dB. Although R is relatively small for a short T, the delay estimate does not converge to the optimal value and it introduces a large delay error. Hence there is a compromise between having a smaller R and a less biased delay estimate. It is seen that all three curves in Figure 3.3 remain constant at $R = 3$ at the beginning, they then start to drop and eventually rise again linearly. The dropping rate is approximately 1.5 per decade of T while the rising rate is 2 per decade of T. The minimum value of R has a value of 1.3 which occurs at approximately $T = 10^2$, $T = 10^4$ and $T = 10^6$ for SNR $= -10$ dB, -20 dB and -30 dB respectively. In general, similar to the high SNR conditions, longer observation time gives a more accurate delay estimate but a larger R and vice versa. Since R changes slowly with T for both high SNR and low SNR, the variance of the ETDE is larger than the CRLB by only a few times for a wide range of observation times and is thus regarded as comparable to the CRLB.

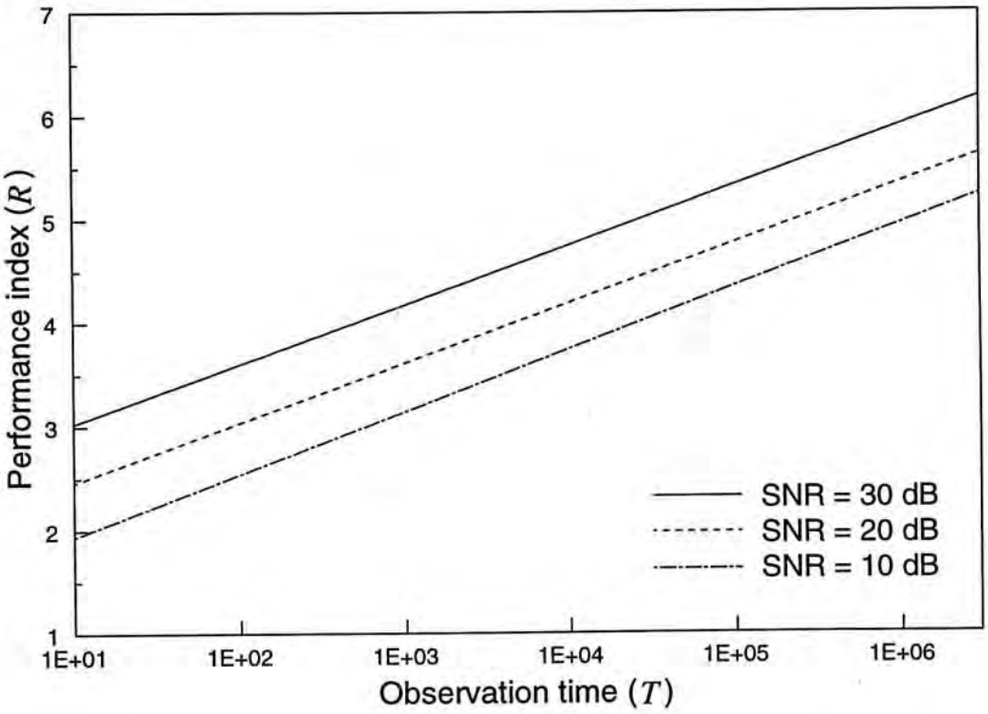

Figure 3.2: Performance comparison with the CRLB at high SNR

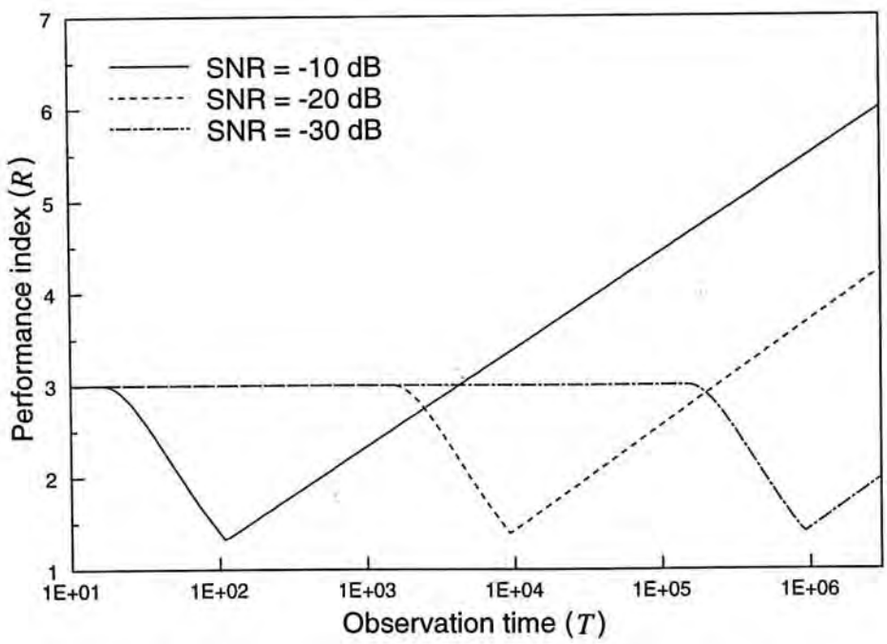

Figure 3.3: Performance comparison with the CRLB at low SNR

3.3 Simulation Results

Extensive computer simulations had been conducted to validate the theoretical derivations and to compare the performance of the ETDE with the LMSTDE and CATDE for time delay estimation between two sensor outputs. The signal $s(k)$ and the noises $n_1(k)$ and $n_2(k)$ were white Gaussian random variables and they were produced by a pseudorandom number generator. The delayed signal $s(k-D)$ was produced by passing $s(k)$ through a 61-tap FIR filter whose transfer function was given by $\sum_{i=-30}^{30} sinc(i-D)z^{-i}$. We made $D \in (-10, 10)$, and unless stated otherwise, P was chosen to be 15 to allow for acceptable truncation error. Different values of SNR were obtained by proper scaling of either the source signal or the random noise sequences. The delay estimate of each method had a resolution of approximately 0.001. To fulfill this requirement, the cosine vector $\vec{M}$ had a length of 513 and the size of the *sinc* table $\mathcal{N}$ was 513×31 in the ETDE and the CATDE while 10 bisections were used in the LMSTDE. The results provided were averages of 200 independent runs.

3.3.1 Corroboration of the ETDE Performance

Figure 3.4 and Figure 3.5 show the learning characteristics of the delay estimate and the mean square delay error respectively, for $D = 1.7$ at high SNR environment. During the first 200 iterations, the LMSTDE was used to determine the round-off integral value of D. The step size parameters μ_w and μ_D were set to 0.0001 and 0.005 respectively while the signal power σ_s^2 had unity value. These values corresponded to a time constant of about 30 and thus the ETDE algorithm was expected to reach steady state in only a few hundred iterations. It can be seen from Figure 3.4 that the convergence behaviours of $\hat{D}(k)$ were independent of the noise power and all experimental results agreed well with the theoretical expectation as given by (3.18). The delay estimates converged to the desired value at approximately the 400th iteration in all cases. From Figure 3.5, it is observed that the transient behaviours of the measured delay errors were very close to $\epsilon(k)$ derived from (3.20). However, there

		Measured variance	Theoretical variance
SNR	∞	4.39×10^{-5}	0
	20 dB	1.50×10^{-4}	1.01×10^{-4}
	10 dB	1.17×10^{-3}	1.10×10^{-3}

Table 3.1: Measured and theoretical delay variances of the ETDE

was slight difference between the experimental results and the theoretical values in steady state. The measured variances of $\hat{D}(k)$, together with their theoretical values as given by (3.25), were tabulated in Table 3.1 which shows that the experimental results conformed with the predicted values especially for a larger noise power.

The trajectories of the delay estimate at SNR = 10 dB for different values of μ_D were illustrated in Figure 3.6. Other settings were kept equal as in the previous experiment. Again, it can be seen that the learning curves of $\hat{D}(k)$ were fairly close to their expected values. Although the measured values of the delay variances were not shown, it had been found that they also agreed well with the theoretical delay variances. It is worthy to note that trials for $\mu_D = 0.005$, $\sigma_s^2 = 0.2$ and $\sigma_s^2 = 0.04$ had also been carried out and almost the same results were obtained.

Figure 3.7 shows the convergence behaviours of $\hat{D}(k)$ for $D = 0.3$ at a SNR of -10 dB. In this test, $\hat{D}(0)$ was assigned to have zero value and no initialization was performed at the beginning. The step size μ_D was assigned with values 0.0002 and 0.00004 while $\sigma_s^2 = 1$. In Figure 3.7, it can be observed that the convergence rates of the delay estimates were similar to those of the expected trajectories. Nevertheless, $\hat{D}(k)$ could not converge to the desired value of 0.3 in both cases. In fact, for $\mu_D = 0.0002$, $\hat{D}(k) \rightarrow 0.24$ at approximately the 4000th iteration and $\hat{D}(k) \rightarrow 0.34$ in about 20000 iterations when $\mu_D = 0.0002$. A possible reason for the bias is due to the last term of (3.14) which cannot be regarded to have zero value when $\sigma_s^2 \ll \sigma_n^2$ and when P was not chosen sufficiently large. To verify this, the experiment was repeated for $P = 150$ and the results are shown in Figure 3.8. It is seen that the delay error was reduced significantly for $\mu_D = 0.00004$ but the estimate of D was

not satisfactory for $\mu_D = 0.0002$. This means that accurate delay estimation can still be provided by the ETDE under very noisy environment if P is selected large enough and if μ_D is chosen sufficiently small. The learning trajectories of the mean square delay error are plotted in Figure 3.9. It can be seen that the learning rate of the experimental results for $P = 15$ and $\epsilon(k)$ were similar but those of $P = 150$ had slower speeds. Moreover, there were some differences in their steady states. When $P = 15$, the measured variances were found to be 3.44×10^{-2} and 4.93×10^{-3} for $\mu_D = 0.0002$ and $\mu_D = 0.00004$ respectively. While they had values of 4.67×10^{-2} and 4.49×10^{-3} for $\mu_D = 0.0002$ and $\mu_D = 0.00004$ respectively when $P = 150$. Although the bias was decreased by using a large P for $\mu_D = 0.00004$, there was almost no reduction in the delay variance. Evaluating the delay variance using (3.25) for $\mu_D = 0.0002$ gives a value of 4.4×10^{-2} and it agreed with the corresponding measurements. Whereas the theoretical value for $\mu_D = 0.00004$ is 8.8×10^{-3} which is larger than the experimental results. The discrepancies are mainly due to the approximations made in deriving (3.20).

The tracking performance of the ETDE for a linearly moving delay, which is given by $D(k) = 0.3 + 0.0001k$, is illustrated in Figure 3.10. The SNR was initially set at 10 dB and then step-changed to -10 dB at the 7000th iteration. The corrupting noises had unity power and the required SNRs were obtained by proper scaling of the source signal. The convergence parameter μ_D was assigned with values of 0.00002 and 0.002 at SNR = 10 dB and SNR = -10 dB respectively so that the time constant of $\hat{D}(k)$ could be kept identical under both SNR conditions. The initial delay estimate $\hat{D}(0)$ was selected to be zero. It is noted that after 3000 iterations, the delay estimate converged and lagged $D(k)$ by 0.077 at SNR = 10 dB and by 0.14 at SNR = -10 dB. Using (3.29), the predicted value of the time lag is calculated as 0.076, which is close to the measured value at SNR = 10 dB but is smaller than that of the low SNR condition. The deviation from the desired value for SNR = -10 dB is mainly due to the approximations in deriving (3.29). Furthermore, the steady state mean square delay errors had been measured and it had values of 6.77×10^{-3} at SNR = 10 dB and 5.85×10^{-2} at SNR = -10 dB. These values are slightly larger

than the expected values as given by (3.31).

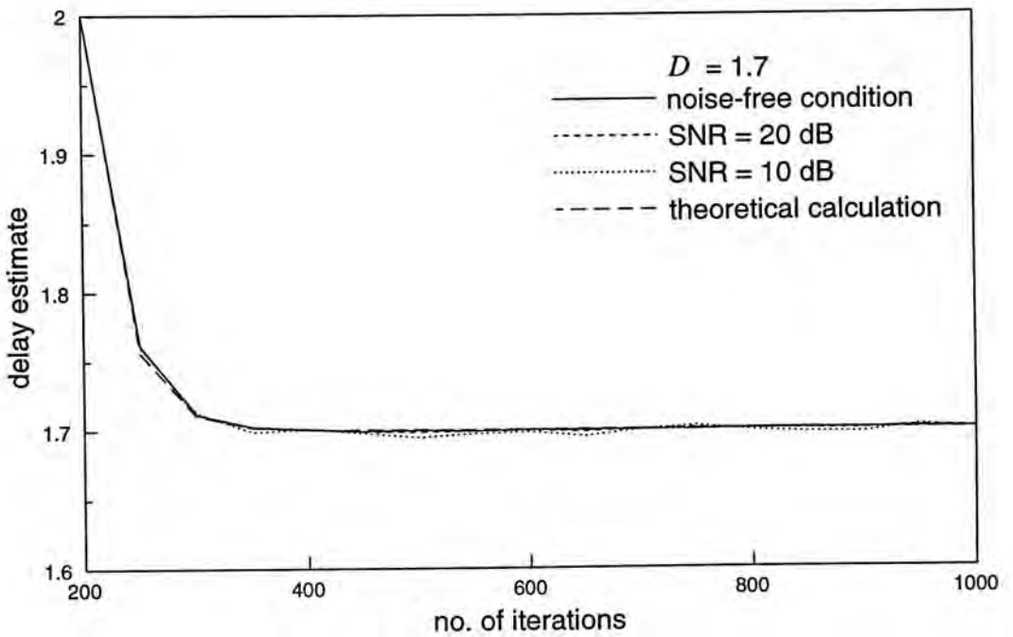

Figure 3.4: Delay estimates of the ETDE for a static delay at high SNR environments

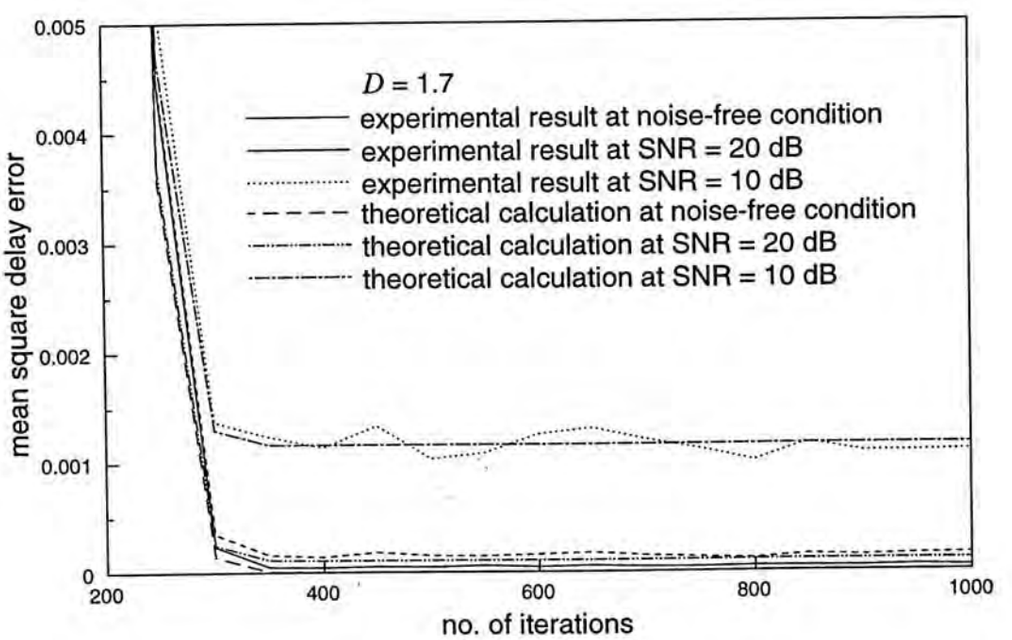

Figure 3.5: Mean square delay errors of the ETDE for a static delay at high SNR environments

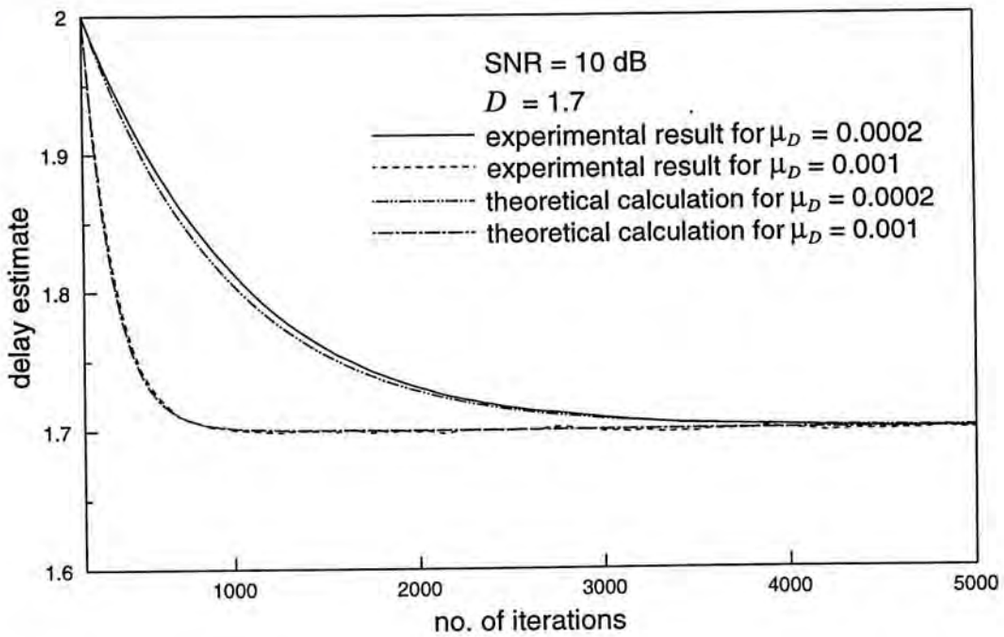

Figure 3.6: Delay estimates of the ETDE for different values of μ_D at SNR = 10 dB

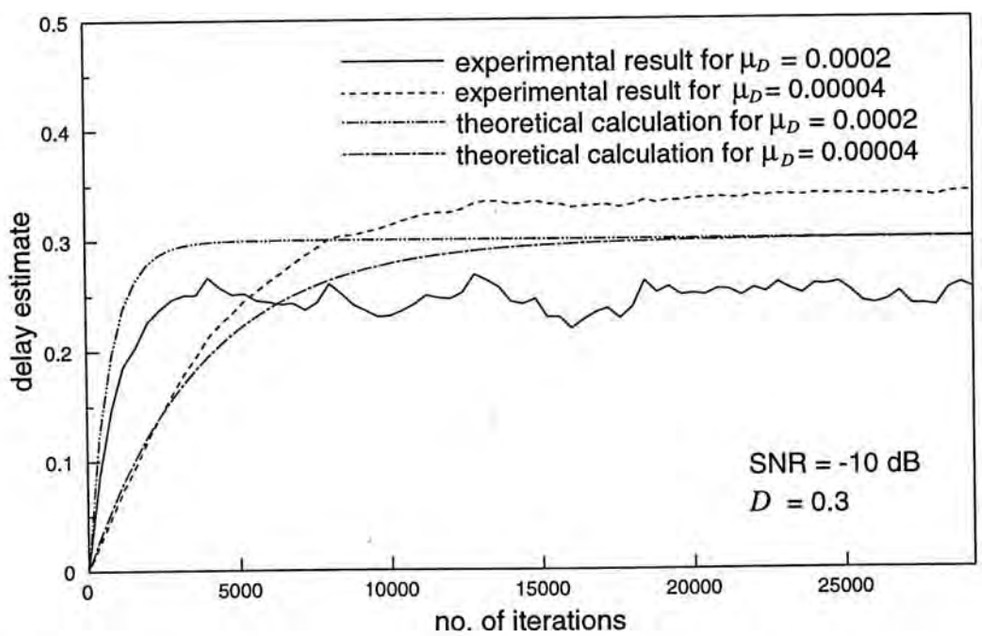

Figure 3.7: Delay estimates of the ETDE for different values of μ_D at SNR = -10 dB

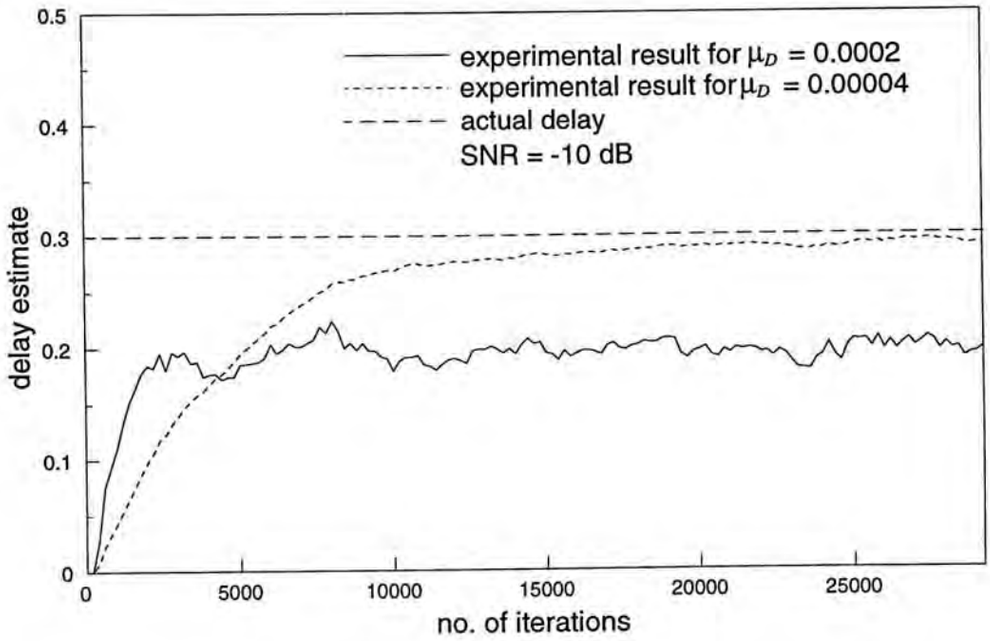

Figure 3.8: Delay estimates of the ETDE for $P = 150$ at SNR $= -10$ dB

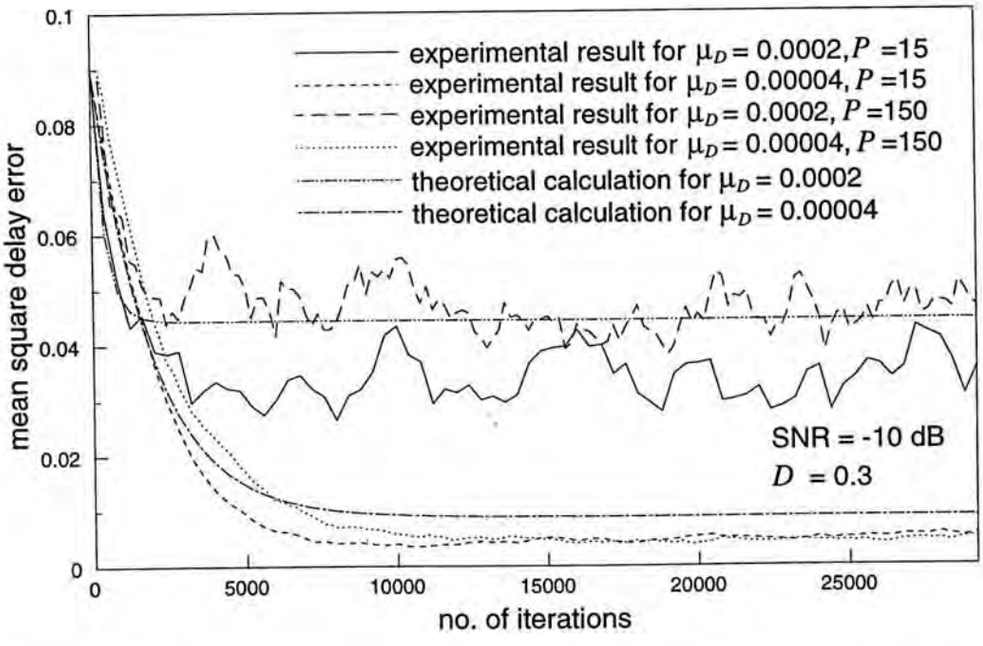

Figure 3.9: Mean square delay errors of the ETDE for different values of μ_D at SNR $= -10$ dB

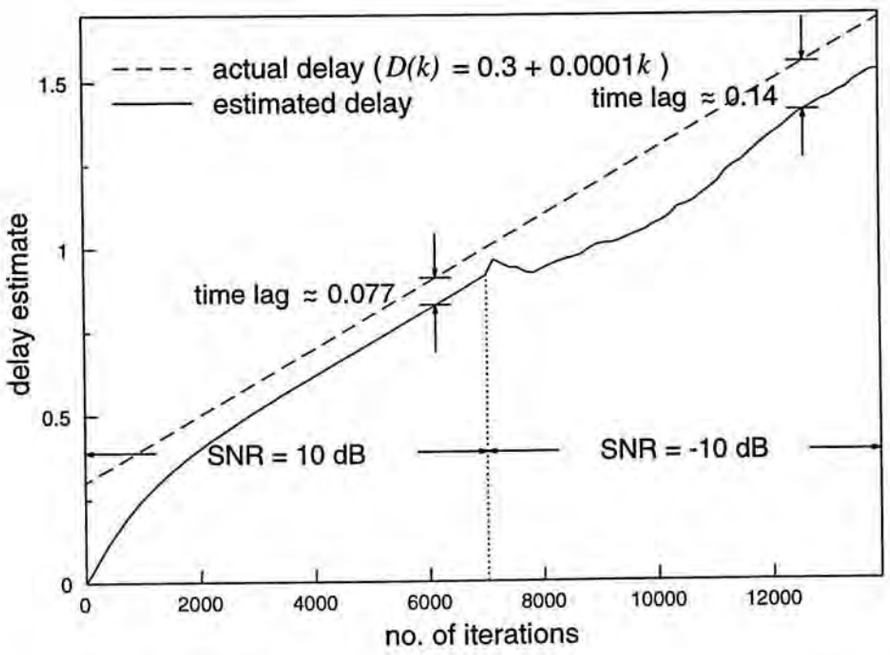

Figure 3.10: Delay estimate of the ETDE for a linearly time-varying delay

3.3.2 Comparative Studies

Figure 3.11 compares the trajectories of the delay estimate of the ETDE, the LMSTDE and the CATDE when SNR = 10 dB and D having a constant value of 0.3. The signal power σ_s^2 was fixed to unity. Two cases were considered for the LMSTDE: (i) all initial filter weights were set to zero and (ii) the initial value of the delay estimate was put to zero by assigning $w_i(0) = \sigma_s^2/(\sigma_s^2 + \sigma_n^2)sinc(i)$. The delay estimates of the ETDE and the CATDE were also assumed to be zero at the beginning. The step sizes μ_D and μ_w had values of 0.0003 and 0.00987 respectively so that the ETDE and the LMSTDE could attain the same delay variance. As given by (3.13), the LMSTDE achieved the fastest convergence in the first experiment and an accurate delay estimate, which was close to 0.3, was obtained after approximately 100 iterations. This supports the rationale of using the LMSTDE to initialize the filter parameters of the ETDE. When the initial estimates of D were set equal, the learning behaviours of the LMSTDE and the ETDE became almost identical. This is confirmed in Figure 3.11 that both methods provided accurate estimates in about 3000 iterations. In addition, at the 3000th iteration, the delay variance of the ETDE

was found to be 7.08×10^{-5} and it was comparable to the CRLB, which has a value of 2.13×10^{-5}. On the other hand, the delay estimate of the CATDE converged to 0.38 in about 2500 iterations. The discrepancy was resulted from the inaccurate mapping using (2.28). The theoretical analysis of this method is verified by noting that $\sigma_s^2/(\sigma_s^2 + \sigma_n^2)sinc(0.3) \approx sinc(0.38)$. At low SNR, the delay estimation error will be much larger. Consequently, the CATDE is only suitable to operate at a noise-free or very high SNR environment.

Performance of the ETDE and the LMSTDE for a step change in the actual delay at SNR $= -10$ dB are illustrated in Figure 3.12. The results of the CATDE are not included since its performance is comparatively poor. The actual delay had a value of 0.3 at the first 40000 iterations and then changed instantaneously to 0.7 afterwards. In order to keep the same delay variance, μ_w was chosen to be 1.32×10^{-5} while $\mu_D = 2 \times 10^{-5}$ in the ETDE. From (2.21) and (3.32), these step sizes corresponded to a delay variance of approximately 0.004. In addition, the filter length of the ETDE was set to 301 in order to reduce the delay bias. It can be seen that the delay estimate of the LMSTDE converged to the desired values at the 15000th and the 55000th iteration while the ETDE gave accurate estimates at the 35000th and the 75000th iteration. Their convergence speeds differed by approximately one half because the time constant of the ETDE was set twice of that of the LMSTDE. Although the delay estimate of the ETDE had a slower convergence rate, it had been found that the measured mean square delay error was approximately equal to 0.002 which was about half of the LMSTDE delay variance. This implies that the performance of these two methods are actually quite similar in this test.

Another test for a fast moving delay at SNR $= -10$ dB was performed and the results are shown in Figure 3.13. In this experiment, the delay D had a value of 0.3 from the first iteration to the 5000th iteration, 0.7 from the 5001th iteration to the 10000th iteration, 0.9 from the 10001th iteration to the 15000th iteration, and 1.2 onwards. In order to track the delay properly, larger values of step sizes were used. In the ETDE, $\mu_D = 2 \times 10^{-4}$ while μ_w had a value of 1.32×10^{-4} in the LMSTDE. In this case, we did not increase P in the ETDE since it has been shown in Section

3.3.1 that the delay bias could not be removed when a large value of μ_D was used. It is observed that the ETDE tracked the step changes after transients. Although it provided biased estimates of D, the measured delay variance of the ETDE was still close to the theoretical value. On the other hand, the performance of the LMSTDE was unsatisfactory and the trajectory of $\hat{D}_w(k)$ fluctuated severely indicating that the delay variance was much larger than the desired value. These fluctuations were due to the occurrence of false peak filter parameters that led to discrepancies in interpolation. A much smaller value of μ_w may be used to locate correct peak weight at each iteration. However, the dynamic behaviour of the LMSTDE will then be greatly degraded.

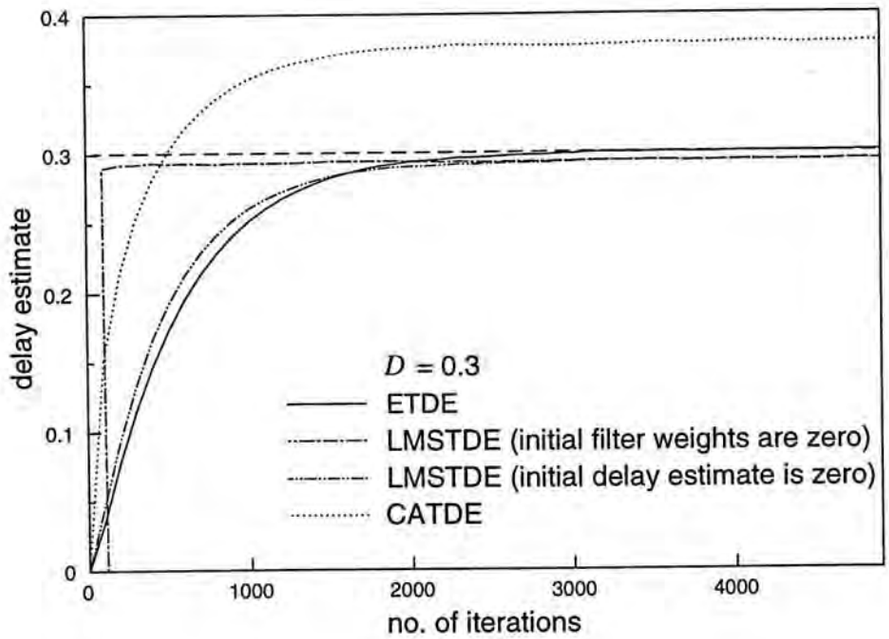

Figure 3.11: Delay estimate of the ETDE and other adaptive TDE methods for a static delay when SNR = 10 dB

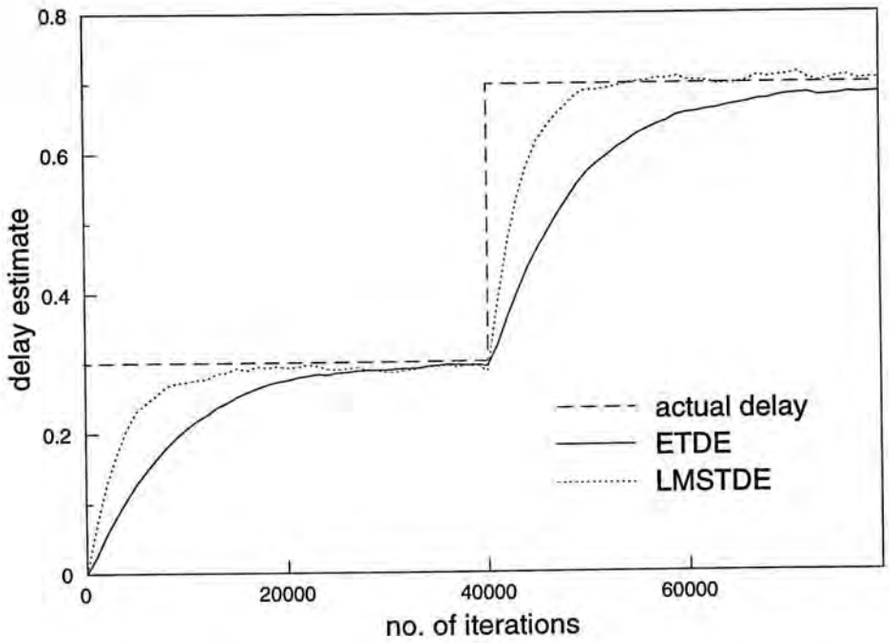

Figure 3.12: Comparison of the ETDE and the LMSTDE for a step change in delay when SNR = -10 dB

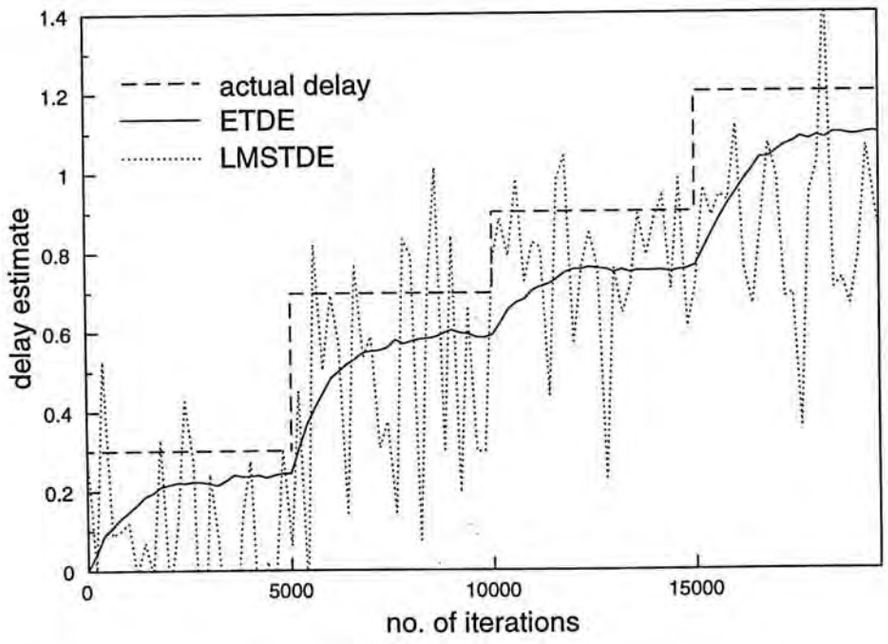

Figure 3.13: Comparison of the ETDE and the LMSTDE for a fast step-changing delay when SNR = -10 dB

3.4 Summary

An adaptive system, which is called the explicit time delay estimator (ETDE), is proposed to estimate the differential delay of a signal received at two spatially separated sensors. Unlike the LMSTDE whose estimated delay is obtained by interpolation of its filter weights, the ETDE updates the time delay estimate explicitly. In addition, the ETDE provides a more reliable approach for delay estimation because wrong estimates due to false peak weight location will not occur. Under very noisy environment, the ETDE gives a satisfactory delay variance although the estimated delay is found to be biased. With the use of lookup tables, the ETDE algorithm is efficient in terms of computational complexity. As an alternative realization of the CATDE, the method provides more accurate and robust delay estimates, particularly under noisy environment. The mean square delay error of the ETDE has been shown to be larger than the CRLB by a few times. The convergence dynamics of the delay estimate and the mean square delay error as well as the delay variance are derived and confirmed by simulations. It is also demonstrated that the ETDE has superior performance than other adaptive TDE methods for both static and nonstationary delays.

Chapter 4

An Improvement to the ETDE

Although the ETDE provides an efficient way to estimate the differential time difference between signals received at two separated sensors, it has been found that its delay estimate is not accurate enough for many practical applications when the signal-to-noise ratio (SNR) is small. In this chapter, an improvement that is based on Wiener solution is made by adding an adaptive gain in series with the ETDE and it is called the explicit time delay and gain estimator (ETDGE) [69],[70]. In Section 4.1, the delay bias of the ETDE is examined and it will be shown that the delay error is a function of the SNR as well as the number of filter taps. Section 4.2 formulates the adaptive algorithm of the ETDGE. It will be proved that this new adaptive system provides unbiased delay estimates for any finite filter length and has smaller delay variances at low SNR. In Section 4.3, performance analysis of the proposed algorithm for both static and linearly moving delays will be given. Whilst experimental results and conclusions are included in Section 4.4 and Section 4.5 respectively.

4.1 Delay Modeling Error of the ETDE

In this section, the delay error of the ETDE for finite filter length is investigated by determining the global minimum of its performance surface which is the mean delay estimate. Equation (3.8) gives the exact value of the ETDE performance surface as

follows,

$$E\{e^2(k)\} = \sigma_s^2 \left(1 + \sum_{i=-P}^{P} sinc^2(i-\hat{D}) - 2\sum_{i=-P}^{P} sinc(i-\hat{D})sinc(i-D)\right) + \sigma_n^2 \left(1 + \sum_{i=-P}^{P} sinc^2(i-\hat{D})\right)$$

Now, differentiating the equation with respect to $\hat{D}$ and then equating the result to zero, the following expression is obtained,

$$\sum_{i=-P}^{P} sinc(i-D)f(i-\bar{D}) = \left(1+\frac{1}{\text{SNR}}\right)\sum_{i=-P}^{P} sinc(i-\bar{D})f(i-\bar{D}) \qquad (4.1)$$

where $\bar{D}$ denotes the global minimum of the error surface. In the absence of noise or if the filter length (2P+1) is infinitely large, then $\bar{D} = D$. As a result, the ETDE is capable to provide unbiased delay measurements. However, for other circumstances, the unbiased property of the ETDE might not necessarily exist. Although a closed form solution for $\bar{D}$ is not available, numerical values of $\bar{D}$ have been found by using Newton-Raphson method to reveal the effect of different SNRs and P on $\bar{D}$. The delay bias, which is defined as $|\bar{D} - D|$, is plotted versus D for SNR $= -10$ dB and $P = 15$ as shown in Figure 4.1. It can be seen that the bias becomes largest at $D = 0.2$, and it has a value of 0.066. Note that the shape of the graph may change as the parameters vary. Figure 4.2 and Figure 4.3 illustrate the delay bias against SNR and P for $D = 0.25$ respectively. It is seen that the delay error decreases when SNR or P increases and becomes almost negligible for high SNR. Moreover, the delay bias can be reduced by one-tenth approximately by either increasing the SNR by 10 dB or by a ten-fold increase of P. In practice, there is a trade-off to use a very large value of P to reduce delay error because the computational complexity of the algorithm, on the other hand, will increase considerably which makes real-time application not feasible. On the contrary, accurate delay estimation may not be achieved by the ETDE if P is not chosen sufficiently large enough or when the estimator is operated at low SNR.

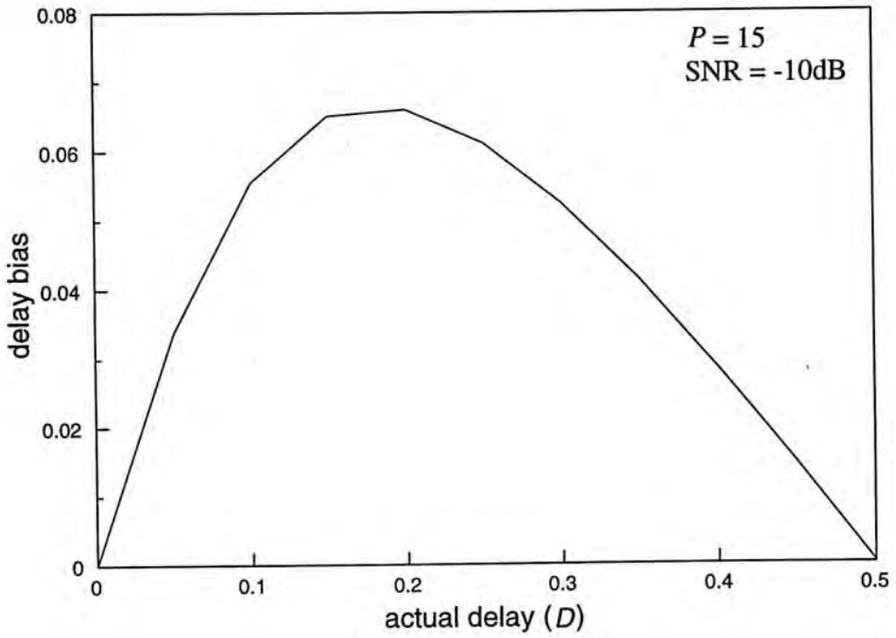

Figure 4.1: Delay bias of the ETDE versus D

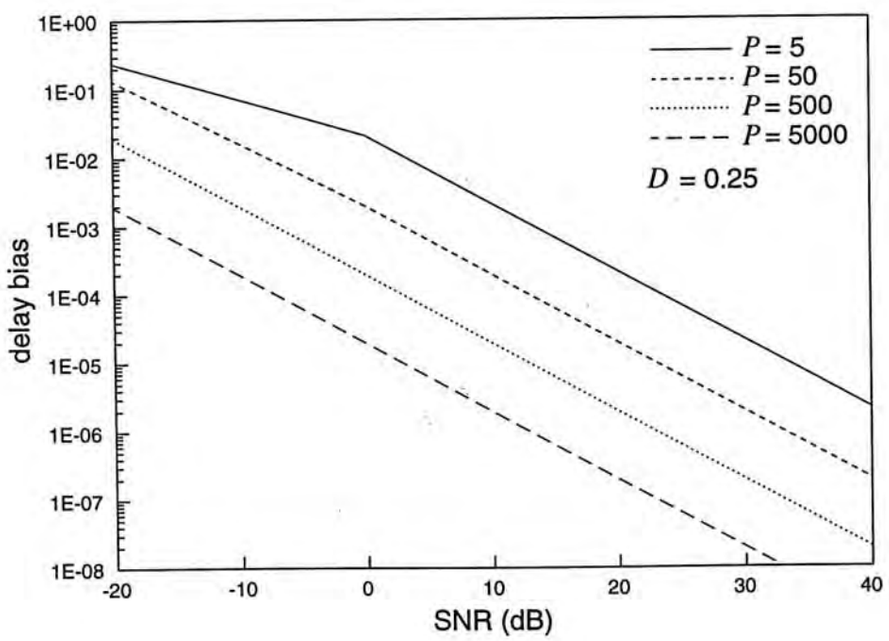

Figure 4.2: Delay bias of the ETDE versus SNR

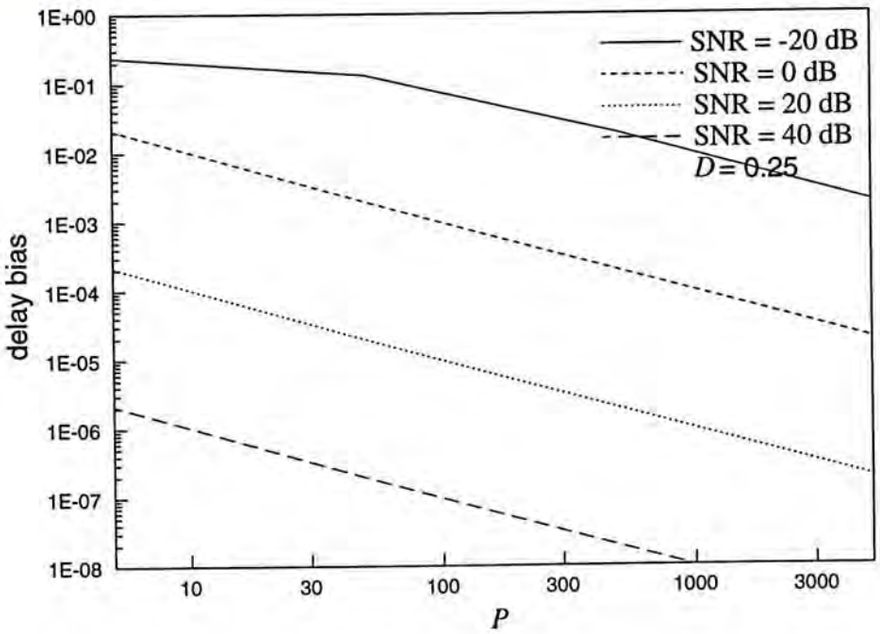

Figure 4.3: Delay bias of the ETDE versus P

4.2 The Explicit Time Delay and Gain Estimator (ETDGE)

From (2.10), the optimal filter weights of a time delay FIR filter are given by

$$w_i^o = \frac{\text{SNR}}{1+\text{SNR}} sinc(i-D), \qquad P \geq i \geq -P$$

In the ETDE, the effect of additive noise is not considered in the adaptation which, however, plays a significant role in the delay estimation process, particularly the steady state delay error. In this section, an improved version of the ETDE is proposed by adding a variable gain, $\hat{g}(k)$, in series with the delay estimator, to adapt the scaling factor SNR/(1 + SNR) separately such that the Wiener solution can be attained. The new system is called the explicit time delay and gain estimator (ETDGE) and its block diagram is depicted in Figure 4.4. The system inputs, $x(k)$ and $y(k)$, are the received signals and the error signal, $e(k)$, in this case can be computed from

$$e(k) = y(k) - \hat{g}(k) \sum_{i=-P}^{P} sinc(i - \hat{D}_G(k)) x(k-i) \tag{4.2}$$

To avoid any confusion, the delay estimate of the ETDGE is denoted by $\hat{D}_G(k)$. While $\hat{D}_G(k)$ is adjusted using (3.3), the gain parameter of the ETDGE is adapted simultaneously and independently according to Widrow's LMS algorithm [45],

$$\begin{aligned}\hat{g}(k+1) &= \hat{g}(k) - \mu_g \frac{\partial e^2(k)}{\partial \hat{g}(k)} \\ &= \hat{g}(k) + 2\mu_g e(k) \sum_{i=-P}^{P} sinc(i - \hat{D}_G(k)) x(k-i) \end{aligned} \tag{4.3}$$

where μ_g is the convergence step size for $\hat{g}(k)$. It is noted that for the same value of P, in terms of computational complexity, the ETDGE requires an extra 1 addition and 4 multiplications.

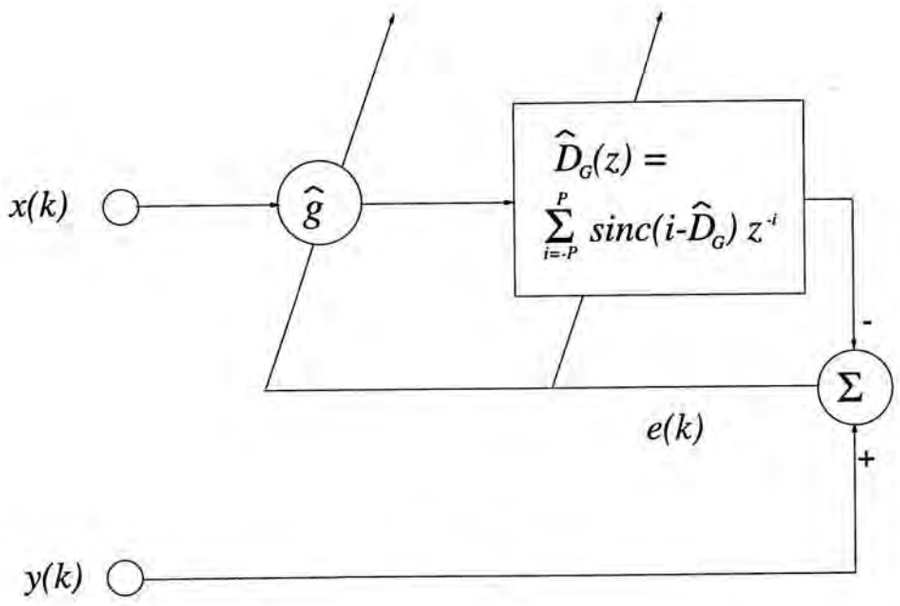

Figure 4.4: System block diagram of the ETDGE

For ease of analysis, we assume that the gain estimate and delay estimate are uncorrelated with the received signals. Squaring both sides of (4.2) and taking expectation, the performance surface of the ETDGE can be found as

$$\begin{aligned} E\{e^2(k)\} &= E\{(s(k-D) + n_2(k) - \hat{g}\tilde{s}(k-\hat{D}_G) - \hat{g}\tilde{n}_1(k-\hat{D}_G))^2\} \\ &= E\{(s(k-D) - \hat{g}\tilde{s}(k-\hat{D}_G))^2\} + E\{(n_2(k) - \hat{g}\tilde{n}_1(k-\hat{D}_G))^2\} \end{aligned}$$

$$= (\sigma_s^2 + \sigma_n^2)\left(1 + \hat{g}^2 \sum_{i=-P}^{P} sinc^2(i - \hat{D}_G)\right)$$
$$-2\hat{g}\sigma_s^2 \sum_{i=-P}^{P} sinc(i - D)sinc(i - \hat{D}_G) \tag{4.4}$$

It is obvious that the ETDGE also has a multimodal performance surface. In order to achieve global convergence, proper initialization of $\hat{D}_G(k)$ is necessary. It is intriguing to note that the initial gain value, $\hat{g}(0)$, can be chosen arbitrarily since $E\{e^2(k)\}$ comprises only the first and second order term of $\hat{g}$. Partial differentiating $E\{e^2(k)\}$ with respect to $\hat{g}$ and $\hat{D}_G$, we get

$$\frac{\partial E\{e^2(k)\}}{\partial \hat{g}} = 2\hat{g}(\sigma_s^2 + \sigma_n^2) \sum_{i=-P}^{P} sinc^2(i - \hat{D}_G)$$
$$-2\sigma_s^2 \sum_{i=-P}^{P} sinc(i - D)sinc(i - \hat{D}_G) \tag{4.5}$$

and

$$\frac{\partial E\{e^2(k)\}}{\partial \hat{D}_G} = (\sigma_s^2 + \sigma_n^2) \cdot \hat{g}^2 \cdot \sum_{i=-P}^{P} 2sinc(i - \hat{D}_G)\frac{\partial sinc(i - \hat{D}_G)}{\partial \hat{D}_G}$$
$$-2\hat{g}\sigma_s^2 \sum_{i=-P}^{P} sinc(i - D)\frac{\partial sinc(i - \hat{D}_G)}{\partial \hat{D}_G}$$
$$= -2\hat{g}^2(\sigma_s^2 + \sigma_n^2) \sum_{i=-P}^{P} sinc(i - \hat{D}_G)f(i - \hat{D}_G)$$
$$+2\hat{g}\sigma_s^2 \sum_{i=-P}^{P} sinc(i - D)f(i - \hat{D}_G) \tag{4.6}$$

Setting the above two equations to zero and consider the second derivatives [71] respectively, it can be shown that the global minimum occurs when

$$\hat{g} = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_n^2} = \frac{\text{SNR}}{1 + \text{SNR}} \tag{4.7a}$$

and

$$\hat{D}_G = D \tag{4.7b}$$

Equations (4.7a) and (4.7b) relate the system variables to the Wiener solution and thus the delay estimate of the ETDGE is unbiased for any finite P. Hence, unlike other parametric methods for time delay estimation [29]-[35],[39],[63] that require a

long filter length in order to reduce truncation error and to avoid biasedness, low order filter may be used for the ETDGE which eventually reduces the computational complexity.

4.3 Performance Analysis

Using (4.2) and (3.3), the expected value of $\hat{D}_G(k)$ of the ETDGE becomes

$$\begin{aligned}
&E\{\hat{D}_G(k+1)\} \\
&= E\{\hat{D}_G(k)\} - 2\mu_D E\left\{(y(k) - \hat{g}(k)\tilde{x}(k - \hat{D}_G(k))) \cdot \sum_{i=-P}^{P} x(k-i) f(i - \hat{D}_G(k))\right\} \\
&= E\{\hat{D}_G(k)\} - 2\mu_D E\{(s(k-D) - \hat{g}(k)\tilde{s}(k - \hat{D}_G(k)) + n_2(k) \\
&\quad -\hat{g}(k)\tilde{n}_1(k - \hat{D}_G(k))) \cdot \sum_{i=-P}^{P} (s(k-i) + n_1(k-i)) f(i - \hat{D}_G(k))\} \\
&= E\{\hat{D}_G(k)\} - 2\mu_D E\left\{\sum_{j=-\infty}^{\infty} s(k-j)sinc(j-D) \sum_{i=-P}^{P} s(k-i) f(i - \hat{D}_G(k))\right\} \\
&\quad +2\mu_D E\left\{\hat{g}(k) \sum_{j=-P}^{P} s(k-j)sinc(j - \hat{D}_G(k)) \sum_{i=-P}^{P} s(k-i) f(i - \hat{D}_G(k))\right\} \\
&\quad +2\mu_D E\left\{\hat{g}(k) \sum_{j=-P}^{P} n_1(k-j)sinc(j - \hat{D}_G(k)) \sum_{i=-P}^{P} n_1(k-i) f(i - \hat{D}_G(k))\right\} \\
&= E\{\hat{D}_G(k)\} - 2\mu_D \sigma_s^2 E\left\{\sum_{i=-P}^{P} sinc(i-D) f(i - \hat{D}_G(k))\right\} \\
&\quad +2\mu_D(\sigma_s^2 + \sigma_n^2) E\left\{\hat{g}(k) \sum_{i=-P}^{P} sinc(i - \hat{D}_G(k)) f(i - \hat{D}_G(k))\right\} \qquad (4.8)
\end{aligned}$$

When P is chosen large enough, we have

$$\sum_{i=P}^{P} sinc(i - \hat{D}_G(k)) f(i - \hat{D}_G(k)) \approx 0 \qquad (4.9)$$

and the last term of (4.8) can then be neglected. Using (3.16) to (3.18), the learning characteristics of the delay estimate of the ETDGE is given by

$$E\{\hat{D}_G(k)\} \approx D + (\hat{D}_G(0) - D)\left(1 - \frac{2}{3}\mu_D \sigma_s^2 \pi^2\right)^k \qquad (4.10)$$

which is identical to that of the ETDE and the parameter $\hat{D}_G(0)$ represents the initial delay estimate. Notice that for small P, $\hat{D}_G(k)$ will still converge to D but

its learning behaviour may be somewhat different from the predicted value.

When $\hat{D}_G(k)$ approaches D, (4.3) is used to derive the learning trajectory of $\hat{g}(k)$,

$$\begin{aligned} E\{\hat{g}(k+1)\} &= E\{\hat{g}(k)\} + 2\mu_g E\{(s(k-D) + n_2(k) - \hat{g}(k)\tilde{s}(k-D) \\ &\quad -\hat{g}(k)\tilde{n}_1(k-D)) \cdot \sum_{i=-P}^{P} sinc(i-D)(s(k-i) + n_1(k-i))\} \\ &= E\{\hat{g}(k)\} + 2\mu_g\sigma_s^2 - 2\mu_g E\{\hat{g}(k)\}(\sigma_s^2 + \sigma_n^2) \end{aligned} \tag{4.11}$$

Solving (4.11) yields

$$E\{\hat{g}(k)\} = g + (\hat{g}(0) - g)(1 - 2\mu_g(\sigma_s^2 + \sigma_n^2))^k \tag{4.12}$$

where

$$g = \frac{\text{SNR}}{1 + \text{SNR}} \tag{4.13}$$

Assume $\hat{g}(k) \to g$ in steady state, the delay variance of the ETDGE, $var(\hat{D}_G)$, is found to be (See Appendix C)

$$\begin{aligned} var(\hat{D}_G) &\triangleq \lim_{k\to\infty}\{(\hat{D}_G(k) - D)^2\} \\ &= \frac{\mu_D\sigma_s^2(1 + 2\text{SNR})}{\text{SNR}^2} \end{aligned} \tag{4.14}$$

Comparing (4.14) with the variance of the ETDE as derived in (3.25), it can be seen that they have the same value of $2\mu_D\sigma_n^2$ when SNR $\gg 1$. However, when SNR $\ll 1$, the delay variance of the ETDGE equals $\mu_D\sigma_n^2/\text{SNR}$, which is only half of that of the ETDE.

On the other hand, the variance of $\hat{g}(k)$, $var(\hat{g})$, can be shown to be (See Appendix D)

$$var(\hat{g}) = \begin{cases} 2\mu_g\sigma_n^2 & , \ \text{SNR} \gg 1 \\ \\ \mu_g\sigma_n^2 & , \ \text{SNR} \ll 1 \end{cases} \tag{4.15}$$

For both SNR $\gg 1$ and SNR $\ll 1$, the value of $var(\hat{g})$ is directly proportional to the convergence parameter μ_g and the noise power σ_n^2.

The tracking behaviour of the delay estimate for a linearly varying delay, $D(k) = D_0 + \lambda k$, where D_0 is the TDOA at $k = 0$ and λ is the Doppler stretch compression, is the same as that of the ETDE, that is,

$$E\{\hat{D}_G(k)\} \approx D_0 + \lambda k + \left(\hat{D}_G(0) - D_0 + \frac{3\lambda}{2\mu_D\sigma_s^2\pi^2}\right)\left(1 - \frac{2}{3}\mu_D\sigma_s^2\pi^2\right)^k - \frac{3\lambda}{2\mu_D\sigma_s^2\pi^2} \tag{4.16}$$

Similar to (3.31), the mean square delay error of the ETDGE, denoted by $\xi(\hat{D}_G)$, is given by

$$\xi(\hat{D}_G) \approx \frac{\mu_D\sigma_s^2(1 + 2\text{SNR})}{\text{SNR}^2} + \frac{9\lambda^2}{4\mu_D^2\sigma_s^4\pi^4} \tag{4.17}$$

From (4.17) and (3.31), it can be seen that the ETDGE in fact has a smaller mean square delay error since its estimated delay variance is smaller.

4.4 Simulation Results

The performance of the ETDGE for both static and nonstationary delays had been studied thoroughly by computer simulations. In our experiments, the signal source $s(k)$ and the additive noises $n_1(k)$ and $n_2(k)$ were generated by a random number generator of Gaussian distribution with a white spectrum. The corrupting noises had unity power and different SNRs were obtained by proper scaling of the source signal. Without loss of generality, the initial values of the variable gain and the estimated delay were set to 1 and 0 respectively. To demonstrate that the ETDGE is unbiased for all filter lengths, P was chosen to be 3. The results provided were averages of 200 independent runs.

The convergence characteristics of $\hat{D}_G(k)$ and $\hat{g}(k)$ for a constant delay $D = 0.3$ were depicted in Figure 4.5 and Figure 4.6 respectively. The SNR was initially set at 10 dB and then step-changed to -10 dB at the 7000th iteration. The step sizes were chosen such that the time constants of $\hat{D}_G(k)$ and $\hat{g}(k)$ were fixed under both SNR conditions. At SNR = 10 dB, μ_D and μ_g were assigned with values of 0.00002 and 0.0001 respectively while $\mu_D = 0.002$ and $\mu_g = 0.001$ for SNR $= -10$ dB. In Figure 4.5, it can be seen that the delay estimate converged to its optimal value

at approximately the 4000th iteration. As expected, the learning curve started to fluctuate after about 7000 iterations due to a large reduction in the SNR. Since (4.10) is derived by using the first order approximation of the f function, there were small discrepancies between the theoretical and experimental transient behaviours of $\hat{D}_G(k)$, but this is not significant. Figure 4.6 shows that $\hat{g}(k)$ converged to 0.91 and 0.09, which were their desired values, at the 4000th and the 10000th iteration respectively. It is observed that at SNR $= -10$ dB, the convergence behaviour of the experimental gain agreed very well with the theoretical value as given by (4.12). However, at the beginning of the adaptation, the learning trajectory of $\hat{g}(k)$ damped and converged at a slower rate than the predicted curve because of the inaccurate estimate of D during transients.

Figure 4.7 illustrates the tracking performance of the ETDGE when the actual delay was a linearly time-varying function given by $D(k) = 0.3 + 0.0001k$. The values of the SNR, μ_D and μ_g were set according to the previous tests. The learning curve of $\hat{g}(k)$ is not shown here since it is very similar to Figure 4.6. It is observed in Figure 4.7 that after 3000 iterations, the delay estimate converged and lagged $D(k)$ by 0.089 at SNR $= 10$ dB and by 0.11 at SNR $= -10$ dB. Evaluating the steady state time lag using (4.17) gives a value of 0.076, which is smaller than both measurements. The difference is mainly due to the approximations made in deriving (4.17).

The mean square errors of $\hat{D}_G(k)$ and $\hat{g}(k)$ for the above two cases are tabulated along with the theoretical values as shown in Table 4.1. It can be seen that for both stationary and nonstationary delays, the experimental results for $var(\hat{g})$ were fairly close to those derived from (4.15). Furthermore, the measured variance of the TDOA also conformed to the expected value when the original delay was a constant. However, the result for $\xi(\hat{D}_G)$ was less satisfactory when the the delay was time-varying. This is mainly due to the difference between the actual and theoretical time lags as illustrated in Figure 4.7. At SNR $= 10$ dB, it is obvious that the value of $\xi(\hat{D}_G)$ was dominated by the time lag. In this case, a larger value of μ_D can be used to attain a smaller mean square delay error. We had also performed the above

experiments for $P = 15$ and all the results were very similar to those of using $P = 3$.

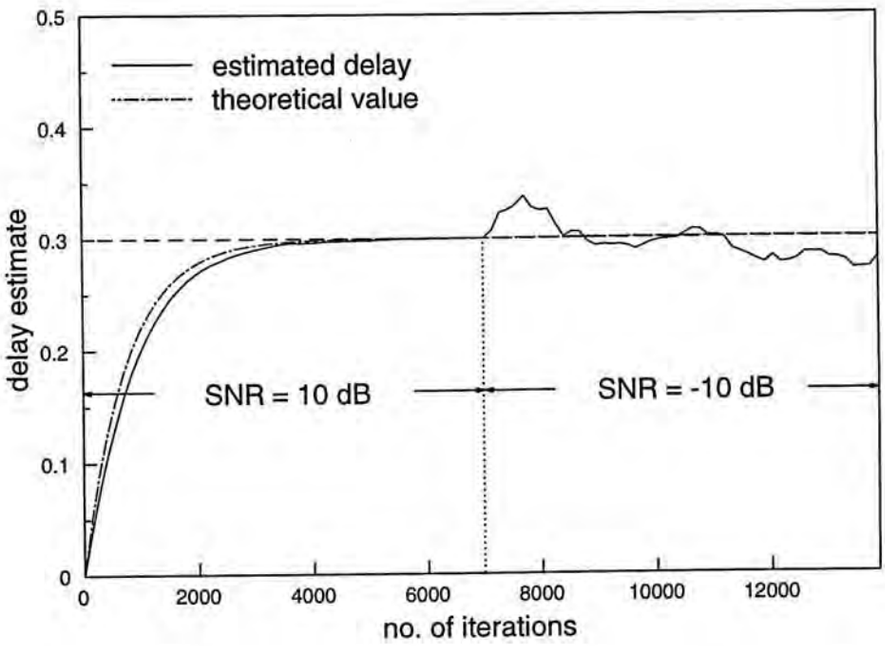

Figure 4.5: Delay estimate of the ETDGE for a static delay

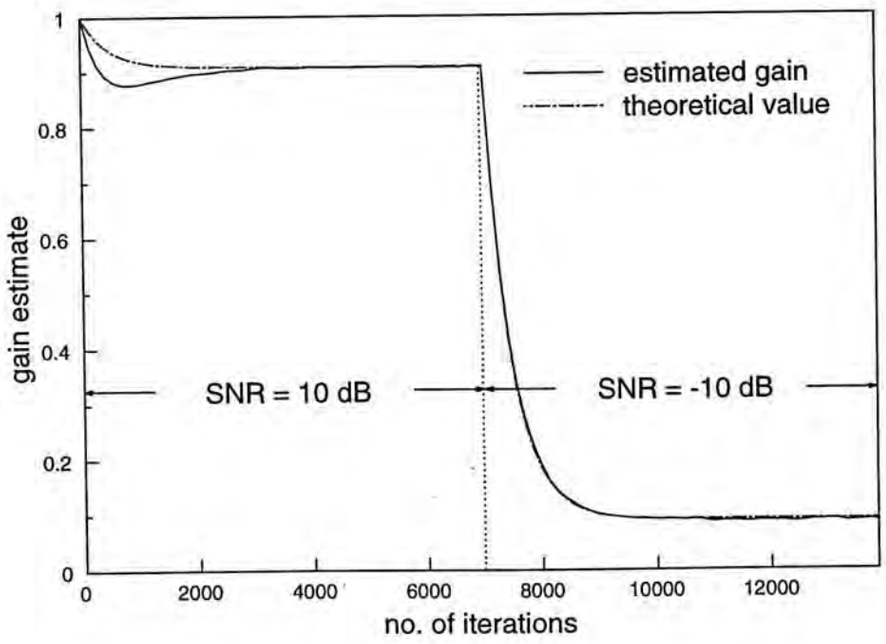

Figure 4.6: Gain estimate of the ETDGE in a nonstationary SNR environment

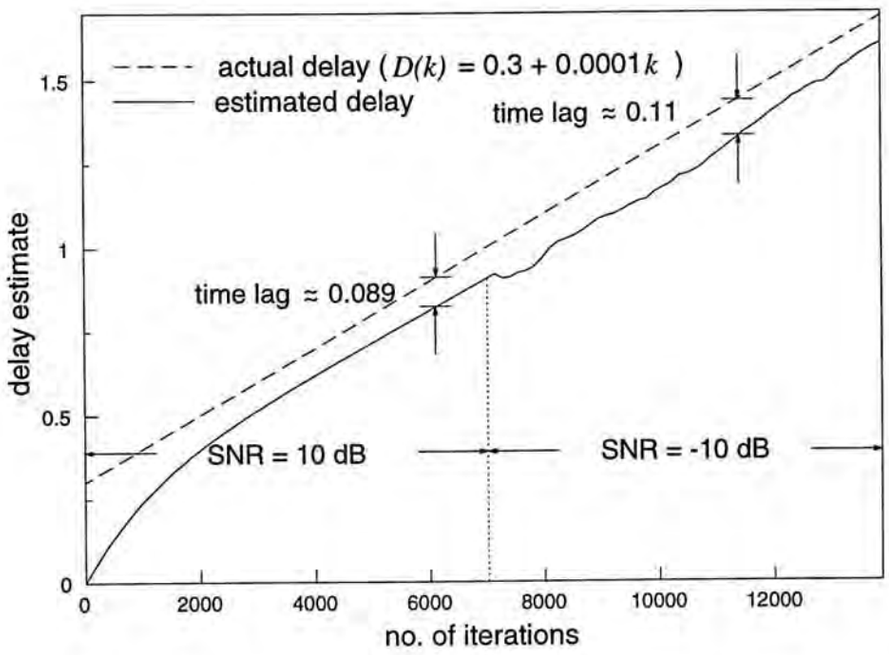

Figure 4.7: Delay estimate of the ETDGE for a linearly moving delay

		SNR=10 dB		SNR=−10 dB	
		Measured value	Theoretical value	Measured value	Theoretical value
$D = 0.3$	$var(\hat{D}_G)$	4.43×10^{-5}	4.20×10^{-5}	2.87×10^{-2}	2.40×10^{-2}
	$var(\hat{g})$	2.00×10^{-4}	2.00×10^{-4}	1.07×10^{-3}	1.00×10^{-3}
$D(k) = 0.3$ $+0.0001k$	$\xi(\hat{D}_G)$	7.53×10^{-3}	5.82×10^{-3}	3.63×10^{-2}	2.98×10^{-2}
	$var(\hat{g})$	2.26×10^{-4}	2.00×10^{-4}	1.05×10^{-3}	1.00×10^{-3}

Table 4.1: Mean square errors of $\hat{D}_G(k)$ and $\hat{g}(k)$ of the ETDGE

4.5 Summary

In conclusion, the ETDE has been proved to be a biased estimator and the corresponding delay error decreases as the number of filter taps or the SNR increases. Based on the Wiener solution, an improved version of the ETDE called the explicit time delay and gain estimator (ETDGE) is proposed by adding an adaptive gain control to the delay estimator. It is shown that the modified algorithm provides smaller delay variances and unbiased delay estimates for all practical filter lengths. The performance of the ETDGE is analyzed for both static and linearly time-varying delays and the theoretical results were confirmed by computer simulations for different SNR conditions.

Chapter 5

TDE in the Presence of Multipath Propagation

So far in our study of the time delay estimation (TDE) problem, it has been assumed that the signal source arrives at each sensor through a single propagation path only. However, multipath transmissions frequently occur in sonar, radar and many other applications [46],[72],[73]. For example, multipaths may come from bottom bounces or reflections of ocean surface in sonar, and reflections of buildings or mountains in radar. When the presence of multipath transmissions is ignored, the delay estimation accuracy based on direct-path-only propagation will be deteriorated. Whereas performance gain can always be achieved if we know the multipath structure since this information contributes additional independent arrays [51]. With the use of the explicit time delay estimator (ETDE) and/or the explicit time delay and gain estimator (ETDGE), two adaptive systems are proposed to extract the time difference of arrival (TDOA) as well as the multipath parameters when each of the two spatially separated sensors receives the radiated source plus its attenuated and delayed replicas under high signal-to-noise (SNR) environment. The first method, namely, the multipath cancellation time delay estimator (MCTDE) [74],[75], consists of two adaptive IIR filters for eliminating the multipath component at each transmission channel and one ETDE to estimate the time difference between the two sensor outputs. Although the multipath canceller of the MCTDE is originally designed to cancel out a replica whose interpath delay is an integral multiple of the sampling interval, it can also be used for any real-valued multipath delay. On

the other hand, the second system, which is called the multipath equalization time delay estimator (METDE) [74],[76], employs only the ETDE and the ETDGE as its building blocks. Based on the concept of multipath equalization, the system parameters of the METDE are adapted through minimization of a multimodal cost function whose global minimum is uniquely related to the actual values of the time delay and the multipath parameters.

In Section 5.1, the problem of TDE in the presence of multipath propagation is formulated and some necessary assumptions are made. Section 5.2 derives the structure and algorithm of the MCTDE. Convergence dynamics of the system parameters are given and effects of additive noises are discussed. In addition, numerical examples are provided for theoretical validation and performance evaluation. The multipath equalization approach, that uses the MCTDE to initialize the estimated parameters, is presented in Section 5.3. Performance analysis of the METDE, together with simulation results, will also be given. The MCTDE and the METDE are summarized and compared in Section 5.4.

5.1 The Multipath TDE problem

The mathematical model of time delay estimation in the presence of multipath transmissions can be stated as follows. Given the received outputs of two spatially separated sensors,

$$x(k) = s(k) + \sum_{p=1}^{M_1} \alpha_{1p} s(k - \Delta_{1p}) + n_1(k) \tag{5.1a}$$

$$y(k) = s(k - D) + \sum_{q=1}^{M_2} \alpha_{2q} s(k - \Delta_{2q}) + n_2(k) \tag{5.1b}$$

where the source signal $s(k)$ and the additive noises $n_1(k)$ and $n_2(k)$ are assumed to be zero-mean, stationary, uncorrelated and white Gaussian processes. As usual, the sampling period is assumed to be unity and the parameter D represents the time delay between the two sensor outputs. The multipath transmissions are characterized by the gain factors α_{1p} and α_{2q} as well as the interpath delays Δ_{1p} and

Δ_{2q}, for $M_1 \geq p \geq 1$ and $M_2 \geq q \geq 1$. Without loss of generality, we let $\alpha_{11} > \alpha_{12} > \cdots > \alpha_{1M_1}$, $\alpha_{21} > \alpha_{22} > \cdots > \alpha_{2M_2}$, $\Delta_{11} < \Delta_{12} < \cdots < \Delta_{1M_1}$ and $\Delta_{21} < \Delta_{22} < \cdots < \Delta_{2M_2}$. Notice that the multipath gains must lie between 0 and 1 while Δ_{1p} and Δ_{2q} should be larger than zero and D respectively. The integers M_1 and M_2 are the numbers of multipaths contained in $x(k)$ and $y(k)$ respectively and they are supposed to be known *a priori*. It is also assumed that the multipath delays are resolvable [49], that is, their differences are all larger than one. The task here is to estimate D, α_{1p}, α_{2q}, Δ_{1p} and Δ_{2q} from the received signals, $x(k)$ and $y(k)$.

5.2 TDE with Multipath Cancellation (MCTDE)

In this section, an adaptive system is proposed to tackle the multipath TDE problem of (5.1) for $M_1 = M_2 = 1$. The sensor outputs are now expressed as

$$x(k) = s(k) + \alpha_{11}s(k - \Delta_{11}) + n_1(k) \tag{5.2a}$$

and

$$y(k) = s(k - D) + \alpha_{21}s(k - D - \Delta'_{21}) + n_2(k) \tag{5.2b}$$

where $\Delta'_{21} = \Delta_{21} - D$. Under this circumstance, we need to estimate the parameters D, α_{11}, α_{21}, Δ_{11}, and Δ'_{21} instead of Δ_{21}. It is noted that an estimate of Δ_{21} can be obtained by summing the estimates of Δ'_{21} and D.

5.2.1 Structure and Algorithm

To simplify the analysis, let us first assume that noise is absent in both sensors and the multipath delays, Δ_{11} and Δ'_{21}, are integral multiples of the sampling period. Let the Z-domain representations of $s(k)$, $x(k)$ and $y(k)$ be $S(z)$, $X(z)$ and $Y(z)$ respectively. Taking Z transform of (5.2) gives

$$X(z) = I_1(z)S(z) \tag{5.3a}$$

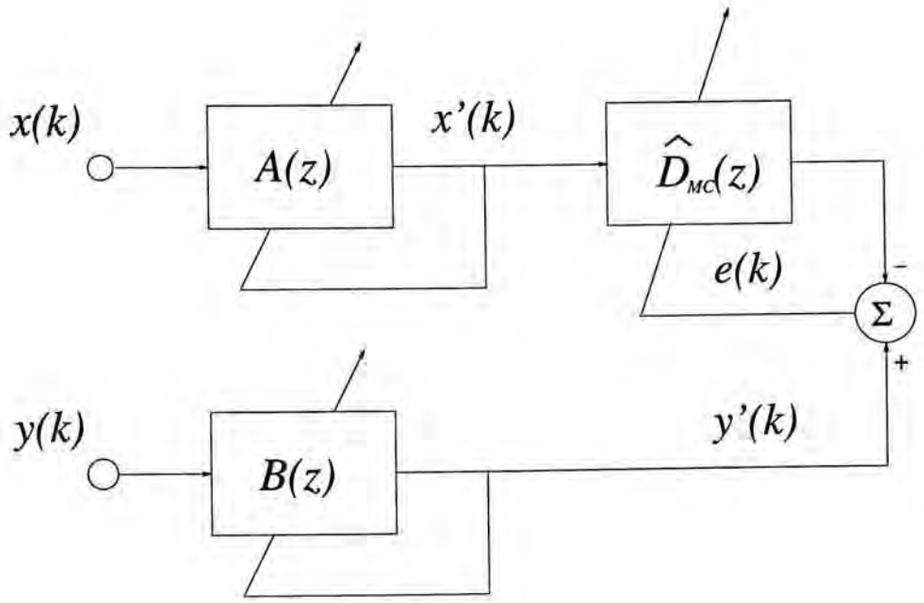

Figure 5.1: System block diagram of the MCTDE

and

$$Y(z) = z^{-D} I_2(z) S(z) \tag{5.3b}$$

where

$$I_1(z) = 1 + \alpha_{11} z^{-\Delta_{11}} \tag{5.4a}$$

$$I_2(z) = 1 + \alpha_{21} z^{-\Delta'_{21}} \tag{5.4b}$$

and

$$z^{-D} = \sum_{n=-\infty}^{\infty} sinc(n - D) z^{-n} \tag{5.4c}$$

The second terms of $I_1(z)$ and $I_2(z)$ essentially represent the multipath effect in $x(k)$ and $y(k)$.

From (5.3), it is trivial that $s(k)$ and $s(k - D)$ can be recovered by passing $x(k)$ and $y(k)$ through $1/I_1(z)$ and $1/I_2(z)$ respectively and these two inverse filters are stable because the zeros of both $I_1(z)$ and $I_2(z)$ are inside the unit circle. When $s(k)$ and $s(k - D)$ are available, the ETDE can be used to model the TDOA between the two filtered signals. The multipath cancellation time delay estimator (MCTDE) [74],[75] adopts this concept and its system block diagram is depicted in Figure 5.1. It comprises two adaptive IIR filters to eliminate the multipath in each channel,

and an ETDE to estimate the time difference D. Of course, the ETDGE can also be used instead of the ETDE. However, it is noticed that the ETDGE provides no significant improvement over the ETDE for the multipath TDE problem in the presence of noise and this will be discussed in Section 5.2.4.

The transfer functions of the two multipath cancellers, $A(z)$ and $B(z)$, are given by

$$A(z) = \frac{1}{1 + \sum_{i=1}^{M} a_i z^{-i}} \tag{5.5a}$$

and

$$B(z) = \frac{1}{1 + \sum_{j=1}^{M} b_j z^{-j}} \tag{5.5b}$$

where $M \geq \max\{\Delta_{11}, \Delta'_{21}\}$. Apparently, $A(z)$ and $B(z)$ have the same filter structure. Upon convergence, we expect $a_i = 0$ for $1 \leq i \leq M$ except $a_{\Delta_{11}} = \alpha_{11}$ and $b_j = 0$ for $1 \leq j \leq M$ except $b_{\Delta'_{21}} = \alpha_{21}$.

Recall from Chapter 3, the transfer function of the delay estimator in the MCTDE, $\hat{D}_{MC}(z)$, is given by

$$\hat{D}_{MC}(z) = \sum_{n=-P}^{P} sinc(n - \hat{D}_{MC}) z^{-n} \tag{5.6}$$

where $\hat{D}_{MC}$ is the estimate of D and the time index is dropped for convenience. As a result, the overall output error, $e(k)$, is of the form

$$e(k) = y'(k) - \sum_{n=-P}^{P} sinc(n - \hat{D}_{MC}(k)) x'(k-n) \tag{5.7}$$

where

$$x'(k) = x(k) - \sum_{i=1}^{M} a_i(k) x'(k-i) \tag{5.8a}$$

and

$$y'(k) = y(k) - \sum_{j=1}^{M} b_j(k) y'(k-j) \tag{5.8b}$$

If $A(z)$, $B(z)$ and $\hat{D}_{MC}(z)$ are adjusted by minimizing the mean square output error, $E\{e^2(k)\}$, the desired solutions may not necessarily be obtained. It is because the

steady state values of the filter coefficients will settle in any real values that satisfy the following relation,

$$\frac{\hat{D}_{MC}(z)A(z)}{B(z)} = \frac{z^{-D}I_2(z)}{I_1(z)} \tag{5.9}$$

To overcome this problem, the filter weights of $A(z)$ and $B(z)$ are adapted by minimizing their respective mean square outputs, $E\{x'^2(k)\}$ and $E\{y'^2(k)\}$. The updating rules are based on Widrow's LMS algorithm [45] with the use of their approximate stochastic gradients. The MCTDE algorithm is, therefore, given by

$$\begin{aligned} a_i(k+1) &\approx a_i(k) - \mu_\alpha \frac{\partial x'^2(k)}{\partial a_i(k)}, \qquad M \geq i \geq 1 \\ &= a_i(k) + 2\mu_\alpha x'(k)x'(k-i) \end{aligned} \tag{5.10}$$

$$\begin{aligned} b_j(k+1) &\approx b_j(k) - \mu_\alpha \frac{\partial y'^2(k)}{\partial b_j(k)}, \qquad M \geq j \geq 1 \\ &= b_j(k) + 2\mu_\alpha y'(k)y'(k-j) \end{aligned} \tag{5.11}$$

$$\hat{D}_{MC}(k+1) = \hat{D}_{MC}(k) - 2\mu_D e(k) \sum_{n=-P}^{P} f(n - \hat{D}_{MC}(k))x'(k-n) \tag{5.12}$$

where μ_α is the adaptation parameter for the multipath cancellers.

It is noteworthy that after some iterations, Δ_{11} and Δ'_{21} can be determined from the peaks of $\{a_i(k)\}$ and $\{b_j(k)\}$ respectively. That is,

$$\Delta_{11} = \arg\max_i \{a_i(k)\} \tag{5.13a}$$

$$\Delta'_{21} = \arg\max_j \{b_j(k)\} \tag{5.13b}$$

Once Δ_{11} and Δ'_{21} are found, we only need to adjust the peak weights in order to reduce the variances of the system parameters as well as the computational complexity of the algorithm.

5.2.2 Convergence Dynamics

We shall next investigate the performance of the multipath cancellers $A(z)$ and $B(z)$. Since the filter structures of $A(z)$ and $B(z)$ are identical, only $A(z)$ is examined here.

The analysis for $B(z)$ will follow suit. The performance surface of $A(z)$ is obtained by taking the mean square value of $x'(k)$,

$$\begin{aligned} E\{x'^2(k)\} &= \frac{1}{2\pi j}\oint \Phi_{x'x'}(z)\frac{dz}{z} \\ &= \frac{\sigma_s^2}{2\pi j}\oint |A(z)(1+\alpha_{11}z^{-\Delta_{11}})|^2\frac{dz}{z} \end{aligned} \tag{5.14}$$

where $j=\sqrt{-1}$ and $\Phi_{x'x'}(z)$ denotes the auto-power spectrum of $x'(k)$. Considering the inequality

$$\frac{1}{2\pi j}\oint |X'(z)-S(z)|^2\frac{dz}{z} \geq 0 \tag{5.15}$$

where $X'(z)$ stands for the Z transform of $x'(k)$, and using [77] yields

$$E\{x'^2(k)\} \geq \sigma_s^2 \tag{5.16}$$

From (5.14), the equality of (5.16) holds only if $|A(z)(1+\alpha_{11}z^{-\Delta_{11}})|^2 = 1$ which implies that the global minimum of $E\{x'^2(k)\}$ occurs when $A(z) = 1/I_1(z)$. It has also been investigated [77] that $E\{x'^2(k)\}$ is unimodal. Consequently, $x'(k)$ will converge to $s(k)$ in steady state.

Since $A(z)$ will approach $1/I_1(z)$, its peak weight can be used to estimate the multipath gain α_{11}. As soon as Δ_{11} has been determined, constrained adaptation is applied, that is, only the largest coefficient $a_{\Delta_{11}}(k)$ is updated whilst all other filter coefficients are set to zero. The adaptive algorithm for $A(z)$ is now simplified to

$$a_{\Delta_{11}}(k+1) = a_{\Delta_{11}}(k) + 2\mu_\alpha x'(k)x'(k-\Delta_{11}) \tag{5.17}$$

where

$$x'(k) = x(k) - a_{\Delta_{11}}(k)x'(k-\Delta_{11}) \tag{5.18}$$

When $x'(k)$ is close to $s(k)$, the learning characteristics of the gain estimate can be obtained by taking the expected value of (5.17) with the assumption that $x'(k)$ is uncorrelated with $a_{\Delta_{11}}(k)$ [57],

$$E\{a_{\Delta_{11}}(k+1)\} = E\{a_{\Delta_{11}}(k)\} + 2\mu_\alpha E\{(x(k)-a_{\Delta_{11}}(k)x'(k-\Delta_{11}))x'(k-\Delta_{11})\}$$

$$\begin{aligned} &\approx E\{a_{\Delta_{11}}(k)\} + 2\mu_\alpha E\{(s(k) + \alpha_{11}s(k-\Delta_{11}) \\ &\quad -a_{\Delta_{11}}(k)s(k-\Delta_{11}))\cdot s(k-\Delta_{11})\} \\ &= E\{a_{\Delta_{11}}(k)\} + 2\mu_\alpha\sigma_s^2(\alpha_{11} - E\{a_{\Delta_{11}}(k)\}) \end{aligned} \tag{5.19}$$

When μ_α is chosen between 0 and $1/\sigma_s^2$, (5.19) can be found as

$$E\{a_{\Delta_{11}}(k)\} = \alpha_{11} + (a_{\Delta_{11}}(0) - \alpha_{11})(1 - 2\mu_\alpha\sigma_s^2)^k \tag{5.20}$$

where $a_{\Delta_{11}}(0)$ is the initial value of $a_{\Delta_{11}}(k)$. As $k \to \infty$, $a_{\Delta_{11}}(k)$ is an unbiased estimate of α_{11}.

The mean square error of the gain estimate, denoted by $\epsilon_a(k)$, can be written as (See Appendix E)

$$\begin{aligned} \epsilon_a(k) &\triangleq E\{(a_{\Delta_{11}}(k) - \alpha_{11})^2\} \\ &\approx \mu_\alpha\sigma_s^2 + ((a_{\Delta_{11}}(0) - \alpha_{11})^2 - \mu_\alpha\sigma_s^2)(1 - 4\mu_\alpha\sigma_s^2)^k \end{aligned} \tag{5.21}$$

provided that μ_α lies within the bound,

$$0 < \mu_\alpha < \frac{1}{2\sigma_s^2} \tag{5.22}$$

Notice that (5.22) is a stringent bound for μ_α since the mean square weight error is confined to be finite [66]. The variance of the gain estimate, $var(a)$, which equals the steady state value of $\epsilon_a(k)$, is thus given by

$$var(a) = \mu_\alpha\sigma_s^2 \tag{5.23}$$

When $A(z)$ has reached its optimal value, (5.18) can be approximated by

$$x'(k) \approx s(k) + (\alpha_{11} - a_{\Delta_{11}}(k))s(k-\Delta_{11}) \tag{5.24}$$

Although the multipath effect cannot be eliminated completely after passing $x(k)$ through $A(z)$, the magnitude of the multipath is reduced significantly since $a_{\Delta_{11}}(k) \to \alpha_{11}$. Similarly, the steady state value of $y'(k)$ is given by

$$y'(k) \approx s(k-D) + (\alpha_{21} - b_{\Delta'_{21}}(k))s(k-D-\Delta'_{21}) \tag{5.25}$$

Since the values of $x'(k)$ and $y'(k)$ are dominated by the first terms of (5.24) and (5.25), we may neglect their second terms. As a result, the learning trajectory of the delay estimate becomes identical to (3.18), that is,

$$E\{\hat{D}_{MG}(k)\} \approx D + (\hat{D}_{MG}(0) - D)\Big(1 - \frac{2}{3}\mu_D\sigma_s^2\pi^2\Big)^k \tag{5.26}$$

where $\hat{D}_{MG}(0)$ is the initial delay estimate and it can be found from (3.12). Considering the second terms of (5.24) and (5.25) as uncorrelated noise sources, the variance of $\hat{D}_{MG}(k)$, $var(\hat{D}_{MG})$, can be computed from (3.26) which has a value

$$var(\hat{D}_{MG}) \approx 2\mu_\alpha\mu_D\sigma_s^4 \tag{5.27}$$

Therefore, the delay variance in this case increases with μ_α, μ_D, and σ_s^2.

5.2.3 The Generalized Multipath Cancellator

When Δ_{11} is not restricted to be an integer, $x(k)$ can be expressed as

$$x(k) = s(k) + \alpha_{11}\sum_{i=-\infty}^{\infty} sinc(i - \Delta_{11})s(k-i) \tag{5.28}$$

It is expected that the transfer function of the optimal multipath canceller, $A^o(z)$, is of the form

$$A^o(z) = \frac{1}{1 + \alpha_{11}\sum_{i=-M}^{M} sinc(i - \Delta_{11})z^{-i}} \tag{5.29}$$

In this case, M should be chosen much larger than Δ_{11} to reduce modeling error. However, it has been shown (See Appendix F) that due to the noncausality of the filter structure, $A^o(z)$ is an unstable system and thus its practical realization is prohibited. It is trivial that a suboptimal solution for the multipath canceller, denoted by $A^s(z)$, can be represented by

$$A^s(z) = \frac{1}{1 + \alpha_{11}\sum_{i=1}^{M} sinc(i - \Delta_{11})z^{-i}} \tag{5.30}$$

As a result, we can still use $A(z)$ to cancel out the replica. In our studies, we replace $a_i(k)$ by $\hat{\alpha}_{11}^s(k)sinc(i - \hat{\Delta}_{11}^s(k))$, for $1 \le i \le M$, where $\hat{\alpha}_{11}^s(k)$ and $\hat{\Delta}_{11}^s(k)$

are the estimates of α_{11} and Δ_{11} respectively. Upon convergence, $\hat{\alpha}^s_{11}(k) \to \alpha_{11}$ and $\hat{\Delta}^s_{11}(k) \to \Delta_{11}$ are desired. Based on the ETDE algorithm, the updating rules for $\hat{\alpha}^s_{11}(k)$ and $\hat{\Delta}^s_{11}(k)$ are derived as

$$\begin{aligned}\hat{\alpha}^s_{11}(k+1) &= \hat{\alpha}^s_{11}(k) - \mu_\alpha \frac{\partial x'^2(k)}{\partial \hat{\alpha}^s_{11}(k)} \\ &= \hat{\alpha}^s_{11}(k) + 2\mu_\alpha x'(k) \sum_{i=1}^{M} x'(k-i) sinc(i - \hat{\Delta}^s_{11}(k)) \end{aligned} \tag{5.31}$$

and

$$\begin{aligned}\hat{\Delta}^s_{11}(k+1) &\approx \hat{\Delta}^s_{11}(k) - \frac{\mu_\Delta}{\hat{\alpha}^s_{11}(k)} \frac{\partial x'^2(k)}{\partial \hat{\Delta}^s_{11}(k)} \\ &= \hat{\Delta}^s_{11}(k) - 2\mu_\Delta x'(k) \sum_{i=1}^{M} f(i - \hat{\Delta}^s_{11}(k)) x'(k-i) \end{aligned} \tag{5.32}$$

where

$$x'(k) = x(k) - \hat{\alpha}^s_{11}(k) \sum_{i=1}^{M} sinc(i - \hat{\Delta}^s_{11}(k)) x'(k-i) \tag{5.33}$$

The parameter μ_Δ is the convergence step size for $\hat{\Delta}^s_{11}(k)$. Notice that (5.31) and (5.32) represent the generalized algorithm for the multipath canceller and they can be used for both integral and non-integral multipath delays. It can be seen that (5.17) is only a special case of (5.31) when Δ_{11} is an integer.

When $\Delta_{11} \gg 0$, the learning characteristics of $\hat{\alpha}^s_{11}(k)$ and $\hat{\Delta}^s_{11}(k)$ are given by (See Appendix G)

$$E\{\hat{\alpha}^s_{11}(k)\} \approx \alpha_{11} + (\hat{\alpha}^s_{11}(0) - \alpha_{11})(1 - 2\mu_\alpha \sigma_s^2)^k \tag{5.34}$$

and

$$E\{\hat{\Delta}^s_{11}(k)\} \approx \Delta_{11} + (\hat{\Delta}^s_{11}(0) - \Delta_{11})\left(1 - \frac{2}{3}\mu_\Delta \alpha_{11} \sigma_s^2 \pi^2\right)^k \tag{5.35}$$

The initial value of the multipath gain estimate, $\hat{\alpha}^s_{11}(0)$, is arbitrarily chosen between 0 and 1. The system starts at the beginning with a few hundred iterations using (5.10) to determine the value of $\hat{\Delta}^s_{11}(0)$ which is given by (5.13a), and the joint delay and gain estimation then follows accordingly.

The variance of $\hat{\alpha}_{11}^s(k)$ is given by (5.23) whereas the variance of $\hat{\Delta}_{11}^s(k)$, $var(\hat{\Delta}_{11}^s)$, can be shown to be (See Appendix H)

$$var(\hat{\Delta}_{11}^s) = \frac{\mu_\Delta \sigma_s^2}{\alpha_{11}} \tag{5.36}$$

We shall investigate the performance of the delay estimate $\hat{D}_{MC}(k)$ when the multipath delays are of real values. Assume that $\Delta_{11} \gg 0$ and $\Delta'_{21} \gg 0$, then the modeling errors in $A(z)$ and $B(z)$ are so small that they can be ignored. In this case, the rate of convergence and the variance of $\hat{D}_{MC}(k)$ can be approximated by (5.26) and (5.27).

However, if Δ_{11} and Δ'_{21} are not large enough, the estimation performance will be degraded. Subsequently, the steady state values of $X'(z)$ and $Y'(z)$ become

$$X'(z) = T_1(z)S(z) \tag{5.37a}$$

and

$$Y'(z) = z^{-D}T_2(z)S(z) \tag{5.37b}$$

where

$$T_1(z) = \frac{1 + \alpha_{11} \sum_{i=-\infty}^{\infty} sinc(i - \Delta_{11})z^{-i}}{1 + \alpha_{11} \sum_{i=1}^{M} sinc(i - \Delta_{11})z^{-i}} \tag{5.38a}$$

and

$$T_2(z) = \frac{1 + \alpha_{21} \sum_{j=-\infty}^{\infty} sinc(j - \Delta'_{21})z^{-j}}{1 + \alpha_{21} \sum_{j=1}^{M} sinc(j - \Delta'_{21})z^{-j}} \tag{5.38b}$$

As a result, the ETDE cannot model the actual delay accurately and the delay estimate is now determined by the following relation,

$$\begin{aligned} \hat{D}_{MC}(z) &= \frac{T_2(z)}{T_1(z)} z^{-D} \\ &\approx \left(1 + \frac{\sum_{i=-M}^{0} (\alpha_{21} sinc(i - \Delta'_{21}) - \alpha_{11} sinc(i - \Delta_{11}))z^{-i}}{(1 + \alpha_{11} \sum_{i=\infty}^{\infty} sinc(i - \Delta_{11})z^{-i})(1 + \alpha_{21} \sum_{j=1}^{M} sinc(j - \Delta'_{21})z^{-j})}\right) z^{-D} \end{aligned} \tag{5.39}$$

It is obvious that if $T_1(z) = T_2(z)$, we can still obtain an accurate delay estimate. Inspection of (5.39) indicates that $T_2(z)/T_1(z)$ will tend to unity when α_{11} and α_{21}

decrease or Δ_{11} and Δ'_{21} increase at the same time. For any other circumstances, the steady state value of $\hat{D}_{MC}(k)$ will deviate from the actual delay. Furthermore, the mean square error of $\hat{D}_{MC}(k)$ will be much larger than the value given by (5.27).

5.2.4 Effects of Additive Noises

The performance of the MCTDE is also studied when corrupting noise sources are present in $x(k)$ and $y(k)$. Again, we shall first examine the effect of the noise to the multipath canceller $A(z)$. Only integral multipath delays are considered and $A(z)$ is adapted according to (5.17). It is assumed that the SNR is high enough such that Δ_{11} can be determined using (5.13a). Therefore, the performance surface of $A(z)$, $E\{x'^2(k)\}$, can be expressed as (See Appendix I)

$$E\{x'^2(k)\} = \frac{(1+\alpha_{11}^2 - 2\alpha_{11}a_{\Delta_{11}})\sigma_s^2 + \sigma_n^2}{1-a_{\Delta_{11}}^2} \tag{5.40}$$

Differentiating (5.40) with respect to $a_{\Delta_{11}}$ yields

$$\frac{\partial E\{x'^2(k)\}}{\partial a_{\Delta_{11}}} = \frac{-2(\alpha_{11}\sigma_s^2 a_{\Delta_{11}}^2 - ((1+\alpha_{11}^2)\sigma_s^2 + \sigma_n^2)a_{\Delta_{11}} + \alpha_{11}\sigma_s^2)}{(1-a_{\Delta_{11}}^2)^2} \tag{5.41}$$

We set (5.41) to zero to evaluate the global minimum point of $E\{x'^2(k)\}$ which can be found to be

$$a^o_{\Delta_{11}} = \frac{(1+\alpha_{11}^2)\sigma_s^2 + \sigma_n^2 - \sqrt{(1-\alpha_{11}^2)^2\sigma_s^4 + \sigma_n^4 + 2(1+\alpha_{11}^2)\sigma_s^2\sigma_n^2}}{2\alpha_{11}\sigma_s^2} \tag{5.42}$$

When $\sigma_n^2 = 0$, $a^o_{\Delta_{11}}$ equals α_{11} and this result has also been obtained in Section 5.2.2. However, under a noisy environment, it can be seen from (5.42) that $a^o_{\Delta_{11}} < \alpha_{11}$. In addition, the difference between $a^o_{\Delta_{11}}$ and α_{11} increases with the noise power.

Using (5.2a) and (5.18), the value of $x'(k)$ in the presence of noise can be approximated by

$$x'(k) \approx s(k) + (\alpha_{11} - a_{\Delta_{11}}(k))s(k-\Delta_{11}) + n_1(k) \tag{5.43}$$

It is similar to (5.24) except the existence of $n_1(k)$. It is noteworthy that the replica will not be eliminated or reduced through the multipath canceller when SNR $\ll 1$. It is because from (5.42), $a^o_{\Delta_{11}} \to 0$ as $\sigma_n^2 \to \infty$. Moreover, it has been shown (See

Appendix J) that the time constant of $a_{\Delta_{11}}(k)$ has a value of $1/(2\mu_\alpha(\sigma_s^2+\sigma_n^2))$ and it is smaller than that of (5.20). However, $a_{\Delta_{11}}(k)$ is no longer an unbiased estimate of α_{11} upon convergence. If we use $a_{\Delta_{11}}(k)$ as an estimate of α_{11}, the steady state mean square gain error, namely, $\xi(a)$, is equal to the sum of $var(a)$ and the bias square,

$$\xi(a) \approx \mu_\alpha \sigma_s^2 + (\alpha_{11} - a^o_{\Delta_{11}})^2 \tag{5.44}$$

On the other hand, $y'(k)$ has a value of

$$y'(k) \approx s(k-D) + (\alpha_{21} - b_{\Delta'_{21}}(k))s(k-D-\Delta'_{21}) + n_2(k) \tag{5.45}$$

As a result, the steady state value of the delay estimate is related to the multipath gains and the additive noises by

$$\hat{D}_{MC}(z) \approx \frac{z^{-D}(1+z^{-\Delta'_{21}}(\alpha_{21} - b^o_{\Delta'_{21}})) + N_2(z)/S(z)}{1+z^{-\Delta_{11}}(\alpha_{11} - a^o_{\Delta_{11}}) + N_1(z)/S(z)} \tag{5.46}$$

where $N_1(z)$ and $N_2(z)$ are the Z transforms of $n_1(k)$ and $n_2(k)$ respectively. The quantity $b^o_{\Delta'_{21}}$ represents the mean estimate of α_{21} and its value is obtained by (5.42) with the substitution of $\alpha_{11} = \alpha_{21}$. Due to incomplete cancellation of multipaths as well as presence of noise components, the delay modeling accuracy of the ETDE will decrease and the mean square error of $\hat{D}_{MC}(k)$ is much larger than $var(\hat{D}_{MC})$. Notice that even if we replace the ETDE by the ETDGE, whose delay estimates are proved to be unbiased under noisy environments, no performance gain can be achieved. It is because the ETDGE will not provide the optimal filter weights when $x'(k)$ and $y'(k)$ consist of noise terms as well as multipath signals.

5.2.5 Simulation Results

Simulation tests had been performed to corroborate the theoretical derivations and to evaluate the performance of the MCTDE for time delay estimation in the presence of multipath transmissions. The signal $s(k)$ and the noises $n_1(k)$ and $n_2(k)$ were white Gaussian random variables and they were produced by a pseudorandom noise generator. The power of $s(k)$ was fixed to unity and different SNRs, where the

SNR was defined as σ_s^2/σ_n^2, were obtained by proper scaling of the random noise sequences. The delayed signal $s(k-D)$ was produced by passing $s(k)$ through an FIR filter of transfer function $\sum_{i=-30}^{30} sinc(i-D)z^{-i}$. Similarly, $s(k-\Delta_{11})$ and $s(k-\Delta_{21})$ were obtained. We fixed $D \in (-10, 10)$ and the values of the multipath delays were less than 10. In order to allow for acceptable truncation errors, M was chosen to be 15 in both $A(z)$ and $B(z)$ while the filter length of the ETDE was 31 ($P = 15$). The initial values of $\{a_i(k)\}$ and $\{b_j(k)\}$ were all set to be zero. In our studies, $A(z)$ and $B(z)$ were updated according to (5.10) and (5.11) respectively at the beginning of the adaptation. At the 600th iteration, we used (5.13) to determine the estimates of Δ_{11} and Δ'_{21}, and $A(z)$ was then adjusted using (5.17) for an integral interpath delay or (5.31) and (5.32) for any real-valued multipath delay. When applying (5.31), the initial value of the multipath gain was arbitrarily selected to 0.5. On the other hand, $\hat{D}_{MC}(0)$ was found by using the LMSTDE for the first 600 iterations and its value was given by (3.12). The step size μ_α had a value of 0.001 and $\mu_\Delta = \mu_D = 0.0005$. All simulation results provided were average of 200 independent runs.

Experiment 1

The learning trajectories of the estimates of the multipath gains and time delay for integral interpath delays under a noise-free condition are shown in Figure 5.2 and Figure 5.3 respectively. In this test, $D = 1.4$, $\alpha_{11} = 0.8$, $\alpha_{21} = 0.4$, $\Delta_{11} = 4$ and $\Delta'_{21} = 5$. After Δ_{11} and Δ'_{21} had been found at the 600th iteration, only $a_4(k)$ and $b_5(k)$ were adapted whilst all other filter weights were set to zero. In Figure 5.2, it can be seen that $a_4(k) \rightarrow 0.8$ and $b_5(k) \rightarrow 0.4$ at the 3000th and the 2000th iteration respectively. Figure 5.3, on the other hand, illustrates that the delay estimate converged to the desired value of 1.4 in about 2500 iterations. The learning curves of $b_5(k)$ and $\hat{D}_{MC}(k)$ agreed well with the predicted trajectories but the convergence time for $a_4(k)$ was slightly longer than its theoretical value. It is mainly because at the beginning of the adaptation, the assumption of $x'(k) \approx s(k)$ that facilitates the derivation of (5.20) was invalid. Upon convergence, the measured variances of $a_4(k)$ and $b_5(k)$ were very close to the theoretical value of 0.001. However, the variance of

$\hat{D}_{MC}(k)$ was found to be 1.16×10^{-5} which was much larger than the value given by (5.27), that is, 1×10^{-6}. The discrepancy existed because the derivation of (5.27) has assumed that the second terms of (5.24) and (5.25) are uncorrelated with the signal source but this is not actually true. We had also attempted to adjust $\{a_i(k)\}$ and $\{b_j(k)\}$ according to (5.10) and (5.11) respectively, and in this case, the convergence behaviours of the gain estimates and the delay estimate were observed to be very similar to the results of Figure 5.2 and Figure 5.3. The variance of $b_5(k)$ had been found to be fairly close to the predicted value, whereas the mean square error of $a_4(k)$ and $\hat{D}_{MC}(k)$ had a value of 1.83×10^{-3} and 1.48×10^{-4} respectively and they were larger than those of the constrained adaptation.

Experiment 2

Another experiment for real-valued multipath delays has been performed and the results are shown in Figure 5.4, Figure 5.5 and Figure 5.6. The setting of the parameters was identical to the previous test except that $\Delta_{11} = 4.5$ and $\Delta'_{21} = 5.7$. In this case when the multipath delays are not integral multiples of the sampling interval, $A(z)$ and $B(z)$ were adjusted according to (5.31) and (5.32). Upon convergence, it can be observed in Figure 5.4 and Figure 5.5 that $\hat{\alpha}^s_{11}(k) \rightarrow 0.76$ and $\hat{\Delta}^s_{11}(k) \rightarrow 4.43$ while the estimates of α_{21} and Δ'_{21}, namely, $\hat{\alpha}^s_{21}(k)$ and $\hat{\Delta}'^s_{21}(k)$, approached 0.40 and 5.69, respectively. The estimates in Figure 5.5 were more accurate since the multipath canceller has a better performance for larger multipath delays. In addition, the convergence rates of $\hat{\Delta}^s_{11}(k)$ and $\hat{\Delta}'^s_{21}(k)$ were similar to their predicted values whereas those of the gain estimates were different from the analytical results. It is because we have assumed that the multipath delay estimate has already approached its optimal value when deriving (5.34). As indicated from (5.35) that the time constant of the multipath delay estimate is inversely proportional to the multipath gain, $\hat{\Delta}^s_{11}(k)$ converged at the 2000th iteration whereas $\Delta'_{21}(k)$ required a longer period of 4000 iterations, to reach its desired value. On the other hand, Figure 5.6 shows that $\hat{D}_{MC}(k)$ converged to 1.35 at the 3000th iteration. The estimate $\hat{D}_{MC}(k)$ was unable to reach the optimal value of 1.4 since the multipath effect was not completely

eliminated, particularly in $x'(k)$. Notice that the rate of convergence of $\hat{D}_{MC}(k)$ was slowed down by the learning processes of the multipath parameter estimates. Consequently, it required a longer time to converge when comparing with Figure 5.3.

The second order statistics of these parameters, together with the theoretical values, are tabulated in Table 5.1. It can be seen that the experimental values of $var(\hat{\alpha}^s_{21})$ and $var(\hat{\Delta}'^s_{21})$, which denote the variances of $\hat{\alpha}^s_{21}(k)$ and $\hat{\Delta}'^s_{21}$ respectively, conformed fairly to the analytical calculations. However, the results for other parameters were not so satisfactory. The measured variances of $\hat{\alpha}^s_{11}(k)$ and $\hat{\Delta}^s_{11}(k)$ were larger than the theoretical values because biases were existed in these estimates. Limited by the inherent modeling errors in the inverse filters of $A(z)$ and $B(z)$, the measured variance of $\hat{D}_{MC}(k)$ was much larger than the desired value.

Experiment 3

Figure 5.7 and Figure 5.8 illustrate the effects of additive noises on the system parameters by repeating Experiment 1 for SNR = 20 dB and SNR = 10 dB. In this test, we did not assume integral multipath delays so that the multipath cancellers were adapted using (5.31) and (5.32). As discussed in Section 5.2.4, it can be roughly seen from Figure 5.7 that the learning speeds of $\hat{\alpha}^s_{11}(k)$ and $\hat{\alpha}^s_{21}(k)$ increased with the noise power. However, the gain estimates were not so precise in the presence of noises. We observe that $\hat{\alpha}^s_{11}(k) \rightarrow 0.77$ and $\hat{\alpha}^s_{21}(k) \rightarrow 0.40$ when SNR = 20 dB while the steady state values of $\hat{\alpha}^s_{11}(k)$ and $\hat{\alpha}^s_{21}(k)$ for SNR = 10 dB were 0.65 and 0.36 respectively. These values agreed well with (5.42). At lower SNRs, we expect that the accuracy of the gain estimates will be worse since $\hat{\alpha}^s_{11}(k)$ and $\hat{\alpha}^s_{21}(k)$ shift away from the desired values as the noise power increases. Although we had not shown the results of the multipath delay estimates, we found that $\hat{\Delta}^s_{11}(k) \rightarrow 4.00$ and $\hat{\Delta}'^s_{21}(k) \rightarrow 5.00$ for both cases of SNR. From Figure 5.8, it can be seen that $\hat{D}_{MC}(k) \rightarrow 1.40$ when SNR = 20 dB and it converged to 1.41 for SNR = 10 dB. The delay error for the second case is due to the additive noises as well as the noticeable multipath components in $x'(k)$ and $y'(k)$ when $\hat{\alpha}^s_{11}(k)$ and $\hat{\alpha}^s_{21}(k)$ did not converge

to the desired values. Due to the inferior performance of the multipath canceller under a noisy environment, it appears that accurate estimates of the TDOA and the multipath parameters can only be attained when SNR $\gg 1$.

Experiment 4

Figure 5.9 to Figure 5.13 demonstrate the ability of the MCTDE to estimate the system parameters under a nonstationary environment in which all five parameters were given step offsets after every 5000 iterations. The noise-free condition and a noisy environment of SNR = 10 dB were investigated. It can be seen that the MCTDE tracked all these step changes in both SNR conditions. In Figure 5.10, Figure 5.12 and Figure 5.13, we observe that the trajectories of $\hat{\Delta}^s_{11}(k)$, $\hat{\Delta}'^s_{21}(k)$ and $\hat{D}_{MC}(k)$ were almost unaffected by the corrupting noises and slightly biased or unbiased estimates of these parameters were found after transients. These biasedness is mainly due to incomplete cancellation of the multipath signals in $A(z)$ and $B(z)$. Whereas Figure 5.9 and Figure 5.11 show that the estimates of the multipath gains were more accurate in the noise-free condition than those under the noisy environment. The discrepancies of $\hat{\alpha}^s_{11}(k)$ and $\hat{\alpha}^s_{21}(k)$ are mainly due to the corrupting noises as well as the modeling errors in $A(z)$ and $B(z)$.

The mean square error of $\hat{D}_{MC}(k)$ had been measured for different multipath parameters in the absence of noise and the results are depicted in Figure 5.14. In this diagram, Δ_{11} was varied from 1.0 to 8.0 and two values of α_{11}, namely, 0.8 and 0.2, were tried while other parameters, α_{21}, Δ'_{21} and D were assigned to have constant values of 0.4, 5.7 and 0.3 respectively. These results agreed with Section 5.2.3 because the mean square delay error decreased as Δ_{11} increased or if α_{11} diminished. On the other hand, it had a large value when $\alpha_{11} = 0.8$ and for small Δ_{11}. It is also noted that when the difference between the integral parts of Δ_{11} and Δ'_{21} is an even number, the mean square delay error is relatively small. This phenomenon can be explained using (5.39). When the integral parts of Δ_{11} and Δ'_{21} are both odd or even, the values of $sinc(i - \Delta_{11})$ and $sinc(i - \Delta'_{21})$ are of same sign and their difference will be relatively small. In these cases, $T_1(z)/T_2(z)$ is close to

unity which implies a smaller error in the delay estimate.

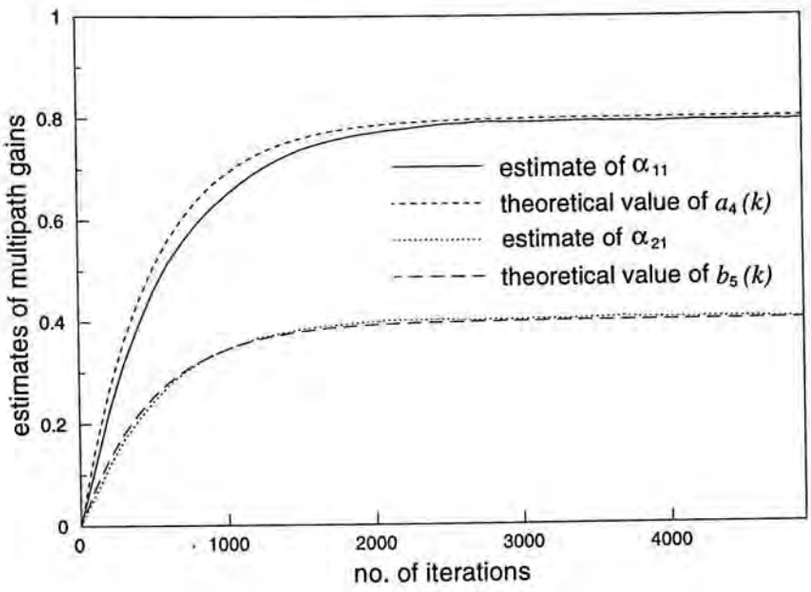

Figure 5.2: Estimates of multipath gains of the MCTDE for integral interpath delays in a noise-free condition

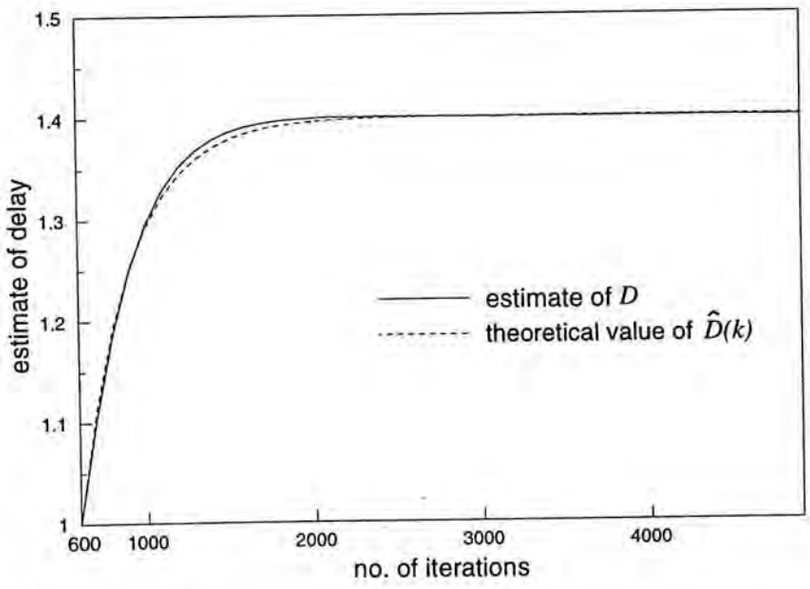

Figure 5.3: Estimate of time delay of the MCTDE for integral interpath delays in a noise-free condition

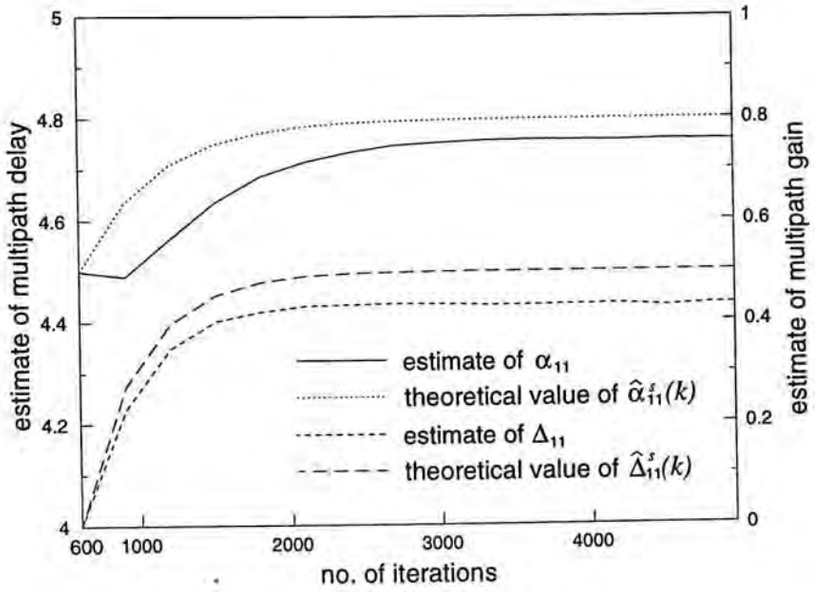

Figure 5.4: Estimate of α_{11} and Δ_{11} of the MCTDE for non-integral interpath delays in a noise-free condition

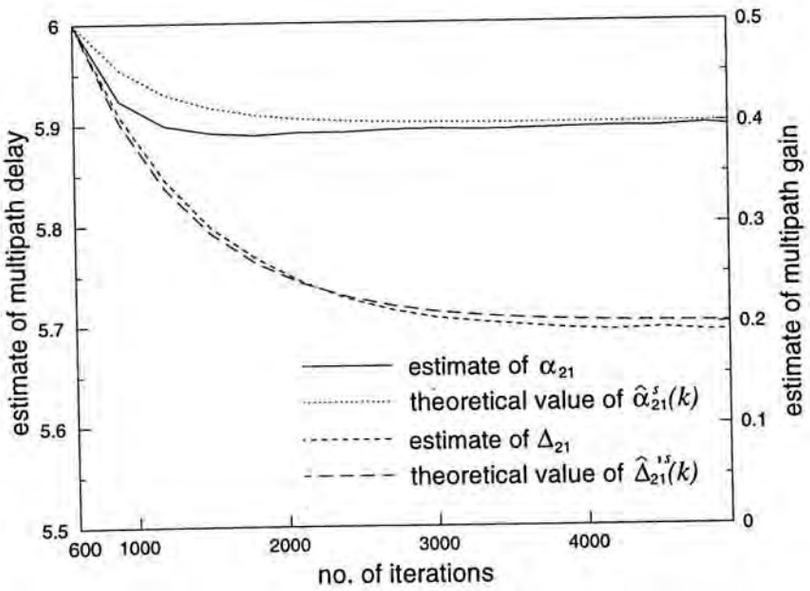

Figure 5.5: Estimate of α_{21} and Δ'_{21} of the MCTDE for non-integral interpath delays in a noise-free condition

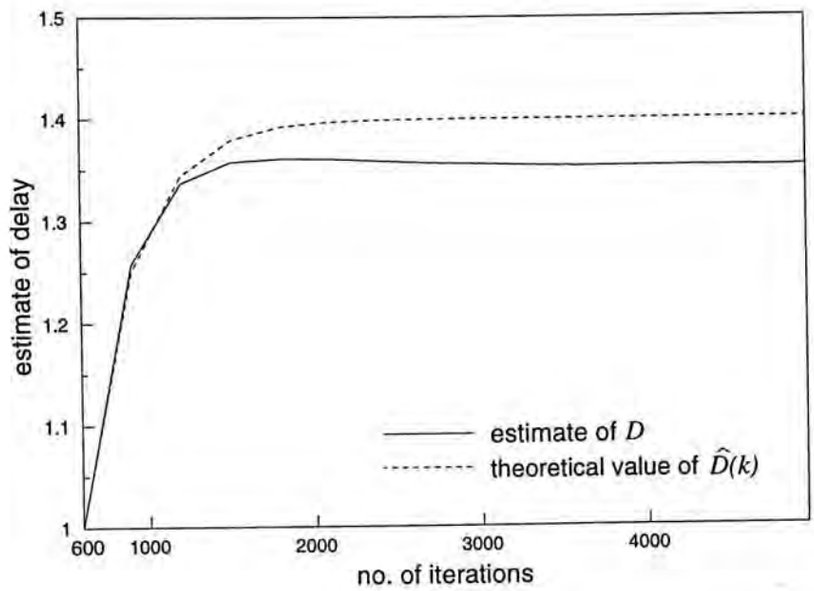

Figure 5.6: Estimate of time delay of the MCTDE for non-integral interpath delays in a noise-free condition

	Measured value	Theoretical value
$var(\hat{\alpha}_{11}^{s})$	2.76×10^{-3}	1.00×10^{-3}
$var(\hat{\Delta}_{11}^{s})$	5.36×10^{-3}	6.25×10^{-4}
$var(\hat{\alpha}_{21}^{s})$	1.04×10^{-3}	1.00×10^{-3}
$var(\hat{\Delta}_{21}^{\prime s})$	1.26×10^{-3}	1.25×10^{-3}
$var(\hat{D}_{MC})$	2.30×10^{-3}	1.00×10^{-6}

Table 5.1: Measured and theoretical variances of the system parameters of the MCTDE for a noise-free condition

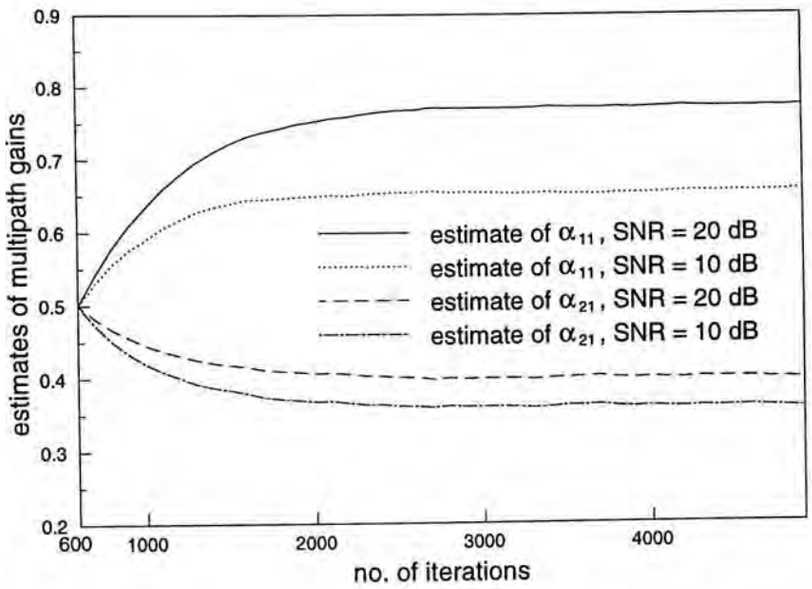

Figure 5.7: Estimates of multipath gains of the MCTDE in noisy environments

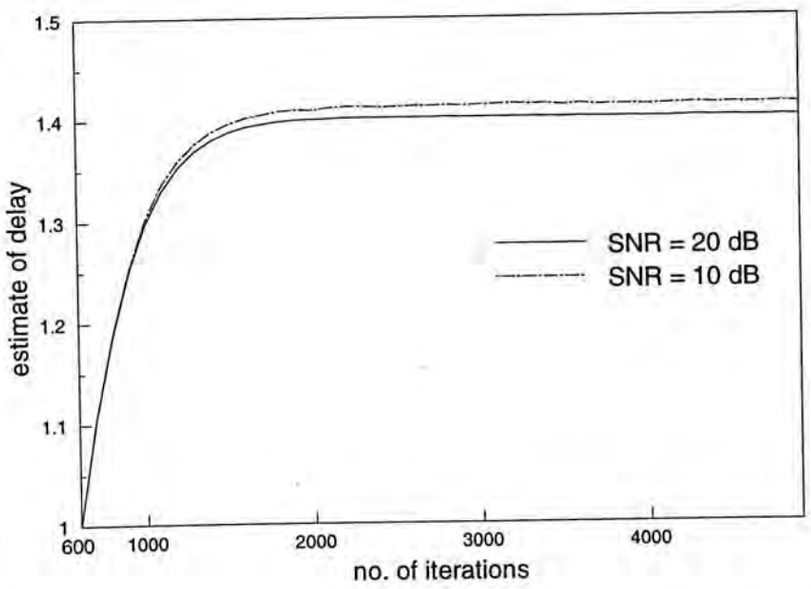

Figure 5.8: Estimates of time delay of the MCTDE in noisy environments

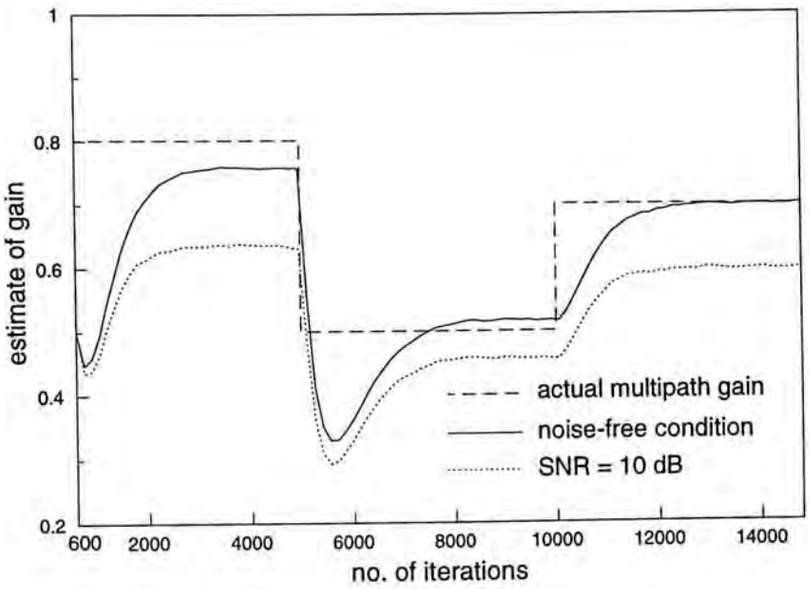

Figure 5.9: Tracking behaviour of $\hat{\alpha}^s_{11}(k)$ of the MCTDE for different SNRs

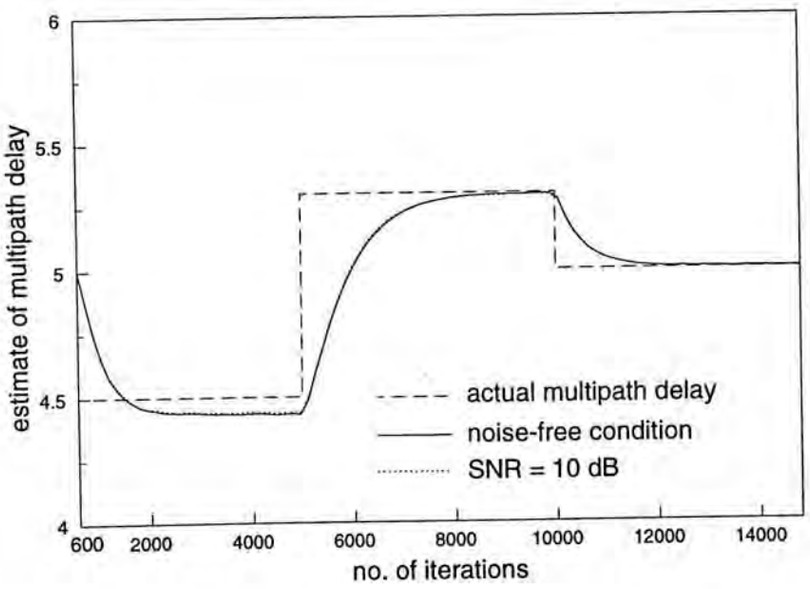

Figure 5.10: Tracking behaviour of $\hat{\Delta}^s_{11}(k)$ of the MCTDE for different SNRs

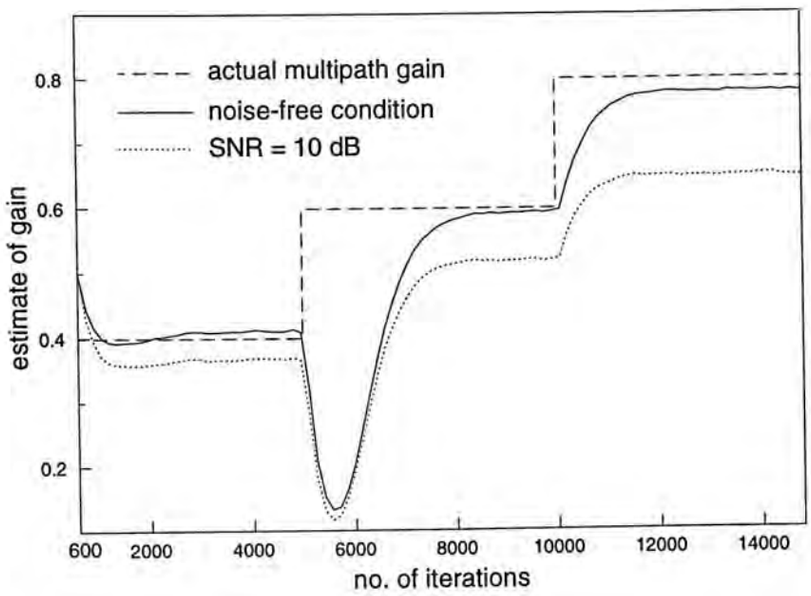

Figure 5.11: Tracking behaviour of $\hat{\alpha}^s_{21}(k)$ of the MCTDE for different SNRs

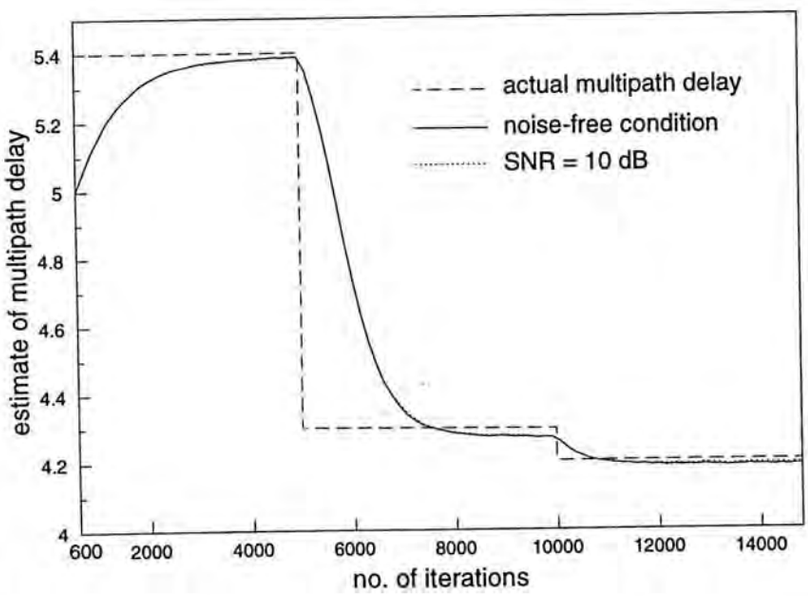

Figure 5.12: Tracking behaviour of $\hat{\Delta}'^s_{21}(k)$ of the MCTDE for different SNRs

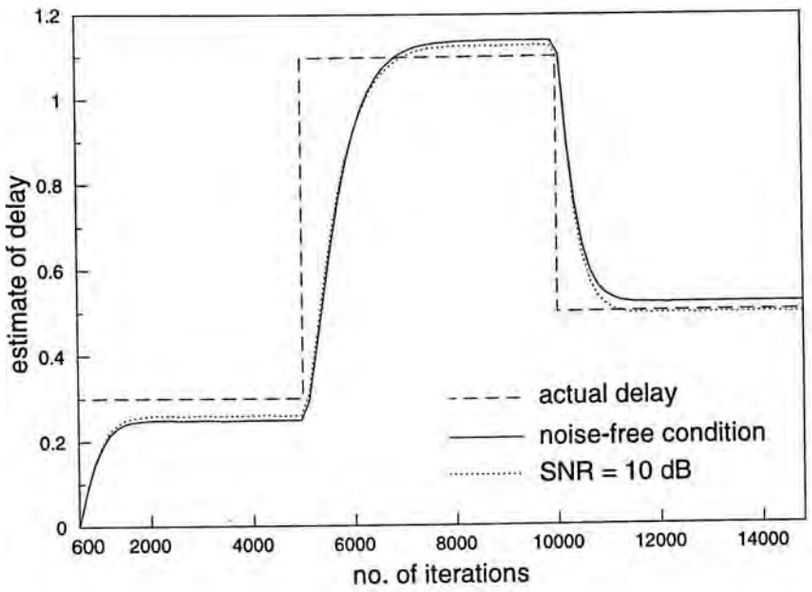

Figure 5.13: Tracking behaviour of $\hat{D}_{MC}(k)$ of the MCTDE for different SNRs

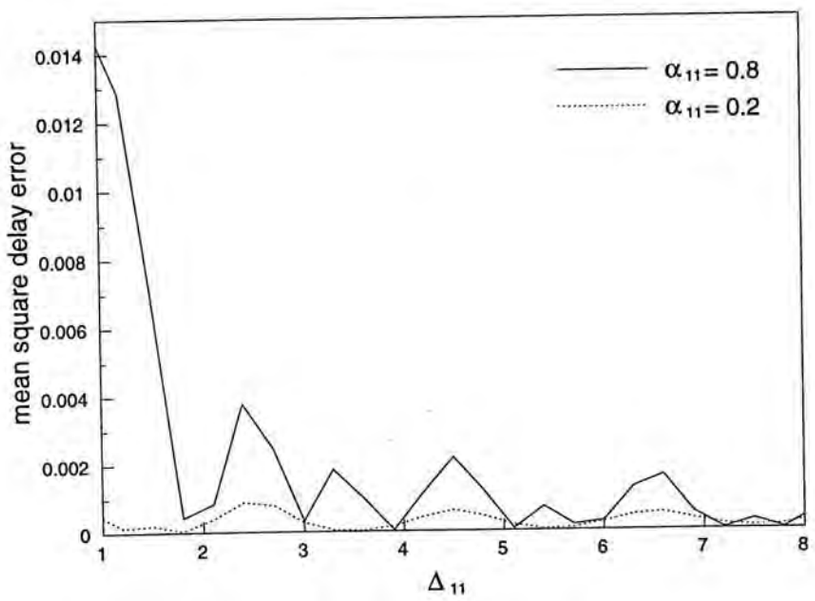

Figure 5.14: Mean square delay error of the MCTDE varies with multipath parameters

5.3 TDE with Multipath Equalization (METDE)

When the multipath delays are fairly small while the multipath gains are relatively large, the performance of the MCTDE will be notably degraded due to incomplete elimination of the multipath signals. In this section, we try to resolve this problem by using another adaptive model which is based on equalization of the multipath effect in each transmission channel. Moreover, the new method can also be applied when there is more than one multipath signal received at each sensor.

5.3.1 The Two-Step Algorithm

Taking Z transform of (5.1), we have

$$X(z) = S(z)\Big(1 + \sum_{p=1}^{M_1} \alpha_{1p} z^{-\Delta_{1p}}\Big) + N_1(z) \tag{5.47a}$$

$$Y(z) = S(z)\Big(z^{-D} + \sum_{q=1}^{M_2} \alpha_{2q} z^{-\Delta_{2q}}\Big) + N_2(z) \tag{5.47b}$$

where

$$z^{-v} = \sum_{n=-\infty}^{\infty} sinc(n-v) z^{-n}, \qquad v = D, \Delta_{1p}, \Delta_{2q} \tag{5.48}$$

If we multiply the bracket term of (5.47b) to (5.47a) and the bracket term of (5.47a) to (5.47b), the two resultant transfer functions will be equal in the absence of noise. Using this idea and from (5.47), we define an error function, $e(k)$, which is given by

$$\begin{aligned} e(k) &= \mathrm{Z}^{-1}\Big\{Y(z)\Big(1 + \sum_{p=1}^{M_1} \hat{\alpha}_{1p}\hat{\Delta}_{1p}(z)\Big) - X(z)\Big(\hat{D}_{ME}(z) + \sum_{q=1}^{M_2} \hat{\alpha}_{2q}\hat{\Delta}_{2q}(z)\Big)\Big\} \\ &= y(k) + \sum_{n=-P}^{P}\sum_{p=1}^{M_1} \hat{\alpha}_{1p} sinc(n - \hat{\Delta}_{1p}) y(k-n) \\ &\quad - \sum_{n=-P}^{P} \Big(sinc(n - \hat{D}_{ME}) + \sum_{q=1}^{M_2} \hat{\alpha}_{2q} sinc(n - \hat{\Delta}_{2q})\Big) x(k-n) \end{aligned} \tag{5.49}$$

where $\mathrm{Z}^{-1}\{v\}$ represents the inverse Z transform of v. The system parameters, $\hat{D}_{ME}$, $\hat{\alpha}_{1p}$, $\hat{\Delta}_{1p}$, $\hat{\alpha}_{2q}$ and $\hat{\Delta}_{2q}$, for $M_1 \geq p \geq 1$ and $M_2 \geq q \geq 1$, denote the estimates of D, α_{1p}, Δ_{1p}, α_{2q} and Δ_{2q} respectively. The transfer function $\hat{v}(z)$ is defined as

$$\hat{v}(z) = \sum_{n=-P}^{P} sinc(n-v) z^{-n}, \qquad \hat{v} = \hat{D}_{ME}, \hat{\Delta}_{1p}, \hat{\Delta}_{2q} \tag{5.50}$$

where $P \gg \max\{\Delta_{1p}, \Delta_{2q}\}$ so that it can model the ideal time shift function z^{-v} fairly accurately. If we minimize $E\{e^2(k)\}$ with respect to these parameters, we expect that perfect estimates of the delay and multipath parameters can be obtained when $n_1(k) = n_2(k) = 0$. At high SNR condition, accurate estimation of D, α_{1p}, Δ_{1p}, α_{2q} and Δ_{2q} is anticipated. However, this is difficult because $E\{e^2(k)\}$ has a multimodal surface. Notice that when $M_1 = M_2 = 0$, $E\{e^2(k)\}$ becomes the performance surface of the ETDE which has been shown to be multimodal.

This multipath equalization based method consists of two steps. The first step is to find initial guesses of the delay and multipath parameters such that they map to a point on $E\{e^2(k)\}$ which is fairly close to the global minimum. By so doing, the system parameters are avoided from being locked at any local minimum. A stochastic gradient descent algorithm is then used to "fine-tune" the estimates of the delay and multipath parameters through the minimization of $E\{e^2(k)\}$. Instead of using the MCTDE, other methods such as correlation techniques [48],[78] can also be used to find good initial values of the multipath parameters.

In our studies, we freely adapt the MCTDE system for a few hundred iterations to initialize the system parameters. Although the MCTDE is originally designed to tackle the multipath TDE problem when $M_1 = M_2 = 1$, it is possible to obtain rough estimates of the TDOA and multipath delays at SNR $\gg 1$. Since $A(z)$ and $B(z)$ aim to whiten the received signals, $x(k)$ and $y(k)$, it is expected that

$$a_i(k) \approx \sum_{p=1}^{M_1} \alpha_{1p} sinc(i - \Delta_{1p}) \tag{5.51a}$$

and

$$b_j(k) \approx \sum_{q=1}^{M_2} \alpha_{2q} sinc(j - \Delta_{2q} + D) \tag{5.51b}$$

in steady state if the filter length of the multipath cancellers, M, is chosen much larger than the maximum of Δ_{1M_1} and Δ_{2M_2}. A good choice of $\hat{\Delta}_{11}(0)$ is given by (5.13a). Moreover, the second peak of $\{a_i(k)\}$ corresponds to $\hat{\Delta}_{12}(0)$ and so on. In a similar manner, the values of $\hat{\Delta}_{21}(0) - \hat{D}_{ME}(0)$, $\hat{\Delta}_{22}(0) - \hat{D}_{ME}(0)$, ... ,$\hat{\Delta}_{2M_2}(0) - \hat{D}_{ME}(0)$ can be determined by the peaks of $\{b_j(k)\}$. It can be deduced from (5.49) that $E\{e^2(k)\}$ becomes unimodal if the time delay and multipath delay parameters

are known *a priori*. This means that the initial estimates of the multipath gains can be arbitrarily selected without affecting the global convergence.

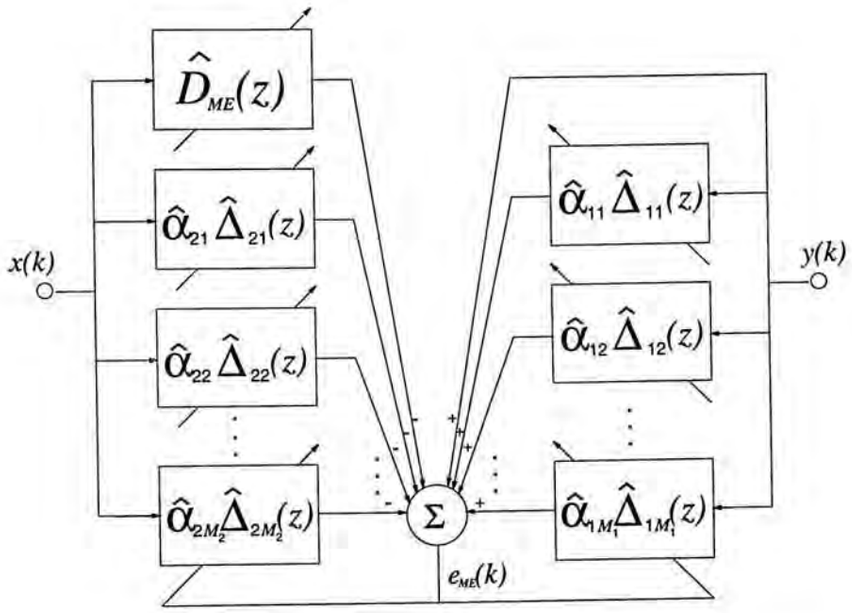

Figure 5.15: System block diagram of the METDE

After initialization, a new adaptive algorithm, namely, the multipath equalization time delay estimator (METDE) [74],[76], is used to obtain accurate estimates of the TDOA and multipath parameters. The system block diagram of the METDE is depicted in Figure 5.15. Basically, it consists of one ETDE and $(M_1 + M_2)$ ETDGE. The output error of the METDE, $e(k)$, is equal to (5.49) except that the system parameters, $\hat{D}_{ME}(k)$, $\hat{\alpha}_{1p}(k)$, $\hat{\Delta}_{1p}(k)$, $\hat{\alpha}_{2q}(k)$ and $\hat{\Delta}_{2q}(k)$, are now time-varying. Therefore, it is an adaptive realization for finding the minimum point of $E\{e^2(k)\}$. The system parameters of the METDE again use stochastic gradient estimates which are obtained by differentiating $e^2(k)$ with respect to the instantaneous TDOA and multipath parameter estimates. The METDE adaptive algorithm is given by

$$\begin{aligned}\hat{D}_{ME}(k+1) &= \hat{D}_{ME}(k) - \mu_D \frac{\partial e^2(k)}{\partial \hat{D}_{ME}(k)} \\ &= \hat{D}_{ME}(k) - 2\mu_D e(k) \sum_{n=-P}^{P} x(k-n) f(n - \hat{D}_{ME}(k)) \quad (5.52)\end{aligned}$$

$$\begin{aligned}\hat{\alpha}_{1p}(k+1) &= \hat{\alpha}_{1p}(k) - \mu_\alpha \frac{\partial e^2(k)}{\partial \hat{\alpha}_{1p}(k)}, \qquad 1 \le p \le M_1 \\ &= \hat{\alpha}_{1p}(k) - 2\mu_\alpha e(k) \sum_{n=-P}^{P} sinc(n - \hat{\Delta}_{1p}(k)) y(k-n) \qquad (5.53)\end{aligned}$$

$$\begin{aligned}\hat{\Delta}_{1p}(k+1) &= \hat{\Delta}_{1p}(k) - \frac{\mu_\Delta}{\hat{\alpha}_{1p}(k)} \frac{\partial e^2(k)}{\partial \hat{\Delta}_{1p}(k)}, \qquad 1 \le p \le M_1 \\ &= \hat{\Delta}_{1p}(k) + 2\mu_\Delta e(k) \sum_{n=-P}^{P} f(n - \hat{\Delta}_{1p}(k)) y(k-n) \qquad (5.54)\end{aligned}$$

$$\begin{aligned}\hat{\alpha}_{2q}(k+1) &= \hat{\alpha}_{2q}(k) - \mu_\alpha \frac{\partial e^2(k)}{\partial \hat{\alpha}_{2q}(k)}, \qquad 1 \le q \le M_2 \\ &= \hat{\alpha}_{2q}(k) + 2\mu_\alpha e(k) \sum_{n=-P}^{P} sinc(n - \hat{D}_{ME}(k)) x(k-n) \qquad (5.55)\end{aligned}$$

$$\begin{aligned}\hat{\Delta}_{2q}(k+1) &= \hat{\Delta}_{2q}(k) - \frac{\mu_\Delta}{\hat{\alpha}_{2q}(k)} \frac{\partial e^2(k)}{\partial \hat{\Delta}_{2q}(k)}, \qquad 1 \le q \le M_2 \\ &= \hat{\Delta}_{2q}(k) - 2\mu_\Delta e(k) \sum_{n=-P}^{P} f(n - \hat{\Delta}_{2q}(k)) x(k-n) \qquad (5.56)\end{aligned}$$

It is easy to show that $((1+M_1+M_2)(2P+1)-1)$ additions and $(3+M_1+M_2)(2P+1)$ multiplications are required to compute $e(k)$ for each sampling interval. The adjustment for each of the $\hat{D}_{ME}(k)$, $\hat{\Delta}_{1p}(k)$ and $\hat{\Delta}_{2q}(k)$ needs $(6P+3)$ additions and $(4P+5)$ multiplications while $(2P+1)$ additions and $(2P+4)$ multiplications are involved in calculating $\hat{\alpha}_{1p}(k)$ or $\hat{\alpha}_{2q}(k)$. Consequently, the amount of computation is roughly proportional to the number of multipaths and the filter length or P.

5.3.2 Performance of the METDE

To analyze the performance of the METDE, a simple case that $M_1 = M_2 = 1$ is first considered. In this situation, only D, α_{11}, Δ_{11}, α_{21} and Δ_{21} are needed to be estimated. It is assumed that good initial estimates of these parameters have been determined such that global convergence is guaranteed, and P is also assumed to be sufficiently large such that $\hat{v}(z)$ can be replaced by z^{-v} in the following derivations.

Assuming that the estimated parameters are uncorrelated with received outputs [57] and $\hat{\alpha}_{11}(k) \to \alpha_{11}$, $\hat{\Delta}_{11}(k) \to \Delta_{11}$, $\hat{\alpha}_{21}(k) \to \alpha_{21}$ and $\hat{\Delta}_{21}(k) \to \Delta_{21}$, we evaluate the learning characteristics of $\hat{D}_{ME}(k)$ by taking the expected value of (5.52),

$$
\begin{aligned}
&E\{\hat{D}_{ME}(k+1)\} \\
&= E\{\hat{D}_{ME}(k)\} - 2\mu_D E\Big\{e(k)\sum_{n=-P}^{P} x(k-n) f(n-\hat{D}_{ME}(k))\Big\} \\
&\approx E\{\hat{D}_{ME}(k)\} - 2\mu_D E\Big\{\big(s(k-D)+\alpha_{11}s(k-D-\Delta_{11})+n_2(k) \\
&\quad +\alpha_{11}n_2(k-\Delta_{11}) - s(k-\hat{D}_{ME}(k)) - \alpha_{11}s(k-\hat{D}_{ME}(k)-\Delta_{11}) \\
&\quad -n_1(k-\hat{D}_{ME}(k)) - \alpha_{21}n_1(k-\Delta_{21})\big) \\
&\quad \cdot \sum_{n=-\infty}^{\infty}\big(s(k-n)+\alpha_{11}s(k-\Delta_{11}-n)+n_1(k-n)\big) f(n-\hat{D}_{ME}(k))\Big\} \\
&= E\{\hat{D}_{ME}(k)\} - 2\mu_D E\Big\{\Big(\sum_{i=-\infty}^{\infty} s(k-i)\cdot(sinc(i-D)+\alpha_{11}sinc(i-D-\Delta_{11}) \\
&\quad -sinc(i-\hat{D}_{ME}(k)) - \alpha_{11}sinc(i-\hat{D}_{ME}(k)-\Delta_{11})) - \sum_{i=-\infty}^{\infty} n_1(k-i) \\
&\quad \cdot(sinc(i-\hat{D}_{ME}(k)) + \alpha_{21}sinc(i-\Delta_{21})\Big)\cdot \sum_{n=-\infty}^{\infty}\Big(s(k-n) \\
&\quad +\alpha_{11}\sum_{j=-\infty}^{\infty} s(k-j)sinc(j-\Delta_{11}-n)+n_1(k-n)\Big) f(n-\hat{D}_{ME}(k))\Big\} \qquad (5.57)
\end{aligned}
$$

Note that $s(k)$, $n_1(k)$ and $n_2(k)$ are white and mutually uncorrelated, and

$$\sum_{i=\infty}^{\infty} sinc(i-a)sinc(i-b) = sinc(b-a)$$

The expression for $E\{\hat{D}_{ME}(k+1)\}$ can be written as

$$
\begin{aligned}
&E\{\hat{D}_{ME}(k+1)\} \\
&= E\{\hat{D}_{ME}(k)\} - 2\mu_D\sigma_s^2 E\Big\{\sum_{n=-\infty}^{\infty} f(n-\hat{D}_{ME}(k))(sinc(n-D) \\
&\quad +\alpha_{11}sinc(n+\Delta_{11}-D)+\alpha_{11}sinc(n-D-\Delta_{11})+\alpha_{11}^2 sinc(n-D) \\
&\quad -sinc(n-\hat{D}_{ME}(k)) - \alpha_{11}sinc(n+\Delta_{11}-\hat{D}_{ME}(k)) - \alpha_{11}^2 sinc(n-\hat{D}_{ME}(k)))\Big\} \\
&\quad +2\mu_D\sigma_n^2\cdot E\Big\{\sum_{n=-\infty}^{\infty} f(n-\hat{D}_{ME}(k))(sinc(n-\hat{D}_{ME}(k)) \\
&\quad +\alpha_{21}sinc(n-\Delta_{21}))\Big\} \\
&= E\{\hat{D}_{ME}(k)\} - 2\mu_D\sigma_s^2 E\{f(D-\hat{D}_{ME}(k)) + \alpha_{11}f(D-\Delta_{11}-\hat{D}_{ME}(k))
\end{aligned}
$$

$$+\alpha_{11}f(D+\Delta_{11}-\hat{D}_{ME}(k))+\alpha_{11}^2 f(D-\hat{D}_{ME}(k))-f(0)-\alpha_{11}f(\Delta_{11})$$
$$-\alpha_{11}f(-\Delta_{11})-\alpha_{11}^2 f(0)\}+2\mu_D\sigma_n^2 E\{f(0)+\alpha_{11}f(\Delta_{21}-\hat{D}_{ME}(k))\} \quad (5.58)$$

Since f is an odd function and $f(0)=0$, (5.58) becomes

$$\begin{aligned} & E\{\hat{D}_{ME}(k+1)\} \\ = \; & E\{\hat{D}_{ME}(k)\}-2\mu_D\sigma_s^2 E\{(1+\alpha_{11}^2)f(D-\hat{D}_{ME}(k))+\alpha_{11}f(D-\hat{D}_{ME}(k)+\Delta_{11}) \\ & +\alpha_{11}f(D-\hat{D}_{ME}(k)-\Delta_{11})\}+2\mu_D\sigma_n^2 E\{\alpha_{11}f(\Delta_{21}-\hat{D}_{ME}(k))\} \end{aligned} \quad (5.59)$$

When $\sigma_s^2 \gg \sigma_n^2$ and when $\hat{D}_{ME}(k)$ is close to D, (5.59) can be approximated by

$$\begin{aligned} E\{\hat{D}_{ME}(k+1)\} &\approx E\{\hat{D}_{ME}(k)\}-2\mu_D(1+\alpha_{11}^2)\sigma_s^2 E\{f(D-\hat{D}_{ME}(k))\} \\ &\approx E\{\hat{D}_{ME}(k)\}-\frac{2}{3}\mu_D(1+\alpha_{11}^2)\sigma_s^2\pi^2 E\{\hat{D}_{ME}(k)-D\} \end{aligned} \quad (5.60)$$

Solving (5.60) yields

$$E\{\hat{D}_{ME}(k)\}\approx D+(\hat{D}_{ME}(0)-D)\Big(1-\frac{2}{3}\mu_D(1+\alpha_{11}^2)\sigma_s^2\pi^2\Big)^k \quad (5.61)$$

provided that, in this case, μ_D satisfies the following condition,

$$0<\mu_D<\frac{3}{(1+\alpha_{11}^2)\sigma_s^2\pi^2} \quad (5.62)$$

Similarly, the learning trajectories of $\hat{\alpha}_{11}(k)$, $\hat{\Delta}_{11}(k)$, $\hat{\alpha}_{21}(k)$ and $\hat{\Delta}_{21}(k)$ have been shown to be (See Appendix K)

$$E\{\hat{\alpha}_{11}(k)\}\approx\frac{\alpha_{11}(\sigma_y^2-\sigma_n^2)}{\sigma_y^2}+\Big(\hat{\alpha}_{11}(0)-\alpha_{11}\frac{(\sigma_y^2-\sigma_n^2)}{\sigma_y^2}\Big)(1-2\mu_\alpha\sigma_y^2)^k \quad (5.63)$$

$$E\{\hat{\Delta}_{11}(k)\}\approx\Delta_{11}+(\hat{\Delta}_{11}(0)-\Delta_{11})\Big(1-\frac{2}{3}\mu_\Delta\alpha_{11}(1+\alpha_{21}^2)\sigma_s^2\pi^2\Big)^k \quad (5.64)$$

$$E\{\hat{\alpha}_{21}(k)\}\approx\frac{\alpha_{21}(\sigma_x^2-\sigma_n^2)}{\sigma_x^2}+\Big(\hat{\alpha}_{21}(0)-\alpha_{21}\frac{(\sigma_x^2-\sigma_n^2)}{\sigma_x^2}\Big)(1-2\mu_\alpha\sigma_x^2)^k \quad (5.65)$$

$$E\{\hat{\Delta}_{21}(k)\}\approx\Delta_{21}+(\hat{\Delta}_{21}(0)-\Delta_{21})\Big(1-\frac{2}{3}\mu_\Delta\alpha_{21}(1+\alpha_{11}^2)\sigma_s^2\pi^2\Big)^k \quad (5.66)$$

where σ_x^2 and σ_y^2 denote the powers of $x(k)$ and $y(k)$ respectively and they are given by

$$\sigma_x^2 = \sigma_s^2(1 + \alpha_{11}^2 + 2\alpha_{11} sinc(\Delta_{11})) + \sigma_n^2 \tag{5.67a}$$

and

$$\sigma_y^2 = \sigma_s^2(1 + \alpha_{21}^2 + 2\alpha_{21} sinc(\Delta_{21} - D)) + \sigma_n^2 \tag{5.67b}$$

It is observed that the unbiasedness of $\hat{D}_{ME}(k)$, $\hat{\Delta}_{11}(k)$ and $\hat{\Delta}_{21}(k)$ can be maintained at high SNR environments. Whereas the gain estimates are biased in the presence of noise although they are unbiased in a noise-free condition. Moreover, the time constants of the TDOA and multipath parameter estimates depend on the actual values of the multipath delays and multipath gains.

The steady state mean square errors of the estimates of the time delay and multipath parameters have been derived and they are approximated by (See Appendix L)

$$\begin{aligned} \xi(\hat{D}_{ME}) &\triangleq \lim_{k\to\infty} E\{(\hat{D}_{ME}(k) - D)^2\} \\ &\approx \mu_D \sigma_{n'}^2 \end{aligned} \tag{5.68}$$

$$\begin{aligned} \xi(\hat{\alpha}_{11}) &\triangleq \lim_{k\to\infty} E\{(\hat{\alpha}_{11}(k) - \alpha_{11})^2\} \\ &\approx \mu_\alpha \sigma_{n'}^2 + \frac{\alpha_{11}^2 \sigma_n^4}{\sigma_y^4} \end{aligned} \tag{5.69}$$

$$\begin{aligned} \xi(\hat{\Delta}_{11}) &\triangleq \lim_{k\to\infty} E\{(\hat{\Delta}_{11}(k) - \Delta_{11})^2\} \\ &\approx \frac{\mu_\Delta \sigma_{n'}^2}{\alpha_{11}} \end{aligned} \tag{5.70}$$

$$\begin{aligned} \xi(\hat{\alpha}_{21}) &\triangleq \lim_{k\to\infty} E\{(\hat{\alpha}_{21}(k) - \alpha_{21})^2\} \\ &\approx \mu_\alpha \sigma_{n'}^2 + \frac{\alpha_{21}^2 \sigma_n^4}{\sigma_x^4} \end{aligned} \tag{5.71}$$

$$\begin{aligned} \xi(\hat{\Delta}_{21}) &\triangleq \lim_{k\to\infty} E\{(\hat{\Delta}_{21}(k) - \Delta_{21})^2\} \\ &\approx \frac{\mu_\Delta \sigma_{n'}^2}{\alpha_{21}} \end{aligned} \tag{5.72}$$

where

$$\sigma_{n'}^2 = \sigma_n^2(2 + \alpha_{11}^2 + \alpha_{21}^2 + 2\alpha_{11}sinc(\Delta_{11})) + 2\alpha_{21}sinc(\Delta_{21} - D)) \qquad (5.73)$$

Basically, the mean square errors of these parameters increase with the step sizes and $\sigma_{n'}^2$. In addition, the values of $\xi(\hat{\Delta}_{11})$ and $\xi(\hat{\Delta}_{21})$ decrease as α_{11} and α_{21} increase. While the steady state error for each of the gain estimates comprises an additional term which equals the squared value of its bias. It is worthy to note that since the inter-dependence of the system parameters has not been considered in the above derivations, the convergence dynamics of these parameters may be different from the theoretical values in actual circumstances. Furthermore, their mean square errors will be greater than those as given by (5.68) to (5.72) and have finite values in the absence of noise.

5.3.3 Simulation Results

The performance of the METDE has been evaluated and contrasted with the MCTDE through computer simulations. The signal source and its delay versions, as well as the corrupting noises and the required SNRs were generated in a similar way as in Section 5.2.5. At the beginning of the adaptation, the MCTDE was freely adjusted for 600 iterations to deduce rough estimates of the TDOA and multipath delays. All initial values of the multipath gain estimates were arbitrarily selected to 0.5. The filter length, $2P + 1$, was chosen be 31 while the step size μ_α had a value of 0.001 and $\mu_\Delta = \mu_D = 0.0005$. The experimental results are obtained by taking ensemble average of 200 independent trials.

Figure 5.16 to Figure 5.18 show the learning trajectories of the estimated delay and multipath parameter estimates for $M_1 = M_2 = 1$ at SNR = 10 dB. The actual parameters were given as follows: $D = 1.4$, $\alpha_{11} = 0.8$, $\alpha_{21} = 0.4$, $\Delta_{11} = 4.5$ and $\Delta_{21} = 7.1$, which were identical to Experiment 2 of Section 5.2.5 except that additive noises were added. It can be seen from Figure 5.16 that $\hat{D}_{ME}(k)$ approaches the optimal value, viz., 1.4, at the 4000th iteration. The multipath variables $\hat{\alpha}_{11}(k)$ and $\hat{\Delta}_{11}(k)$ converged to 0.73 and 4.50 at the 2500th and the 4500th iteration respectively

in Figure 5.17 while $\hat{\alpha}_{21}(k) \rightarrow 0.38$ and $\hat{\Delta}_{21}(k) \rightarrow 7.11$ in about 4000 iterations as observed from Figure 5.18. As discussed in Section 5.3.2, the estimates of the TDOA and multipath delays were close to the actual values whereas $\hat{\alpha}_{11}(k)$ and $\hat{\alpha}_{21}(k)$ were biased due to the corrupting noises. Moreover, we see that the convergence behaviour of $\hat{\alpha}_{11}(k)$ agreed with the predicted trajectory but the other parameters had slower learning rates as compared with their theoretical values. The discrepancies are mainly caused by the inter-dependence of the system parameters which have not been taken into account in the analytical derivations. Because when deriving the learning trajectory of any of these parameters in the analysis, it is assumed that other parameters have already converged to the optimal values. But this assumption is not true necessarily in practice.

The mean square errors of $\hat{D}_{ME}(k)$, $\hat{\alpha}_{11}(k)$, $\hat{\Delta}_{11}(k)$, $\hat{\alpha}_{21}(k)$ and $\hat{\Delta}_{21}(k)$ are tabulated in Table 5.2. Again, since the inter-dependence of the system parameters has not been considered in the theoretical study, the experimental results for all parameters were larger than the desired values. However, the mean square errors of these estimates had the same order with the theoretical expectations. This implies that (5.68) to (5.72) can still be used to evaluate the performance of the METDE.

Experiment 4 of Section 5.2.5 was repeated for the METDE and the results are shown in Figure 5.19 to Figure 5.23. In this experiment, the time delay and multipath parameters were made time-varying and they were given step offsets after every 5000 iterations for both noise-free and SNR = 10 dB cases. In the absence of noise, it can be seen that the system was capable to track these step changes and provided accurate estimates of D, Δ_{11} and Δ_{21} after the transient. The tracking capability for $\hat{\alpha}_{11}(k)$ and $\hat{\alpha}_{21}(k)$ was also satisfactory except that there were minor errors during the last 5000 iterations. These biases can be removed by using longer filter lengths in the METDE. This test had been carried out for $P = 30$ and we had found that all estimates of the time delay and multipath parameters converged to the optimum values. At SNR = 10 dB, the TDOA and multipath delay estimates were also close to the correct values. Small biases existed in these estimates because the actual values of the TDOA and multipath parameters do not map exactly to the

global minimum of the METDE performance surface in the presence of noises. On the other hand, $\hat{\alpha}_1(k)$ and $\hat{\alpha}_2(k)$ did not converge to their desired values due to the additive noises and this has been conferred in (5.63) and (5.65). Comparing these results with Figure 5.9 to Figure 5.13, we see that the METDE is superior to the MCTDE in the noise-free environment. However, they had comparable performance at SNR = 10 dB.

The estimation performance of the METDE in the presence of four multipaths was also investigated. In this experiment, $M_1 = M_2 = 2$, $D = 0.7$, $\alpha_{11} = 0.8$, $\alpha_{21} = 0.5$, $\alpha_{12} = 0.6$, $\alpha_{22} = 0.4$, $\Delta_{11} = 3.3$, $\Delta_{21} = 4.8$ $\Delta_{12} = 6.4$ and $\Delta_{22} = 8.7$. Two METDE systems were considered. The first METDE was an insufficient multipath equalization model which assumed that there was only one multipath at each sensor. Whereas the second one provided exact modeling and it consisted of one ETDE and four ETDGEs. The estimates of these two METDE configurations at a noise-free condition and SNR = 10 dB were tabulated in Table 5.3. It can be seen that for the insufficient model, accurate estimation of D was still achieved in both tests but the results for other parameters were not as satisfactory. On the other hand, we observe that precise estimates of the time delay and multipath parameters were obtained by using an exact equalization model when noise was absent. While in the presence of corrupting noises, the METDE gave good estimates of the TDOA and multipath delays but the gain estimates were relatively less precise. This is expected since (5.63) and (5.65) indicate that the gain estimates are biased under a noisy environment.

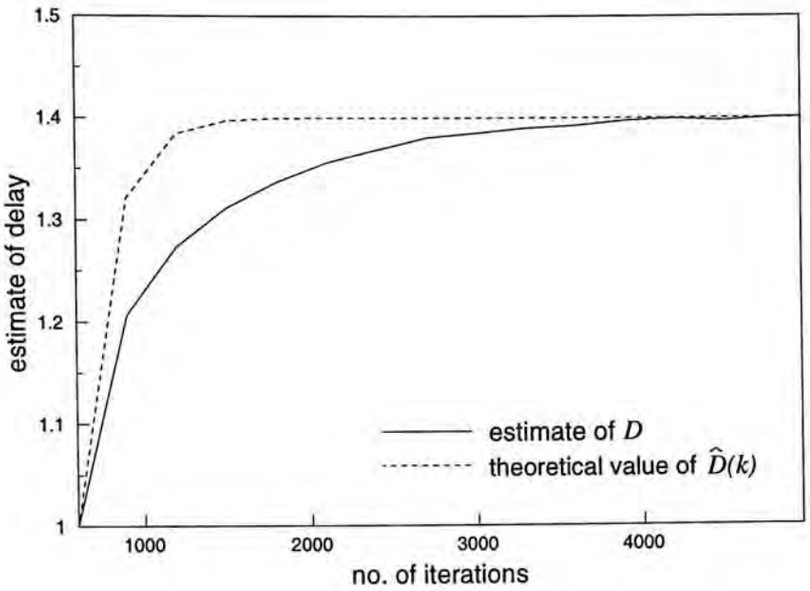

Figure 5.16: Estimate of D of the METDE for SNR = 10 dB

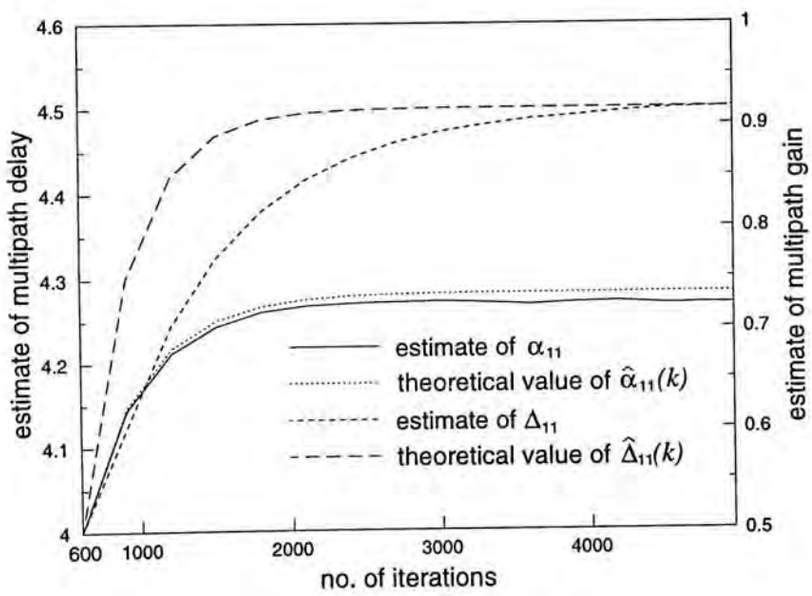

Figure 5.17: Estimate of α_{11} and Δ_{11} of the METDE for SNR = 10 dB

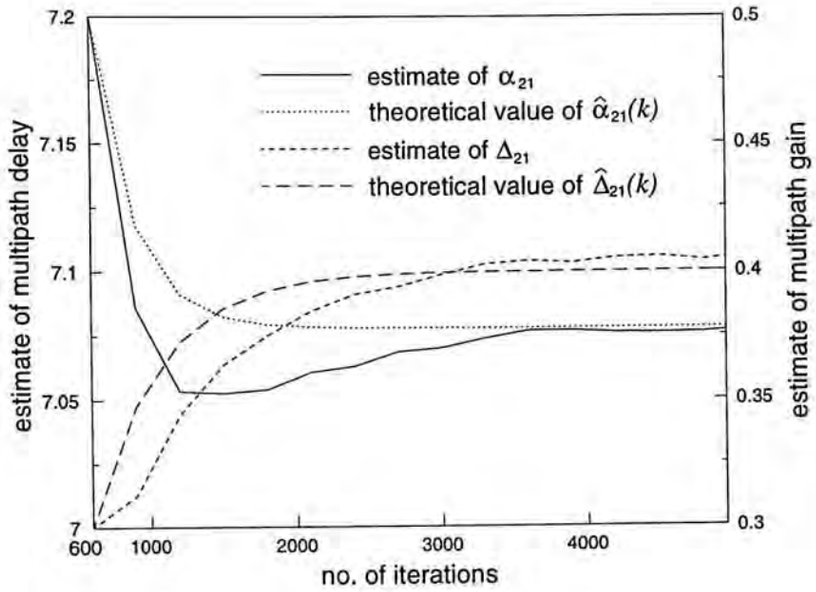

Figure 5.18: Estimate of α_{21} and Δ_{21} of the METDE for SNR = 10 dB

	Measured value	Theoretical value
$var(\hat{D}_{ME})$	2.00×10^{-4}	1.44×10^{-4}
$var(\hat{\alpha}_{11})$	5.96×10^{-3}	4.56×10^{-3}
$var(\hat{\Delta}_{11})$	2.11×10^{-4}	1.80×10^{-4}
$var(\hat{\alpha}_{21})$	8.20×10^{-4}	7.54×10^{-4}
$var(\hat{\Delta}_{21})$	5.00×10^{-4}	3.60×10^{-4}

Table 5.2: Measured and theoretical variances of the system parameters of the METDE for SNR = 10 dB

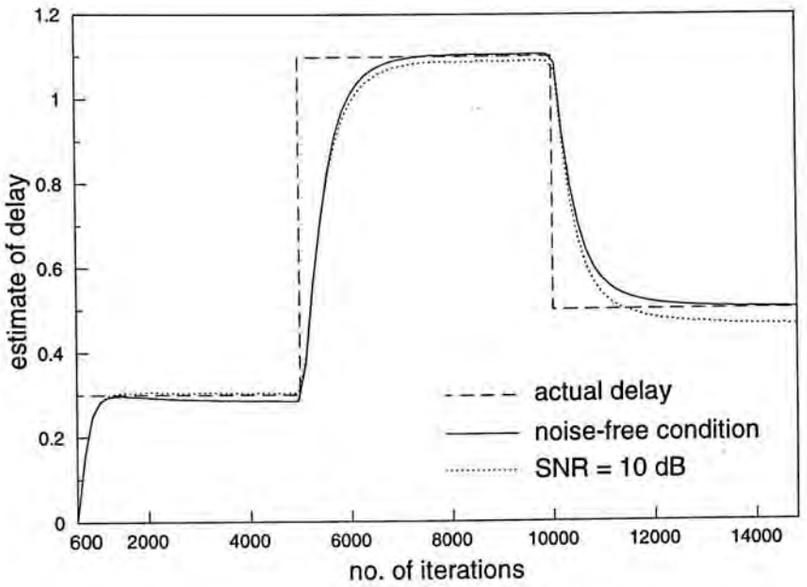

Figure 5.19: Tracking behaviour of $\hat{D}_{ME}(k)$ of the METDE for different SNRs

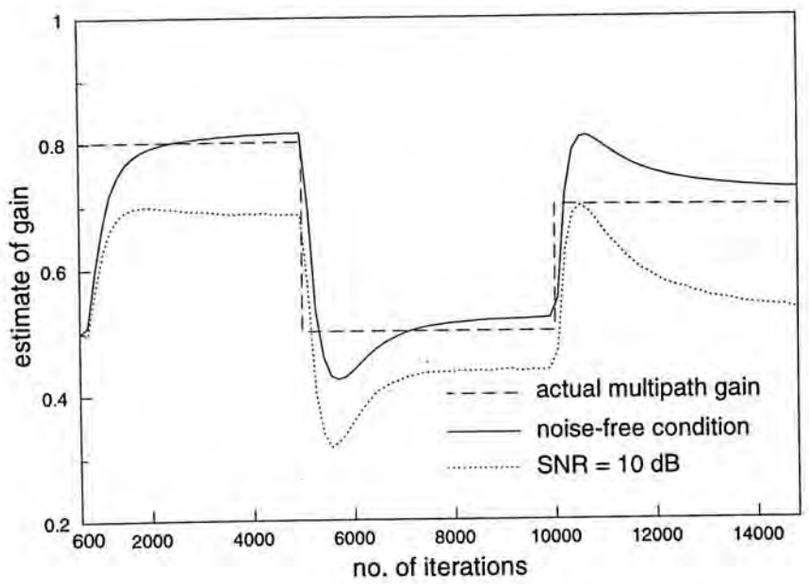

Figure 5.20: Tracking behaviour of $\hat{\alpha}_{11}(k)$ of the METDE for different SNRs

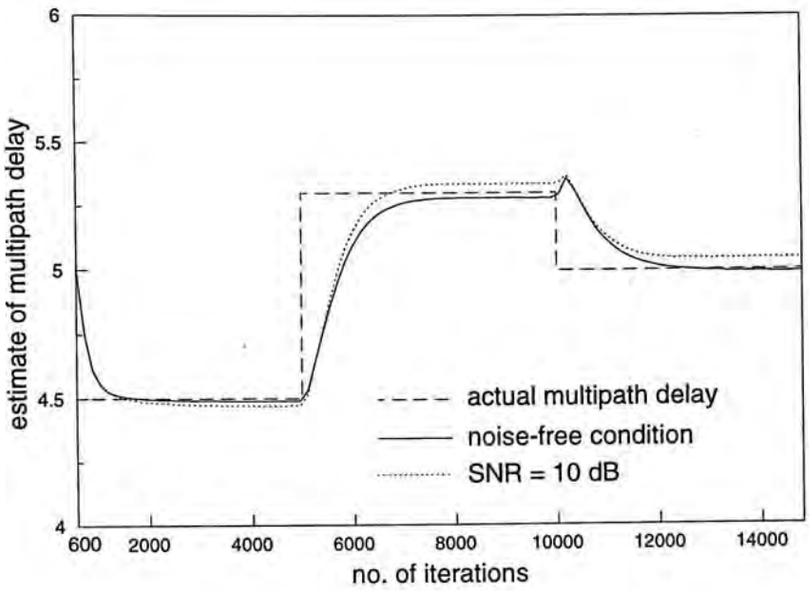

Figure 5.21: Tracking behaviour of $\hat{\Delta}_{11}(k)$ of the METDE for different SNRs

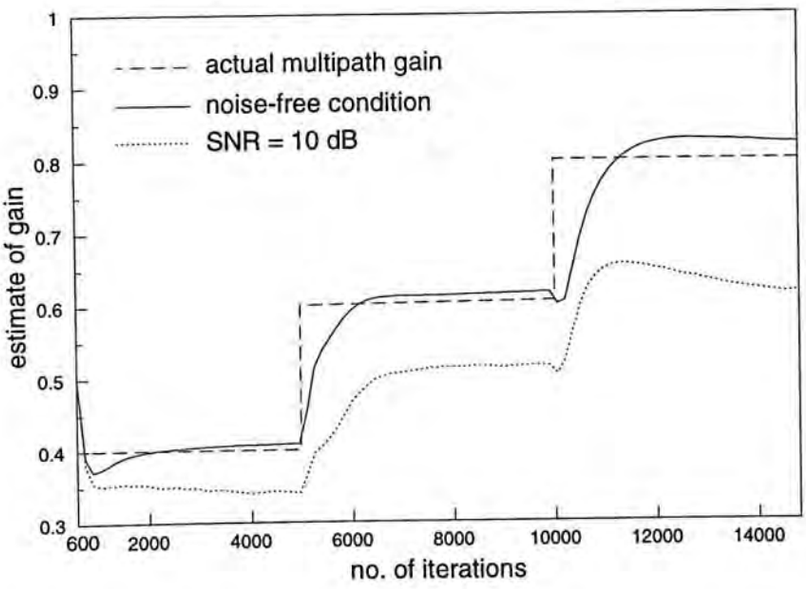

Figure 5.22: Tracking behaviour of $\hat{\alpha}_{21}(k)$ of the METDE for different SNRs

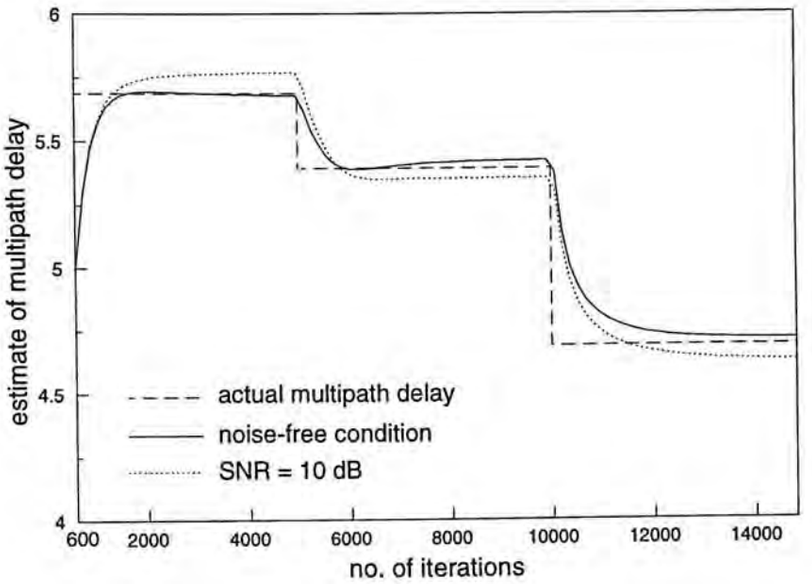

Figure 5.23: Tracking behaviour of $\hat{\Delta}_{21}(k)$ of the METDE for different SNRs

		SNR=∞		SNR=10 dB	
	Optimum value	Insufficient modeling	Sufficient modeling	Insufficient modeling	Sufficient modeling
$E\{\hat{D}_{ME}\}$	0.7	0.693	0.686	0.701	0.693
$E\{\hat{\alpha}_{11}\}$	0.8	0.608	0.765	0.562	0.643
$E\{\hat{\Delta}_{11}\}$	3.3	2.99	3.25	3.01	3.16
$E\{\hat{\alpha}_{21}\}$	0.5	0.524	0.462	0.488	0.383
$E\{\hat{\Delta}_{21}\}$	4.8	5.27	4.84	5.28	5.01
$E\{\hat{\alpha}_{12}\}$	0.6	–	0.584	–	0.455
$E\{\hat{\Delta}_{12}\}$	6.4	–	6.40	–	6.36
$E\{\hat{\alpha}_{22}\}$	0.4	–	0.374	–	0.275
$E\{\hat{\Delta}_{22}\}$	8.7	–	8.62	–	8.54

Table 5.3: Estimates of the METDE for four multipaths using different models

5.4 Summary

Two adaptive systems for multipath time delay estimation between two spatially separated sensors at high SNR environments are proposed. They are called the multipath cancellation time delay estimator (MCTDE) and the multipath equalization time delay estimator (METDE) respectively. The MCTDE comprises two adaptive IIR filters for canceling out the multipath signal at each channel and one ETDE to extract the time difference between the sensor outputs. Although the MCTDE is designed to provide complete multipath cancellation for interpath delays which are integral multiples of the sampling interval, it can also give accurate estimates of multipath parameters when the gains are small or if the delays are large. On the other hand, the METDE is a parallel-split filter whose subunit is either the ETDE or the ETDGE. It equalizes the effect of the two sensor outputs through minimization of a cost function whose global minimum maps to the actual values of the TDOA and multipath parameters. Experimental results are presented to corroborate and evaluate the performance of these two algorithms. Due to the inherent modeling errors in the IIR filters of the MCTDE, its performance is inferior to the METDE particularly in situation that the multipath delays are small and the multipath gains are large. However, it is found that the TDOA and multipath delay estimates of the MCTDE are more tolerant of additive noises. Whereas the METDE provides an exact model for multipath equalization in a noise-free condition and it can tackle the problem when there is more than one multipath at each sensor.

Chapter 6

Conclusions and Suggestions for Future Research

6.1 Conclusions

Estimation of time delay between signals received at two spatially separated sensors has found important applications in many areas such as sonar, radar and biomedicine. One promising approach for the time delay estimation (TDE) problem is offered by adaptive filtering techniques. Two adaptive TDE algorithms, namely, the least mean square time delay estimator (LMSTDE) and the the adaptive digital delay-lock discriminator (ADDLD), are first briefly reviewed. Despite their limitations, these methods have two major advantages over the conventional implementations of generalized cross correlators. Firstly, it can track time-varying delays functions. Secondly, no spectral estimation is required.

A novel TDE adaptive algorithm, called the explicit time delay estimator (ETDE), is developed to circumvent the shortcomings of other adaptive TDE methods. The idea of the ETDE is to model the delay by an FIR filter whose coefficients are constrained to be samples of a *sinc* function. The delay estimate is updated directly and iteratively using an LMS-type algorithm. As a result, the ETDE is more computational efficient than the LMSTDE whose estimated delay is obtained by interpolating its filter weights. Furthermore, the ETDE provides a more reliable approach for delay estimation because wrong estimates due to false peak weights

will not occur. Although both the ADDLD and the ETDE give explicit delay measurements, the ETDE has the advantage over the ADDLD since it can estimate delays that are non-integral multiples of the sampling period. As an alternative realization of the constrained adaptive time delay estimation (CATDE) algorithm developed by Ching and Chan, the ETDE provides more accurate delay estimates, particularly under noisy environment. The mean square delay error of the ETDE has been shown to be larger than the Cramér-Rao lower bound (CRLB) by a few times.

However, the ETDE is proved to be biased for finite filter lengths and the delay error increases when the signal-to-noise ratio (SNR) or the number of filter taps decreases. Under very noisy environment, the ETDE is unable to provide accurate delay estimates with a reasonable filter length. Based on Wiener solution, an adaptive gain control is added to the ETDE so that optimal filter weights for time delay estimation can be achieved. It has been shown that this improved ETDE, namely, the explicit time delay and gain estimator (ETDGE), gives a smaller delay variance at low SNR and an unbiased delay estimate for short filter lengths.

In the presence of multipath transmissions, it is very difficult to achieve accurate delay estimation though not totally impossible. Two adaptive systems based on the ETDE and the ETDGE are proposed to tackle the multipath TDE problem for high SNR conditions. They are called the multipath cancellation time delay estimator (MCTDE) and the multipath equalization time delay estimator (METDE) respectively. The MCTDE can be realized as a pair of multipath cancellers, which are adaptive IIR filters, followed by an ETDE. When the multipath delays are integral multiples of the sampling interval, the MCTDE attains the desired performance. Otherwise, the multipath cancellers will suffer inherent modeling errors and thus complete multipath cancellation is not possible. In these circumstances, the accuracy of the delay and multipath parameter estimates increases when the SNR increases, the multipath gains decrease or the interpath delays increase. On the other hand, the METDE, which consists of the ETDE and the ETDGE only, allows exact modeling of the time difference of arrival (TDOA) and multipath parameters.

The concept of the METDE is to equalize the two sensor outputs through minimization of a cost function whose global minimum maps to the actual values of the time delay and multipath parameters. In addition, the METDE provides accurate estimates even when there is more than one multipath at each sensor.

6.2 Suggestions for Future Research

1. Multipath time delay estimation under noisy environment

In the presence of multipath propagation, both MCTDE and METDE have the same limitation to operate at high SNR conditions. From (5.42), it can be observed that the multipath cancellers of the MCTDE will become ineffective when noise dominates. While the global minimum of the cost function in the METDE shifts away from the desired value in the presence of noise. Based on the METDE, a possible solution for multipath TDE under noisy environment is suggested by minimizing the mean square value of (5.49) subject to some constraints such that the effect of noise is eliminated. For example, we may maintain the noise in the cost function to a constant value which does not affect the minimization process [79]. If this is feasible, unbiased TDOA and multipath parameter estimates can be obtained.

2. Time delay estimation for monochromatic signals

In radar and certain types of underwater acoustics signals, the radiated signal can be modeled as a pure sinusoid [80]. Using the techniques in deriving the ETDE, efficient algorithms may be designed for estimating and tracking the phase between sinusoids received at two spatially separated sensors. Apart from phase information, the amplitude as well as the frequency of the sinusoids may also be estimated.

3. Joint time delay estimation and system identification

When the transmitting media in time delay estimation are some unknown transfer functions, it is desired to estimate the TDOA as well as the impulse responses of

unknown channels [43]. The proposed ETDE or ETDGE may be applied in such situations. Notice that we have already applied the ETDE and the ETDGE to the multipath channels.

4. New structures and algorithms for time delay estimation

Other novel structures and algorithms may be developed for time delay estimation to obtain faster convergence rate, better tracking capability and smaller delay variance. All the adaptive algorithms developed in this thesis belong to LMS family. Based on the philosophy of the ETDE and/or the ETDGE, different types of LMS methods or even new recursive least square (RLS) algorithms [56] may be designed for time delay estimation.

Appendix A

Derivation of (3.20)

Equation (3.20) is derived mathematically as follows. Using (2.1) and (3.2), (3.3) can be expressed as

$$\hat{D}(k+1) = \hat{D}(k) - 2\mu_D(T_1 + T_2 + T_3 + T_4) \tag{A.1}$$

where

$$T_1 = (s(k-D) - \tilde{s}(k-\hat{D}(k))) \cdot \sum_{i=-P}^{P} s(k-i) f(i-\hat{D}(k))$$

$$T_2 = (s(k-D) - \tilde{s}(k-\hat{D}(k))) \cdot \sum_{i=-P}^{P} n_1(k-i) f(i-\hat{D}(k))$$

$$T_3 = (n_2(k) - \tilde{n}_1(k-\hat{D}(k)) \cdot \sum_{i=-P}^{P} s(k-i) f(i-\hat{D}(k))$$

$$T_4 = (n_2(k) - \tilde{n}_1(k-\hat{D}(k)) \cdot \sum_{i=-P}^{P} n_1(k-i) f(i-\hat{D}(k))$$

Subtracting both sides of (A.1) by D, squaring both sides and then taking the expected value yields

$$\begin{aligned}\epsilon(k+1) \;=\; & \epsilon(k) - 4\mu_D(E\{(\hat{D}(k)-D)T_1\} + E\{(\hat{D}(k)-D)T_2\} \\ & + E\{(\hat{D}(k)-D)T_3\} + E\{(\hat{D}(k)-D)T_4\}) + 4\mu_D^2(E\{T_1^2\} + E\{T_2^2\} \\ & + E\{T_3^2\} + E\{T_4^2\} + 2E\{T_1T_2\} + 2E\{T_1T_3\} + 2E\{T_1T_4\} + 2E\{T_2T_3\} \\ & + 2E\{T_2T_4\} + 2E\{T_3T_4\})\end{aligned} \tag{A.2}$$

Since $s(k)$, $n_1(k)$, $n_2(k)$ and $\hat{D}(k)$ are assumed to be mutually uncorrelated, $E\{(\hat{D}-D)T_2\}$, $E\{(\hat{D}(k)-D)T_3\}$, $E\{T_1T_2\}$, $E\{T_1T_3\}$, $E\{T_2T_4\}$ and $E\{T_3T_4\}$ are all equal to zero. The other terms of (A.2) are then evaluated one by one,

$$
\begin{aligned}
&E\{(\hat{D}(k)-D)T_4\} \\
&= E\left\{(\hat{D}(k)-D)(n_2(k)-\tilde{n}_1(k-\hat{D}(k)))\cdot\sum_{i=-P}^{P} n_1(k-i)f(i-\hat{D}(k))\right\} \\
&= -E\left\{(\hat{D}(k)-D)\sum_{j=-P}^{P} n_1(k-j)sinc(j-\hat{D}(k))\cdot\sum_{i=-P}^{P} n_1(k-i)f(i-\hat{D}(k))\right\} \\
&= -\sum_{i=-P}^{P}\sum_{j=-P}^{P} E\{n_1(k-i)n_1(k-j)\}\cdot E\{(\hat{D}(k)-D)sinc(j-\hat{D}(k)) \\
&\quad \cdot f(i-\hat{D}(k))\} \\
&= -\sigma_n^2 E\left\{(\hat{D}(k)-D)\sum_{i=-P}^{P} sinc(j-\hat{D}(k))f(i-\hat{D}(k))\right\} \\
&\approx -\sigma_n^2 E\{(\hat{D}(k)-D)\cdot f(0)\}
\end{aligned}
$$

Since $f(0)=0$, $E\{(\hat{D}(k)-D)T_4\}$ becomes

$$E\{(\hat{D}(k)-D)T_4\}=0 \tag{A.3}$$

Similarly, $E\{T_1T_4\}$ and $E\{T_2T_3\}$ can also be shown to have zero values. For the term $E\{(\hat{D}(k)-D)T_1\}$, it can be computed from

$$
\begin{aligned}
&E\{(\hat{D}(k)-D)T_1\} \\
&= E\{(\hat{D}(k)-D)\cdot\left(\sum_{j=-\infty}^{\infty} s(k-j)sinc(j-D)-\sum_{j=-P}^{P} s(k-j)sinc(j-\hat{D}(k))\right) \\
&\quad \cdot\sum_{i=-P}^{P} s(k-i)f(i-\hat{D}(k))\} \\
&= \sum_{i=-P}^{P}\sum_{j=-P}^{P} E\{s(k-i)s(k-j)\}E\{(\hat{D}(k)-D)(sinc(j-D)-sinc(j-\hat{D}(k))) \\
&\quad \cdot f(i-\hat{D}(k))\} \\
&\approx \sigma_s^2\cdot E\{(\hat{D}(k)-D)\cdot f(D-\hat{D}(k))\}
\end{aligned}
$$

Expanding $f(D-\hat{D}(k))$ using Taylor's series up to the first order term, $E\{(\hat{D}(k)-D)T_1\}$ is approximated by

$$E\{(\hat{D}(k)-D)T_1\}\approx\frac{1}{3}\pi^2\sigma_s^2\epsilon(k) \tag{A.4}$$

For sufficiently large P, $E\{T_1^2\}$ can be approximated by

$$\begin{aligned} E\{T_1^2\} \approx & \sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty}\sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty} E\{s(k-i)s(k-j)s(k-m)s(k-n)\} \\ & \cdot E\{(sinc(i-D)-sinc(i-\hat{D}(k)))(sinc(j-D)-sinc(j-\hat{D}(k))) \\ & \cdot f(m-\hat{D}(k))f(n-\hat{D}(k))\} \end{aligned}$$

It has been shown that [81] if x_1, x_2, x_3 and x_4 denote four different random variables of a stationary Gaussian process, we may write

$$E\{x_1x_2x_3x_4\} = E\{x_1x_2\}\cdot E\{x_3x_4\} + E\{x_1x_3\}\cdot E\{x_2x_4\} + E\{x_1x_4\}\cdot E\{x_2x_3\}$$

Thus

$$\begin{aligned} & E\{T_1^2\} \\ = & \sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty}\sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty} (E\{s(k-i)s(k-j)\}\cdot E\{s(k-m)s(k-n)\} \\ & +E\{s(k-i)s(k-m)\}\cdot E\{s(k-j)s(k-n)\} + E\{s(k-i)s(k-n)\} \\ & \cdot E\{s(k-j)s(k-m)\})\cdot E\{(sinc(i-D)-sinc(i-\hat{D}(k))) \\ & \cdot(sinc(j-D)-sinc(j-\hat{D}(k)))\cdot f(m-\hat{D}(k))f(n-\hat{D}(k))\} \\ = & \sigma_s^4 E\left\{\sum_{i=-\infty}^{\infty}(sinc(i-D)-sinc(i-\hat{D}(k)))^2\cdot\sum_{m=-\infty}^{\infty}f^2(m-\hat{D}(k))\right\} \\ & +2\sigma_s^4\sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty} E\{(sinc(m-D)-sinc(m-\hat{D}(k)))f(m-\hat{D}(k)) \\ & \cdot(sinc(n-D)-sinc(n-\hat{D}(k)))f(n-\hat{D}(k))\} \end{aligned}$$

From [58]

$$\sum_{m=-\infty}^{\infty} f^2(m) = \frac{\pi^2}{3}$$

therefore, we obtain

$$\begin{aligned} E\{T_1^2\} \approx & \sigma_s^4 E\left\{(2-2\sum_{i=-\infty}^{\infty}(sinc(i-D)sinc(i-\hat{D}(k)))\cdot\frac{\pi^2}{3}\right\} \\ & +2\sigma_s^4 E\left\{\left(\sum_{m=-\infty}^{\infty}(sinc(m-D)-sinc(m-\hat{D}(k)))f(m-\hat{D}(k))\right)^2\right\} \\ = & \frac{2}{3}\pi^2\sigma_s^4 E\{1-sinc(D-\hat{D}(k))\} + 2\sigma_s^4 E\{f^2(D-\hat{D}(k))\} \\ \approx & \frac{1}{3}\pi^4\sigma_s^4\epsilon(k) \end{aligned} \tag{A.5}$$

In a similar manner, it can be shown that

$$E\{T_2^2\} = \frac{1}{9}\pi^4\sigma_s^2\sigma_n^2\epsilon(k) \tag{A.6}$$

$$E\{T_3^2\} = \frac{2}{3}\pi^2\sigma_s^2\sigma_n^2 \tag{A.7}$$

and

$$E\{T_4^2\} = \frac{2}{3}\pi^2\sigma_n^4 \tag{A.8}$$

Putting (A.3) to (A.8) into (A.2) yields

$$\epsilon(k+1) = \epsilon(k)\left(1 - \frac{4}{3}\mu_D\pi^2\sigma_s^2\left(1 - \mu_D\pi^2\sigma_s^2 - \frac{1}{3}\mu_D\pi^2\sigma_n^2\right)\right) + \frac{8}{3}\mu_D^2\pi^2\sigma_n^2(\sigma_s^2 + \sigma_n^2) \tag{A.9}$$

Solving (A.9) gives

$$\epsilon(k) \approx \alpha^k(\hat{D}(0) - D)^2 + \beta\left(\frac{1-\alpha^k}{1-\alpha}\right)$$

where α and β are given by (3.21) and (3.22) respectively.

Appendix B

Derivation of (3.29)

The derivation of (3.29) is given as follows. Let

$$r = 1 - \frac{2}{3}\pi^2 \mu_D \sigma_s^2$$

Then the first k equations of (3.28) can be written as

$$\begin{aligned}
&E\{\hat{D}(k)\} = rE\{\hat{D}(k-1)\} + (1-r)(D_0 + \lambda(k-1)) \\
&E\{\hat{D}(k-1)\} = rE\{\hat{D}(k-2)\} + (1-r)(D_0 + \lambda(k-2)) \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
&E\{\hat{D}(2)\} = rE\{\hat{D}(1)\} + (1-r)(D_0 + \lambda) \\
&E\{\hat{D}(1)\} = r\hat{D}(0) + (1-r)D_0
\end{aligned} \tag{B.1}$$

Multiplying r to the second equation of (B.1), r^2 to the third equation and so on, and then summing the results together gives

$$\begin{aligned}
E\{\hat{D}(k)\} &= r^k\hat{D}(0) + D_0(1-r)(1+r+r^2+\cdots+r^{k-1}) \\
&\quad +(1-r)\lambda((k-1)+\lambda(k-2)+\cdots+\lambda^{k-2}) \\
&= r^k\hat{D}(0) + D_0(1-r^k) + (1-r)\lambda\left(\sum_{i=0}^{k-2}\lambda^i + \sum_{i=0}^{k-1}\lambda^i + \cdots + 1\right) \\
&= r^k\hat{D}(0) + D_0(1-r^k) + (1-r)\lambda\left(\frac{1-r^{k-1}}{1-r} + \frac{1-r^{k-2}}{1-r} + \cdots + 1\right) \\
&= D_0 + \lambda k + \left(\hat{D}(0) - D_0 + \frac{\lambda}{1-r}\right) - \frac{\lambda}{1-r}
\end{aligned}$$

which is identical to (3.29).

Appendix C

Derivation of (4.14)

The derivation of (4.14) follows a similar approach as in Appendix A. The updating rule for $\hat{D}_G(k)$ of the ETDGE algorithm can be written as

$$\hat{D}_G(k+1) = \hat{D}_G(k) - 2\mu_D(T_1 + T_2 + T_3 + T_4) \tag{C.1}$$

where

$$T_1 = (s(k-D) - \hat{g}(k)\tilde{s}(k - \hat{D}_G(k))) \cdot \sum_{i=-P}^{P} s(k-i) f(i - \hat{D}_G(k))$$

$$T_2 = (s(k-D) - \hat{g}(k)\tilde{s}(k - \hat{D}_G(k))) \cdot \sum_{i=-P}^{P} n_1(k-i) f(i - \hat{D}_G(k))$$

$$T_3 = (n_2(k) - \hat{g}(k)\tilde{n}_1(k - \hat{D}_G(k)) \cdot \sum_{i=-P}^{P} s(k-i) f(i - \hat{D}_G(k))$$

$$T_4 = (n_2(k) - \hat{g}(k)\tilde{n}_1(k - \hat{D}_G(k)) \cdot \sum_{i=-P}^{P} n_1(k-i) f(i - \hat{D}_G(k))$$

Same as the ETDE, the iterative equation of the mean square delay error of the ETDGE is given by

$$\begin{aligned} \epsilon_G(k+1) \;=\; & \epsilon_G(k) - 4\mu_D(E\{(\hat{D}_G(k) - D)T_1\} + E\{(\hat{D}_G(k) - D)T_2\} \\ & + E\{(\hat{D}_G(k) - D)T_3\} + E\{(\hat{D}_G(k) - D)T_4\}) + 4\mu_D^2(E\{T_1^2\} + E\{T_2^2\} \\ & + E\{T_3^2\} + E\{T_4^2\} + 2E\{T_1T_2\} + 2E\{T_1T_3\} + 2E\{T_1T_4\} + 2E\{T_2T_3\} \\ & + 2E\{T_2T_4\} + 2E\{T_3T_4\}) \end{aligned} \tag{C.2}$$

where

$$\epsilon_G(k) \;\triangleq\; E\{(\hat{D}_G(k) - D)^2\}$$

Examining the steady state value of (C.2) with the assumption $\hat{g}(k) \to g$ and noting that the terms $E\{(\hat{D}_G(k)-D)T_2\}$, $E\{(\hat{D}_G(k)-D)T_3\}$, $E\{(\hat{D}_G(k)-D)T_4\}$, $E\{T_1T_2\}$, $E\{T_1T_3\}$, $E\{T_1T_4\}$, $E\{T_2T_3\}$, $E\{T_2T_4\}$ and $E\{T_3T_4\}$ all have zero values, we obtain

$$\lim_{k\to\infty} E\{(\hat{D}_G(k) - D)T_1\} = \mu_D \lim_{k\to\infty} \left(E\{T_1^2\} + E\{T_2^2\} + E\{T_3^2\} + E\{T_4^2\}\right) \quad \text{(C.3)}$$

Following the derivation of (A.4), the L.H.S. of (C.3) is found to be

$$\lim_{k\to\infty} E\{(\hat{D}_G(k) - D)T_1\} \approx \frac{1}{3}\pi^2\sigma_s^2 var(\hat{D}_G) \quad \text{(C.4)}$$

Similarly, using (A.5) to (A.8), we get

$$\lim_{k\to\infty} E\{T_1^2\} \approx \frac{1}{3}\pi^2\sigma_s^4(1-g)^2 + \frac{\pi^4\sigma_s^4}{9}(g+2)var(\hat{D}_G) \quad \text{(C.5)}$$

$$\lim_{k\to\infty} E\{T_2^2\} \approx \frac{1}{3}\pi^2\sigma_s^2\sigma_n^2(1-g)^2 + \frac{1}{9}\pi^4\sigma_s^2\sigma_n^2 var(\hat{D}_G) \quad \text{(C.6)}$$

$$\lim_{k\to\infty} E\{T_3^2\} \approx \frac{1}{3}\pi^2\sigma_s^2\sigma_n^2(1+g^2) \quad \text{(C.7)}$$

$$\lim_{k\to\infty} E\{T_4^2\} \approx \frac{1}{3}\pi^2\sigma_n^4(1+g^2) \quad \text{(C.8)}$$

Substituting (C.4) to (C.8) into (C.3) gives

$$var(\hat{D}_G)\left(\frac{\pi^2}{3}\sigma_s^2 - \frac{\pi^4\mu_D\sigma_s^2}{9}(\sigma_s^2(g+2)+\sigma_n^2)\right) = \frac{\pi^2\mu_D}{3}(\sigma_s^2(\sigma_s^2+\sigma_n^2)(1-g)^2 + \sigma_n^2(\sigma_s^2+\sigma_n^2)(1+g^2) \quad \text{(C.9)}$$

When the adaptation constant μ_D is chosen sufficiently small, (C.9) can be approximated by

$$\begin{aligned} var(\hat{D}_G) &\approx \frac{\mu_D}{\sigma_s^2}\left(\sigma_s^2(\sigma_s^2+\sigma_n^2)(1-g)^2 + \sigma_n^2(\sigma_s^2+\sigma_n^2)(1+g^2)\right) \\ &= \mu_D\sigma_s^2\left(\left(1+\frac{1}{\text{SNR}}\right)(1-g)^2 + \left(\frac{1}{\text{SNR}}+\frac{1}{\text{SNR}^2}\right)(1+g^2)\right) \\ &= \mu_D\sigma_s^2\left(\frac{1}{\text{SNR}(1+\text{SNR})} + \frac{1+2\text{SNR}(1+\text{SNR})}{\text{SNR}^2(1+\text{SNR})}\right) \\ &= \frac{\mu_D\sigma_s^2(1+2\text{SNR})}{\text{SNR}^2} \end{aligned}$$

which is (4.14).

Appendix D

Derivation of (4.15)

The derivation of (4.15) is given as follows. Subtracting g from both sides of (4.3), squaring both sides and then taking the expected value gives

$$\begin{aligned} E\{(\hat{g}(k+1)-g)^2\} &= E\{(\hat{g}(k)-g)^2\} + 4\mu_g E\{(\hat{g}(k)-g)e(k)\tilde{x}(k-\hat{D}_G(k))\} \\ &\quad +4\mu_g^2 E\{e^2(k)\tilde{x}^2(k-\hat{D}_G(k))\} \end{aligned} \tag{D.1}$$

Considering the steady state condition and assuming that $\hat{D}_G(k) \to D$, (D.1) becomes

$$\lim_{k\to\infty} E\{(\hat{g}(k)-g)e(k)\tilde{x}(k-D)\} = -\mu_g \lim_{k\to\infty} E\{e^2(k)\tilde{x}^2(k-D)\} \tag{D.2}$$

Using (4.11), the L.H.S. of (D.2) equals

$$\lim_{k\to\infty} E\{(\hat{g}(k)-g)e(k)\tilde{x}(k-D)\} = -var(\hat{g})(\sigma_s^2+\sigma_n^2) \tag{D.3}$$

where

$$var(\hat{g}) \triangleq \lim_{k\to\infty} E\{(\hat{g}(k)-g)^2\}$$

is the variance of the gain estimate. On the other hand, when P is chosen sufficiently large, the steady state value of $E\{e^2(k)\tilde{x}^2(k-D)\}$ can be computed from

$$\begin{aligned} &\lim_{k\to\infty} E\{e^2(k)\tilde{x}^2(k-D)\} \\ &\quad \approx \lim_{k\to\infty} E\{((1-\hat{g}(k))s(k-D)+n_2(k)-\hat{g}(k)n_1(k-D))^2 \cdot (s(k-D) \\ &\qquad +n_1(k-D))^2\} \\ &\quad = \lim_{k\to\infty} E\{(1-\hat{g}(k))^2 s^4(k-D)+(1-\hat{g}(k))^2 s^2(k-D)n_1^2(k-D) \end{aligned}$$

$$
\begin{aligned}
&+s^2(k-D)n_2^2(k)+n_1^2(k-D)n_2^2(k)+\hat{g}^2(k)s^2(k-D)n_1^2(k-D)\\
&+\hat{g}^2(k)n_1^4(k-D)-4\hat{g}(k)(1-\hat{g}(k))s^2(k-D)n_1^2(k-D)\}\\
&=\lim_{k\to\infty}E\{3\sigma_s^4(1-\hat{g}(k))^2+\sigma_s^2\sigma_n^2((1-\hat{g}(k))^2+1+\hat{g}^2(k)\\
&-4\hat{g}(k)(1-\hat{g}(k)))+\sigma_n^4(1+3\hat{g}^2(k))\}\\
&=3\sigma_s^4(1-2g+g^2+var(\hat{g}))+\sigma_s^2\sigma_n^2(2-6g+6g^2+6var(\hat{g}))\\
&+\sigma_n^4(1+3g^2+3var(\hat{g}))
\end{aligned}
\tag{D.4}
$$

Substituting (D.3) and (D.4) into (D.2) and considering μ_g is chosen sufficiently small, the variance of $\hat{g}(k)$ has a value of

$$
\begin{aligned}
var(\hat{g}) &\approx \frac{\mu_g}{\sigma_s^2+\sigma_n^2}(3\sigma_s^4(1-2g+g^2)+\sigma_s^2\sigma_n^2(2-6g+6g^2)+\sigma_n^4(1+3g^2))\\
&= \mu_g\sigma_n^2\left(\frac{3\mathrm{SNR}^2+\mathrm{SNR}(2(1+\mathrm{SNR})^2-6\mathrm{SNR})+1+2\mathrm{SNR}+4\mathrm{SNR}^2}{(1+\mathrm{SNR})^3}\right)\\
&= \mu_g\sigma_n^2\left(\frac{1+4\mathrm{SNR}+5\mathrm{SNR}^2+2\mathrm{SNR}^3}{(1+\mathrm{SNR})^3}\right)
\end{aligned}
\tag{D.5}
$$

As a result, for the cases of high SNR and low SNR, the values of $var(\hat{g})$ are given by

$$
var(\hat{g})=\begin{cases}2\mu_g\sigma_n^2 & , \quad \mathrm{SNR}\gg 1\\ \\ \mu_g\sigma_n^2 & , \quad \mathrm{SNR}\ll 1\end{cases}
$$

Appendix E

Derivation of (5.21)

The derivation of (5.21) is shown as follows. Recall the updating rule for the gain estimate as given by (5.17)

$$a_{\Delta_{11}}(k+1) = a_{\Delta_{11}}(k) + 2\mu_\alpha x'(k)x'(k-\Delta_{11}) \tag{E.1}$$

Subtracting α_{11} from both sides of (E.1), squaring the results, and then taking expectation yields

$$\begin{aligned}\epsilon_a(k+1) &= \epsilon_a(k) + 4\mu_\alpha E\{(a_{\Delta_{11}}(k)-\alpha_{11})(x(k)-a_{\Delta_{11}}(k)x'(k-\Delta_{11}) \\ &\quad \cdot x'(k-\Delta_{11})\} + 4\mu_\alpha^2 E\{x'^2(k)x'^2(k-\Delta_{11})\}\end{aligned} \tag{E.2}$$

where

$$\epsilon_a(k) \triangleq E\{(a_{\Delta_{11}}(k)-\alpha_{11})^2\}$$

As $x'(k) \to s(k)$, (E.2) can be approximated by

$$\begin{aligned}\epsilon_a(k+1) &= \epsilon_a(k) + 4\mu_\alpha E\{(a_{\Delta_{11}}(k)-\alpha_{11})(s(k)+\alpha_{11}s(k-\Delta_{11}) \\ &\quad -a_{\Delta_{11}}(k)s(k-\Delta_{11}))s(k-\Delta_{11})\} + 4\mu_\alpha^2 E\{s^2(k)s^2(k-\Delta_{11})\} \\ &= \epsilon_a(k) + 4\mu_\alpha\sigma_s^2 E\{-(a_{\Delta_{11}}(k)-\alpha_{11})^2\} + 4\mu_\alpha^2\sigma_s^4\end{aligned} \tag{E.3}$$

Solving (E.3) gives

$$\epsilon_a(k) = \mu_\alpha\sigma_s^2 + ((a_{\Delta_{11}}(0)-\alpha_{11})^2 - \mu_\alpha\sigma_s^2)(1-4\mu_\alpha\sigma_s^2)^k$$

Appendix F

Proof of unstablity of $A^o(z)$

We attempt to show here that the optimal multipath canceller $A^o(z)$ as given by (5.29) is an unstable system by applying Rouché's theorem [82] which can be stated as follows.

Let u and v be functions which are analytic inside and on a positively oriented simple closed contour C. If $|u(z)| > |v(z)|$ at each point z on C, the functions $u(z)$ and $u(z)+v(z)$ have the same number of zeros, counting multiplicities, inside C.

Based on this theorem, two functions, $u(z)$ and $v(z)$, which are analytic on the contour $|z| \leq 1$, are defined as

$$u(z) = z^M$$

and

$$v(z) = \alpha_{11} \sum_{i=-M}^{M} sinc(i - \Delta_{11}) z^{M-i}$$

such that $(u(z)+v(z))z^{-M}$ equals the demoninator of $A^o(z)$. We then evaluate the magnitudes of $u(z)$ and $v(z)$ when $|z| = 1$,

$$|u(z)| = |z|^M = 1$$

and

$$\begin{aligned} |v(z)| &= |\alpha_{11} \sum_{i=-M}^{M} sinc(i - \Delta_{11}) z^{M-i}| \\ &= |\alpha_{11}| \cdot |\sum_{i=-M}^{M} sinc(i - \Delta_{11}) z^{-i}| \cdot |z^M| \\ &\leq \alpha_{11} \cdot |\sum_{i=-M}^{M} sinc(i - \Delta_{11})| \end{aligned}$$

From [34], the discrete Fourier series representation of $e^{j\omega\Delta}$ is given by

$$e^{j\omega\Delta} = \sum_{i=-\infty}^{\infty} sinc(\Delta + i)e^{-j\omega i} \tag{F.1}$$

where ω represents the frequency and Δ is a real number. Setting $\omega = 0$ in (F.1), we obtain

$$\sum_{i=-\infty}^{\infty} sinc(\Delta + i) = 1$$

Thus for sufficiently large M, the value of $v(z)$ on the contour $|z| = 1$ is bounded by

$$|v(z)| \leq \alpha_{11}$$

It can be seen that $|u(z)| > |v(z)|$ for $|z| = 1$. Using Rouché's theorem, since $u(z)$ has M zeros interior to the unit circle, so does $u(z)+v(z)$. This implies that there are also M zeros, which are outside or on the unit circle, in $u(z) + v(z)$. Consequently, $A^o(z)$ only has M poles lying inside the unit circle and thus its unstability is proved. It is noteworthy that, in a similar manner, the causal IIR filter

$$A^s(z) = \frac{1}{1 + \alpha_{11} \sum_{i=1}^{M} sinc(i - \Delta_{11})z^{-i}}$$

can be shown to have M zeros which implies its stability.

Appendix G

Derivation of (5.34)-(5.35)

In this appendix we shall derive equations (5.34) and (5.35). Substituting (5.33) into (5.32), and then taking the expected value gives

$$\begin{aligned} E\{\hat{\Delta}_{11}^s(k+1)\} &= E\{\hat{\Delta}_{11}^s(k)\} - 2\mu_\Delta E\{(s(k) + \hat{\alpha}_{11}^s(k)s(k-\hat{\Delta}_{11}^s(k)) - \hat{\alpha}_{11}^s(k) \\ &\quad \cdot \sum_{j=1}^{M} sinc(j-\hat{\Delta}_{11}^s(k))x'(k-j)) \cdot \sum_{i=1}^{M} f(i-\hat{\Delta}_{11}^s(k))x'(k-i)\} \end{aligned} \tag{G.1}$$

Assume that $x'(k)$ is reasonably close to $s(k)$ and $M \gg \hat{\Delta}_{11}^s(k) \gg 1$. Then, (G.1) can be simplified as

$$\begin{aligned} E\{\hat{\Delta}_{11}^s(k+1)\} &\approx E\{\hat{\Delta}_{11}^s(k)\} - 2\mu_\Delta E\{(s(k) + \alpha_{11}s(k-\Delta_{11}) - \hat{\alpha}_{11}^s(k) \\ &\quad \cdot \sum_{j=1}^{M} sinc(j-\hat{\Delta}_{11}^s(k))s(k-j)) \cdot \sum_{i=1}^{M} f(i-\hat{\Delta}_{11}^s(k))s(k-i)\} \\ &= E\{\hat{\Delta}_{11}^s(k)\} - 2\mu_\Delta E\{(\alpha_{11} \sum_{j=-\infty}^{\infty} sinc(j-\Delta_{11})s(k-j) - \hat{\alpha}_{11}^s(k) \\ &\quad \cdot \sum_{j=1}^{M} sinc(j-\hat{\Delta}_{11}^s(k))s(k-j)) \cdot \sum_{i=1}^{M} f(i-\hat{\Delta}_{11}^s(k))s(k-i) \\ &= E\{\hat{\Delta}_{11}^s(k)\} - 2\mu_\Delta\alpha_{11}\sigma_s^2 E\{\sum_{i=1}^{M} sinc(i-\Delta_{11})f(i-\hat{\Delta}_{11}^s(k))\} \\ &\quad +2\mu_\Delta\sigma_s^2 E\{\hat{\alpha}_{11}^s(k) \cdot \sum_{i=1}^{M} sinc(i-\hat{\Delta}_{11}^s(k))f(i-\hat{\Delta}_{11}^s(k))\} \\ &\approx E\{\hat{\Delta}_{11}^s(k)\} - 2\mu_\Delta\alpha_{11}\sigma_s^2 E\{f(\Delta_{11}-\hat{\Delta}_{11}^s(k))\} \\ &\quad +2\mu_\Delta\sigma_s^2 E\{\hat{\alpha}_{11}^s(k) \cdot f(0)\} \end{aligned} \tag{G.2}$$

Since $f(0) = 0$ and using the first order approximation of $f(\Delta_{11} - \hat{\Delta}^s_{11}(k))$, (G.2) becomes

$$\begin{aligned} E\{\hat{\Delta}^s_{11}(k+1)\} &\approx \hat{\Delta}^s_{11}(k)\} - 2\mu_\Delta \alpha_{11}\sigma_s^2 E\left\{\left(f(0) - \frac{\pi^2}{3}(\Delta_{11} - \hat{\Delta}^s_{11}(k))\right)\right\} \\ &= E\{\hat{\Delta}^s_{11}(k)\}\left(1 - \frac{2}{3}\mu_\Delta \alpha_{11}\sigma_s^2\pi^2\right) + \frac{2}{3}\mu_\Delta \alpha_{11}\sigma_s^2\pi^2\Delta_{11} \end{aligned} \quad (G.3)$$

Solving (G.3) gives the learning trajectory of the multipath delay estimate which is of the form

$$E\{\hat{\Delta}^s_{11}(k)\} \approx \Delta_{11} + (\hat{\Delta}^s_{11}(0) - \Delta_{11})\left(1 - \frac{2}{3}\mu_\Delta \alpha_{11}\sigma_s^2\pi^2\right)^k \quad (G.4)$$

As $\hat{\Delta}^s_{11}(k) \to \Delta_{11}$ and using previous assumptions, the expected value of (5.31) can be approximated by

$$\begin{aligned} E\{\hat{\alpha}^s_{11}(k+1)\} &\approx E\{\hat{\alpha}^s_{11}(k)\} + 2\mu_\alpha E\{(s(k) + (\alpha_{11} - \hat{\alpha}^s_{11}(k))s(k - \Delta_{11})) \\ &\quad \cdot \sum_{i=1}^{M} sinc(i - \Delta_{11})s(k-i)\} \\ &= E\{\hat{\alpha}^s_{11}(k)\} + 2\mu_\alpha \sigma_s^2 E\{(\alpha_{11} - \hat{\alpha}^s_{11}(k))\sum_{i=1}^{M} sinc^2(i - \Delta_{11})\} \end{aligned} \quad (G.5)$$

Since

$$\sum_{i=-\infty}^{\infty} sinc^2(i - \Delta_{11}) = sinc(0) = 1$$

(G.5) becomes

$$E\{\hat{\alpha}^s_{11}(k+1)\} \approx E\{\hat{\alpha}^s_{11}(k)\} + 2\mu_\alpha \sigma_s^2 E\{\alpha_{11} - \hat{\alpha}^s_{11}(k)\} \quad (G.6)$$

and solving (G.6) will give the following solution,

$$E\{\hat{\alpha}^s_{11}(k)\} \approx \alpha_{11} + (\hat{\alpha}^s_{11}(0) - \alpha_{11})(1 - 2\mu_\alpha \sigma_s^2)^k$$

Hence the proofs of (5.34) and (5.35) are completed.

Appendix H

Derivation of variance of $\hat{\alpha}_{11}^{s}(k)$ and $\hat{\Delta}_{11}^{s}(k)$

We attempt to find the variances of the multipath gain and multipath delay estimates for non-integral interpath delays analytically. Assuming that $M \gg \hat{\Delta}_{11}^{s}(k) \gg 1$, (5.31) can be approximated by

$$\hat{\alpha}_{11}^{s}(k+1) = \hat{\alpha}_{11}^{s}(k) + 2\mu_{\alpha} x'(k) x'(k - \hat{\Delta}_{11}^{s}(k))$$

Denote

$$\epsilon_{\alpha}(k) = E\{(\hat{\alpha}_{11}^{s}(k) - \alpha_{11})^2\}$$

Then we get

$$\begin{aligned} \epsilon_{\alpha}(k+1) &= \epsilon_{\alpha}(k) + 4\mu_{\alpha} E\{(\hat{\alpha}_{11}^{s}(k) - \alpha_{11}) x'(k) x'(k - \hat{\Delta}_{11}^{s}(k))\} \\ &\quad + 4\mu_{\alpha}^2 E\{x'^2(k) x'^2(k - \hat{\Delta}_{11}^{s}(k))\} \end{aligned} \tag{H.1}$$

By considering the steady state condition when $\hat{\Delta}_{11}^{s}(k) \to \Delta_{11}$ and $x'(k)$ is sufficiently close to $s(k)$, (H.1) becomes

$$\begin{aligned} \epsilon_{\alpha}(k+1) &\approx \epsilon_{\alpha}(k) + 4\mu_{\alpha} E\{(\hat{\alpha}_{11}^{s}(k) - \alpha_{11})(s(k) + \alpha_{11} s(k - \Delta_{11}) \\ &\quad - \hat{\alpha}_{11}^{s}(k) s(k - \Delta_{11})) s(k - \Delta_{11})\} + 4\mu_{\alpha}^2 E\{s^2(k) s^2(k - \Delta_{11})\} \\ &\approx \epsilon_{\alpha}(k) - 4\mu_{\alpha} E\{(\hat{\alpha}_{11}^{s}(k) - \alpha_{11})^2\} E\{s^2(k - \Delta_{11})\} \\ &\quad + 4\mu_{\alpha}^2 E\{s^2(k)\} E\{s^2(k - \Delta_{11})\} \\ &= \epsilon_{\alpha}(k) - 4\mu_{\alpha}\sigma_s^2 \epsilon_{\alpha}(k) + 4\mu_{\alpha}^2 \sigma_s^4 \end{aligned} \tag{H.2}$$

As k goes to infinity, the variance of $\hat{\alpha}_{11}^s(k)$, denoted by $var(\hat{\alpha}_{11}^s)$, can be obtained from (H.2),

$$\begin{aligned} var(\hat{\alpha}_{11}^s) &\triangleq \lim_{k\to\infty} \epsilon_\alpha(k) \\ &= \mu_\alpha \sigma_s^2 \end{aligned}$$

which is identical to (5.23).

From (5.32), the updating equation for the multipath delay estimate $\hat{\Delta}_{11}^s(k)$ is given by

$$\hat{\Delta}_{11}^s(k+1) = \hat{\Delta}_{11}^s(k) - 2\mu_\Delta x'(k) \sum_{i=1}^{M} f(i - \hat{\Delta}_{11}^s(k)) x'(k-i) \tag{H.3}$$

Subtracting Δ_{11} from both sides of (H.3), squaring the results and then taking the expected value, we have

$$\begin{aligned} \epsilon_\Delta(k+1) &= \epsilon_\Delta(k) - 4\mu_\Delta E\{(\hat{\Delta}_{11}^s(k) - \Delta_{11}) x'(k) \sum_{i=1}^{M} f(i - \hat{\Delta}_{11}^s(k)) x'(k-i)\} \\ &\quad + 4\mu_\Delta^2 E\{(x'(k) \sum_{i=1}^{M} f(i - \hat{\Delta}_{11}^s(k)) x'(k-i))^2\} \end{aligned} \tag{H.4}$$

where

$$\epsilon_\Delta(k) = E\{(\hat{\Delta}_{11}^s(k) - \Delta_{11})^2\}$$

Again, we use the assumptions of $x'(k) \approx s(k)$ and $M \gg \hat{\Delta}_{11}^s(k) \gg 1$, therefore (H.4) can be approximated by

$$\begin{aligned} \epsilon_\Delta(k+1) &\approx \epsilon_\Delta(k) - 4\mu_\Delta E\{(\hat{\Delta}_{11}^s(k) - \Delta_{11})(s(k) + \alpha_{11} s(k - \Delta_{11}) \\ &\quad - \hat{\alpha}_{11}^s(k) s(k - \hat{\Delta}_{11}^s(k))) \sum_{i=1}^{M} f(i - \hat{\Delta}_{11}^s(k)) s(k-i)\} \\ &\quad + 4\mu_\Delta^2 E\{(s(k) \sum_{i=1}^{M} f(i - \hat{\Delta}_{11}^s(k)) s(k-i))^2\} \\ &\approx \epsilon_\Delta(k) - 4\mu_\Delta E\{(\hat{\Delta}_{11}^s(k) - \Delta_{11})(\alpha_{11} \sum_{i=-\infty}^{\infty} sinc(j - \Delta_{11}) s(k-j) \\ &\quad - \hat{\alpha}_{11}^s(k) \cdot \sum_{i=-\infty}^{\infty} sinc(j - \hat{\Delta}_{11}^s(k)) s(k-j)) \cdot \sum_{i=1}^{M} f(i - \hat{\Delta}_{11}^s(k)) \\ &\quad \cdot sinc(i - \hat{\Delta}_{11}^s(k))) + 4\mu_\Delta^2 E\{s^2(k)\} E\{(\sum_{i=1}^{M} f(i - \hat{\Delta}_{11}^s(k)) s(k-i))^2\} \end{aligned}$$

$$
\begin{aligned}
&\approx \epsilon_\Delta(k) - 4\mu_\Delta\sigma_s^2 E\{(\hat{\Delta}_{11}^s(k) - \Delta_{11})(\alpha_{11} f(\Delta_{11} - \hat{\Delta}_{11}^s(k)) \\
&\quad -\hat{\alpha}_{11}^s(k) f(0)\} + 4\mu_\Delta^2\sigma_s^4 \sum_{i=1}^{M} E\{f^2(i - \hat{\Delta}_{11}^s(k))\} \\
&\approx \epsilon_\Delta(k) - 4\mu_\Delta\alpha_{11}\sigma_s^2 E\{(\hat{\Delta}_{11}^s(k) - \Delta_{11}) f(\Delta_{11} - \hat{\Delta}_{11}^s(k))\} \\
&\quad +4\mu_\Delta^2\sigma_s^4 \sum_{i=-\infty}^{\infty} E\{f^2(i - \hat{\Delta}_{11}^s(k))\} \qquad \text{(H.5)}
\end{aligned}
$$

Since [58]

$$
\sum_{i=-\infty}^{\infty} f^2(i) = \frac{\pi^2}{3}
$$

and using the first order approximation of the f function, (H.5) becomes

$$
\epsilon_\Delta(k+1) \approx \epsilon_\Delta(k) - \frac{4}{3}\mu_\Delta\alpha_{11}\sigma_s^2\pi^2\epsilon_\Delta(k) + \frac{4}{3}\mu_\Delta^2\sigma_s^4\pi^2 \qquad \text{(H.6)}
$$

The variance of the multipath delay estimate, $var(\hat{\Delta}_{11}^s)$, is then given by the steady state value of (H.6), which is,

$$
\begin{aligned}
var(\hat{\Delta}_{11}^s) &\triangleq \lim_{k\to\infty} \epsilon_\Delta(k) \\
&= \frac{\mu_\Delta\sigma_s^2}{\alpha_{11}}
\end{aligned}
$$

Appendix I

Derivation of (5.40)

The performance surface of $A(z)$ for integral multipath delays in the presence of noise is derived here. Assume that the multipath delay Δ_{11} has been correctly determined and $A(z)$ is adapted according to (5.17). The output of $A(z)$, $x'(k)$, is now expressed as

$$x'(k) = s(k) + \alpha_{11}s(k-\Delta_{11}) + n_1(k) - a_{\Delta_{11}}x'(k-\Delta_{11}) \tag{I.1}$$

where the time index for $a_{\Delta_{11}}$ is dropped for convenience. Squaring both sides of (I.1) and then taking the expected value, we obtain

$$\begin{aligned} E\{x'^2(k)\} &= (1+\alpha_{11}^2)\sigma_s^2 + \sigma_n^2 + a_{\Delta_{11}}^2 E\{x'^2(k)\} \\ &\quad -2a_{\Delta_{11}}E\{(s(k)+\alpha_{11}s(k-\Delta_{11})+n_1(k))x'(k-\Delta_{11})\} \end{aligned} \tag{I.2}$$

Using (I.1), we find that $x'(k-\Delta_{11})$ is uncorrelated with $s(k)$ and $n_1(k)$ while $E\{s(k-\Delta_{11})x'(k-\Delta_{11})\} = \sigma_s^2$. As a result, (I.2) becomes

$$E\{x'^2(k)\} = (1+\alpha_{11}^2)\sigma_s^2 + \sigma_n^2 + a_{\Delta_{11}}^2 E\{x'^2(k)\} - 2a_{\Delta_{11}}\alpha_{11}\sigma_s^2$$

$$\Rightarrow E\{x'^2(k)\} = \frac{(1+\alpha_{11}^2 - 2\alpha_{11}a_{\Delta_{11}})\sigma_s^2 + \sigma_n^2}{1-a_{\Delta_{11}}^2}$$

which is (5.40).

Appendix J

Derivation of time constant of $a_{\Delta_{11}}(k)$

We shall prove that the time constant of the gain estimate in the presence of noise is given by $1/(2\mu_\alpha(\sigma_s^2+\sigma_n^2))$. Substituting (5.2a) into (5.17) and taking expectation on both sides, we obtain

$$\begin{aligned} E\{a_{\Delta_{11}}(k+1)\} &= E\{a_{\Delta_{11}}(k)\} + 2\mu_\alpha E\{(s(k)+\alpha_{11}s(k-\Delta_{11})+n_1(k) \\ &\quad -a_{\Delta_{11}}(k)x'(k-\Delta_{11}))x'(k-\Delta_{11})\} \end{aligned} \tag{J.1}$$

Using the approximate value of $x'(k)$ in (5.43), (J.1) becomes

$$\begin{aligned} E\{a_{\Delta_{11}}(k+1)\} &= E\{a_{\Delta_{11}}(k)\} + 2\mu_\alpha E\{(s(k)+\alpha_{11}s(k-\Delta_{11})+n_1(k) \\ &\quad -a_{\Delta_{11}}(k)(s(k-\Delta_{11})+(\alpha_{11}-a_{\Delta_{11}}(k-\Delta_{11}))s(k-2\Delta_{11}) \\ &\quad +n_1(k-\Delta_{11})))(s(k-\Delta_{11})+(\alpha_{11}-a_{\Delta_{11}}(k-\Delta_{11})) \\ &\quad \cdot s(k-2\Delta_{11})+n_1(k-\Delta_{11}))\} \\ &\approx E\{a_{\Delta_{11}}(k)\} + 2\mu_\alpha E\{(\alpha_{11}-a_{\Delta_{11}}(k))s^2(k-\Delta_{11}) \\ &\quad -a_{\Delta_{11}}(k)n^2(k-\Delta_{11})\} \\ &= E\{a_{\Delta_{11}}(k)\}(1-2\mu_\alpha(\sigma_s^2+\sigma_n^2)) + 2\mu_\alpha\alpha_{11}\sigma_s^2 \\ &= \alpha_{11}\frac{\sigma_s^2}{\sigma_s^2+\sigma_n^2} + (a_{\Delta_{11}}(0)-\alpha_{11}\frac{\sigma_s^2}{\sigma_s^2+\sigma_n^2})(1-2\mu_\alpha(\sigma_s^2+\sigma_n^2))^{k+1} \end{aligned} \tag{J.2}$$

Obviously, the time constant of $a_{\Delta_{11}}(k)$ is $1/(2\mu_\alpha(\sigma_s^2+\sigma_n^2))$. Notice that from (J.2), the gain estimate converges to $\alpha_{11}\sigma_s^2/(\sigma_s^2+\sigma_n^2)$ which is close to the exact value as given by (5.42).

Appendix K

Derivation of (5.63)-(5.66)

The convergence behaviours of the multipath parameter estimates are examined here. In the following derivations, it is assumed [57] that $\hat{\alpha}_{11}(k)$, $\hat{\Delta}_{11}(k)$, $\hat{\alpha}_{21}(k)$, $\hat{\Delta}_{21}(k)$ and $\hat{D}_{ME}(k)$ are uncorrelated with the received signals $x(k)$ and $y(k)$. From (5.53), the updating rule for $\hat{\alpha}_{11}(k)$ is of the form

$$\hat{\alpha}_{11}(k+1) = \hat{\alpha}_{11}(k) - 2\mu_{\alpha} e(k) \sum_{n=-P}^{P} sinc(n - \hat{\Delta}_{11}(k)) y(k-n) \qquad \text{(K.1)}$$

When $M_1 = M_2 = 1$ and if P is chosen sufficiently large, the error function $e(k)$ is expressed as

$$\begin{aligned} e(k) &= y(k) + \hat{\alpha}_{11}(k) \sum_{n=-P}^{P} sinc(n - \hat{\Delta}_{11}(k)) y(k-n) \\ &\quad - \sum_{n=-P}^{P} sinc(n - \hat{D}_{ME}(k)) x(k-n) - \hat{\alpha}_{21}(k) \sum_{n=-P}^{P} sinc(n - \hat{\Delta}_{21}(k)) x(k-n) \\ &\approx y(k) + \hat{\alpha}_{11}(k) y(k - \hat{\Delta}_{11}(k)) - x(k - \hat{D}_{ME}(k)) - \hat{\alpha}_{21}(k) x(k - \hat{\Delta}_{21}(k)) \end{aligned}$$

where

$$x(k) = s(k) + \alpha_{11} s(k - \Delta_{11}) + n_1(k)$$

and

$$y(k) = s(k - D) + \alpha_{21} s(k - \Delta_{21}) + n_2(k)$$

Assume that the other parameters have converged to their optimal values, that is, $\hat{\Delta}_{11}(k) \to \Delta_{11}$, $\hat{\alpha}_{21}(k) \to \alpha_{21}$, $\hat{\Delta}_{21}(k) \to \Delta_{21}$ and $\hat{D}_{ME}(k) \to D$. In this case, $e(k)$ can be simplified to

$$e(k) = y(k) + \hat{\alpha}_{11}(k) y(k - \Delta_{11}) - x(k - D) - \alpha_{21} x(k - \Delta_{21})$$

$$= (\hat{\alpha}_{11}(k) - \alpha_{11})y(k - \Delta_{11}) + n_2(k) + \alpha_{11}n_2(k - \Delta_{11}) - n_1(k - D)$$
$$-\alpha_{21}n_1(k - \Delta_{21}) \qquad \text{(K.2)}$$

Using (K.2) and

$$\sum_{n=-\infty}^{\infty} sinc(n - \Delta_{11})y(k - n) = y(k - \Delta_{11})$$

the expected value of (K.1) is approximated by

$$\begin{aligned} E\{\hat{\alpha}_{11}(k+1)\} &\approx E\{\hat{\alpha}_{11}(k)\} - 2\mu_\alpha E\{((\hat{\alpha}_{11}(k) - \alpha_{11})y(k - \Delta_{11}) + n_2(k) \\ &\quad +\alpha_{11}n_2(k - \Delta_{11}) - n_1(k - D) - \alpha_{21}n_1(k - \Delta_{21}))y(k - \Delta_{11})\} \\ &= E\{\hat{\alpha}_{11}(k)\} - 2\mu_\alpha E\{(\hat{\alpha}_{11}(k) - \alpha_{11})\} \cdot E\{y^2(k - \Delta_{11})\} \\ &\quad -2\mu_\alpha E\{n_2(k) \cdot \sum_{i=-\infty}^{\infty} n_2(k - i)sinc(i - \Delta_{11}) + n_2^2(k - \Delta_{11})\} \\ &= E\{\hat{\alpha}_{11}(k)\} - 2\mu_\alpha\sigma_y^2 E\{(\hat{\alpha}_{11}(k) - \alpha_{11})\} \\ &\quad -2\mu_\alpha\sigma_n^2(sinc(\Delta_{11}) + \alpha_{11}) \qquad \text{(K.3)} \end{aligned}$$

where

$$\sigma_y^2 = \sigma_s^2(1 + \alpha_{21}^2 + 2\alpha_{21}sinc(\Delta_{21} - D)) + \sigma_n^2$$

is the power of $y(k)$. For large Δ_{11} and when $\sigma_s^2 \gg \sigma_n^2$, we may ignore the term $2\mu_\alpha\sigma_n^2 sinc(\Delta_{11})$. As a result, (K.3) becomes

$$E\{\hat{\alpha}_{11}(k+1)\} = E\{\hat{\alpha}_{11}(k)\}(1 - 2\mu_\alpha\sigma_y^2) + 2\mu_\alpha\alpha_{11}(\sigma_y^2 - \sigma_n^2) \qquad \text{(K.4)}$$

Solving (K.4) yields the learning trajectory of $\hat{\alpha}_{11}(k)$,

$$E\{\hat{\alpha}_{11}(k)\} = \frac{\alpha_{11}(\sigma_y^2 - \sigma_n^2)}{\sigma_y^2} + \Big(\hat{\alpha}_{11}(0) - \alpha_{11}\frac{(\sigma_y^2 - \sigma_n^2)}{\sigma_y^2}\Big)(1 - 2\mu_\alpha\sigma_y^2)^k$$

From (5.54), the iterative equation for $\hat{\Delta}_{11}(k)$ is of the form

$$\hat{\Delta}_{11}(k+1) = \hat{\Delta}_{11}(k) + 2\mu_\Delta e(k) \sum_{n=-P}^{P} f(n - \hat{\Delta}_{11}(k))y(k - n) \qquad \text{(K.5)}$$

If we assume $\hat{\alpha}_{11}(k) \to \alpha_{11}$, $\hat{\alpha}_{21}(k) \to \alpha_{21}$, $\hat{\Delta}_{21}(k) \to \Delta_{21}$ and $\hat{D}_{ME}(k) \to D$, then the value of $e(k)$ can be approximated by

$$\begin{aligned} e(k) &\approx \alpha_{11}s(k - D - \hat{\Delta}_{11}(k)) + \alpha_{11}\alpha_{21}s(k - \Delta_{21} - \hat{\Delta}_{11}(k)) + \alpha_{11}n_2(k - \hat{\Delta}_{11}(k)) \\ &\quad -\alpha_{11}s(k - D - \Delta_{11}) - \alpha_{11}\alpha_{21}s(k - \Delta_{21} - \Delta_{11}) + n_2(k) - n_1(k - D) \\ &\quad -\alpha_{21}n_1(k - \Delta_{21}) \qquad \text{(K.6)} \end{aligned}$$

Substituting (K.6) into (K.5) and taking the expected value, we get

$$
\begin{aligned}
&E\{\hat{\Delta}_{11}(k+1)\} \\
&= E\{\hat{\Delta}_{11}(k)\} + 2\mu_\Delta E\{(\alpha_{11}s(k-D-\hat{\Delta}_{11}(k)) + \alpha_{11}\alpha_{21}s(k-\Delta_{21}-\hat{\Delta}_{11}(k)) \\
&\quad +\alpha_{11}n_2(k-\hat{\Delta}_{11}(k)) - \alpha_{11}s(k-D-\Delta_{11}) - \alpha_{11}\alpha_{21}s(k-\Delta_{21}-\Delta_{11}) \\
&\quad +n_2(k) - n_1(k-D) - \alpha_{21}n_1(k-\Delta_{21}))\cdot \sum_{n=-P}^{P} f(n-\hat{\Delta}_{11}(k)) \\
&\quad \cdot(s(k-D-n) + \alpha_{21}s(k-\Delta_{21}-n) + n_2(k-n))\} \\
&= E\{\hat{\Delta}_{11}(k)\} + 2\mu_\Delta E\{(\sum_{i=-\infty}^{\infty} s(k-i)\cdot(\alpha_{11}sinc(i-D-\hat{\Delta}_{11}(k)) \\
&\quad +\alpha_{11}\alpha_{21}sinc(i-\Delta_{21}-\hat{\Delta}_{11}(k)) - \alpha_{11}sinc(i-D-\Delta_{11}) \\
&\quad -\alpha_{11}\alpha_{21}sinc(i-\Delta_{21}-\Delta_{11})) + \alpha_{11}\sum_{i=-\infty}^{\infty} n_2(k-i)sinc(i-\hat{\Delta}_{11}(k)) + n_2(k)) \\
&\quad \cdot \sum_{n=-P}^{P} f(n-\hat{\Delta}_{11}(k))\cdot(\sum_{j=-\infty}^{\infty} s(k-j)\cdot(sinc(j-D-n) \\
&\quad +\alpha_{21}sinc(j-\Delta_{21}-n)) + n_2(k-n)) \qquad \text{(K.7)}
\end{aligned}
$$

Since the signal source and the corrupting noises are white and

$$\sum_{i=\infty}^{\infty} sinc(i-a)sinc(i-b) = sinc(b-a)$$

(K.7) becomes

$$
\begin{aligned}
&E\{\hat{\Delta}_{11}(k+1)\} \\
&= E\{\hat{\Delta}_{11}(k)\} + 2\mu_\Delta\sigma_s^2 E\{\sum_{n=-P}^{P} f(n-\hat{\Delta}_{11}(k))\cdot(\alpha_{11}sinc(n-\hat{\Delta}_{11}(k)) \\
&\quad +\alpha_{11}\alpha_{21}sinc(n+\Delta_{21}-D-\hat{\Delta}_{11}(k)) + \alpha_{11}\alpha_{21}sinc(n+D-\Delta_{21}-\hat{\Delta}_{11}(k)) \\
&\quad +\alpha_{11}\alpha_{21}^2 sinc(n-\hat{\Delta}_{11}(k)) - \alpha_{11}sinc(n-\Delta_{11}) - \alpha_{11}\alpha_{21}sinc(n+\Delta_{21}-D \\
&\quad -\Delta_{11}) - \alpha_{11}\alpha_{21}sinc(n+D-\Delta_{21}-\Delta_{11}) - \alpha_{11}\alpha_{21}^2 sinc(n-\Delta_{11}))\} \\
&\quad +2\mu_\Delta\sigma_n^2 E\{\alpha_{11}\sum_{n=-P}^{P} f(n-\hat{\Delta}_{11}(k))sinc(n-\hat{\Delta}_{11}(k)) + f(-\hat{\Delta}_{11}(k))\} \\
&\approx E\{\hat{\Delta}_{11}(k)\} + 2\mu_\Delta\sigma_s^2 E\{\alpha_{11}f(0) + \alpha_{11}\alpha_{21}f(D-\Delta_{21}) + \alpha_{11}\alpha_{21}f(\Delta_{21}-D) \\
&\quad +\alpha_{11}\alpha_{21}^2 f(0) - \alpha_{11}f(\Delta_{11}-\hat{\Delta}_{11}(k)) - \alpha_{11}\alpha_{21}f(D-\Delta_{21}+\Delta_{11}-\hat{\Delta}_{11}(k)) \\
&\quad -\alpha_{11}\alpha_{21}f(\Delta_{21}-D+\Delta_{11}-\hat{\Delta}_{11}(k)) - \alpha_{11}\alpha_{21}^2 f(\Delta_{11}-\hat{\Delta}_{11}(k))\} \\
&\quad +2\mu_\Delta\sigma_s^2 E\{\alpha_{11}f(0) + f(-\hat{\Delta}_{11}(k))\} \qquad \text{(K.8)}
\end{aligned}
$$

Noting that f is an odd function and considering when the SNR is high enough such that the term $2\mu_\Delta \sigma_n^2 f(\hat{\Delta}_{11}(k))$ can be neglected, (K.8) can be simplified to

$$E\{\hat{\Delta}_{11}(k+1)\} = E\{\hat{\Delta}_{11}(k)\} - 2\mu_\Delta \alpha_{11}(1+\alpha_{21}^2)\sigma_s^2 E\{f(\Delta_{11} - \hat{\Delta}_{11}(k))\} \quad \text{(K.9)}$$

Expanding $f(\Delta_{11} - \hat{\Delta}_{11}(k))$ up to the first order term and then solving (K.9) yields

$$E\{\hat{\Delta}_{11}(k)\} \approx \Delta_{11} + (\hat{\Delta}_{11}(0) - \Delta_{11})\Big(1 - \frac{2}{3}\mu_\Delta \alpha_{11}(1+\alpha_{21}^2)\sigma_s^2\pi^2\Big)^k$$

which is identical to (5.64).

In a similar manner, the learning trajectory of $\alpha_{21}(k)$ and $\Delta_{21}(k)$ can be shown to be

$$E\{\hat{\alpha}_{21}(k)\} \approx \frac{\alpha_{21}(\sigma_x^2 - \sigma_n^2)}{\sigma_x^2} + \Big(\hat{\alpha}_{21}(0) - \alpha_{21}\frac{(\sigma_x^2 - \sigma_n^2)}{\sigma_x^2}\Big)(1 - 2\mu_\alpha \sigma_x^2)^k$$

and

$$E\{\hat{\Delta}_{21}(k)\} \approx \Delta_{21} + (\hat{\Delta}_{21}(0) - \Delta_{21})\Big(1 - \frac{2}{3}\mu_\Delta \alpha_{21}(1+\alpha_{11}^2)\sigma_s^2\pi^2\Big)^k$$

Appendix L

Derivation of (5.68)-(5.72)

The mean square errors of the TDOA and multipath parameter estimates are derived as follows. From (5.52), the updating equation for the delay estimate $\hat{D}_{ME}(k)$ is given by

$$\hat{D}_{ME}(k+1) = \hat{D}_{ME}(k) - 2\mu_D e(k) \sum_{n=-P}^{P} x(k-n) f(n - \hat{D}_{ME}(k)) \tag{L.1}$$

Subtracting D from both sides of (L.1), squaring the results and taking the expected value yields

$$\begin{aligned} & E\{(\hat{D}_{ME}(k+1) - D)^2\} \\ &= E\{(\hat{D}_{ME}(k) - D)^2\} - 4\mu_D E\{(\hat{D}_{ME}(k) - D) \cdot e(k) \sum_{n=-P}^{P} x(k-n) \\ &\quad \cdot f(n - \hat{D}_{ME}(k))\} + 4\mu_D^2 E\{e^2(k) \cdot (\sum_{n=-P}^{P} x(k-n) f(n - \hat{D}_{ME}(k)))^2\} \end{aligned} \tag{L.2}$$

We consider the steady state condition of (L.2) to obtain the following equation

$$\begin{aligned} & \lim_{k\to\infty} E\{(\hat{D}_{ME}(k) - D) \cdot e(k) \sum_{n=-P}^{P} x(k-n) f(n - \hat{D}_{ME}(k))\} \\ &\qquad = \mu_D \lim_{k\to\infty} E\{e^2(k) \cdot (\sum_{n=-P}^{P} x(k-n) f(n - \hat{D}_{ME}(k)))^2\} \end{aligned} \tag{L.3}$$

From (5.60), it has been shown that

$$E\{e(k) \sum_{n=-P}^{P} x(k-n) f(n - \hat{D}_{ME}(k))\} \approx \frac{1}{3}(1 + \alpha_{11}^2)\sigma_s^2 \pi^2 E\{\hat{D}_{ME}(k) - D\}$$

Thus, the L.H.S. of (L.3) can be approximated by

$$\lim_{k\to\infty} E\{(\hat{D}_{ME}(k)-D)\cdot e(k)\sum_{n=-P}^{P} x(k-n)f(n-\hat{D}_{ME}(k))\}$$
$$\approx \frac{1}{3}(1+\alpha_{11}^2)\sigma_s^2\pi^2 \lim_{k\to\infty} E\{(\hat{D}_{ME}(k)-D)^2\} \quad \text{(L.4)}$$

On the other hand, since $e(k)$ is dominated by the corrupting noises in steady state while $\sigma_s^2 \gg \sigma_n^2$ in $x(k)$, the R.H.S. of (L.3) is evaluated as

$$\begin{aligned}
&\mu_D \lim_{k\to\infty} E\{e^2(k)\cdot(\sum_{n=-P}^{P} x(k-n)f(n-\hat{D}_{ME}(k)))^2\} \\
&= \mu_D \lim_{k\to\infty} E\{(n_2(k)+\alpha_{11}n_2(k-\Delta_{11})-n_1(k-\hat{D}_{ME}(k))-\alpha_{21}n_1(k-\Delta_{21}))^2 \\
&\quad \cdot(\sum_{n=-P}^{P} f(n-\hat{D}_{ME}(k))(s(k-n)+\alpha_{11}s(k-\Delta_{11}-n))^2\} \\
&= \mu_D \lim_{k\to\infty} E\{(n_2(k)+\alpha_{11}n_2(k-\Delta_{11})-n_1(k-\hat{D}_{ME}(k))-\alpha_{21}n_1(k-\Delta_{21}))^2\} \\
&\quad \cdot E\{\sum_{m=-P}^{P}\sum_{n=-P}^{P} f(m-\hat{D}_{ME}(k))f(n-\hat{D}_{ME}(k))(s(k-m)+\alpha_{11}s(k-m-\Delta_{11})) \\
&\quad \cdot(s(k-n)+\alpha_{11}s(k-n-\Delta_{11}))\} \\
&\approx \mu_D\sigma_n^2(2+\alpha_{11}^2+\alpha_{21}^2+2\alpha_{11}sinc(\Delta_{11})+2\alpha_{21}sinc(\Delta_{21}-D)) \\
&\quad \cdot \lim_{k\to\infty} E\{\sum_{n=-P}^{P} f^2(n-\hat{D}_{ME}(k))\cdot(s^2(k-n)+\alpha_{11}^2s^2(k-n-\Delta_{11}))\} \\
&= \mu_D(1+\alpha_{11}^2)\sigma_s^2\sigma_n^2(2+\alpha_{11}^2+\alpha_{21}^2+2\alpha_{11}sinc(\Delta_{11})+2\alpha_{21}sinc(\Delta_{21}-D)) \\
&\quad \cdot \sum_{n=-P}^{P} f^2(n-\hat{D}_{ME}(k)) \qquad \text{(L.5)}
\end{aligned}$$

Define

$$\sigma_{n'}^2 = \sigma_n^2(2+\alpha_{11}^2+\alpha_{21}^2+2\alpha_{11}sinc(\Delta_{11})+2\alpha_{21}sinc(\Delta_{21}-D))$$

and use [58]

$$\sum_{n=-\infty}^{\infty} f^2(n) = \frac{\pi^2}{3}$$

we can further simplify (L.5) to

$$\mu_D \lim_{k\to\infty} E\{e^2(k)\cdot(\sum_{n=-P}^{P} x(k-n)f(n-\hat{D}_{ME}(k)))^2\} \approx \frac{1}{3}\mu_D(1+\alpha_{11}^2)\sigma_s^2\sigma_{n'}^2\pi^2 \quad \text{(L.6)}$$

Equating (L.4) and (L.6), we get the variance of the TDOA estimate as given by (5.68),

$$\lim_{k\to\infty} E\{(\hat{D}_{ME}(k)-D)^2\} \approx \mu_D\sigma_{n'}^2$$

To derive the mean square error of $\hat{\alpha}_{11}(k)$, we first rewrite (5.69) as

$$\xi(\hat{\alpha}_{11}) = \lim_{k\to\infty} E\{(\hat{\alpha}_{11}(k) - \bar{\alpha}_{11})^2\} + (\bar{\alpha}_{11} - \alpha_{11})^2 \tag{L.7}$$

where

$$\bar{\alpha}_{11} = \frac{\alpha_{11}(\sigma_y^2 - \sigma_n^2)}{\sigma_y^2}$$

which is the steady state mean value of the gain estimate $\hat{\alpha}_{11}(k)$. Notice that

$$(\bar{\alpha}_{11} - \alpha_{11})^2 = \frac{\alpha_{11}^2\sigma_n^4}{\sigma_y^4}$$

Similar to (L.3), a steady state equation can be obtained with the use of (K.1) and is of the form

$$\lim_{k\to\infty} E\{(\hat{\alpha}_{11}(k) - \bar{\alpha}_{11})e(k) \sum_{n=-P}^{P} sinc(n - \hat{\Delta}_{11}(k))y(k-n)\}$$
$$= \mu_\alpha \lim_{k\to\infty} E\{e^2(k)(\sum_{n=-P}^{P} sinc(n - \hat{\Delta}_{11}(k))y(k-n))^2\} \tag{L.8}$$

From (K.4), we have

$$\begin{aligned} E\{e(k) \sum_{n=-P}^{P} sinc(n - \hat{\Delta}_{11}(k))y(k-n)\} &= E\{\sigma_y^2(\hat{\alpha}_{11}(k) - \alpha_{11}) + \sigma_n^2\alpha_{11} \\ &= E\{\sigma_y^2(\hat{\alpha}_{11}(k) - \bar{\alpha}_{11})\} \end{aligned}$$

Thus the L.H.S. of (L.8) becomes

$$\lim_{k\to\infty} E\{(\hat{\alpha}_{11}(k) - \bar{\alpha}_{11})e(k) \sum_{n=-P}^{P} sinc(n - \hat{\Delta}_{11}(k))y(k-n)\}$$
$$= \sigma_y^2 \lim_{k\to\infty} E\{(\hat{\alpha}_{11}(k) - \bar{\alpha}_{11})^2\} \tag{L.9}$$

For sufficiently large P, the R.H.S. of (L.8) can be approximated by

$$\begin{aligned} &\mu_\alpha \lim_{k\to\infty} E\{e^2(k)(\sum_{n=-P}^{P} sinc(n - \hat{\Delta}_{11}(k))y(k-n))^2\} \\ &\approx \mu_\alpha \lim_{k\to\infty} E\{(n_2(k) + \alpha_{11}n_2(k-\Delta_{11}) - n_1(k-D) - \alpha_{21}n_1(k-\Delta_{21})^2\} \\ &\quad \cdot E\{(s(k-D-\Delta_{11}) + \alpha_{21}s(k-\Delta_{21}-\Delta_{11}))^2\} \\ &= \mu_\alpha\sigma_{n'}^2(\sigma_y^2 - \sigma_n^2) \end{aligned} \tag{L.10}$$

Equating (L.9) and (L.10), we get

$$\begin{aligned}\lim_{k\to\infty} E\{(\hat{\alpha}_{11}(k)-\bar{\alpha}_{11})^2\} &\approx \frac{\mu_\alpha \sigma_{n'}^2(\sigma_y^2-\sigma_n^2)}{\sigma_y^2} \\ &\approx \mu_\alpha \sigma_{n'}^2 \end{aligned} \tag{L.11}$$

Substituting (L.11) into (L.7) yields

$$\lim_{k\to\infty} E\{(\hat{\alpha}_{11}(k)-\alpha_{11})^2\} \approx \mu_\alpha \sigma_{n'}^2 + \frac{\alpha_{11}^2\sigma_n^4}{\sigma_y^4}$$

which is identical to (5.69).

From (K.5), we obtain a steady state equation which is given by

$$\begin{aligned}&\lim_{k\to\infty} E\{(\hat{\Delta}_{11}(k)-\Delta_{11})\cdot e(k)\sum_{n=-P}^{P} y(k-n)f(n-\hat{\Delta}_{11}(k))\} \\ &\qquad = \mu_\Delta \lim_{k\to\infty} E\{e^2(k)\cdot(\sum_{n=-P}^{P} y(k-n)f(n-\hat{\Delta}_{11}(k)))^2\}\end{aligned} \tag{L.12}$$

Using (K.9), we have

$$\begin{aligned}&\lim_{k\to\infty} E\{(\hat{\Delta}_{11}(k)-\Delta_{11})\cdot e(k)\sum_{n=-P}^{P} y(k-n)f(n-\hat{\Delta}_{11}(k))\} \\ &= \lim_{k\to\infty} E\{(\hat{\Delta}_{11}(k)-\Delta_{11})\alpha_{11}(1+\alpha_{21}^2)\sigma_s^2 f(\Delta_{11}-\hat{\Delta}_{11}(k))\} \\ &\approx \frac{1}{3}\alpha_{11}(1+\alpha_{21}^2)\sigma_s^2\pi^2 \lim_{k\to\infty} E\{(\hat{\Delta}_{11}(k)-\Delta_{11})^2\}\end{aligned} \tag{L.13}$$

Following (L.5), the R.H.S. of (L.13) can be approximated by

$$\mu_\Delta \lim_{k\to\infty} E\{e^2(k)\cdot(\sum_{n=-P}^{P} y(k-n)f(n-\hat{\Delta}_{11}(k)))^2\} \approx \mu_\Delta \sigma_{n'}^2 \cdot \frac{\pi^2}{3}\cdot(1+\alpha_{21}^2)\sigma_s^2 \tag{L.14}$$

The variance of $\hat{\Delta}_{11}(k)$ is thus found by equating (L.13) and (L.14), which has a value of

$$\lim_{k\to\infty} E\{(\hat{\Delta}_{11}(k)-\Delta_{11})^2\} \approx \frac{\mu_\Delta \sigma_{n'}^2}{\alpha_{11}}$$

In a similar manner, the mean square errors of $\hat{\alpha}_{21}(k)$ and $\hat{\Delta}_{21}(k)$ can be shown to be

$$\lim_{k\to\infty} E\{(\hat{\alpha}_{21}(k)-\alpha_{21})^2\} \approx \mu_\alpha \sigma_{n'}^2 + \frac{\alpha_{21}^2\sigma_n^4}{\sigma_x^4}$$

and

$$\lim_{k\to\infty} E\{(\hat{\Delta}_{21}(k)-\Delta_{21})^2\} \approx \frac{\mu_\Delta \sigma_{n'}^2}{\alpha_{21}}$$

References

[1] "Special issue on time delay estimation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.29, June 1981.

[2] G.C.Carter, *Coherence and Time Delay Estimation: An Applied Tutorial for Research, Development, Test, and Evaluation Engineers*, IEEE Press, 1993.

[3] A.H.Quazi, "An overview on the time delay estimate in active and passive systems for target localization," *IEEE Trans, Acoust., Speech, Signal Processing*, vol.29, no.3, pp.527-533, June 1981.

[4] G.C.Carter and P.B.Abraham, "Estimation of source motion from time delay and time compression measurements," *J. Acoust. Soc. Amer.*, vol.67, no.3, pp.830-832, March 1980.

[5] G.C.Carter, "Time delay estimation for passive sonar signal processing," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.29, no.3, pp.463-470, June 1981.

[6] W.W.Smith and P.G.Steffes, "Time-delay techniques for satellite interference location system," *IEEE Trans. Aerospace and Elect. Sys.*, vol.25, no.2, pp.224-230, March 1989.

[7] K.C.Ho and Y.T.Chan, "Solution and performance analysis of geolocation by TDOA," *IEEE Trans. Aerospace and Elect. Sys.*, vol.29, no.4, Oct. 1993.

[8] M.J.Marcus, "Satellite security: legacy of Captain Midnight," *Telecommunications*, vol.21, pp.61-62, June 1987.

[9] G.Johnson, *Personal communication*, Contel Corporation, Altanta, Ga., 1987.

[10] V.S.Jasrotia and P.A.Parker, "Matched filters in nerve conduction velocity estimation," *IEEE Trans. Biomed. Eng.*, vol.30, no.1, pp.1-9, Jan. 1983.

[11] G.N.McVicar and P.A.Parker, "Spectrum dip estimator of nerve conduction velocity," *IEEE Trans. Biomed. Eng.*, vol.35, no.12, pp.1069-1076, Dec. 1988.

[12] N.J.I.Mars, "Time delay estimator for EEG analysis based on information theory," *Proc. ICASSP-82*, vol.2, pp.733-735, 1982.

[13] P.Laguna, R.Jane and P.Caminal, "A time delay estimator based on the signal integral: theoretical performance and testing on ECG signals," *IEEE Trans. Signal Processing*, vol.42, pp.3224-3223, Nov. 1994.

[14] M.M.Gibson, "Delay estimation of disturbances on the basilar membrane," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.29, no.3, pp.621-623, June 1981.

[15] M.Simaan, "A frequency-domain method for time-shift estimation and alignment of seismic signals," *IEEE Trans. Geoscience and Remote Sensing*, vol.23, no2, pp.132-138, 1985.

[16] H.Meyr, "Delay-lock tracking of stochastic signals," *IEEE Trans. Commun.*, vol.24, no.3, March 1976.

[17] P.Bolon and Jean-Louis Lacoume, "Speed measurement by cross correlation - theoretical aspects and applications in the paper industry", *IEEE Trans, Acoust., Speech, Signal Processing*, vol.31, no.6, pp.1374-1377, Dec. 1983.

[18] J.E.Salt, B.L.F.Daku and H.C.Wood, "Noncontact velocity measurement using time delay estimation techniques," *IEEE Trans. Signal Processing*, vol.41, pp.288-295, Jan. 1993.

[19] N.C.Mohanty, "Estimation of delay of M PPM signals in Laguerre communications," *IEEE Trans. Commun.*, vol.22, no.5, pp.713-714, 1974.

[20] A.Petropulu, C.L.Nikias and J.G.Proakis, "Cummulant cepstrum of FM signals and high-resolution time delay estimation," *Proc. ICASSP-88*, vol.5, pp.2642-2645, 1988.

[21] W.Stammler and F.Negretti, "Mobile radio bit synchronization and time delay estimation measurement with a digital phase locked loop," *Proc. ICASSP-90*, vol.3, pp.1695-1698, 1990.

[22] T.Takeuchi, M.Sako and S.Yoshida, "Multipath delay estimation for indoor wireless communication," *40th IEEE Vehicular Technology Conference*, pp.401-406, Orlando, 1990.

[23] A.N.Netravali and J.D.Robbins, "Motion-compensated television coding: part I," *Bell System Technical Journal*, vol.58, no.3, pp.631-670, March 1979.

[24] V.Nimier, G.Jourdain and B.Faure, "Use of high resolution methods in time delay estimation for tomography," *Proc. ICASSP-89*, vol.4, pp.2708-2711, 1989.

[25] D.Mauuary and G.Jourdain, "Bayesian time delay estimation for ocean acoustic tomography," *Proc. ICASSP-94*, vol.2, pp.341-344, 1994.

[26] C.H.Knapp and G.C.Carter, "The generalized correlation method for estimation of time delay," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.24, pp.320-327, Aug. 1976.

[27] J.C.Hassab and R.E.Boucher, "Optimum estimation of time delay by a generalized correlator," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.27, no.4, pp.373-380, Aug. 1979.

[28] A.O.Hero and S.C.Schwartz, "A new generalized cross correlator," *IEEE Trans. Acoust, Speech, Signal Processing*, vol.33, no.1, pp.38-45, 1985.

[29] F.A.Reed, P.L.Feintuch and N.J.Bershad, "Time delay estimation using the LMS adaptive filter - static behaviour," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.29, no.3, pp.561-571, June 1981.

[30] P.L.Feintuch, N.J.Bershad and F.A.Reed, "Time delay estimation using the LMS adaptive filter - dynamic behaviour," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.29, no.3, pp.571-576, June 1981.

[31] D.H.Youn, N.Ahmed and G.C.Carter, "On using the LMS algorithm for time delay estimation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.30, no.5, pp.798-801, Oct. 1982.

[32] D.H.Youn, N.Ahmed and G.C.Carter, "An adaptive approach for time delay estimation of band-limited signals," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.31, no.3, pp.780-784, June 1983.

[33] D.H.Youn and V.J.Mathews, "Adaptive realizations of the maximum likelihood processor for time delay estimation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.32, no.4, pp.938-940, Aug. 1984.

[34] Y.T.Chan, J.M.Riley and J.B.Plant, "A parameter estimation approach to time-delay estimation and signal detection," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.28, no.1, pp.8-16, Feb. 1980.

[35] Y.T.Chan, J.M.F.Riley and J.B.Plant, "Modeling of time-delay and its application to estimation of nonstationary delays," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.29, no.3, pp.577-581, June 1981.

[36] P.C.Ching and Y.T.Chan, "Adaptive time delay estimation with constraints," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.36, no.4, pp.599-602, April 1988.

[37] K.C.Ho and P.C.Ching, "A new constrained least mean square time-delay estimation system," *IEEE Trans. Circuits and Systems*, vol.37, no.8, pp.1060-1064, Aug. 1990.

[38] K.C.Ho, P.C.Ching and Y.T.Chan, "A new configuration for convergence speedup in adaptive time-delay estimation," *IEEE Trans. Signal Processing*, vol.40, no.11, pp.2683-2691, Nov. 1992.

[39] K.C.Ho, Y.T.Chan, P.C.Ching, "Adaptive time-delay estimation in nonstationary signal and/or noise power environments," *IEEE Trans. Signal Processing*, vol.41, no.2, pp.592-601, July 1993.

[40] D.M.Etter and S.D.Stearns, "Adaptive estimation of time delays in sampled data systems," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.29, no.3, pp.582-587, June 1981.

[41] J.O.Smith and B.Friedlander, "Adaptive interpolated time-delay estimation," *IEEE Trans. Aerospace and Elect. Sys.*, vol.21, no.2, pp.180-199, March 1985.

[42] H.Messer and Y.Bar-Ness, "Closed-loop least mean square time-delay estimator," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.35, no.4, pp.413-424, April 1987.

[43] D.Boudreau and P.Kabal, "Joint time-delay estimation and adaptive recursive least squares filtering," *IEEE Trans. Signal Processing*, vol.41, pp.2289-2299, Feb. 1993.

[44] B.Widrow, J.M.McCool, M.G.Larimore and C.R.Johnson, "Stationary and nonstationary learning characteristics of the LMS adaptive filter," *Proc. IEEE*, vol.64, no.8, pp.1151-1162, Aug. 1976.

[45] B.Widrow and S.D.Stearns, *Adaptive Signal Processing*, Englewood Cliffs, NJ: Prentice-Hall, 1985.

[46] N.J.Nilsson, "On the optimum range resolution of radar signals in noise," *IRE Trans. Inform. Theory*, vol.11, pp.245-253, 1961.

[47] R.C.Kemerait and D.G.Childers, "Signal detection and extraction by cepstrum techniques," *IEEE Trans. Inform. Theory*, vol.22, pp.444-454, July 1976.

[48] J.P.Ianniello, "Large and small error performance limits for multipath time delay estimation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.34, no.2, pp.245-251, April 1986.

[49] J.P.Ianniello, "High-resolution multipath time delay estimation for broad-band random signals," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.36, no.3, pp.320-327, March 1988.

[50] J.O.Smith and B.Friedlander, "Adaptive Multipath Delay Estimation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, pp.812-822, August 1985.

[51] M.J.D.Rendas and J.M.F.Moura, "Cramér-Rao Bound for Location Systems in Multipath Environment," *IEEE Trans. Signal Processing*, vol.39, pp.2593-2610, Dec. 1991.

[52] J.C.Hassab, "Passive tracking of a moving source by a single observer in shallow water," *J. Sound and Vibration*, vol.44(1), pp.127-145, 1976.

[53] P.C.Ching, Y.T.Chan and K.C.Ho, "Constrained Adaptation for Time Delay Estimation with Multipath Propagation,", *IEE Proceeding-F*, vol. 138, no.5, pp.453-458, 1991.

[54] Simon Haykin, *Adaptive Filter Theory*, Englewood Cliffs, NJ: Prentice-Hall, 1991.

[55] P.M.Clarkson, *Optimal and Adaptive Signal Processing*, CRC Press, Inc., 1993.

[56] N.Kalouptsidis and S.Theodoridis, *Adaptive System Identification and Signal Processing Algorithms*, Englewood Cliffs, NJ: Prentice-Hall, 1993.

[57] J.R.Treichler, "Transient and convergent behaviour of the adaptive line enhancer," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.27, no.1, pp.53-62, Feb. 1979.

[58] I.S.Gradshtegn and I.M.Ryzhik, *Tables of Integrals, Series and Products*, New York: Academic Press, 1980, p.7.

[59] J.Krolik, M.Joy, S.Pasupathy and M.Eizenman, "A comparative study of the LMS adaptive filter versus generalized correlation methods for time delay estimation," *Proc. ICASSP-84*, vol.1, pp.15.11/1-4, March 1984.

[60] D.H.Youn and N.Ahmed, "Comparison of two adaptive methods for time delay estimation," *IEEE Trans. Aerospace and Elect. Sys.*, vol.20, no.5, Sep. 1984.

[61] R.Cusani, "Fast techniques for time delay estimation," *Proc. MELECON'89* pp.177-180, April 1989.

[62] G.Jacovitti and G.Scarano, "Discrete time techniques for time delay estimation," *IEEE Trans. Signal Processing*, vol.41, no.2, pp.525-533, Feb. 1993.

[63] H.C.So, P.C.Ching and Y.T.Chan, "A new algorithm for explicit adaptation of time delay," *IEEE Trans. Signal Processing*, vol.42, pp.1816-1820, July 1994.

[64] H.C.So and P.C.Ching, "Comparative performances of LMSTDE and ETDE for delay and Doppler estimation," *Proc. of the 28th Annual Asilomar Conference on Signals, Systems, and Computers.*, vol.2, pp.1501-1505, Oct. 1994.

[65] H.L.Van Trees, *Detection, Estimation, and Modulation Theory, Part I*, New York: Wiley, 1968.

[66] A.Feuer and E.Weinstein, "Convergence analysis of LMS filters with uncorrelated Gaussian data," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, pp.222-230, Feb. 1985.

[67] E.Weinstein and D.Kletter, "Delay and Doppler estimation by time-space partition of the array data," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, pp.1523-1535, Dec. 1983.

[68] V.H.MacDonald and P.M.Schultheiss, "Optimum passive bearing estimation in a spatially incoherent noise environment," *J. Acoust. Soc. Am. 46*, pp.37-43, 1969.

[69] H.C.So, P.C.Ching and Y.T.Chan, "An improvement to the explicit time delay estimator," *Proc. ICASSP-95*, vol.5, pp.3151-3154, May 1995, Detroit, Michigan.

[70] H.C.So and P.C.Ching, "An efficient and unbiased TDOA estimator for geolocation," *submitted to IEEE Trans. Signal Processing, under 2nd review.*

[71] Tom M.Apostol, *Calculus, Volume II*, John Wiley & Sons, Inc., 1969.

[72] C.B.Officier, *Introduction to the Theory of Sound Transmission, with Applications to the Ocean*, New York: McGraw-Hill, 1958

[73] R.J.Urick, *Principles of Underwater Sound for Engineers*, New York: McGraw-Hill, 1967.

[74] P.C.Ching and H.C.So, "Two adaptive algorithms for multipath time delay estimation," *IEEE Journal of Oceanic Engineering*, vol.19, no.3, pp.458-463, July 1994.

[75] P.C.Ching, K.C.Ho, H.C.So and Y.T.Chan, "An adaptive algorithm for multipath time delay estimation," *Proc. of the 3rd International Symposium on Signal Processing and its Applications*, pp.421-424, August 1992, Australia.

[76] H.C.So, P.C.Ching, K.C.Ho and Y.T.Chan, "A novel constrained algorithm for delay estimation in the presence of multipath transmissions," *Proc. ICASSP-93*, vol.1, pp.313-316, April 1993, Minneapolis, USA.

[77] K.J.Astrom and T.Soderstrom, "Uniqueness of the maximum likelihood estimates of the parameters of an ARMA model," *IEEE Trans. Automat. Contr.*, vol.19, pp.769-773, Dec. 1974.

[78] B.M.Bell and T.E.Ewart, "Separating Multipaths by Global Optimization of a Multidimensional Matched Filter," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, no.5, Oct. 1986.

[79] K.C.Ho and Y.T.Chan, "Bias removal in equation-error adaptive IIR filters," *IEEE Trans. Signal Processing*, vol.43, pp.51-62, Jan. 1995.

[80] A.J.Weiss, "Bounds on time-delay estimation for monochromatic signals," *IEEE Trans. Aerospace and Elect. Sys.*, vol.23, no.6, pp.798-808, Nov. 1987.

[81] K.Sam Shanmugan and A.M.Breipohl, *Random Signals: Detection, Estimation and Data Analysis*, John Wiley & Sons, Inc., 1988.

[82] R.V.Churchill and J.W.Brown, *Complex Variables And Applications* New York: McGraw-Hill, 1984.