The Hong Kong Polytechnic University

Department of Building Services Engineering

Analysis of Exponentially Decaying Pulse Signals and Weak Unsteady Signals using Statistical Approach

CHAN, Chun Man

A thesis submitted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy

December 2007

Declaration

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, non material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

Signature : ____________________________

Name : ____________________________

Abstract of Thesis

Using statistical test in engineering is becoming common in these few years. In this thesis, several statistical techniques are investigated to solve acoustical engineering problems. A stochastic volatility model incorporated the exponential power distributions and Student – t distributions are adopted in this thesis to analyze exponentially decay pulses in the presence of background noises of various magnitudes. It is found that the present stochastic volatility model can retrieve the instant of the pulse initiation and the decay constant within engineering tolerance even when the noise is slightly stronger that the pulse amplitude. The results suggest that both these distributions can give accurate recovery of the instants when the abrupt changes take place if the background noise level is lower than that of the changes by 3dB. They also indicate that the exponential-power distribution is more useful when the signal-to-noise ratio falls below 0dB. The results are compared with those obtained by the conventional short-time Fourier transform and its performance is considerably better than that of the latter when the frequency of the decay pulse fluctuates.

To recover the initialization of an exponentially growing wave embedded inside a stationary background noise is very important especially in building services engineering where the early detection of very small alien signal is crucial to the smooth operation of machines. A parameter derived from two statistical tests, namely the Jarque-Bera and the D'Agostino's tests, which are used for checking data normality, is introduced. It is found that the newly derived parameter is very sensitive to the change incurred by the wave to the background noise statistics and is very helpful in locating the instant of the wave initialization even when the signal-to-noise ratio drops to -30dB. The corresponding accuracy of the recovery can be as low as 3 time steps. A simple numerical function together with the Fourier Transform analysis in the detection of very weak sinusoidal signals embedded in a non-stationary random broadband background noise is proposed in the present study. Its performance is studied through the use of two numerical examples. It is found that the present method enables good recovery of the sinusoidal signals and the instants of their initiations even when the signal-to-noise ratio is down to -17dB.

Publications Arising from this Thesis

Chan, C.M., S.K. Tang and H. Wong (2006) On analysis of exponentially decaying pulse signals using stochastic volatility model.
Journal of Acoustical Society of America. **119** (3), pp.1519-1526.

Chan, C.M., S.K. Tang (2006) On analysis of exponentially decaying pulse signals using stochastic volatility model : Part II Student-t Distribution.
Journal of Acoustical Society of America. **120** (4), pp.1783-1786.

Chan, C.M., S.K. Tang, H. Wong and W.L. Lee (2007) On Weak Unsteady Signal Detection using Statistical Tests
To appear in Mathematical Modelling and Applied Computing

W. L. Lee, S.K. Tang, C.M. Chan (2007) Detecting weak sinusoidal signals embedded in a non-stationary random broadband noise – A simulation study.
Journal of Sound and Vibration. **304**, pp. 831-844.

Chan, C.M., S.K. Tang and H. Wong (2008) A Simple Statistical Model for Jumps and Decaying Pulses in the Presence of a Background Noise.
To appear in Applied Acoustic.

Acknowledgements

I would like to express my gratitude to my chief supervisor Prof. Shiu-Keung Tang for his guidance throughout the research project and co-supervisor Dr. Heung Wong for his contribution and enlightenment.

Last but not least, I would like to give special thanks to my God, Bonnie, Pan-Lam, Mother, Father, Brother, Uncle Miao, Dr. Boris Choy and in memory of my Grandaunt.

CONTENTS

Chapter 1

Introduction

1.1 Motivation and Different Algorithms of Signal Detection

Since most of these signals are not stationary, short-time Fourier transform (STFT) [1] and the wavelet transforms (WT) [2] have been proposed and tested in the past few decades. These signals can be found in many branches of sciences, economics, finances and engineering. Typical examples in economics include the modeling of exchange rates, return data or interest rate, while the typical examples in engineering include the vibration of a diesel engine crank-shaft and the turbulent pressure fluctuations within an air jet. During the last decade, time-frequency analysis are commonly used in the field of time series analysis and condition monitoring in different aspects. Many algorithms for its implementation and the applications have also been developed. The early application of signal detection in engineering problems begins with the work of Forrester [3] who applied the Wigner-Ville Distribution (WVD) to average vibration signals, and showed that different faults can be detected in the WVD plot. Few years later, McFadden and Wang [4] applied the normal WVD and the weighted version of the WVD to detect early abnormal structures in signals. Staszewski and Tomlinson [5] applied the wavelet transformation (WT) and neural networks for classification of different conditions with abnormal behaviors, such as sudden jumps in the signal data. Yesilyurt [6] applied wavelet analysis and the instantaneous power spectrum (IPS) and showed that the progression of a signal could be found from the contour plots. Although the IPS can recover the relatively stronger frequency components from the vibration signals, it gives rise to significant error when confronting higher frequency components. About the detection procedure using time-frequency method, it is usually based on visual observation of the time-frequency contour plots. The progression of a abnormal structure in a signal can be monitored by observing changes in the features of the distribution in the contour plots. Different problems may require the use of different time-frequency techniques. For the reverberation time measurement in a noisy environment, the maximum length sequence [7] gives satisfactory results provided that the signal is not too weak when compared to the noise. At high S/N, both the STFT and the WT perform well in the measurement of the decay constant [8].

On the other hand, wavelet transforms have attracted extensive interest, especially in applications to time series data. However, like IPS, it is found that it can only recover the relatively stronger frequency components from the vibration signals, and gives rise to significant error when confronting higher frequency components.

For more details, please see Wong, Ip and Li [9]. The latter try to apply the wavelet to detect the jumps in heteroscedastic autogressive models.

1.2 Review of some Statistical Methods used in Signal Detection

In modern spectral analysis, autoregressive (AR) modeling method is proven appropriate for the estimation of power spectra with sharp peaks. The spectrum of signal vibration may be precisely classified using this method. Since the early 1980's, various model-based approaches have been introduced to machine condition monitoring, mainly for the diagnosis of malfunctions in manufacturing and processing equipment. A number of parametric methods are available for modeling systems. These include autoregressive (AR) modeling, autoregressive moving average (ARMA) modeling, and stochastic volatility modeling. We will discuss the AR and ARMA modeling here and the latter will be discussed in subsequent chapters.

1.2.1 AR Model

AR is developed as a result of the demand for high-resolution spectral estimation. It is a parametric modeling method with a rational transfer function. The AR model is appropriate for representing spectra with sharp peaks, which is the case for time series signals or engineering signals, and is particularly useful for modeling sinusoidal data. For *M* real sinusoids, a *2M* order model has been shown suitable.

A deterministic random process is one that is perfectly predictable based on the infinite past. This means that a time series x[n] can be expressed by an infinite linear combination of all its preceding points, i.e.

$$x[n] = -\sum_{k=1}^{\infty} a[k] \cdot x[n-k] \tag{1.1}$$

The time series may be approximated using its finite (p) preceding values. This model is expressed by a linear regression on itself plus an error term,

$$x[n] = -\sum_{k=1}^{\infty} a[k] \cdot x[n-k] + e[n], \tag{1.2}$$

where *p* is the model order, a[k] are termed the autoregressive coefficients and e[n] is a Gaussian white noise series with zero mean and variance σ^2. If we consider Equation [1.2] as a linear system where e[n] is the input and x[n] the output, the transfer function of the system, H(z), will be rational, that is

$$H(z) = \frac{1}{1 + \sum_{k=1}^{p} a[k] \cdot z^{-k}}, \tag{1.3}$$

where *z* is the forward shift operator. This is called an AR or an *all-pole* model of order *p* and usually denoted as an *AR(p)* model. The power spectral density (PSD) of the time series x[n] in Equation [1.2] is

$$P(f) = \frac{\sigma^2}{|A(f)|^2}, \tag{1.4}$$

where $|A(f)|^2$ represents the PSD of the AR coefficients, a[n], *n=1,2,…, p*. Since the estimation of AR parameters only involves linear Equations, there is a well-established method in estimating the AR coefficients, called the Levinson-Durbin recursion.

1.2.2 ARMA Model

As we have remarked, dependence is very common in time series observations. To model this time series dependence, we start with univariate ARMA models. To motivate the model, there are basically two lines of thinking. First, for a series x_t, we can model dependence of the level of its current observation on the level of its lagged observations. For example, if we observe a high signal realization in this quarter, we would expect that the signal in the next few quarters are good as well. This way of thinking can be represented by an AR model. The AR(1) (autoregressive of order one) can be written as:

$$x_t = \phi x_{t-1} + \varepsilon_t , \tag{1.5}$$

where $\varepsilon_t \sim WN(0, \sigma^2)$, and *WN* denotes White Noise.

Similarly, AR(p) (autoregressive of order p) can be written as:

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + \phi_p x_{t-p} + \varepsilon_t \tag{1.6}$$

In the second way of thinking, we can model that the observations of a random variable at time *t* are not only affected by the shock at time t, but also the shocks that have taken place before time *t*. For example, if we observe a negative shock to the economy, say, a catastrophic earthquake, then we would expect that this negative effect affects the economy not only for the time it takes place, but also for the near future. This kind of thinking can be represented by an Moving Average (MA) model. The MA(1) (moving average of order one) and MA(q) (moving average of order q) can be written as

$$x_t = \varepsilon_t + \theta \varepsilon_{t-1} \tag{1.7}$$

and

$$x_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}\,, \tag{1.8}$$

respectively.

If we combine these two models, we get a general ARMA(p, q) model,

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + \phi_p x_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} \tag{1.9}$$

The ARMA model provides one of the basic tools in time series modeling. In the next few sections, we will discuss how to draw inferences using a univariate ARMA model.

1.3 Review of Monte Carlo Methods

In this sector, we review the Monte Carlo methods, including the basic theories of Markov Chain Monte Carlo and different sampling procedures, like

Metropolis-Hastings algorithm and Gibbs sampler.

Markov chain Monte Carlo (MCMC) methodology provides enormous scope for realistic statistical modeling. MCMC is essentially Monte Carlo numerical integration using Markov chains. With the Bayesian approach, it is common to integrate the posterior distribution over undesired model parameters, to make inference about other desired model parameters or to make predictions, given the data. Often, it is also necessary to evaluate expectations, a process which also require integration. As most realistic data models lead to multivariate and highly non-linear posterior distributions that are not analytically workable, numerical methods are attractive alternatives. Markov Chain Monte Carlo integration draws samples from the required distribution by running a cleverly constructed Markov chain, and then forms sample averages to approximate expectations.

Another strategy to obtain the samples is to observe the states of a Markov chain whose limiting distributions is our distribution $\pi(x)$ of interest. Each state of the chain represents a bin of the histogram. These algorithms have nice properties, such as guaranteed convergence and insensitivity to the initial values, which, in some circumstances, might outweigh the computational burden, and can be significant in some cases.

1.3.1 Background on Markov chains

A Markov chain is a stochastic process described in terms of states. Denoting the set of possible state values that the process x_i can take by $A = \{S_0, S_1, \cdots, S_q\}$, then the Markov chain is said to be in state *k* when $x_i = S_k$.

For a first-order Markov process, the probability of the next state x_{i+1}, given all previous value of the process, depends only on the present state, i.e.,

$$P[x_{i+1} \in A \mid x_0, x_1, \cdots, x_i] = P(x_{i+1} \in A \mid x_i], \qquad (1.10)$$

where $P[\cdot \mid \cdot]$ denotes a conditional probability. Suppose the one-step transition probability from another state is $P_{k|j}[x_{i+1}]$, i.e.

$$P_{k|j}[x_{i+1}] = P[x_{i+1} = S_k \mid x_i = S_j]. \qquad (1.11)$$

Assuming that the transition probabilities are stationary over time, it is possible to represent the complete set of these probabilities by a transition matrix as follows

$$T = \begin{bmatrix} P_{1|1} & P_{1|2} & \cdots & P_{1|q} \\ P_{2|1} & P_{2|2} & \cdots & P_{2|q} \\ & \vdots & & \\ P_{q|1} & P_{q|2} & \cdots & P_{q|q} \end{bmatrix} \qquad (1.12)$$

such that we can write a general expression for the probability of x_i being in state S_k from state S_j after n iterations as follows

$$P_{k,j}^{(n)}[x_i] = P[x_i = S_k \mid x_{i-n} = S_j] = T^n P[x_{i-n} \mid S_j]. \qquad (1.13)$$

If the kernel satisfies certain conditions, the Markov chain will converge toward a limiting distribution π, which is independent of the initial state x_{n0}, as follows

$$\pi = \lim_{n \to \infty} P[x_i] = \lim_{n \to \infty} T^{n-n_0} P[x_{n_0}] \tag{1.14}$$

However, how long it takes the chain to reach the equilibrium state depends on a number of factors and, in particular, the number of states that must be discarded at the initial stage, a transient period known as the "burn-in" of a Markov chain. Thus, if the limiting distribution π of a Markov chain is the posterior distribution of interest, the stages of the chain become the samples from the distribution.

1.3.2 Properties of Markov chains

Although not all Markov chains have a limiting distribution, many algorithms exist to set up Markov chains that will converge to the desired density function. For these algorithms to perform as intended, some criteria related to invariance, reversibility, irreductibility, aperiodicity, and recurrence must be satisfied:

- Invariance – The invariance property means that all states in a Markov chain have reached limiting distribution and are distributed according to the distribution of the resulting chain of the Markov chain.
- Reversibility – The reversibility means that the probability of a transition of a Markov chain from one state to another is equal to the probability of a transition in the reverse direction. Reversibility is a sufficient condition for the states of the chain to be in their limiting distributions.

- Irreducibility – The irreducibility condition means that from all starting points the Markov chain can reach any non-empty set with positive probability, in some finite number of iterations.
- Aperiodicity – The aperiodicity means that a Markov chain has kernels that do not induce a periodic behaviour in the states.
- Recurrence – The recurrence means that from all starting points all states can be reached infinitely often.

All MCMC algorithms have been designed to satisfy these constraints. Next we will present two of them: the Metropolis-Hasting (M-H) algorithm and the Gibbs sampler.

1.3.3 Metropolis- Hastings Algorithm

The Metropolis-Hasting (M-H) algorithm [10] is a very flexible method to provide a random sequence of a sample from a given density. Suppose a candidate function $q(\cdot)$ sampled from $s(\cdot)$. This candidate function is chosen to be easy to sample from. One major advantage of this algorithm is that the knowledge of the normalizing constant of the posterior distribution is not required. The posterior distribution is only present in ratios, where this unknown normalizing constant will cancel out, assuming that it remains constant.

Assuming that the chain is in state *x*, we can obtain candidate x^* for the next state by sampling $q(\cdot)$, which in the general case is conditional on *x*. This candidate will be accepted with probability α defined as

$$\alpha(x^*, x) = \min\{r(x^*, x), 1\}, \tag{1.15}$$

with the acceptance ratio *r* is defined as

$$r(x^*, x) = \frac{s(x^*)q(x \mid x^*)}{s(x)q(x^* \mid x)} \qquad (1.16]$$

If the candidate is accepted, the chain takes the new state x^*; otherwise the chain remains at the current state *x*.

In order to get a well-mixed chain, the candidate function $q(\cdot)$ should allow the chain to explore the entire probability space, but with substantial probability of being accepted. To satisfy the irreductibility and the aperiodicity properties, the M-H algorithm required that $q(\cdot)$ be continuous and strictly positive on the support of $s(\cdot)$. It can be shown that the acceptance probability defined in Equation (1.15) and Equation (1.16) guarantees the reversibility requirements.

1.3.4 Gibbs Sampler

Being a special case of the M-H algorithm, the Gibbs sampler allows one to break down the problem of drawing samples from a multivariate density into one of drawing successive samples from densities of smaller dimensionality.

Given a random vector *x* of length *K*, the Gibbs sampler samples each parameter, one at that time, according to the full conditional distributions when all the other parameters are fixed. Let $q(x_k \mid x_{-k})$, $k = 1, \cdots, K$ denote the full conditional density of the k th component of the vector x, where

$$q(x_k \mid x_{-k}) = q(x_i \mid x_1, x_2, \cdots, x_{k-1}, x_{k+1}, \cdots, x_K).$$

Instead of sampling from a complex K dimensional distribution, the problem is reduced to sampling K times from one dimensional conditional distribution. As soon as a variate is drawn, it is inserted into the full conditional probability density function, and it remains there until the next iteration. For this algorithm to be a viable option, all the full conditional posterior distributions must be available in their analytical form.

The rate of convergence of the Gibbs sampler is governed by the posterior correlations between the different parameters and the dimensionality of the parameter space. One way to improve the rate of convergence is to jointly sample highly correlated variables by creating partitions. Also, it might be beneficial to randomly vary the order of the components.

Suppose $\theta_1, \cdots, \theta_k$ are (not necessarily univariate) components of θ. Given a realization at stage t, $\theta^{(t)} = (\theta_1^{(t)}, \cdots, \theta_k^{(t)})$, y is the signal data, the Gibbs sampler proceeds by successively making random drawings :

$\theta_1^{(t+1)}$ from $p(\theta_1 \mid \theta_2^{(t)}, \theta_3^{(t)}, \cdots, \theta_k^{(t)}, y)$,

$\theta_2^{(t+1)}$ from $p(\theta_2 \mid \theta_1^{(t+1)}, \theta_3^{(t)}, \cdots, \theta_k^{(t)}, y)$,

$$\vdots$$

$\theta_k^{(t+1)}$ from $p(\theta_k \mid \theta_1^{(t+1)}, \theta_2^{(t+1)}, \cdots, \theta_{k-1}^{(t+1)}, y)$

thus completing a move from state $\theta^{(t)}$ to state $\theta^{(t+1)}$.

1.3.5 Metropolis – Hastings One-at-a-time Algorithm

In the case where x is of high dimension, it becomes very difficult to select a good candidate function that would lead to a reasonable acceptance rate and allow the chain to mix. To address this problem, the *Metropolis-Hasting one-at-the-time* algorithm [11] in a similar fashion to the Gibbs sampler samples each component (or partition), conditionally on the other components, using a set of candidate functions. Obviously, this algorithm includes the Gibbs sampler as a special case for which the candidate functions are the full conditional distributions and the candidates are always accepted.

In the M-H one-at-the-time algorithm, only one element x_k^* for $k = 1,2,\cdots,K$ at a time will be sampled from $q(\cdot)$, and this candidate vector x^*, defined as

$$x^* = [x_1, \cdots, x_{k-1}, x_k^*, x_{k+1}, \cdots, x_K]^T,$$

will be accepted as the next state with a probability as in Equation [1.14].

1.4 Structure of Thesis

In this thesis, I will adopt some Markov chain Monte Carlo methods, especially the Gibbs sampler and Metropolis-Hastings algorithm to perform random variates simulations from conditional distributions throughout Chapter 2 to 4. Besides, a class of more advanced parametric models, namely the stochastic volatility models (SV), which have been widely used in modeling time series volatilities, are expected to be useful in the modeling of the pulse signal with a sharp pulse followed by a decay embedded in a random noise. The properties of these models and their formulation in the present study will be discussed in the Chapter 2. Recently, these models have been used together with efficient Bayesian computational technique in analyzing time series in economic studies. However, they are rarely applied, at least to the knowledge of the author, to deal with engineering problems. In this chapter, the performance of a specific SV model on retrieving the properties of exponentially decaying pulses in the presence of random noise will be presented. It is showed that the results provide useful information for the future enhancement of signal detection and machine diagnosis.

In Chapter 3, the performance of the SV model developed in Chapter 2 incorporating the Student-t distribution in analyzing decaying pulses is investigated. The results supplement those discussed in Chapter 2. The Student-t distribution is a conventional distribution and has been used in the detection of outlying observations in a time series. It can be expressed in a two-stage scale mixture representation and can in principle be an alternative to the Exponential Power distribution.

In Chapter 4, the effectiveness of statistical distributions in modeling pulses and jumps in the presence of background noises of various magnitudes is investigated. Special attentions are paid on locating the initiation of an exponentially decaying harmonic pulse and the jump. Two heavy-tailed distributions, namely, the Exponential-Power distribution (EP) and the Student-t distribution are adopted in the present study. The effects of various parameters of these distributions on the performance of modeling signals are investigated.

When there is a sudden change to a normally running system, a jump or an abrupt change to the signals from the system will result. Successful and early detection of such changes is very important as some changes can be detrimental to the system. Sensitivities of the tests in manifesting statistical changes are important issues in Chapter 5, especially for the process of machine condition monitoring. Two statistical tests will be introduced. Their performance in locating the instant of change under different signal-to-noise ratios will be investigated.

Though much has been proposed in Chapter 2 to Chapter 5 for improving signal detection and analysis, it failed to give satisfactory results when the signal-to-noise ratio (S/N) goes down to -17, which is the case in the early development of a fault in the building services equipment. The focus of Chapter 6 is on enhancing the detection of weak signals embedded in a stronger non-stationary signal / noise. A better retrieval of the weak signal after the time-frequency procedure is further proposed. Finally, a conclusion is given in Chapter 7.

Chapter 2

On Analysis of Exponentially Decaying Pulse Signals using Stochastic Volatility Model using Exponential Power Distribution

2.1 Introduction

A sharp pulse followed by a decay embedded in a random noise, provided that its magnitude is not too weak when compared to that of the noise, will create a relatively higher volatility in the signal fluctuations at the instant the pulse is introduced. A more advanced parametric model, namely the stochastic volatility models (SV) [12], which have been widely used in modeling time series volatilities, are expected to be useful in the modeling of the abovementioned pulse signal. The properties of these models and their formulation in the present study will be discussed in the next section. In this chapter, the performance of a specific SV model on retrieving the properties of exponentially decaying pulses in the presence of random noises will be presented. It is hoped that the results will provide useful information for the future enhancement of signal detection and machine diagnosis.

2.2 Stochastic Volatility Model

The SV model formulates the volatility by a numerical process which allows the latter to vary stochastically [12,13]. It involves the use of a statistical distribution so chosen to fit the time series / signal to be analyzed. Both the Gaussian distribution and the Student-*t* distribution have been employed in determining outlying observations in financial time series [14,15]. Recently, Choy and his co-workers have investigated the properties of the SV model which incorporates the exponential power distributions (EP) [14,15,16]. In their studies, the EP distribution is expressed as a scale mixture representation, which can highly reduce the computational time in the numerical simulation study. Also, the mixing parameters in the representation can be used to identify the extreme data very accurately. Therefore, the EP distributions are adopted in the present study.

Theoretically, stochastic volatility (SV) models are an alternative version of the autoregressive conditional heteroscedasticity (ARCH) models, which are commonly used to model asset returns. The conditional variance of the ARCH models is assumed to be a function of the previous observations and past variances, since in real situations, the variance of the asset returns varies over time. Instead, the conditional variance is modeled with a stochastic process in the SV models and hence the estimation procedure of the SV models is noticeably harder than the ARCH family of models. Recently, a number of literatures attempt to produce efficient estimation of the SV models.

Let r_t be our signals of interest at time $t = 0,1,2,\cdots,n$. The mean adjusted y_t at time t is defined as

$$y_t = \ln\left(\frac{r_t}{r_{r-1}}\right) - \frac{1}{n}\sum_{i=1}^{n}\ln\left(\frac{r_i}{r_{i-1}}\right), \qquad t = 1,2,\cdots,n$$

The simplest SV model for the signal y_t and log-volatility h_t is specified by

$$y_t = \beta \exp(h_t / 2)\varepsilon_t, \qquad t = 1,2,\cdots,n$$
$$h_t = \begin{cases} \sigma_\eta \eta_1 / \sqrt{1-\phi^2} & t = 1 \\ \phi h_{t-1} + \sigma\eta_t & t > 1 \end{cases}$$

where ε_t and η_t are independent standard Gaussian processes. Hence, β is a constant factor that represents the model instantaneous volatility which is usually set to unity in many literatures. σ is the variance of the log-volatility and ϕ is the persistence of the volatility which takes a value within the interval (-1,1) to satisfy the stationarity condition.

This SV model can be easily implemented using either likelihood or Bayesian approaches. However, in many situations, normality assumption for the distribution of signals may be inappropriate. Many statisticians and engineers may use heavy-tailed distributions such as the Student-t and symmetric stable distributions for modeling signals. However, this extension increases the computational effort substantially. By representing the Student-t distribution as a scale mixtures of normals, the use of mixture densities can speed up the computational effort of Bayesian methods.

2.2.1. Exponential Power Distributions

The exponential power family of distributions provides both heavier- and/or lighter- tailed distributions than the normal Gaussian one. Let $\{y\}$ be the data set, θ its mean, σ its scale parameter and $\beta \in (0,2]$ the kurtosis parameter that controls the thickness of the tails, the density function of the EP distribution EP(y,θ,σ,β) is given by [17]

$$\mathrm{EP}(y \mid \theta,\sigma,\beta) = \frac{1}{2^{(1+\beta/2)}\,\Gamma(1+\beta/2)\sigma}\exp\left(-\frac{1}{2}\left|\frac{y-\theta}{\sigma}\right|^{\frac{2}{\beta}}\right), \tag{2.1}$$

where Γ is the gamma function. The corresponding mean and variance are equal to θ and $2^{\beta}\sigma^{2}\,\Gamma(3\beta/2)/\Gamma(\beta/2)$ respectively. The EP distribution has been studied thoroughly by a number of researchers for statistical modeling and Bayesian robustness (for instance, Choy and Chan [15] and Choy and Walker [18]). Choy and Smith [16] adopted the normal scale mixtures property of the EP density for Bayesian inference using Markov chain Monte Carlo methods with $1 < \beta \le 2$. Recently, Walker and Gutierrez-Pena [19] discovered the following uniform scale mixtures representation for the EP density:

$$\mathrm{EP}(y \mid \theta,\sigma,\beta) = \int_0^\infty U(y \mid \theta - \sigma u^{0.5\beta}, \theta + \sigma u^{0.5\beta})G(u \mid 1+0.5\beta, 0.5)du\,, \tag{2.2}$$

where $U(y \mid a,b)$ is the uniform density function defined on the interval (a,b) and

$G(y \mid c,d)$ is the gamma density function with mean c/d. This representation is valid for the entire range of β and also allows re-writing the EP distribution into the following hierarchical form

$$y \mid u \sim U(\theta - \sigma u^{0.5\beta}, \theta + \sigma u^{0.5\beta}) \quad \text{and} \quad u \sim G(1+0.5\beta, 0.5), \qquad (2.3)$$

where u is referred to as the mixing parameter of the uniform scale mixtures representation. It should be noted that the normal and the Laplace distributions are special cases of the EP family with $\beta = 1$ and 2 respectively.

2.2.2. Bayesian EP SV Model

In this chapter, the family of EP distributions is considered as a generalization of the normal family to model the signal data. This family provides both leptokurtic and platykurtic shapes of distributions which the normal, Student-*t* and stable families cannot offer. From a practical point of view, the EP distribution is believed to be appropriate to model certain types of data and it is worthwhile to develop efficient methods for statistical analysis. A Gibbs sampling approach using the uniform scale mixtures is discussed in Section 2.3 of Chapter 2.

The usual choice of the normal distribution for the white noise ε_t of the SV model is replaced by the EP distribution with known kurtosis parameter β:

$$y_t \mid h_t \sim \mathrm{EP}(0, e^{0.5h_t}, \beta), \qquad (2.4)$$

where t is the index and t = 1,2,…,n. h is the log-volatility. Equation (2.4) can be re-written into the following hierarchical form :

$$y_t \mid h_t, u_t \sim U(-e^{0.5h_t} u_t^{0.5\beta}, e^{0.5h_t} u_t^{0.5\beta}) \text{ at } u_t \sim G(1+0.5\beta, 0.5)\,, \tag{2.5}$$

where u is the mixing parameter of the SV model. Normality assumption is still valid for the conditional and marginal distributions of h in the present study :

$$h_t \mid h_{t-1}, \phi, \sigma^2 \sim N(\phi h_{t-1}, \sigma^2) \text{ and } h_t \mid \phi, \sigma^2 \sim N(0, \sigma^2/(1-\phi^2))\,, \tag{2.6}$$

where $N(0,\alpha)$ is the normal distribution with mean 0 and standard deviation α, and ϕ the persistence parameter. The above SV model with EP white noise and normal log-volatility is referred as the EP-N SV model. In order to complete a full Bayesian framework for the SV model, the followings are assigned priors to other model parameters :

$$\sigma^2 \sim G_{inv}(a_\sigma, b_\sigma) \text{ and } \phi + 1 \sim 2Be(a_\phi, b_\phi)\,, \tag{2.7}$$

where $Be(a,b)$ is the beta distribution with mean $a/(a+b)$, G_{inv} the inverse gamma distribution, and a_σ, b_σ , a_ϕ and b_ϕ are pre-specified constants. The prior distribution for ϕ is a shifted beta distribution with density

$$p(\phi) \propto (1+\phi)^{a_\phi - 1}(1-\phi)^{b_\phi - 1}, \ |\phi| < 1\,. \tag{2.8}$$

2.3 Gibbs Sampler for EP-N SV Models

The simulation-based Gibbs sampling approach [20] is one of the standard methods for carrying out statistical analysis of complicated Bayesian models. The Gibbs sampler allows us to study posterior characteristics via a sequence of iteratively simulated values drawn from a system of full conditional distributions. The efficiency of the Gibbs sampler can be substantially increased if the required samples are drawn from distributions of some standard forms. Gibbs sampling has also been used in underwater acoustic application [21].

The joint distribution of $\vec{y} = (y_1, y_2, \cdots, y_n)$, $\vec{h} = (h_1, h_2, \cdots, h_n)$, $\vec{u} = (u_1, u_2, \cdots, u_n)$, ϕ and σ^2 is given by

$$p(\vec{y}, \vec{h}, \vec{u}, \phi, \sigma^2) = \prod_{t=1}^{n} p(y_t \mid h_t, u_t)\, p(h_1 \mid \phi, \sigma^2)\, p(h_t \mid h_{t-1}, \phi, \sigma^2)\, p(\vec{u})\, p(\phi)\, p(\sigma^2) \quad (2.9)$$

Then, the Gibbs sampling scheme performs successive random variate generation from the following conditional distributions. The duration of the Gibbs sampler computation varies from a few minutes to about 25 minutes on a Pentium V personal computer, depending on the kurtosis parameter of the EP distribution. The number of iterations adopted is 12000 with the first 2000 iterations as the burn-in period. In the foregoing discussions, $\vec{h}_{-t} = (h_1, \cdots, h_{t-1}, h_{t+1,} \cdots, h_n)$ and $\vec{u}_{-t} = (u_1, \cdots, u_{t-1,} u_{t+1}, \cdots, u_n)$.

2.3.1 Full Conditional Densities of *h*

Full conditional density of *h* is given by

$$p(h_t \mid \vec{y}, \vec{h}_{-t}, \vec{u}, \phi, \sigma^2) \propto p(y_t \mid h_t, u_t)\, p(h_t \mid h_{t-1}, u_t, \phi, \sigma^2)\, p(h_{t+1} \mid h_t, u_{t+1}, \phi, \sigma^2). \qquad (2.10)$$

One can then show that these full conditional distributions are truncated normal of the form

$$h_t \mid \vec{y}, \vec{h}_{-t}, \vec{u}, \phi, \sigma^2 \sim \begin{cases} N(\phi h_{t+1} - \sigma^2/2, \sigma^2) & t = 1 \\ N\left(\dfrac{\phi(h_{t-1} + h_{t+1}) - \sigma^2/2}{1+\phi^2}, \dfrac{\sigma^2}{1+\phi^2}\right) & 2 \le t \le n-1 \\ N(\phi h_{t-1} - \sigma^2/2, \sigma^2) & t = n \end{cases} \qquad (2.11)$$

subject to $h_t > \ln y_t^2 - \ln \alpha^2 - \beta \ln u_t$. I have adopted an algorithm for generating random variates from the truncated normal distribution.

2.3.2 Full Conditional Densities of *u* and σ^2

By representing the EP density into a uniform scale mixtures form, one can show that the full conditional distribution of the mixing parameter *u* is a truncated exponential distribution of the form

$$u_t \mid \vec{y}, \vec{h}, \vec{u}_{-t},\ \phi, \sigma^2 \sim \mathrm{Exp}(u|0.5) = 0.5e^{-0.5u}, \qquad (2.12)$$

subject to the condition

$$u_t^{\beta} > \frac{y_t^2}{\alpha^2} e^{-h_t}. \qquad (2.13)$$

Inversion method can be used to sample random variate from the truncated exponential distribution. For σ^2, the use of conjugate prior lead to an inverse gamma full conditional distribution and σ^2 can then be directly sampled from

$$\sigma^2 \mid \vec{y},\vec{h},\vec{u},\phi \sim G_{inv}\left(a_\sigma + \frac{n}{2}, b_\sigma + \frac{1}{2}\left((1-\phi^2)h_1^2 + \sum_{t=2}^{n}(h_t - \phi h_{t-1})^2\right)\right). \qquad (2.14)$$

2.3.3 Full Conditional Densities of ϕ

The full conditional density of ϕ is given by

$$p(\phi \mid \vec{y},\vec{h},\vec{u},\sigma^2) \propto p(h_1 \mid \phi,\sigma^2)\prod_{t=1}^{n} p(h_t \mid h_{t-1},\phi,\sigma^2)\, p(\phi). \qquad (2.15)$$

It can easily be verified that $\prod_{t=2}^{n} p(h_t \mid h_{t-1},\phi,\sigma^2)$ is proportional to a normal density of ϕ with mean μ_ϕ and variance σ_ϕ^2 where

$$\mu_\phi = \frac{\sum_{t=1}^{n-1} h_t h_{t+1}}{\sum_{t=1}^{n-1} h_t^2} \text{ and } \sigma_\phi^2 = \frac{\sigma^2}{\sum_{t=1}^{n-1} h_t^2}. \qquad (2.16)$$

One can express the full conditional density by a product of a truncated normal and a shifted beta density functions :

$$p(\phi \mid \vec{y},\vec{h},\vec{u},\sigma^2) \propto N\left(\frac{\sum_{t=1}^{n-1} h_t h_{t+1}}{\sum_{t=2}^{n-1} h_t^2}, \frac{\sigma^2}{\sum_{t=2}^{n-1} h_t^2}\right)(1+\phi)^{a_\phi - 1/2}(1-\phi)^{b_\phi - 1/2} \qquad (2.17)$$

for $|\phi < 1|$. Sampling random variates from this full conditional density can be done easily by using the Metropolis-Hastings method.

2.4 Numerical Examples

An exponentially decaying harmonic wave is chosen for the illustrations. This kind of signal can be regarded as the simplest (but important) signal in acoustics and vibration studies. The noise signals $n(\tau)$, where τ denotes measurement time, used in the foregoing illustrations are white noises with zero mean values and are Gaussian distributed fluctuations. The combined signal $y(\tau)$ is

$$y(\tau) = s(\tau) + n(\tau)\,, \tag{2.18}$$

where $s(\tau)$ contains the pulses. The S/N in dB is defined as

$$\text{S/N} = 10\log_{10}\left(|s|_{\max} / |n|_{\max}\right). \tag{2.19}$$

Suppose the pulse is initiated at $\tau = \tau_o$ and let η be the decay constant, the signal $y(\tau)$ is

$$y(\tau) = e^{-\eta(\tau-\tau_o)}\cos(2\pi f(\tau-\tau_o))H(\tau-\tau_o) + n(\tau)\,, \tag{2.20}$$

where H denotes the Heavside step function and f the frequency of the harmonic wave. The present investigation focuses on the determination of the instant of the pulse initiation and the decay constant η. The S/N ranges from +∞ to -10dB. Without loss of generality, η is fixed at 2 and f is normally set at 50Hz, but is allowed to have a narrow ±10% time fluctuation in some cases discussed later. τ_o = 1s throughout the present investigation. Figures 2.1(a) and 2.1(b) illustrate the hypothetical signals with S/N = 3dB with constant f (≡50Hz) and time varying f (= 50Hz + 10% time fluctuation) respectively. One can notice that the signal with fluctuating f does not show a clearly exponential decay feature [Figure 2.1(b)].

The parameter β in the SV model determines the shapes of the EP distributions and thus has significant impact on the modeling of the decaying signal by the SV model. The EP distributions at various β are given in Figure 2.2. A large β gives rise to a thick tail of the distribution and is in general not good for detecting changes as a thick tail distribution tends to down-weight the extreme values and thus affect the detection of pulses in the presence of a random noise.

Two parameters are important in the SV model. They are the mixing parameter u and the log-volatility h. The former illustrates fluctuation and thus should be able to suggest the instants of rapid changes in a signal. The latter should follow the shapes of the decaying pulses. In the analysis of a decaying pulse, these two parameters should not be considered in isolation.

2.4.1 Constant *f*

Figure 2.3 shows the time variations of u at different β for S/N = 10dB. The signal is considerably stronger than the background noise in this case. One can notice a prominent sharp peak within a continuous background spikes in each case illustrated. This peak appears around the instant of the pulse initiation. In fact, the magnitudes of the background spikes before the peak, which are due to the background Gaussian noise, relative to that of the sharp peak increase with increasing β. It is found from a closer look at the data at around τ = 1s that the magnitudes of the spikes continue to decrease when τ increases towards 1s as shown Figure 2.3(a) with S/N = 10dB and β = 0.25. A clear reversal is observed at τ equals 1s. Similar phenomenon is also observed at other values of β investigated in the present study under this signal-to-noise ratio.

The time variations of the log-volatility h for S/N = 10dB with various β are presented in Figure 2.4 The patterns of the h variations in general follow a linear decay but the increase in β appears to have smoothen the pulse identity. The linear decay also becomes less obvious as β increases. Therefore, only results at β = 0.1 will be presented in the foregoing discussions.

A decrease in the S/N results in a rocky decay of h and the exponential decay becomes not really traceable at S/N = -7dB though a decaying pulse is still suggested by the time variation of h [Figure 2.5]. However, a clear reversal in u at τ ~ 1s remains prominent even up to a S/N of 0dB and some indications of an abrupt signal jump are still observable at S/N = -7dB [Figure 2.6]. One should note that the instant of the peak u does not necessarily collapse with that of the pulse initiation. The focus should be on the instant of the reversal. It should also be noted that the magnitude of the u peak does not carry much meaning in engineering applications. Table 2.1 summarizes the instant of the u reversal under different S/Ns with $\beta = 0.1$. Though one can anticipate ambiguity in determining the instant of this major u reversal at strong background noise magnitude and thus the error in locating the instant of the pulse initiation, the prominent sharp rise of u, such as that observed in Figure 2.6(c), indicates together with the variation of h that some important changes are embedded in a random signal. This shows the versatility of the present SV model in detecting changes.

Apart from detecting the instant of the pulse initiation, the determination of the decay constant is also an important task in the signal analysis [8]. The method shown in the Appendix 2.1 is adopted to minimize the effects of the background noise in the process.

The average h for the background noise can be obtained using 800 data points starting from $\tau = 0$s. Data points close to the major u reversal should be avoided. The decay constant η estimated under different S/Ns are shown in Table 2.1. ηs from the SV model are obtained using the 1000 data points after the major u reversal and the correlation coefficients R^2 of the regression are in general greater than 0.9 for S/N $\geq$ -3dB. The corresponding values obtained from the WT using the fourth derivative of the Morlet-Grossmann wavelet [22] (with a scale from 1 to 32) and the

STFT (with a frequency resolution of 3.9Hz and 60% data overlapping) are also presented in Table 2.1 for the sake of comparison. Allowing for error in the estimation of the slopes of the decay curves, the two sets of results are comparable. However, the WT and the STFT are not able to provide the time resolution one can obtain from the present SV model unless the frequency resolution is lowered to 40Hz and the S/N is high. Thus, the WT and the STFT are less suitable for the determination of τ_o.

2.4.2 Time Varying *f*

The introduction of a ±10% Gaussian fluctuation in *f* does not affect much the pattern of the time varying mixing parameter *u*. The fluctuation in *f* does increase the log-volatility *h* when the background noise is strong relative to the signal as shown in Figure 2.5. The τ_o and η determined from the present SV model with EP distribution having $\beta = 0.1$ are tabulated in Table 2.2. The randomness in *f* does not produces observable deterioration of the SV model performance, except that the error in locating τ_o becomes slightly higher than that in the constant *f* case when S/N drops below 0dB.

The randomness in *f*, however, results in a broadband time-frequency distribution as computed by the STFT [Figure 2.7]. In order to cover the range of the frequency fluctuation, the frequency resolution here is taken to be 15.6Hz. Sixty percent data overlapping is again adopted in the STFT and the time resolution is 0.024s. One can notice from Table 2.2 that the WT and the STFT do not provide reliable estimation of η. The ambiguity of τ_o determination is high when S/N is less than 3dB. Though η is estimated to be 1.74 and 1.85 by the WT and STFT respectively at S/N = 3dB, the actual decay curve is problematic as shown in Figure 2.8 (all corrected for background noise). The linear decay is not so reflected by the STFT and the WT, while that obtained from the present SV model is still acceptable.

2.5 Conclusion

A stochastic volatility model incorporated the exponential power distribution is used in the present study to capture the initiation and decay of exponentially decaying signals in the presence of random background noises. Its performance is compared with that of the short-time Fourier analysis.

When the pulse is a decaying harmonic wave of constant frequency, the accuracy of decay constant estimation using the present stochastic volatility model is very good when the signal-to-noise ratio is not less than 3dB. Such accuracy deteriorates as the signal-to-noise ratio decreases, but it is still comparable to that of the short-time-Fourier transform. The present model is able to detect the instant of the pulse initiation even up to a signal-to-noise ratio of -7dB within engineering tolerance. The short-time Fourier transform does not provide a comparable time resolution for such detection unless the frequency resolution is scarified and the signal-to-noise ratio is high.

The introduction of a time varying frequency (±10%) does not produce significant effect to the performance of the stochastic volatility model, though it rises up the volatility when the background noise is relatively strong, making the detection of pulse difficult once the signal-to-noise ratio drops below 0dB. However, the performance of the short-time Fourier transform is substantially worsened even in the absence of the background noise.

In order to check for the various samplers performed, I have checked for the Ergodic averages of the samplers and the autocorrelation function as the prior study. Although the results are not shown here, I notice that there has no convergence problem and the autocorrelation functions dies out relatively slow and I therefore believe that the selected values from the Markov chain are quite independent.

Appendix 2.1

In this chapter, a hypothetical signal y made up of an exponential decaying signal d initiated at $\tau = \tau_o$ (> 0) and a continuous white noise (normally distributed) n is considered. The SV model suggests

$$y = d + n = H_t^{1/2}\varepsilon_t\,, \tag{A2.1}$$

where ε_t is a time fluctuating white noise and H_t the volatility. Suppose one can find $d = H_{d,t}^{1/2}\varepsilon_t$ (no noise) and $n = H_{n,t}^{1/2}\varepsilon_t$, one then has the following approximate relationship.

$$H_t^{1/2} = H_{d,t}^{1/2} + H_{n,t}^{1/2}\,. \tag{A2.2}$$

Let $h_t = \log H_t$, $h_{d,t} = \log H_{d,t}$ and $h_{n,t} = \log H_{n,t}$, one can then find

$$h_{d,t} = 2\log\left(e^{h_t/2} - e^{h_{n,t}/2}\right). \tag{A2.3}$$

At $\tau > \tau_o$ and excluding the sinusoidal time fluctuation, one obtains $\left|d\right| \sim Ae^{-\alpha(\tau-\tau_o)}$ and $H_{d,t}^{1/2} \sim Ae^{\alpha\tau_o}e^{-\alpha\tau}$ so that $h_{d,t} = \log H_{d,t} = -2\alpha\tau + 2\log(Ae^{\alpha\tau_o})$, and

$$\log(Ae^{\alpha\tau_o}) - \alpha\tau = \log\left(e^{h_t/2} - e^{h_{n,t}/2}\right). \tag{A2.4}$$

Since n is normally distributed and thus can be treated as ε_t multiplied by a constant, it is straight-forward to conclude that

$$n = A(\mathrm{S/N})\varepsilon_t \Rightarrow e^{h_{n,t}/2} \sim A \times \mathrm{S/N}\,. \tag{A2.5}$$

One can normalize y by A such that A = 1 in the above derivation. $h_{n,t}$ can be obtained from the data at $\tau < \tau_o$. Usually, an average from less than 400 points is good enough to get a reliable estimate of $h_{n,t}$ when S/N ≥ 0.

Table 2.1 Estimated τ_o and η (constant f)

	SV Model		STFT	Wavelet
S/N (dB)	τ_o (s)	η (s^{-1})	η (s^{-1})	η (s^{-1})
+∞	1.000	2.00	2.00	2.00
10	1.000	2.05	2.05	2.06
3	1.000	2.16	1.81	1.78
0	1.001	2.42	1.58	1.51
-3	1.002	2.36	1.30	1.23
-7	1.022	--	--	--

Table 2.2 Estimated τ_o and η (fluctuating f)

	SV Model		STFT		Wavelet	
S/N (dB)	τ_o (s)	η (s^{-1})	t_o (s)	η (s^{-1})	t_o (s)	η (s^{-1})
+∞	1.000	2.00	0.96	2.53	0.95	2.58
10	1.000	2.04	0.96	2.83	0.93	2.91
3	1.000	2.30	--	1.85	--	1.74
0	1.010	2.48	--	1.10	--	1.04
-3	1.001	2.37	--	--	--	--
-7	--	--	--	--	--	--

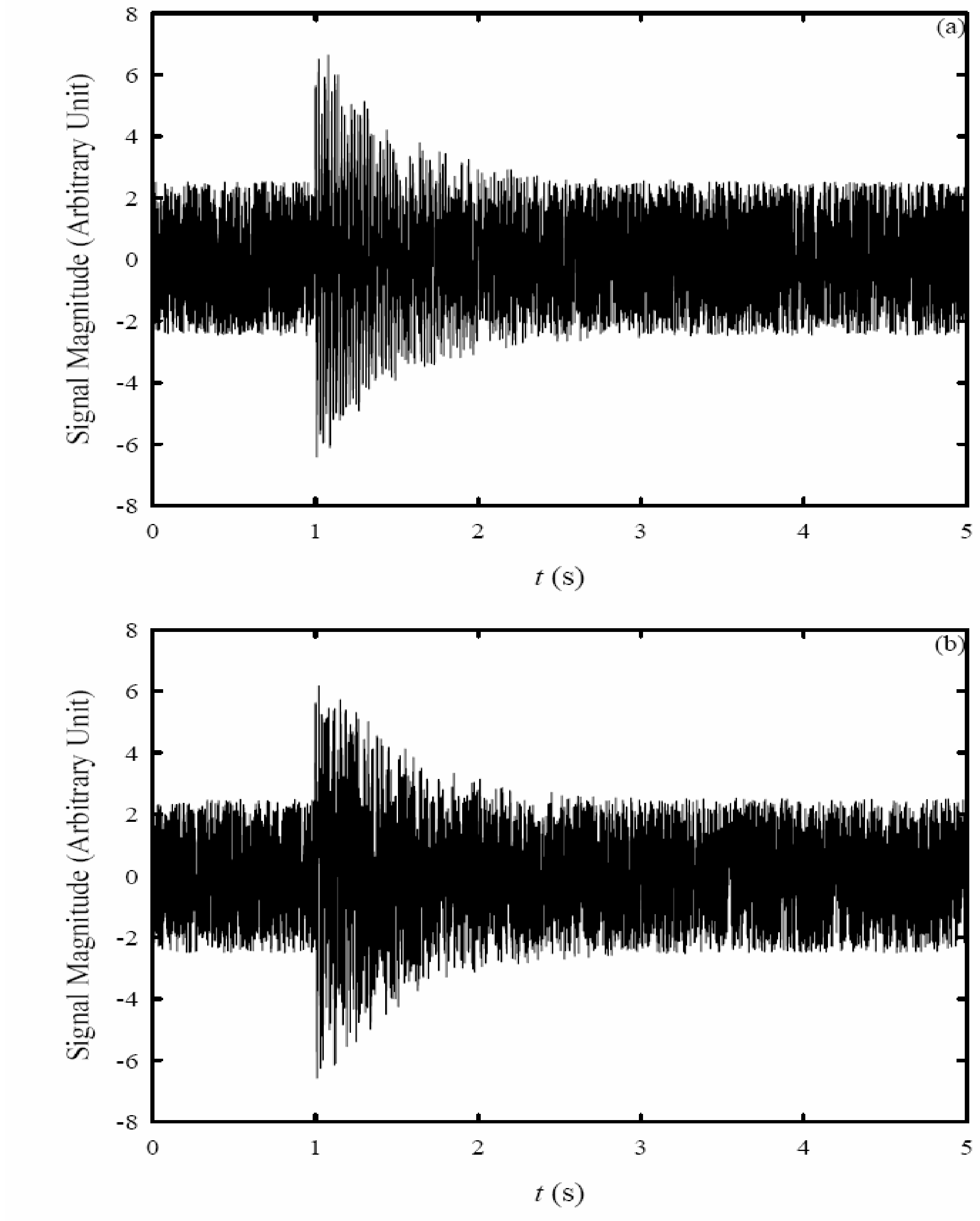

Figure 2.1 Examples of exponentially decaying signals with S/N = 10dB.

(a) Constant *f*; (b) time varying *f*.

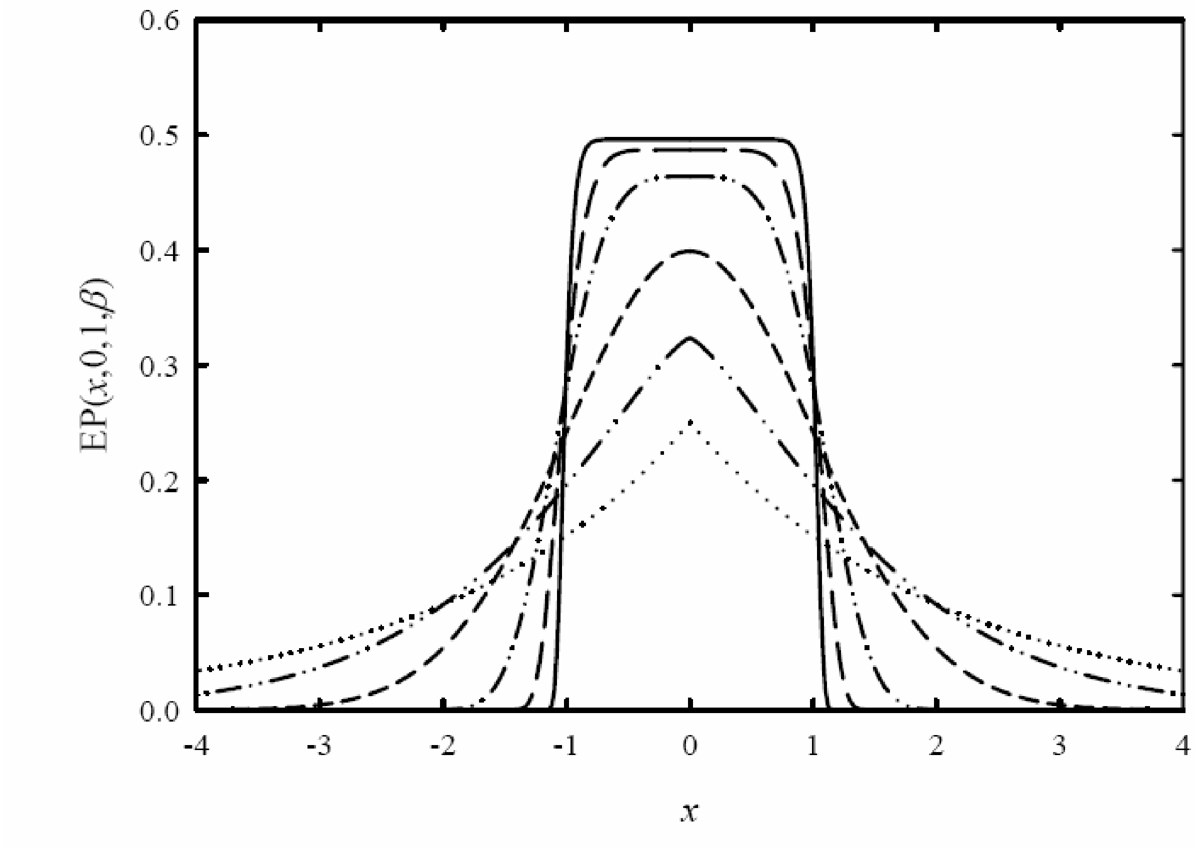

Figure 2.2 Effects of β on the shape of an EP distribution.

——— : β = 0.1; — — — : β = 0.25; — · — : β = 0.5; – – – – : β = 1;

— · · — : β = 1.5; · · · · · · · · · : β = 2.

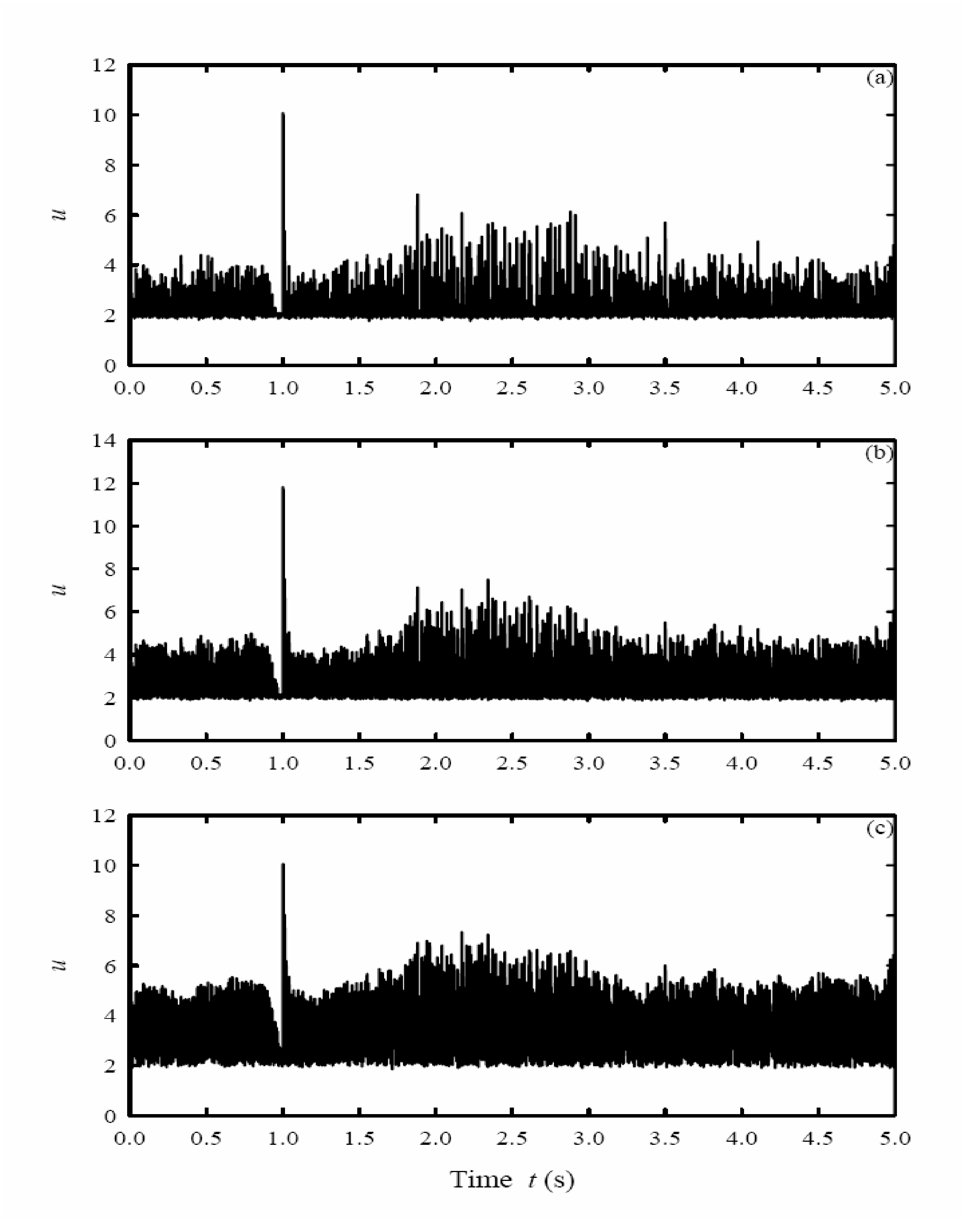

Figure 2.3 Effects of β on the time variation of the mixing parameter u.

(a) $\beta = 0.25$; (b) $\beta = 0.75$; (c) $\beta = 1.5$; (d) $\beta = 2.0$.

S/N = 10dB.

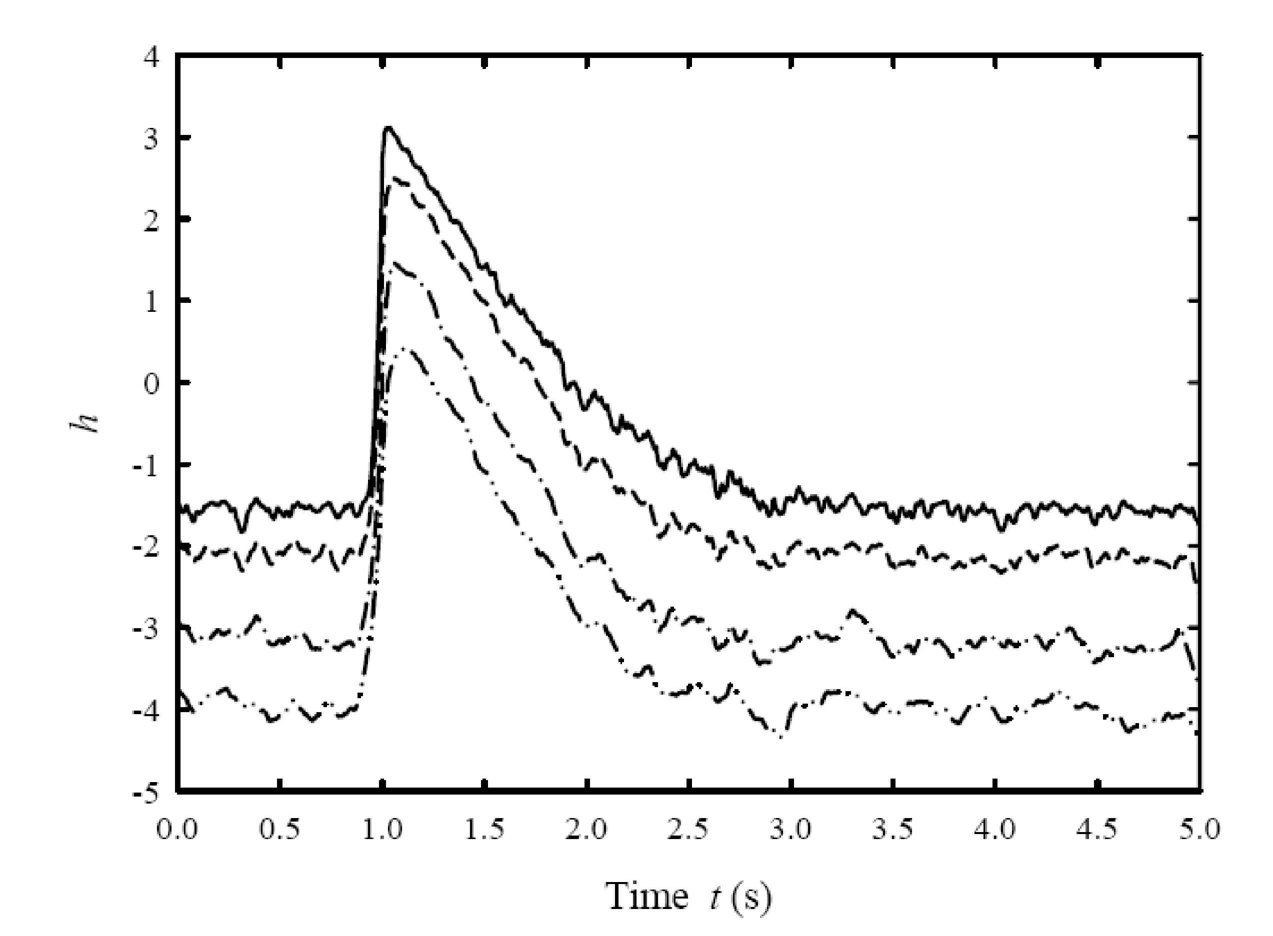

Figure 2.4 Effects of β on the time variation of the log-volatility *h*.

——— : $\beta = 0.25$; – – – – : $\beta = 0.75$; — · — : $\beta = 1.5$; — · · — : $\beta = 2.0$.

S/N = 10dB.

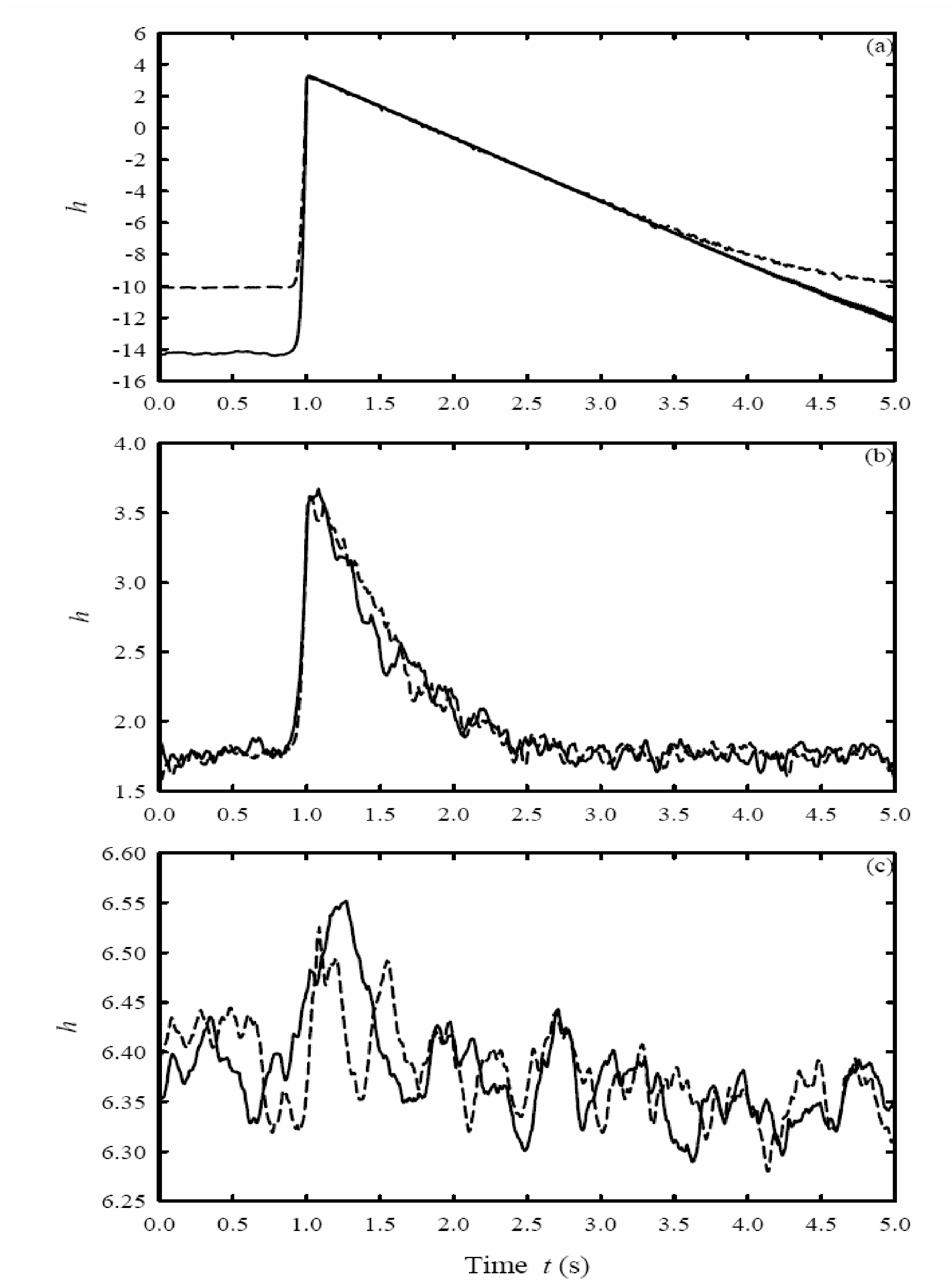

Figure 2.5 Time variations of h at increased noise magnitude with $\beta = 0.1$.

(a) S/N = +∞dB; (b) S/N = 3dB; (c) S/N = -7dB.

——— : Constant f; — — : time varying f.

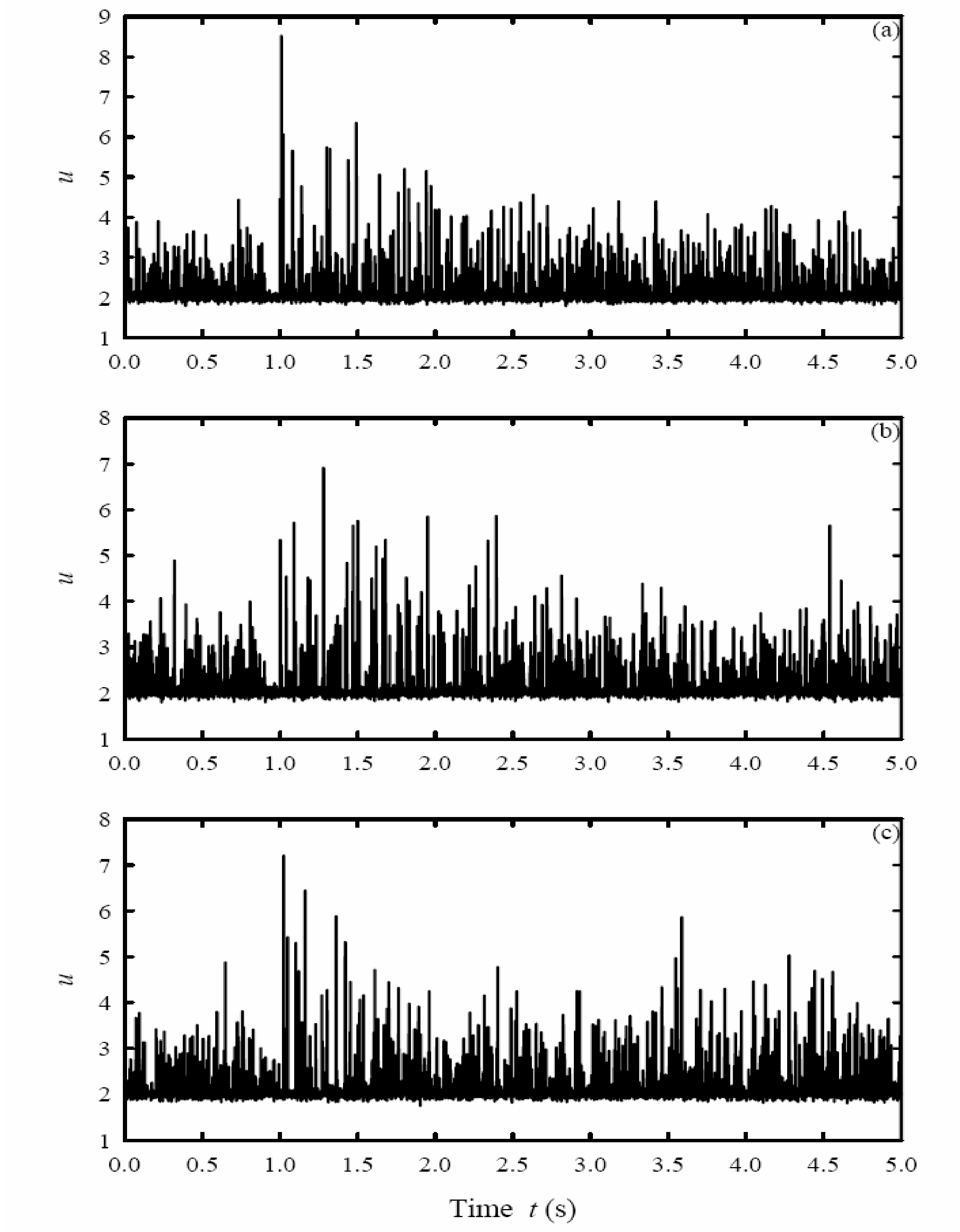

Figure 2.6 Time variations of u at increased noise magnitude with $\beta = 0.1$

(a) S/N = 3dB; (b) S/N = 0dB; (c) S/N = -7dB.

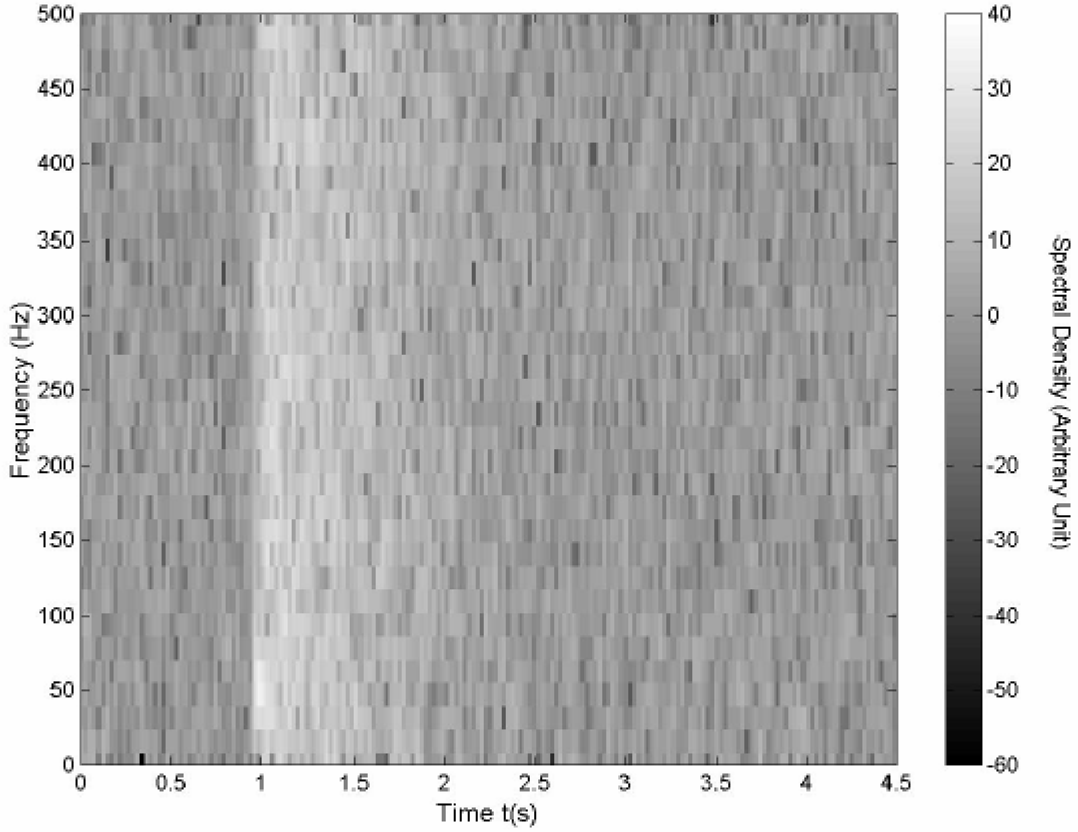

Figure 2.7 Time-frequency plot of an exponential decaying signal with time varying *f*. S/N = 10dB, frequency resolution : 15.6Hz.

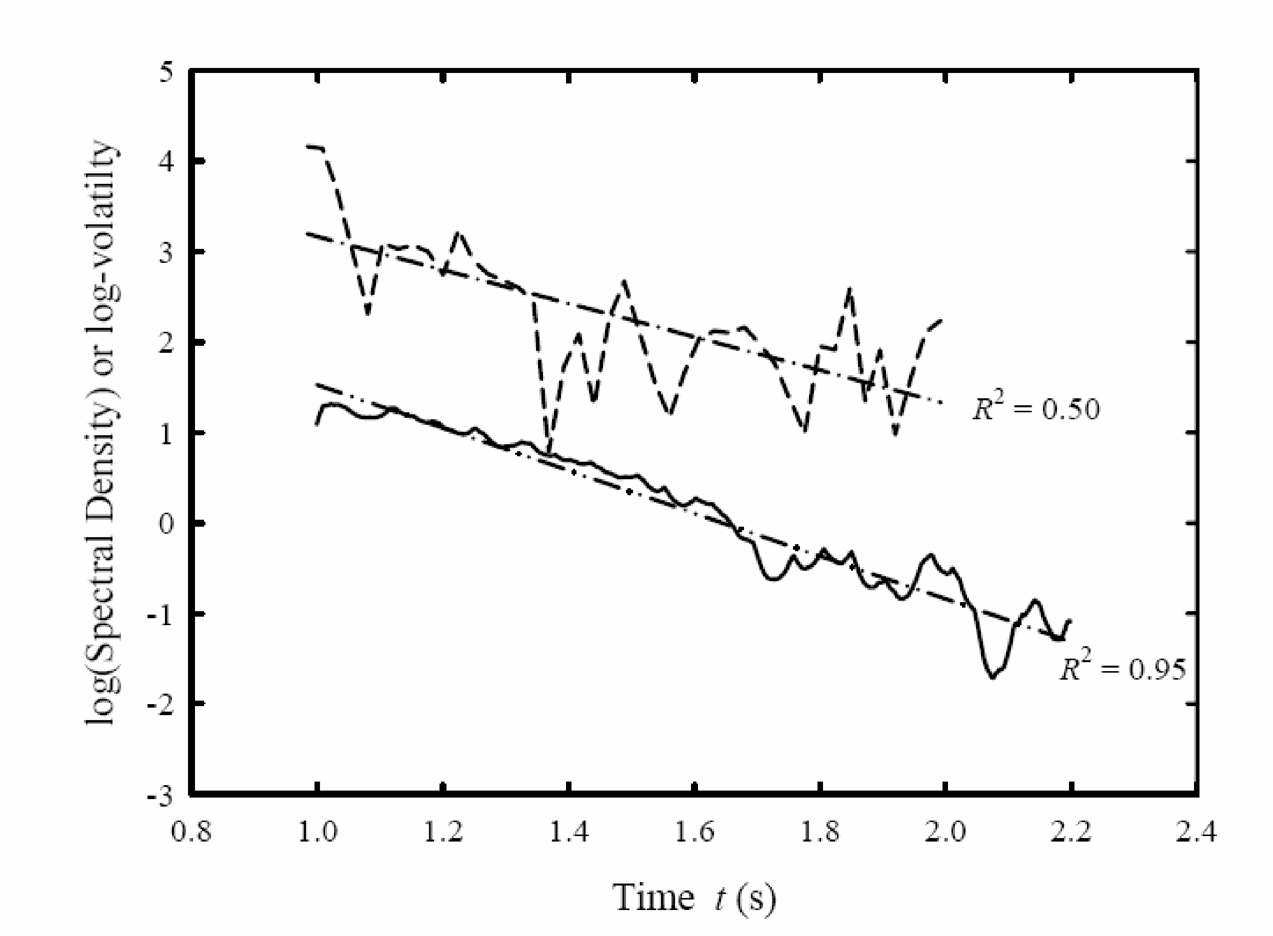

Figure 2.8 Decay curves with time varying f at S/N = 3dB.

——— : Log-volatility h ($\beta = 0.1$) — · · — : regression line for h;

— — : log(Spectral Density); — · — : regression line for log(Spectral Density);

Chapter 3

On Analysis of Exponentially Decaying Pulse Signals using Stochastic Volatility Model using Student-*t* Distribution

3.1 Introduction

Results in Chapter 2 show that the stochastic volatility (SV) model incorporating the exponential power distribution (EP) is able to retrieve the instant of the pulse initiation and the decay constant within engineering tolerance even when the background noise magnitude is comparable to that of the pulse. Its performance is better than that of the conventional short-time Fourier transform when there is a small fluctuation in the frequency of the decaying pulse.

The Student-*t* distribution is a conventional distribution form and has been used in the detection of outlying observations in a time series [23]. It can be expressed in a two-stage scale mixtures representation and can in principle be an alternative to the EP distribution. In this chapter, the performance of the SV model incorporating the Student-*t* distribution in analyzing decaying pulses is investigated.

3.2 TWO-STAGE SCALE MIXTURES REPRESENTATION

A standard random variable *X* having the normal scale mixtures representation can be expressed in the form of $X = Z \times \lambda$ where *Z* is the standard normal random variate and λ is a positive random variate known as the mixing variable having a probability density function *g*, which can be either continuous or discrete.

Let θ and σ be the location and scale parameters of the scale mixtures of the normal random variable *X* respectively. The probability density function of *X* takes the mixture form :

$$f(x) = \int_{\Re^+} N(x \mid \theta, \lambda\sigma^2) g(\lambda) d\lambda, \tag{3.1}$$

where $N(\cdot \mid \cdot)$ denotes the normal density defined on $\Re^+ = (0, \infty)$. In the Bayesian framework, the mixture density in Equation (1) can be expressed into a two-stage hierarchy of the form

$$X \mid \theta, \sigma^2, \lambda \sim N(\theta, \lambda\sigma^2), \quad \lambda \sim g(\lambda). \tag{3.2}$$

The Student-*t* distribution with a degree of freedom α corresponds to an inverse gamma mixing distribution :

$$\lambda \sim g(\lambda) = G_{inv}(0.5\alpha, 0.5\alpha), \tag{3.3}$$

where $G_{inv}(a,b)$ is the inverse gamma distribution with density ($a > 0$ and $b > 0$) :

$$g(\lambda) = b^a e^{-b/\lambda} \lambda^{-(a+1)} \big/ \Gamma(a). \tag{3.4}$$

To facilitate an efficient computation for the SV models, use is made of the class of scale mixtures of uniform representation for the normal density. Since X is a normal random variable with mean θ and variance σ^2, its density function can be rewritten into

$$N(x \mid \theta, \sigma^2) = \int_{\left(\frac{x-\theta}{\sigma}\right)^2}^{\infty} \frac{1}{2\sigma\sqrt{u}} G(u \mid 1.5, 0.5)\, du \,, \tag{3.5}$$

where $G(u \mid a, b)$ is the gamma density function with parameter *a* and *b*. The Student-*t* distribution with a degree of freedom α can be expressed into the following hierarchy

$$X \mid \theta, \sigma^2, \lambda, u \sim U(\theta - \sigma\sqrt{\lambda u}, \theta + \sigma\sqrt{\lambda u}),\ \lambda \sim G_{inv}(0.5\alpha, 0.5\alpha) \text{ and}$$

$$u \sim G(1.5, 0.5), \tag{3.6}$$

where $U(a, b)$ is a uniform distribution defined on the interval (a, b).

3.3 Bayesian Student-t SV Model and Gibbs Sampling

Let H_t and h_t be the volatilities and log-volatilities respectively. In the SV model, the signal data, y_t (where $t = 1, 2, \cdots, n$), is defined as

$$y_t = \beta\sqrt{H_t}\varepsilon_t \quad \text{and} \quad h_t = \begin{cases} \sigma\eta_1 / \sqrt{1-\phi^2} & t = 1 \\ \phi h_{t-1} + \sigma\eta_t & t > 1 \end{cases}, \tag{3.7}$$

where $\{\varepsilon_t\}$ and $\{\eta_t\}$ are independent standard Gaussian processes. β is a constant representing the model instantaneous volatility, σ the variance of the log-volatilities and $\phi \in (-1,1)$ the persistence of the volatility.

In this chapter, the signal data set is modeled by a Student-*t* distribution while the log-volatility is assumed to follow a normal distribution :

$$y_t \mid h_t \sim t_\alpha(0, \beta^2 H_t) \tag{3.8}$$

and $$h_t \mid h_{t-1}, \phi, \sigma^2 \sim N(\phi h_{t-1,} \sigma^2), \tag{3.9}$$

while the marginal distribution is $h_t \mid \phi, \sigma^2 \sim N(0, \sigma^2 / (1-\phi^2))$.

To complete the Bayesian framework, the shifted beta and inverse gamma distributions are assigned to be the independent priors for ϕ and σ^2 respectively :

$$\phi \sim 2Be(a_\phi, b_\phi) - 1 \text{ and } \sigma^2 \sim G_{inv}(a_\sigma, b_\sigma). \tag{3.10}$$

The SV model can then be rewritten hierarchically as

$$y_t \mid h_t, \lambda_t, u_t \sim U(-\beta H_t^{1/2} \lambda_t^{1/2} u_t^{1/2}, \beta H_t^{1/2} \lambda_t^{1/2} u_t^{1/2}) \text{ and } h_t \mid h_{t-1}, \phi, \sigma^2 \sim N(\phi h_{t-1}, \sigma^2) \tag{3.11}$$

with $\lambda_t \sim G_{inv}(0.5\alpha, 0.5\alpha), u_t \sim G(1.5, 0.5), \phi \sim 2Be(a_\phi, b_\alpha) - 1, \sigma^2 \sim G_{inv}(a_\sigma, b_\sigma)$.

It is implemented using the Gibbs sampling approach with the variables $\vec{y} = (y_1, .., y_n)$, $\vec{h} = (h_1, .., h_n)$, $\vec{\lambda} = (\lambda_1, .., \lambda_n)$, $\vec{u} = (u_1, .., u_n)$, $\vec{h}_{-t} = (h_1, .., h_{t-1}, h_{t+1}, .., h_n)$, $\vec{\lambda}_{-t} = (\lambda_1, .., \lambda_{t-1}, \lambda_{t+1}, .., \lambda_n)$ and $\vec{u}_{-t} = (u_1, .., u_{t-1}, u_{t+1}, .., u_n)$. With arbitrarily chosen starting values for $\vec{h}, \vec{\lambda}, \vec{u}, \sigma^2$ and ϕ, the Gibbs sampler iteratively sample random variates from a system of full conditional distributions and the resulting simulations are used to mimic a random sample from the targeted joint posterior distribution. In the present study, β is fixed at unity.

Now the system of full conditionals is given below:

1. Full conditional distribution for $h_t, t = 1, \cdots, n$

$$h_t \mid \vec{h}_{-t}, \vec{\lambda}, \vec{u}, \sigma^2, \phi, \vec{y} \sim N(a_t, b_t\sigma^2) \qquad h_t > \ln y_t^2 - 2\ln\beta - \ln\lambda_t - \ln u_t$$

where

$$a_t = \begin{cases} \phi h_{t+1} - \sigma^2/2 & t=1 \\ (1+\phi^2)^{-1}(\phi(h_{t-1}+h_{t+1}) - \sigma^2/2 & 2 \le t \le n-1 \\ \phi h_{t-1} - \sigma^2/2 & t=n \end{cases}$$

and

$$b_t = \begin{cases} 1 & t=1,n \\ (1+\phi^2)^{-1} & 2 \le t \le n-1 \end{cases}$$

2. Full conditional distribution for σ^2 :

$$\sigma^2 \mid \vec{h}, \vec{\lambda}, \vec{u}, \phi, \vec{y} \sim IG\left(a_\sigma + \frac{n}{2}, b_\sigma + \frac{1}{2}\left[(1-\phi^2)h_1^2 + \sum_{t=2}^{n}(h_t - \phi h_{t-1})^2 \right]\right)$$

3. Full conditional distribution for $\lambda_t,\ t = 1,\cdots,n$:

$$\lambda_t \mid \vec{h}, \vec{\lambda}_{-t}, \vec{u}, \phi, \sigma^2, \vec{y} \sim IG(\frac{\alpha+1}{2}, \frac{\alpha}{2}) \qquad \lambda_t > \frac{y_t^2}{\beta^2 H_t u_t}$$

4. Full conditional distribution for $u_t, t = 1,\cdots,n$:

$$u_t \mid \vec{h}, \vec{\lambda}, \vec{u}_{-t}, \phi, \sigma^2, \vec{y} \sim Exp(\frac{1}{2}) \qquad u_t > \frac{y_t^2}{\beta^2 H_t \lambda_t}$$

5. Full conditional distribution for ϕ :

$$\phi \mid \vec{h}, \vec{\lambda}, \vec{u}, \sigma^2, \vec{y} \sim N\left(\phi \mid \frac{\sum_{t=2}^{n} h_{t-1}h_t}{\sum_{t=2}^{n} h_t^2}, \frac{\sigma^2}{\sum_{t=2}^{n} h_t^2} \right)(1+\phi)^{a_\phi - 1/2)}(1-\phi)^{b_\phi - 1/2} \qquad |\phi| \le 1.$$

3.4 Numerical Examples

The artificial signals in Section 2.4 of chapter 2 are adopted again in this chapter. They consist of an exponentially decaying harmonic wave, *s*, and Gaussian white noises, υ, of various levels :

$$y(\tau) = e^{-\eta(\tau-\tau_o)} \cos(2\pi f(\tau - \tau_o))H(\tau - \tau_o) + \upsilon(\tau)\,, \tag{3.12}$$

where *H* denotes the Heavside step function, *f* the frequency of the wave, η the decay constant, τ_o the instant of pulse initiation and τ the time in second. The signal-to-noise ratio S/N in dB is defined as

$$\text{S/N} = 10\log_{10}\left(|s|_{\max} \big/ |\upsilon|_{\max}\right). \tag{3.13}$$

Without loss of generality, η is fixed at 2 and *f* is normally set at 50Hz in this chapter. The parameter α in the SV model determines the shapes of the Student-*t* distributions and thus has significant impact on the modeling of the decaying signal by the SV model. A small α gives rise to a thick tail distribution and is in general not good for detecting changes as a thick tail distribution tends to down-weigh extreme values, affecting adversely the detection of pulses in the presence of noises. However, since the shape of the Student-*t* distribution does not vary much for α > 30 and the computation with large α is in general impractical as it is very demanding on computing resources, the largest α included in the present investigation is 30. The incorporation of the Student-*t* distribution into the SV model gives rise to two mixing parameters, namely u and λ, and a log-volatility h as shown in the previous section. The two mixing parameters illustrate fluctuation while the latter follows the envelope of the signal magnitude.

Figure 3.1 illustrates the time variations of u, λ and h for S/N = 10dB and $\alpha = 30$. One can observe that the variation patterns of u and λ are opposite. The initiation of the pulse results in a prominent upward and downward spike in u and λ respectively. The present u is basically the same as those obtained using the EP distribution with a kurtosis of 0.75 in Chapter 2 (Figure 2.3). The variation pattern of h is also very close to that shown in Chapter 2 (Figure 2.4). Further increasing α may result in a slightly better performance, but the very computer resources demanding calculation makes it impractical in reality.

The performance of the SV model deteriorates when the background noise level increases. For S/N = -3dB with $\alpha = 30$, both u and λ are unable to indicate without ambiguity the instant of the pulse initiation [Figures 3.2(a) and 3.2(b)]. The log-volatility h does not even suggest the presence of a decaying pulse [Figure 3.2(c)]. However, the pulse initiation and its eventual decay are still observable with the EP distribution at this signal-to-noise ratio level as shown in Chapter 2 (Figures 2.3 and 2.4)

Table 3.1 summarizes the performance of the present SV model and has it compared with that of the previous study in Chapter 2. One can find that the use of the EP distribution in the SV model results in a better analysis when the S/N drops below 3dB. The slightly worse performance of the Student-*t* distribution overall is probably due to the relatively longer tail of the Student-*t* distribution compared to those of the EP family even when the degree of freedom is large.

3.5 Conclusions

A stochastic volatility model which incorporates the Student-*t* distribution is used in the present study to retrieve the instant of initiation and the decay constant of an exponentially decaying signal in the presence of random background noises. It is found that the performance of the Student-*t* distribution is comparable to that of the EP distribution for a signal-to-noise ratio higher than or equal to 3dB. It deteriorates quickly when this ratio falls below 0dB.

Table 3.1 Estimated τ_o and η.*

S/N (dB)	τ_o (s)	η (s^{-1})
$+\infty$	1.000 (1.000)	1.95 (2.00)
10	1.000 (1.000)	2.30 (2.05)
3	1.000 (1.000)	2.20 (2.16)
0	1.010 (1.001)	2.60 (2.42)
-3	1.022 (1.002)	--- (2.36)

*Numbers in parenthesis are those obtained with the EP distribution.[3]

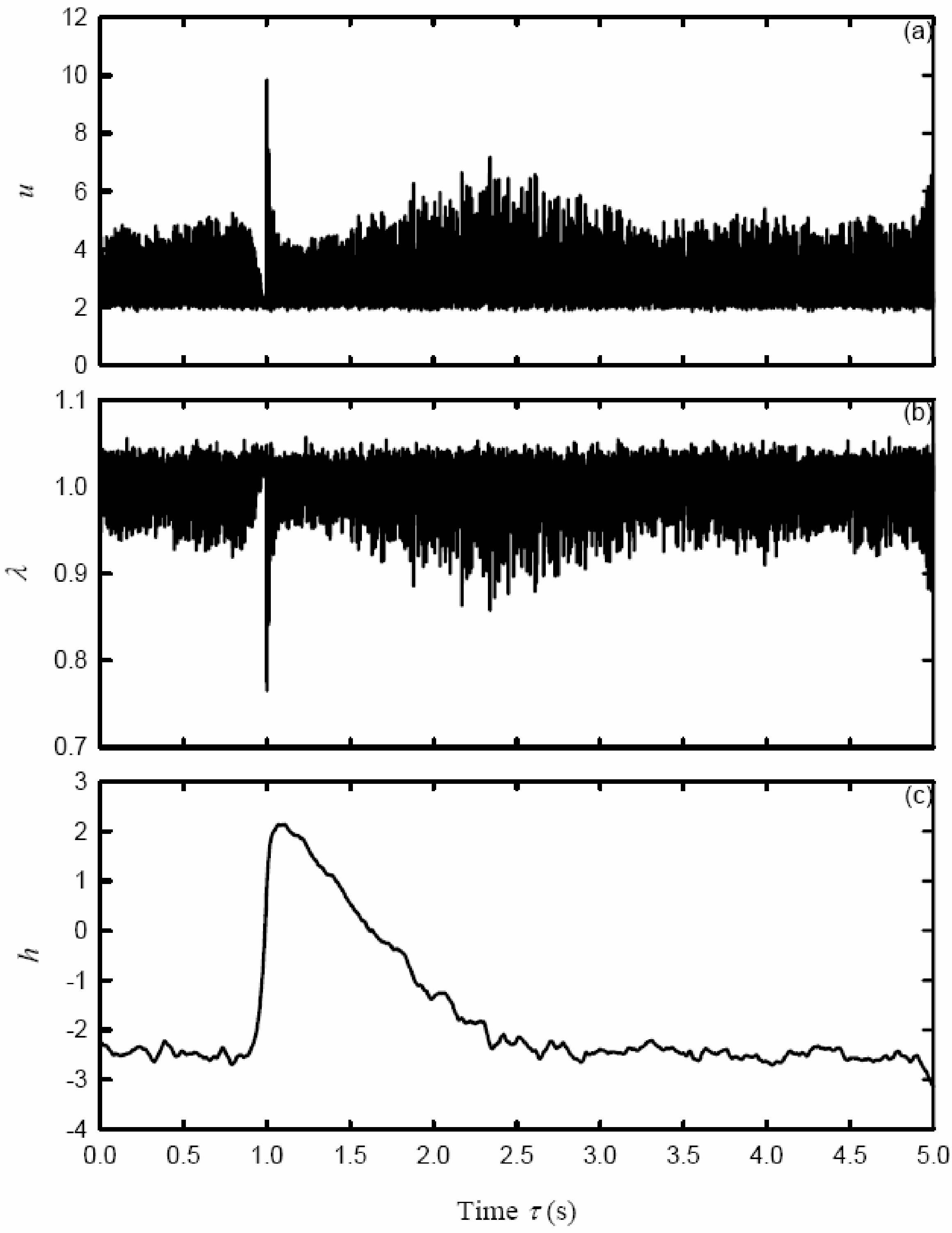

Figure 3.1 Time variations of the mixing parameters for S/N = 10dB and $\alpha = 30$.

(a) u; (b) λ; (c) h.

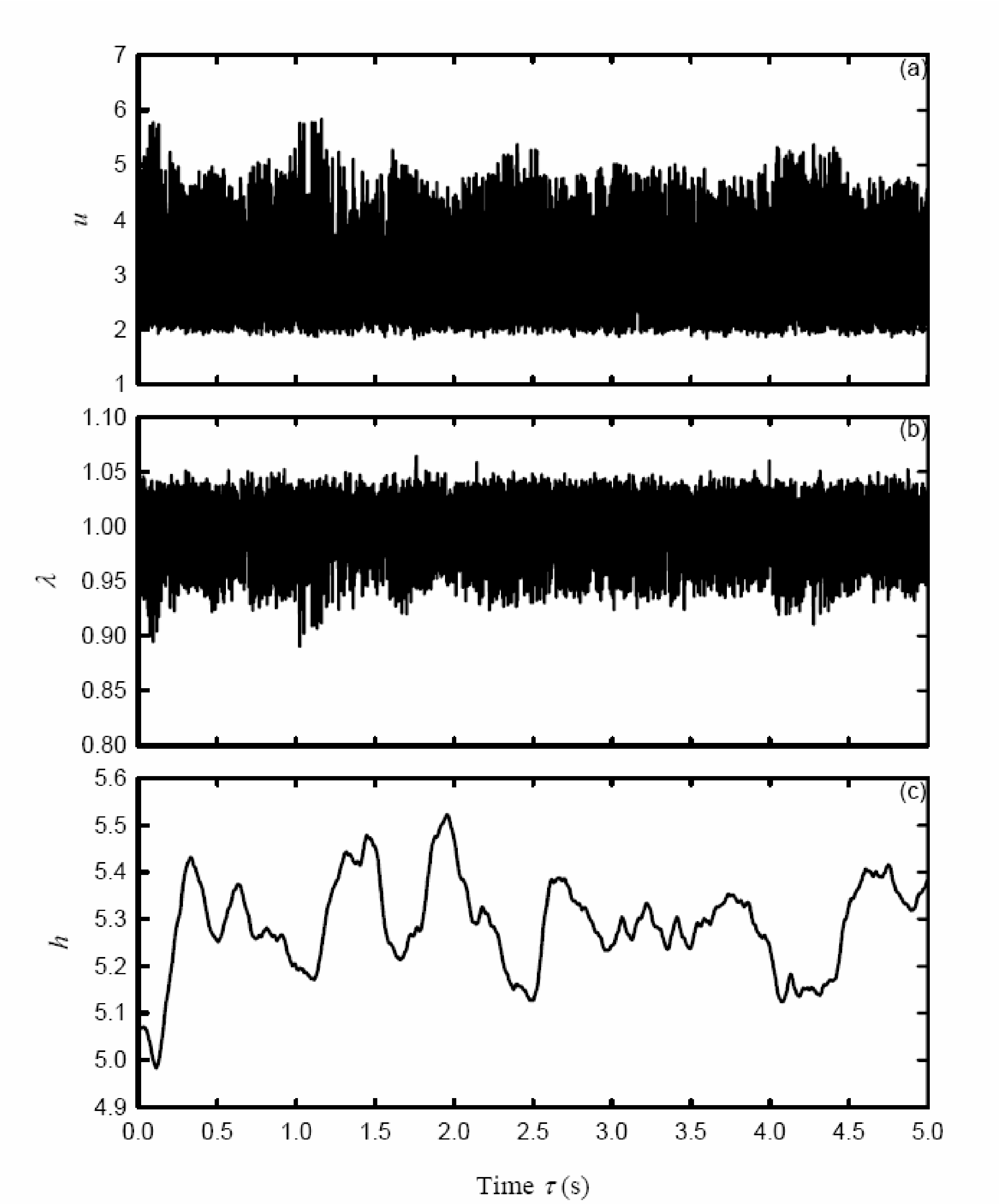

Figure 3.2 Time variations of the mixing parameters for S/N = -3dB and α = 30.

(a) u; (b) λ; (c) h.

Chapter 4

A Statistical Model for Jumps and Decaying Pulses in the Presence of Background Noise

4.1 Introduction

In this chapter, the effectiveness of statistical distributions in modeling pulses and jumps in the presence of background noises of various magnitudes are investigated. Special attentions are paid on locating the instant of initiation of an exponentially decaying harmonic pulse and the jump. Two heavy-tailed distributions, namely, the Exponential-Power distribution (EP) and the Student-t distribution are adopted in this chapter. The effects of various parameters of these distributions on the signal modeling performance are investigated.

4.2 Gibbs Sampling for The Models

The aim of the this chapter is to determine the instant of pulse initiation in the presence of background noise and also to demonstrate how our model can be used to determine the width of a rectangular waveform.

In the following, we assume x_i to be the non-stationary signal and we shall try to model the signal by the heavy-tailed distributions, namely Exponential-power distribution, Student-t distribution and Normal distribution. The proposed model and the corresponding posterior conditional distributions are shown as follows. Throughout this chapter, "rest" represents all parameters but excluding the parameters of interest. a_1 and b_1 are hyper-parameters of inverse-gamma distribution and are pre-specified.

4.2.1 Student-t Distribution

When the signal is assumed to be of the Student-t distribution with a degree of freedom α, the mean and variance are of the Normal and Inverse Gamma distributions respectively. The mixing parameter λ is of the Gamma distribution. The joint distribution of the parameter $(\vec{x}, \mu, \tau^2, \vec{\lambda})$ is $P(\vec{x}, \mu, \tau^2, \vec{\lambda})$

$$\begin{aligned} &P(\vec{x} \mid \mu, \tau^2, \vec{\lambda}) P(\mu) P(\tau^2) P(\vec{\lambda}) \\ &= \prod_i N(x_i \mid \mu, \tau^2, \lambda_i) N(\mu \mid \theta, \sigma^2) G_{inv}(\tau^2 \mid a_1, b_1) G(\lambda_i \mid \alpha/2, \alpha/2), \end{aligned} \tag{4.1}$$

while the posterior conditional distributions of the mean μ, the variance τ^2 and

the mixing

parameter λ_i are

$$\begin{aligned} [\mu \mid rest] &= P(\vec{x} \mid \mu, \tau^2, \vec{\lambda}) P(\mu) = \prod_i N(x_i \mid \mu, \frac{\tau^2}{\lambda_i}) N(\mu \mid \theta, \sigma^2) \\ &\propto \exp\left[\frac{-\lambda_i}{2\tau^2} \sum (x_i - \mu)^2\right] N(\mu \mid \theta, \sigma^2) = \exp\left[\frac{-\lambda_i}{2\tau^2} (\sum (x_i - \bar{x})^2 + n(\bar{x} - \mu)^2)\right] N(\mu \mid \theta, \sigma^2) \\ &\propto \exp\left[\frac{-n\lambda_i}{2\tau^2} (\bar{x} - \mu)^2\right] N(\mu \mid \theta, \sigma^2) \propto N\left(\mu \mid \bar{x}, \frac{\tau^2}{n\lambda_i}\right) N(\mu \mid \theta, \sigma^2) \\ &\sim N\left(\frac{n\lambda_i \bar{x}/\tau^2 + \theta/\sigma^2}{n\lambda_i/\tau^2 + 1/\sigma^2}, \frac{1}{n/\tau^2 + 1/\sigma^2}\right), \end{aligned}$$

(4.2)

$$
\begin{aligned}
[\tau^2 \mid rest] &= P(\vec{x} \mid \mu, \tau^2, \vec{\lambda})P(\tau^2) \propto \prod_i N\left(x_i \mid \mu, \frac{\tau^2}{\lambda_i}\right) G_{inv}(\tau^2 \mid a_1, b_1) \\
&\propto \left(\frac{\lambda_i}{\tau^2}\right)^{\frac{n}{2}} \exp\left[\frac{-1}{2\tau^2}\sum_i \lambda_i (x_i - \mu)^2\right]\left(\frac{1}{\tau^2}\right)^{a_1+1} \exp\left(\frac{-b_1}{\tau^2}\right) \\
&\propto (\frac{1}{\tau^2})^{\frac{n}{2}+a_1+1} \exp\left[\frac{-1}{2\tau^2}\sum_i \lambda_i (x_i - \mu)^2 - \frac{b_1}{\tau^2}\right] \\
&\sim G_{inv}\left(a_1 + \frac{n}{2}, b_1 + \frac{\sum_i \lambda_i (x_i - \mu)^2}{2}\right)
\end{aligned}
$$

(4.3)

And

$$
\begin{aligned}
[\lambda_i \mid rest] &= P(\vec{x} \mid \mu, \tau^2, \vec{\lambda})P(\lambda_i) \propto \prod_i N\left(x_i \mid \mu, \frac{\tau^2}{\lambda_i}\right) G(\lambda_i \mid 0.5\alpha, 0.5\alpha) \\
&\propto \left(\frac{\lambda_i}{\tau^2}\right)^{\frac{1}{2}} \exp\left[-\frac{\lambda_i (x_i - \mu)^2}{2\tau^2}\right]\lambda_i^{\frac{\alpha}{2}-1} \exp\left(\frac{-\alpha}{2}\lambda_i\right) \\
&\propto \lambda_i^{\frac{\alpha-1}{2}} \exp\left\{-\left[\frac{(x_i - \mu)^2}{2\tau^2} + \frac{\alpha}{2}\right]\lambda_i\right\} \sim G\left(\frac{\alpha+1}{2}, \frac{(x_i - \mu)^2}{2\tau^2} + \frac{\alpha}{2}\right),
\end{aligned}
$$

respectively.

(4.4)

4.2.2 EP Distribution

When the signal is of the EP distribution, the mean is calculated from an indicator function and the variance is of truncated inverse gamma distribution. The mixing parameter is of truncated exponential distribution. Throughout the rest of the chapter, the indicator function means that the parameters are only valid or defined in the corresponding region in the bracket valid only. Denoting $c_o = \Gamma(3\beta/2)/\Gamma(\beta/2)$.

The joint distribution of the parameters $(\vec{x}, \mu, \tau^2, \vec{\lambda})$ in this case is $P(\vec{x}, \mu, \tau^2, \vec{\lambda})$, where

$$
\begin{aligned}
& P(\vec{x} \mid \mu, \tau^2, \vec{\lambda}) P(\mu) P(\tau^2) P(\vec{\lambda}) \\
& = \prod_i U\left(x_i \mid \mu - \frac{\tau}{\sqrt{2c_o}} \lambda_i^{\beta/2}, \mu + \frac{\tau}{\sqrt{2c_o}} \lambda_i^{\beta/2} \right) N(\mu \mid \theta, \sigma^2) G_{inv}(\tau^2 \mid a_1, b_1) G(\lambda_i \mid 1 + 0.5\beta, 2^{-1/\beta}).
\end{aligned}
\tag{4.5}
$$

The corresponding posterior conditional distributions of μ, τ^2 and λ_i are, respectively,

$$
\begin{aligned}
& [\mu \mid rest] = P(\vec{x} \mid \mu, \tau^2, \vec{\lambda}) P(\mu) = \prod_i U\left(x_i \mid \mu - \frac{\tau}{\sqrt{2c_o}} \lambda_i^{\beta/2}, \mu + \frac{\tau}{\sqrt{2c_o}} \lambda_i^{\beta/2} \right) N(\mu \mid \theta, \sigma^2) \\
& \sim N(\theta, \sigma^2) I\left(\max(x_i - \frac{\tau}{\sqrt{2c_o}} \lambda_i^{\beta/2}) < \mu < \min(x_i + \frac{\tau}{\sqrt{2c_o}} \lambda_i^{\beta/2}) \right),
\end{aligned}
\tag{4.6}
$$

$$
\begin{aligned}
& [\tau^2 \mid rest] = P(\vec{x} \mid \mu, \tau^2, \vec{\lambda}) P(\tau^2) = \prod_i U\left(x_i \mid \mu - \frac{\tau}{\sqrt{2c_o}} \lambda_i^{\beta/2}, \mu + \frac{\tau}{\sqrt{2c_o}} \lambda_i^{\beta/2} \right) G_{inv}(\tau^2 \mid a_1, b_1) \\
& \quad \sim G_{inv}(a_1 + \frac{n}{2}, b_1) I\left(\tau^2 > \max \frac{2c_o (x_i - \mu)^2}{\lambda_i^{\beta}} \right)
\end{aligned}
\tag{4.7}
$$

And

$$[\vec{\lambda} \mid rest] = P(\vec{x} \mid \mu, \tau^2, \vec{\lambda})P(\vec{\lambda}) = \prod_i U\left(x_i \mid \mu - \frac{\tau}{\sqrt{2c_o}} \lambda_i^{\beta/2}, \mu + \frac{\tau}{\sqrt{2c_o}} \lambda_i^{\beta/2} \right) G\left(\lambda_i \mid 1 + \frac{\beta}{2}, 2^{-1/\beta} \right)$$

$$\sim Exp(2^{-1/\beta}) I\left(\lambda_i \geq 2c_o \left(\frac{(x_i - \mu)^2}{\tau^2} \right)^{1/\beta} \right).$$

Again, α is the degree of freedom in the Student-t distribution. The prior distribution of μ is assumed to be non-informative, in Bayesian context, implying that we have no past information or experience of this parameter. The priors for σ^2 and τ^2 are assumed to be the inverse gamma distribution, while the former is the variance of the normal distribution and the latter is the variance of the inverse gamma distribution.

Since all the conditional distributions are of standard forms, it is easy to perform random generation from truncated normal, truncated gamma and truncated exponential distribution. The simulation algorithms of the first two are suggested in Robert [24]. In our setting, we assume flat prior and $G_{inv}(a_1, b_1)$ to be the prior distributions of μ and τ^2 respectively, while a_1 and b_1 are both set to 0.01 in order to achieve non-informative prior for τ^2. The Gibbs sampler method will be adopted while a single chain of 12000 iterations is run. The first 2000 iterations are taken as a burn-in period and a random sample is picked every 10 iterations.

4.3 Illustrative Examples and Discussions

4.3.1 Exponentially Decaying Harmonic Pulses

The exponentially decaying harmonic wave described in Section 2.4 of Chapter 2. Figure 4.1 gives an example of such decaying pulse with S/N = 3dB. The time resolution is 0.001s. Figure 4.2 illustrates the time variations of the mixing parameter λ at different kurtosis parameter β for S/N = 10dB. The EP distributions are adopted. A sharp rise in λ at the instant of the pulse initiation t_o is observed for all the β adopted, but the decay tail is shorter and the λ peak is higher at smaller β. One should note that a shorter decay tail is beneficial to the modeling of more complicated signals, such as the repeated echoes in building acoustics [25]. Therefore, one can expect a better modeling when β is reduced. In the foregoing discussions on the performance of the EP distribution, β is set at 0.25.

The time variations of λ at different degree of freedom α of the Student-t distribution for S/N = 10dB are shown in Figure 4.3. A sharp and abrupt drop of λ can be observed at $t \sim t_o$ for all α. However, the magnitude of the drop decreases while the decay tail is shortened as α increases. The former is expected to lower the performance of the distribution on resolving the pulse initiation when the background noise becomes stronger. The drawback of the decay tail has been discussed in the last paragraph. It is also found that the magnitudes of the initial noises in λ (at $t < t_o$)

are larger at smaller α. It is not certain on how the variation of α is affecting the modeling in the presence of stronger background noises. It should be noted that the shape of the Student-t distribution becomes close to that of the Normal distribution when α > 30. One can then assume that the performance of the Student-t distribution for very large α will resemble that for α = 100 (Figure 4.3c).

The increase in the background noise magnitude does not deteriorate the performance of the EP distribution in locating t_o as far as S/N ≥ 0dB (Figure 4.4). The corresponding data for S/N = 10dB has been given in Figure 4.2a. It should be reminded that highest λ spike needs not to be at t_o. Ambiguity is observed when S/N exceeds 0dB as shown in Figure 4.4d, where the maximum error is around 24 time steps. The corresponding results obtained using the Student-t distribution are provided in Figure 4.5. It can be observed that a relatively stronger downward spike occurs at the instant t_o for S/N ≥ 0dB as in the case of the EP distribution. However, these spikes are much less prominent than those in the previous EP cases and their strengths relative to the noises decrease more quickly as S/N decreases. The situation becomes worse when S/N exceeds 0dB. One can notice from Figure 4.5d that the spike due to the pulse initiation is completely drowned by the noises for S/N = -3dB. The result shown in Figure 4.4d suggests that the EP distribution is a better choice for the modeling in the presence of a relatively stronger background noise.

It can be concluded from Figures 4.4 and 4.5 that though both the EP and Student-t distributions can retrieve the instant of the pulse initiation accurately when the S/N is greater than or equal to 0dB, the EP distribution performs better in the presence of background noises, especially when the background noise is stronger than the pulse. The Student-t distribution produces a relatively longer decay tail, which is an undesirable feature in signal modeling.

4.3.2 Rectangular Waveforms and Jumps

The second signal to be discussed in this chapter is the rectangular waveform. It basically contains two jumps and takes the form :

$$y(t) = n(t) + A\left[H(t - t_s) - H(t - t_e)\right], \tag{4.8}$$

where t_s and t_e are the instants of initiation and termination respectively. The width of the waveform, w, is therefore equal to $t_e - t_s$. A and t_s are fixed at 3 and 1s respectively in the foregoing discussions. Figure 4.6 illustrates the waveform with S/N = 6dB and w = 0.4s. Unless otherwise stated, w is fixed at 0.4s in the foregoing discussions.

Some examples of the time variations of λ obtained using the EP distribution at various β for S/N = 6dB are shown in Figure 4.7. The two critical instants t_s and t_e should correspond to those at which an abrupt rise and fall of λ are observed respectively. One can notice that better modeling can be achieved at larger β. This phenomenon is opposite to that found for the decaying pulses. When the Student-t distribution is used, it seems that a smaller α will produce better modeling effects. The corresponding results are given in Figure 4.8. However, similar to the cases for the decaying pulses, the magnitudes of the λ spikes due to the background noises become larger as α decreases. The presence of stronger background noises may thus deteriorate the recovery of the two time scales at small degree of freedom.

It is expected that the decrease in the S/N results in less satisfactory recovery of t_s and t_e. Figure 4.9 illustrates the time variations of λ at various S/N obtained using the EP distribution with $\beta = 2$. The ambiguity in locating the jumps introduces error in the estimating the width of the waveform w. Table 4.1 summarizes the estimated t_s, t_e and w at various S/N studied. The performance of the EP distribution is very good as far as the S/N remains higher than or equal to 3dB. The absolute discrepancy does not depend on w.

The performance of the Student-t distribution is more complicated as it depends not only on the S/N, but also on the degree of freedom α. Some examples of the corresponding time variations of λ for various S/N with $\alpha = 1$ and 100 are given in Figures 4.10 and 4.11 respectively. For $\alpha = 1$, the values of λ at $1.0 < t < 1.4$ is already close to those resulted from the background noise when S/N = 0dB (Figure

4.10c), making the estimation of t_s amd t_e un-reliable. However, the corresponding λ jumps obtained with α = 100 can still be distinctly observed (Figure 4.11c). The Student-t distribution fails to indicate signal jumps for S/N = -6dB as shown in Figures 4.10d and 11d. Table 4.2 illustrates its performance at α = 1, 30 and 100. The accuracy of the recovery of the two critical times is satisfactory for S/N > 0dB. The Student-t distribution gives very bad recovery when S/N is further decreased, while the EP distribution with β = 2 can still indicate λ jumps near to the instants t_s and t_e, though the errors are considerable.

Results in Tables 4.1 and 4.2 also suggest that both the EP and Student-t distributions are able to retrieve the location of a unit pulse when the S/N is not less than 3dB. A unit pulse is a fundamental signal form in studies of acoustics and electronics which has a width of one time step [26].

4.4 Conclusions

The modeling of jumps and exponentially time decaying pulses in the presence of background noises of various magnitudes is investigated in the present study by using statistical distributions. The focus is on the recovery of the instants when the jumps and pulses are created. The exponential-power and Student-t distributions, which are two heavily-tailed distributions, are chosen for the investigation. The uniform scale mixture representations of these distributions are derived. The Gibbs sampling technique is adopted in this chapter.

It is found that both the exponential-power and the Student-t distributions can recover the instant of the jumps and the initiation of the decaying pulses very well when the signal-to-noise ratio is higher than 3dB. For the decaying pulses, one requires a small β for the exponential-power distribution or a large degree of freedom for the Student-t distribution in order to achieve good recovery performance. The opposite is observed when the jumps in the rectangular waveforms are concerned.

The performance of both distributions deteriorates when the signal-to-noise ratio falls below 0dB. For this range of signal-to-noise ratio, a larger degree of freedom of the Student-t distribution tends to give better results, but this distribution type fails to indicate the initiation of the pulses and the location of jumps in the rectangular waveform for S/N < 0dB and -6dB respectively. The present results indicate that the exponential-power distribution appears to be a more robust distribution for the present purposes.

S/N (dB)	t_s (s)	t_e (s)	W (s)
+∞	1.000	1.400	0.400
+10	1.000	1.400	0.400
+6	1.000	1.400	0.400
+3	1.000	1.400	0.400
0	1.004	1.400	0.396
-3	1.007	1.392	0.385
-6	1.012	1.377	0.365

Table 4.1. Estimation of t_s and t_e for the rectangular waveform using the EP distributions ($\beta = 2$)

S/N (dB)	$\alpha = 1$			$\alpha = 30$			$\alpha = 100$		
	t_s (s)	t_e (s)	w (s)	t_s (s)	t_e (s)	w (s)	t_s (s)	t_e (s)	w (s)
+∞	1.000	1.400	0.400	1.000	1.400	0.400	1.000	1.400	0.400
+10	1.000	1.400	0.400	1.000	1.400	0.400	1.000	1.400	0.400
+6	1.000	1.400	0.400	1.000	1.400	0.400	1.000	1.400	0.400
+3	1.000	1.400	0.400	1.000	1.400	0.400	1.000	1.400	0.400
0	1.004	1.400	0.396	1.004	1.400	0.396	1.004	1.400	0.396
-3	-	-	-	1.010	1.392	0.382	1.012	1.392	0.380
-6	-	-	-	-	-	-	-	-	-

Table 4.2. Estimation of t_s and t_e for the rectangular waveform using the Student-t distributions

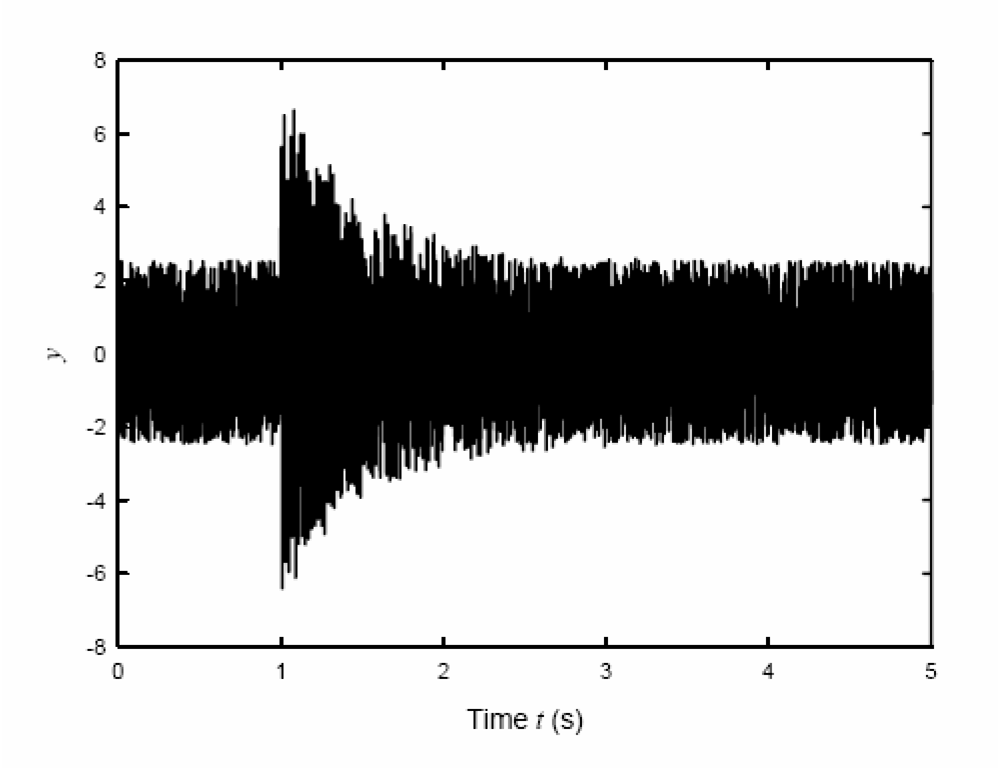

Figure 4.1 Decaying harmonic pulse in the presence of a background noise of S/N = 3dB.

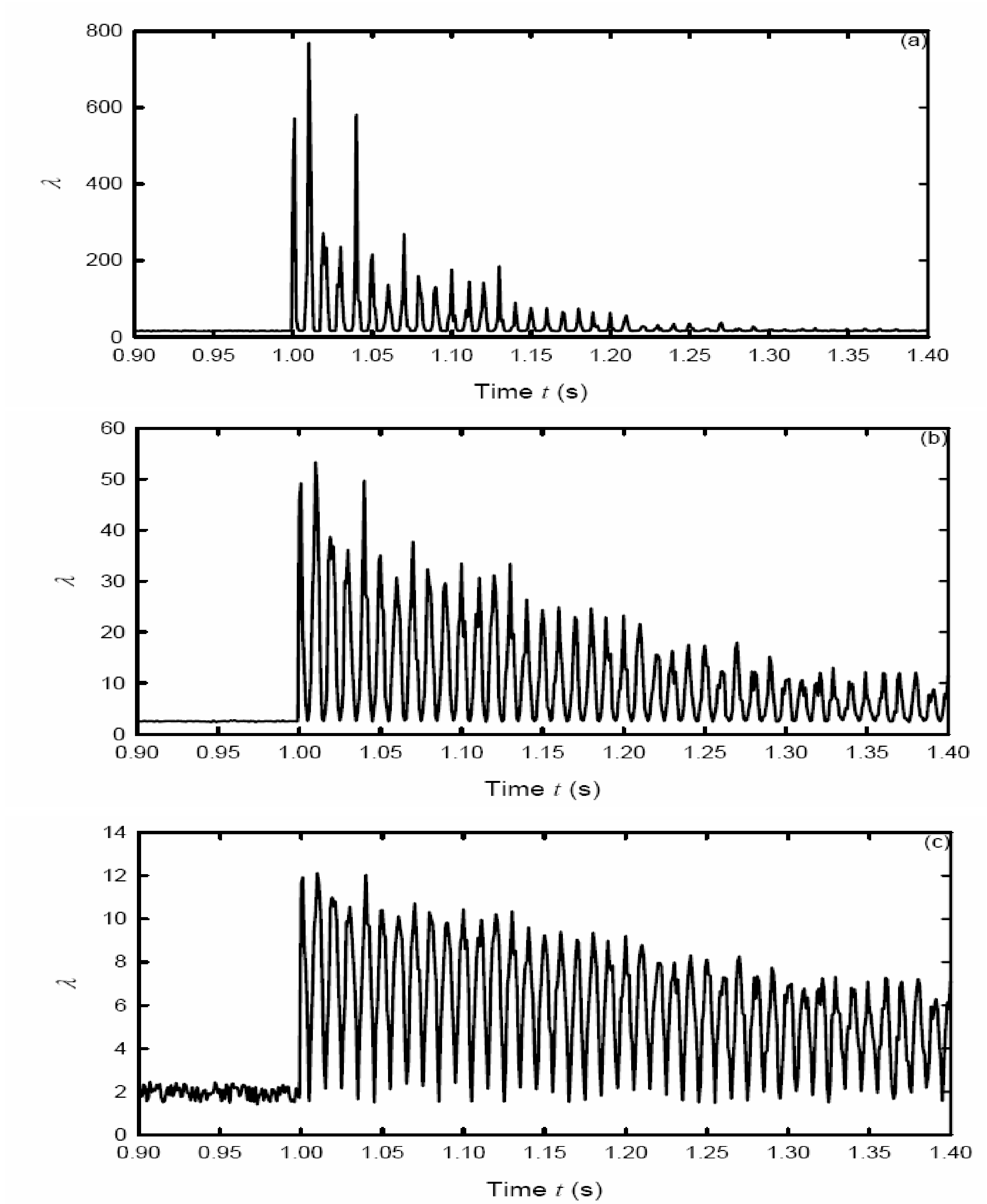

Figure 4.2 Time variations of λ for decaying pulse obtained using the EP distributions. (a) $\beta = 0.25$; (b) $\beta = 0.75$; (c) $\beta = 2$.

S/N = 10dB.

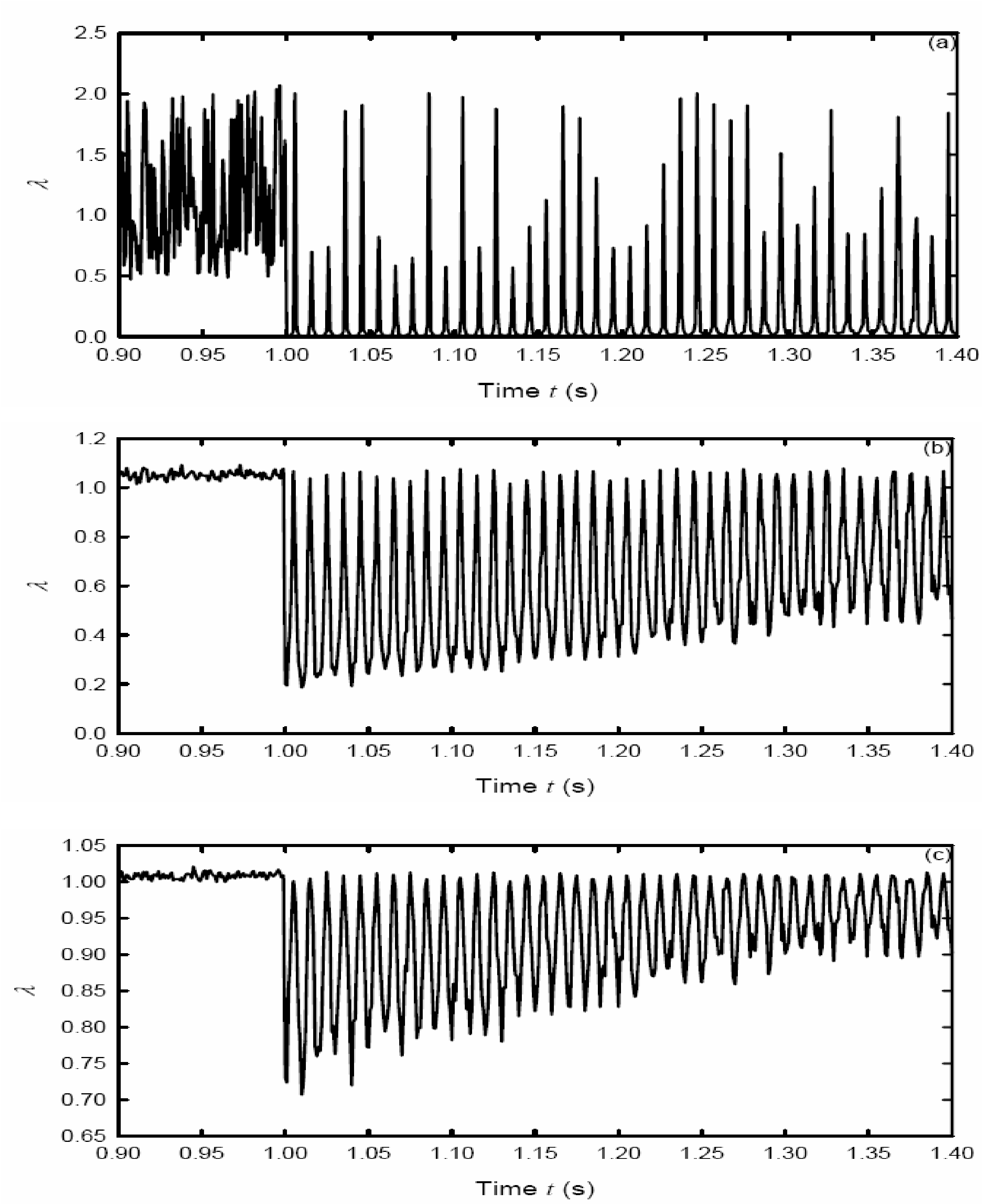

Figure 4.3 Time variations of λ for decaying pulse obtained using the Student-t distributions. (a) $\alpha = 1$; (b) $\alpha = 15$; (c) $\alpha = 100$. S/N = 10dB.

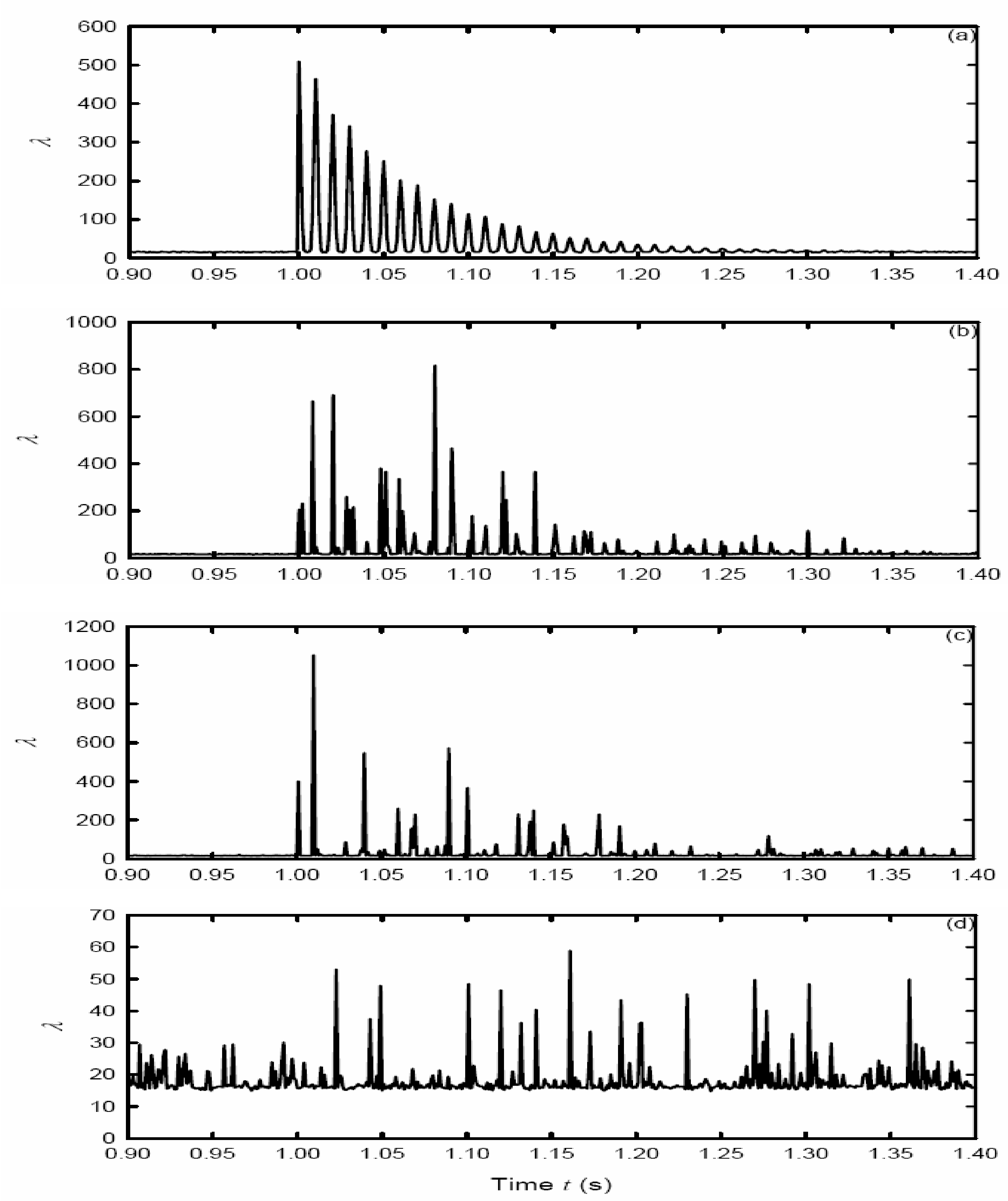

Figure 4.4 Effects of S/N on the time variations of λ for decaying pulse obtained using the EP distribution with $\beta = 0.25$.

(a) S/N = +∞dB; (b) S/N = 3dB; (c) S/N = 0dB; (d) S/N = -3dB.

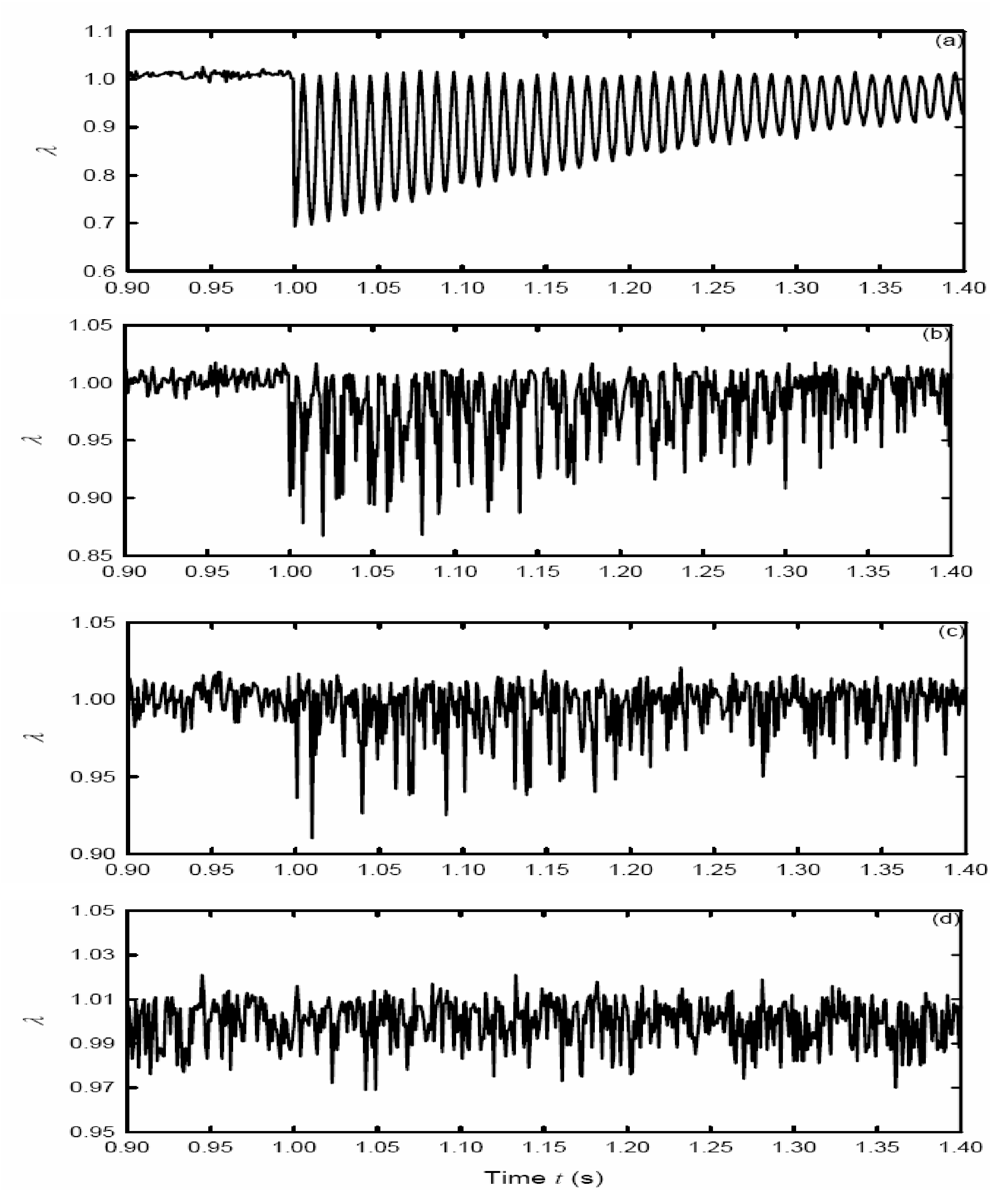

Figure 4.5 Effects of S/N on the time variations of λ for decaying pulse obtained using the Student-t distribution with $\alpha = 100$.

(a) S/N = +∞dB; (b) S/N = 3dB; (c) S/N = 0dB; (d) S/N = -3dB.

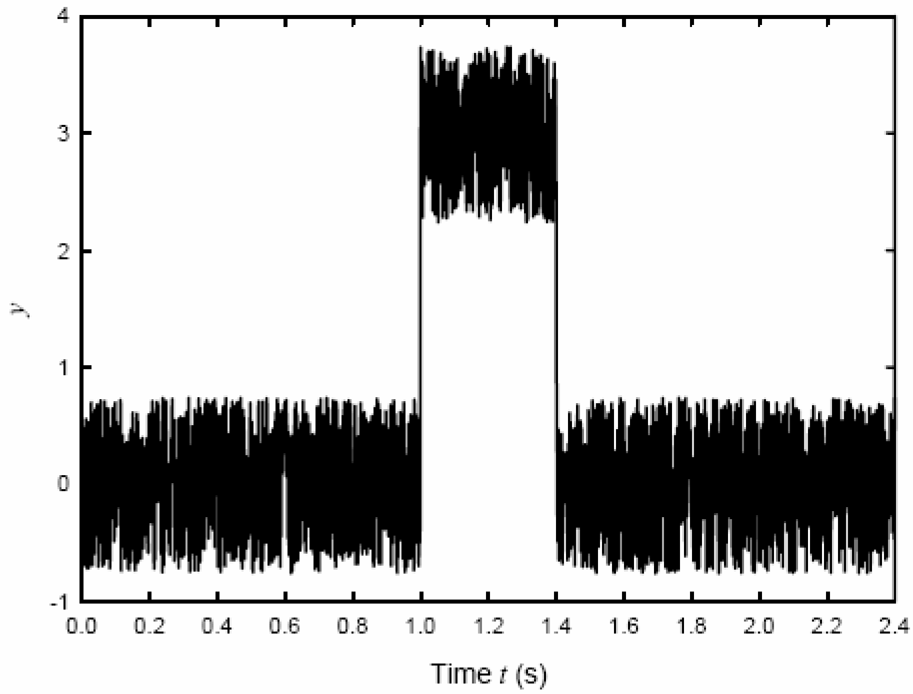

Figure 4.6 A rectangular waveform of width 0.4s with a background noise of S/N = 6dB.

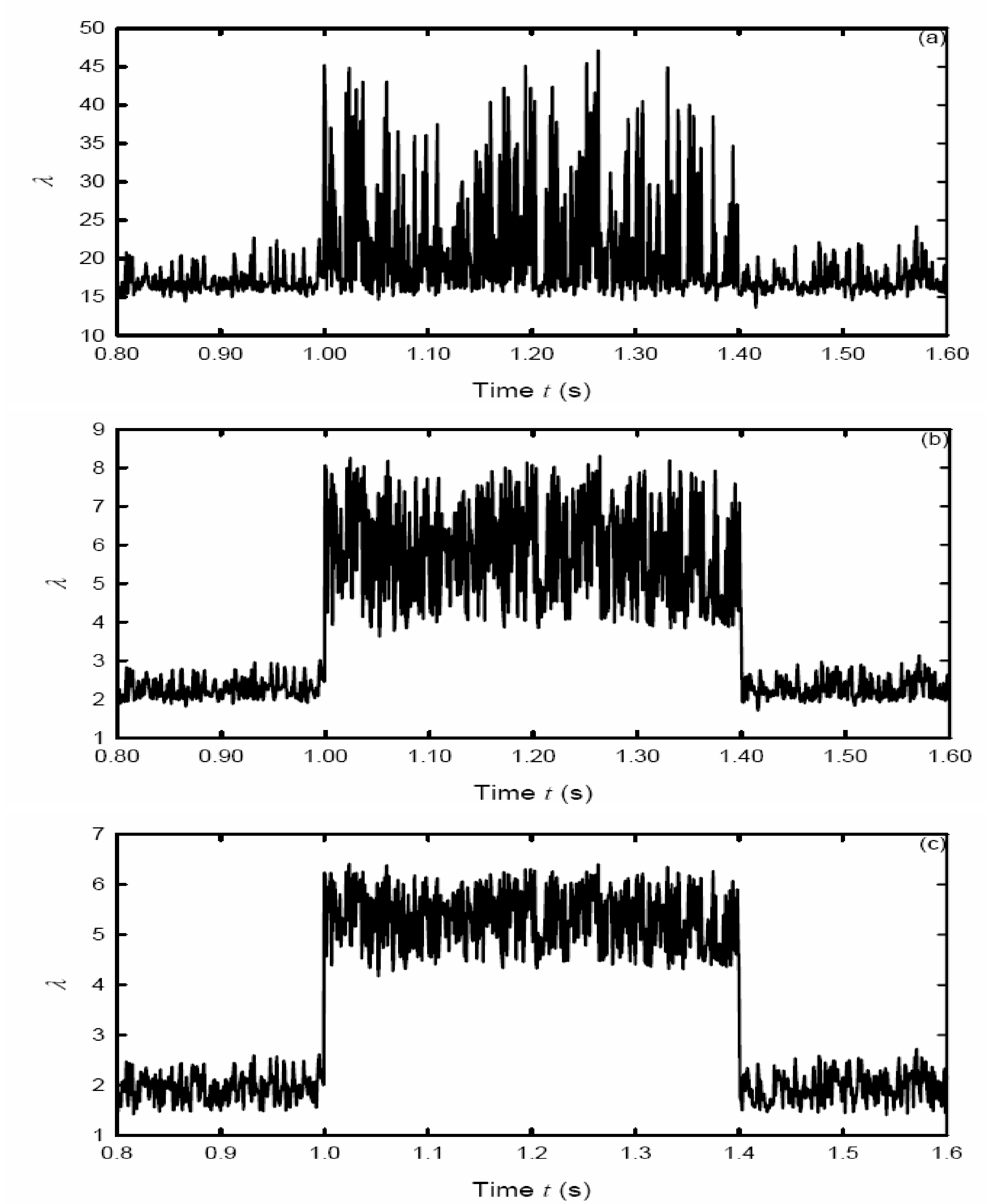

Figure 4.7 Time variations of λ for rectangular waveform obtained using the EP distributions. (a) $\beta = 0.25$; (b) $\beta = 1$; (c) $\beta = 2$. S/N = 6dB.

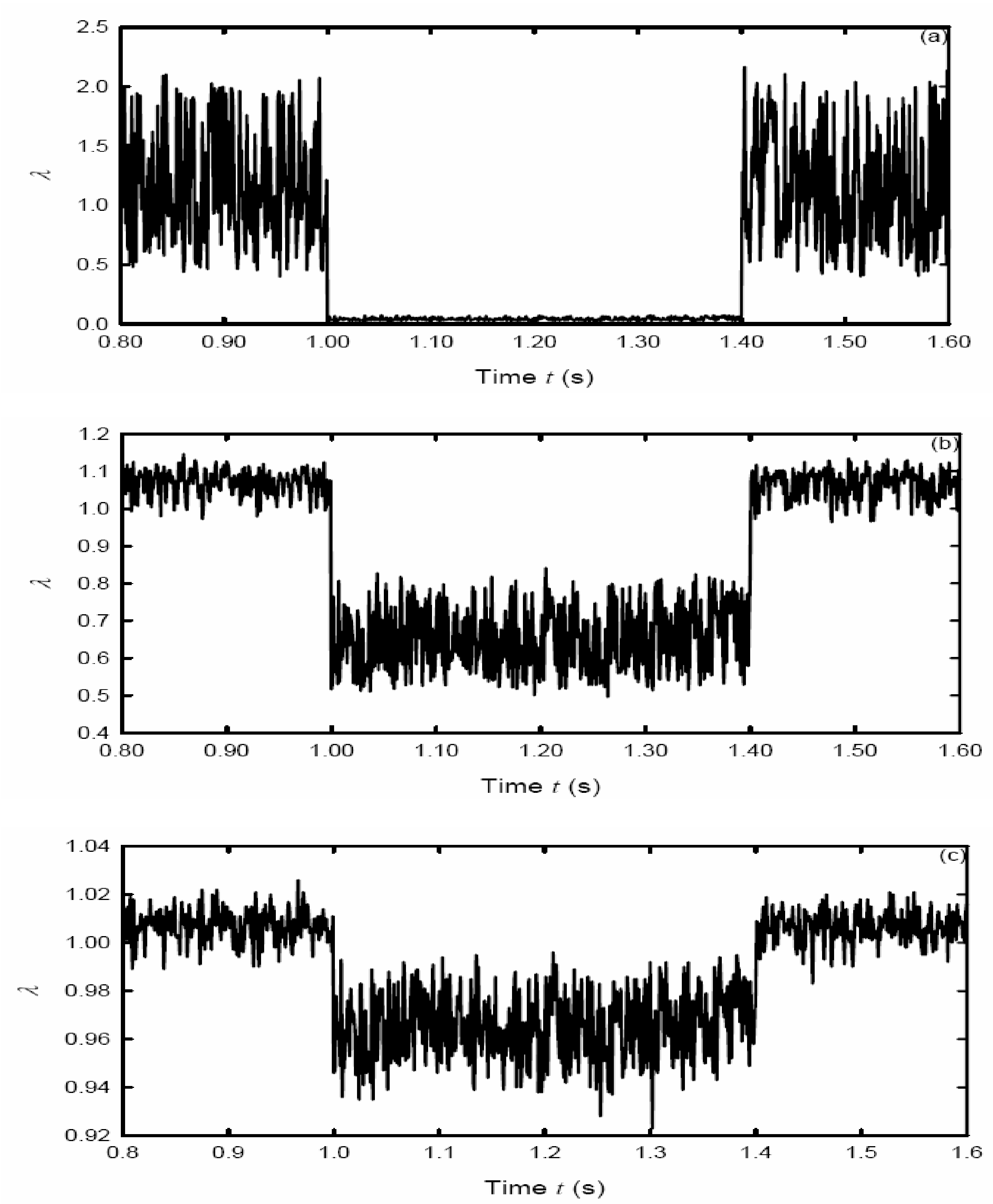

Figure 4.8 Time variations of λ for rectangular waveform obtained using the Student-t distributions. (a) $\alpha = 1$; (b) $\alpha = 10$; (c) $\alpha = 100$. S/N = 6dB.

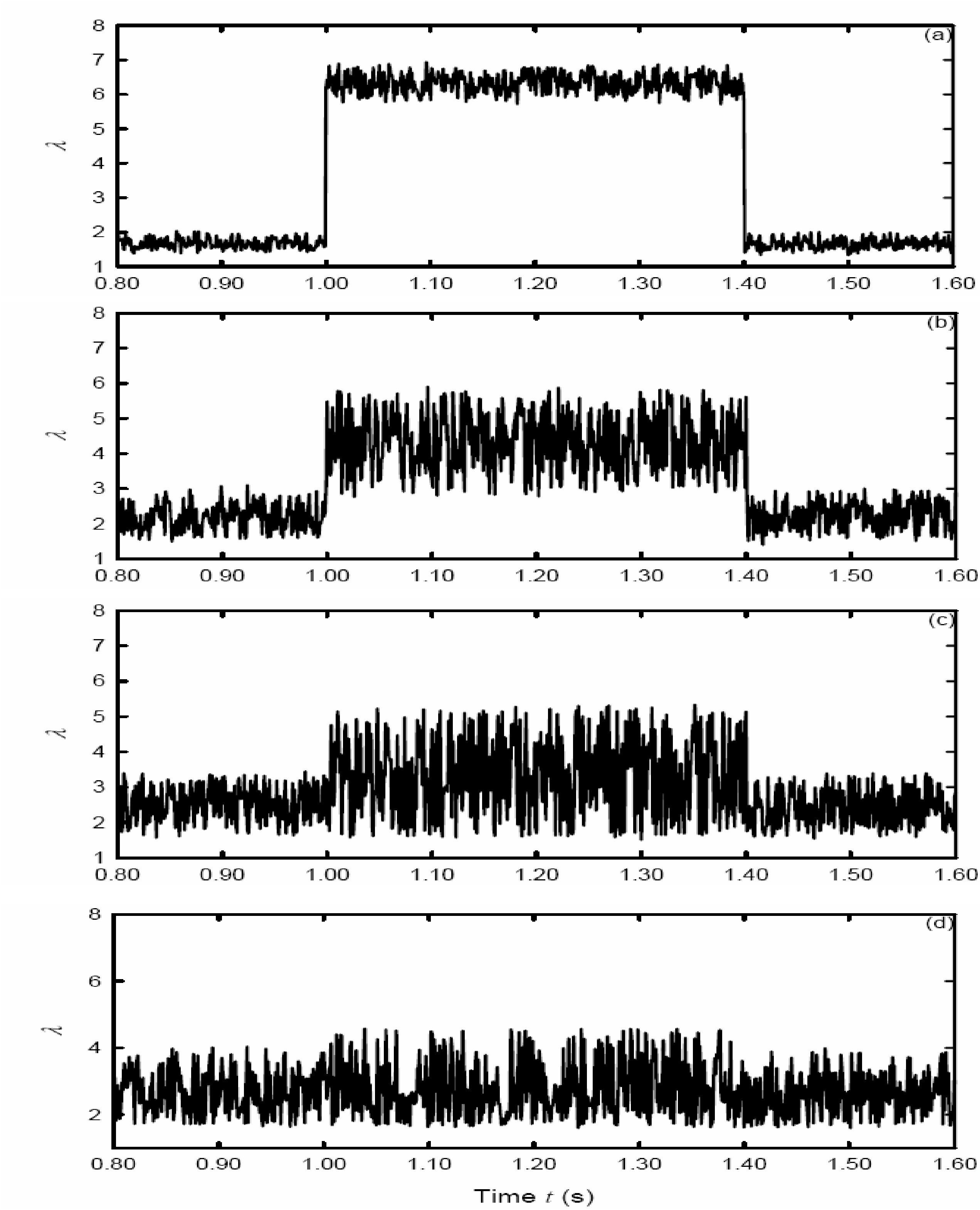

Figure 4.9 Effects of S/N on the time variations of λ for rectangular waveform obtained using the EP distribution with β = 2.S/N = 10dB; (b) S/N = 3dB; (c) S/N = 0dB; (d) S/N = -6dB.

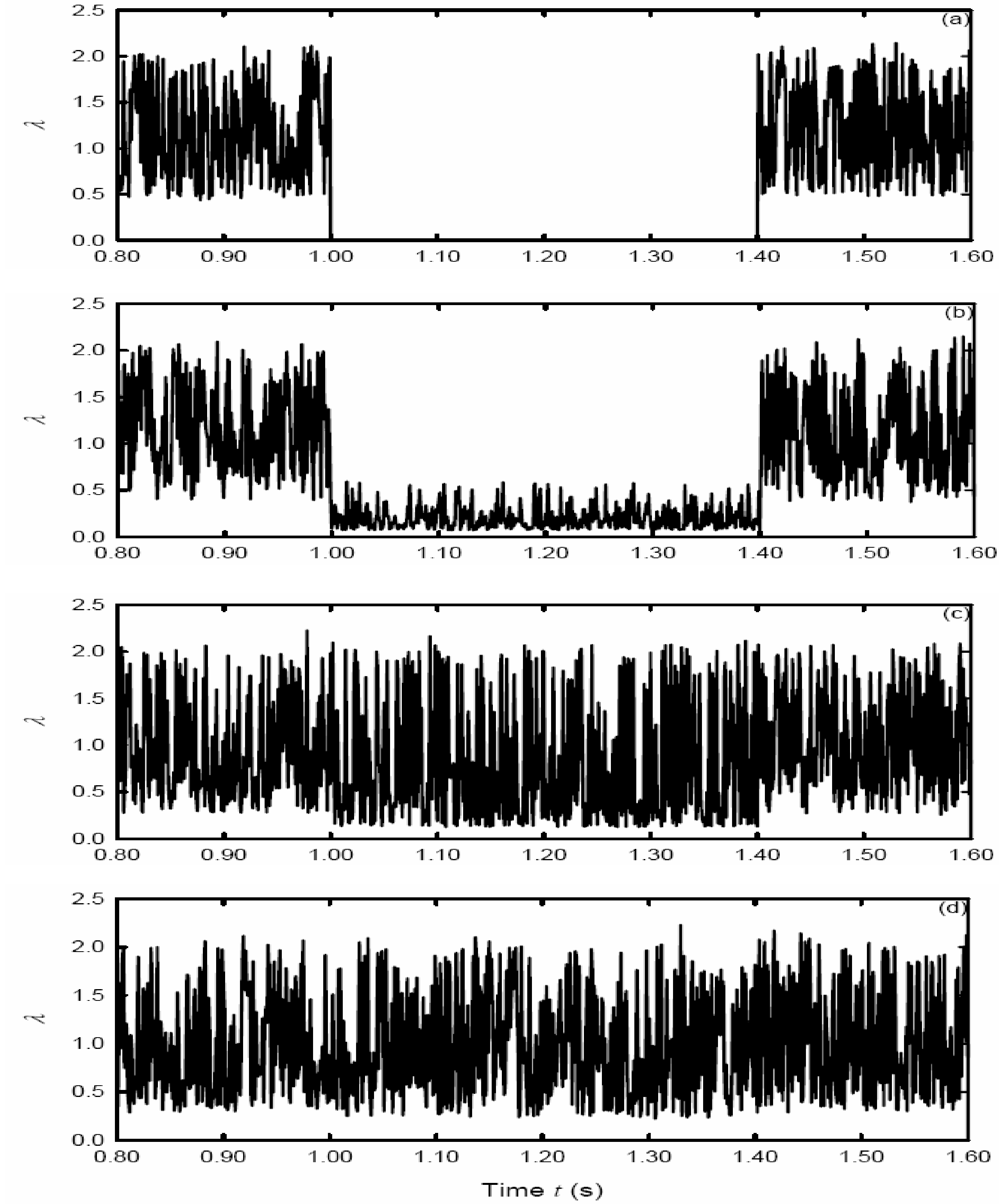

Figure 4.10 Effects of S/N on the time variations of λ for rectangular waveform obtained using the Student-t distribution with $\alpha = 1$.

(a) S/N = 10dB; (b) S/N = 3dB; (c) S/N = 0dB; (d) S/N = -6dB.

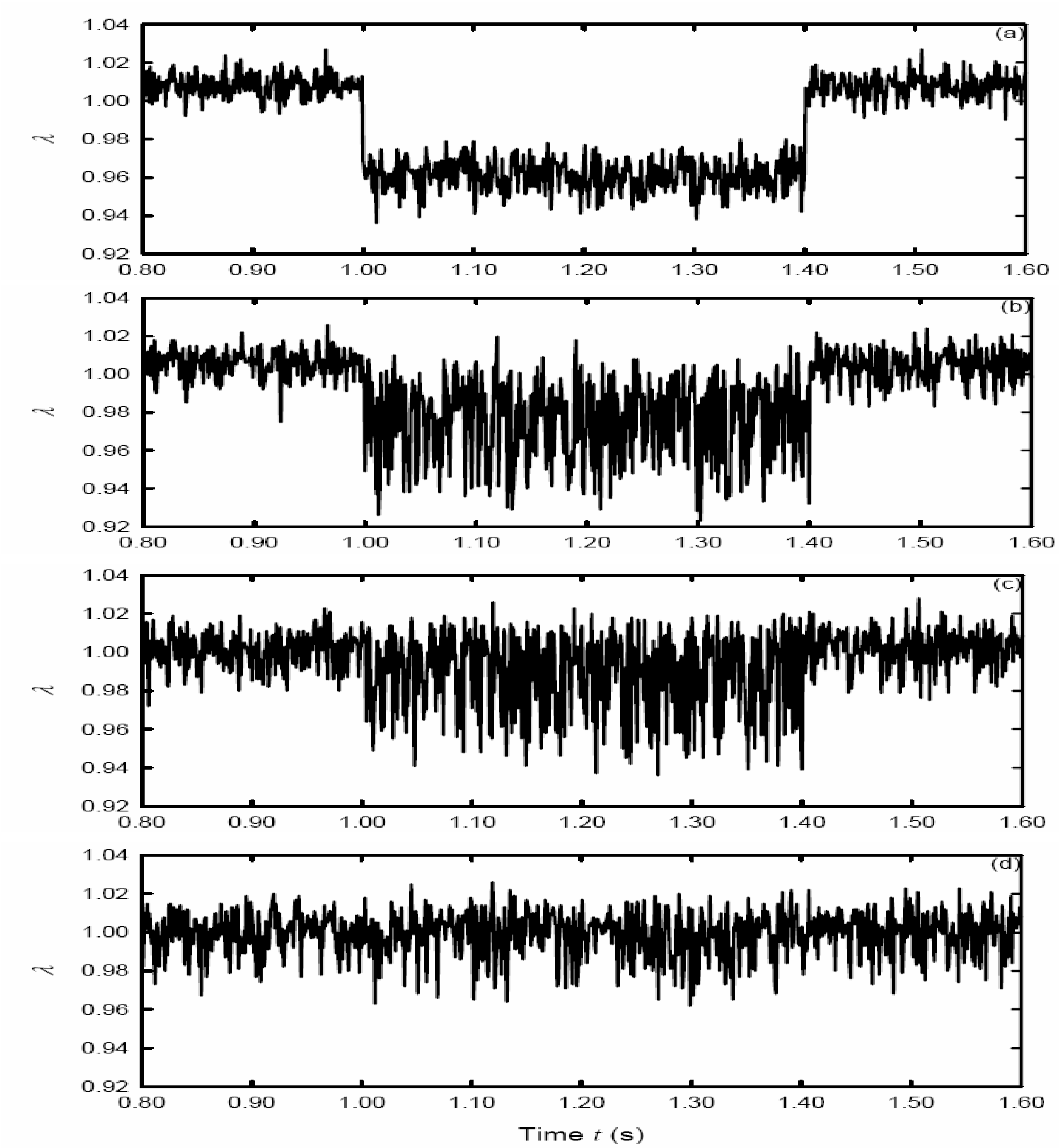

Figure 4.11 Effects of S/N on the time variations of λ for rectangular waveform obtained using the Student-t distribution with $\alpha = 100$.

(a) S/N = 10dB; (b) S/N = 3dB; (c) S/N = 0dB; (d) S/N = -6dB.

Chapter 5

On Weak Unsteady Signal Detection using Statistical Tests

5.1 Introduction

When there is a sudden change to a normally running system, a jump or an abrupt change to the signals from the system will be resulted. Successful and early detection of such changes is very important as some changes can be detrimental to the system. It is particularly true in building services engineering when some parts of a piece of building services equipment start to go wrong but the initial fault signals can be very weak compared to other signals within the same equipment [27]. The fault then grows in size (magnitude) and finally causes damage to or even destruction of the equipment.

It is believed that the statistical structure of a signal will change when an additional signal is introduced into it. Therefore, it is expected that statistical tests will be very useful in detecting these changes. Sensitivities of the tests in manifesting statistical changes are also an important issue, especially for the process of machine condition monitoring. In this chapter, two statistical tests developed by Jarque and Bera [28] and D'Agostino [29] will be adopted. Their performance in locating the instant of change under different signal-to-noise ratios will be investigated. The background noise is assumed to be stationary. This is basically acceptable at least in building services engineering where the various parts of services equipment are operated in pre-determined cycles and thus the background signals usually have regular time patterns.

5.2 The Statistical Tests and Indicators

A statistical analysis of the problem of detecting the presence of unknown weak signals in noisy observations is adopted. The approach is based on Jarque-Bera Test (*JB*) and D'Agostino's Test (*D*). The *JB* test is proposed by Jarque and Bera [28]. It is based on the difference between the skewness and kurtosis of the data set and those of the assumed normal distribution. The *D* test is developed by D'Agostino [29] and has been used for testing normality.

5.2.1. Jarque-Bera test (JB)

In statistics, the *JB* test is a goodness-of-fit measure of departure from normality, based on the sample kurtosis and skewness of the data set and those of the assumed normal distribution [28,30]. The *JB* test evaluates the hypothesis that the data set has a normal distribution with unspecified mean and variance, against the alternative that the data set does not have a normal distribution. The test is based on the sample skewness and kurtosis of the data. For a true normal distribution, the sample skewness should be near to 0 and the sample kurtosis should be around 3. The *JB* test determines whether the sample skewness and kurtosis are unusually different than their expected values, as measured by a chi-square statistic.

The test for normality works well in the general case. The test is based, first, on independent random variables and, second, on the residuals in the classical linear regression. The power of the *JB* test is good for distributions with short tails, especially if the shape is bi-modal.

The test is well known to have very good power properties in the testing for normality. It is clearly easy to compute and is commonly used in the regression context in econometrics. One limitation of the test is that it is designed only for testing normality, while the empirical likelihood ratio test can be applied to test for any types of underlying distribution with some appropriate modification to the moment equations. The JB test statistic is

$$JB = n\left(\frac{\alpha_3^2}{6} + \frac{(\alpha_4 - 3)^2}{24}\right), \tag{5.1}$$

where $\alpha_3 \equiv s^{-3}n^{-1}\sum_{i=1}^{n}(y_i - \bar{y})^3, \alpha_4 \equiv s^{-4}n^{-1}\sum_{i=1}^{n}(y_i - \bar{y})^4, s^2 \equiv n^{-1}\sum_{i=1}^{n}(y_i - \bar{y})^2$, $\bar{y}$ is the sample mean and s^2, α_3 and α_4 are the second, third and fourth sample moments about the mean respectively. The JB statistics has an asymptotic distribution, which is $\chi^2_{(2)}$ under the null hypothesis.

5.2.2 D'Agostino's (D) Test

The D test of normality is based on the ratio of Gini's mean difference to the sample standard deviation [29]. Although the test is known to be inconsistent for some skew alternatives, a systematic study has yet to be made for normal mixtures. The test for skewness is good at detecting non-normality caused by asymmetry. The test is a powerful test for normality. It is not as powerful for detecting skewness.

Suppose $y_1, y_2, \ldots, y_n$ is the data set. $y_{1,n}, y_{2,n}, \cdots, y_{n,n}$ are the ordered observations, where $y_{1,n} \le y_{2,n} \le \cdots \le y_{n,n}$. The D test statistics takes the form :

$$D = T/(n^2 s), \tag{5.2}$$

where s is the sample standard deviation and $s^2 \equiv n^{-1}\sum_{i=1}^{n}(y_i - \bar{y})^2$, and $T = \sum_{i=1}^{n}\left\{i - \frac{n+1}{2}\right\}y_{i,n}$. If the sample is drawn from a normal distribution, then for sufficiently large sample size n,

$$E(D)=\frac{(n-1)\Gamma\left(\frac{n}{2}-\frac{1}{2}\right)}{2\sqrt{2n\pi}\Gamma\left(\frac{n}{2}\right)}\approx\frac{1}{2\sqrt{\pi}}\approx 0.28209479\,, \tag{5.3}$$

where Γ is the Gamma function. The asymptotic standard deviation of the D test statistics is :

$$asd(D)\approx 0.02998598\big/\sqrt{n}\ . \tag{5.4}$$

The standardized D test statistics, D^*, is

$$D^*=\frac{D-E(D)}{asd(D)}\,. \tag{5.5}$$

Under the null hypothesis, D^* is asymptotically distributed as $N(0,1)$. If the sample is drawn from a distribution other than normal, $E(D^*)$ tends to differ from zero. If the underlying distribution has greater than normal kurtosis, then $E(D^*) < 0$. If it has less than normal kurtosis, then $E(D^*) > 0$.

D test is a test of normality based on order statistics from sample data. It is a modification of the Shapiro-Wilk test [30], and it is readily calculated without the coefficients of the order statistics. It is based on the ratio of a linear unbiased estimator of the standard deviation (using order statistics) to the usual mean square estimator. The test was originally proposed for moderate sample sizes and can detect departures from normality both for skewness and kurtosis.

The Shapiro-Wilk test and the D test are goodness-of-fit statistical tests recommended for testing if the underlying probability density function for a data set has a normal (or lognormal) distribution. The test is a goodness-of-fit test that can be used for data sets larger than 50, which is the upper limit for the Shapiro-Wilk test. The D test can be used with data sets numbering between 50 and 1000 samples [31].

It should also be noted that the power of the empirical likelihood ratio test studied by Dong and Giles [32] is even lower than that of the *D* test for small sample sizes.

5.3 Numerical Examples and Discussions

The performance of the two statistical test indicators (*JB* and *D**) on detecting weak signal is illustrated through using a data set consists of an abruptly generated exponentially growing sinusoidal wave and a stationary background noise. The signal series {*y*} takes the form :

$$y_i = AH(i-i_s)e^{\gamma(i-i_s)}\cos(\omega(i-i_s)) + \theta\varphi_i \quad \text{for } i = 1 \text{ to } n, \tag{5.6}$$

where i_s is the location index of wave initiation, *H* the unit step function, γ the growth rate of the wave and ω the angular speed of the wave. {φ} is a stationary noise with magnitude bounded between ±1 and θ is a number which determines the strength of this noise. Without lost of generality, *A*, γ, i_s and ω are fixed at 0.01, 0.002, 1001 and 0.1π respectively.

Each stationary noise {φ} used in the present study is created by repeating one set of random data generated by MATLAB and $\overline{\varphi_i} = 0$. The signal-to-noise ratio S/N is defined as $10\log_{10}(A/\theta)$ and the unit is dB. This signal type {*y*} is of special implication to the engineering field, especially in the situation when a fault is being developed inside a building services system [26]. The early detection of such change is very important for machine condition monitoring [27]. The focus here is to determine the instant i_s at which the wave, which will jeopardize system operation when it grows to some magnitude level, is created. The effects of S/N on the signal detection will be discussed.

Figure 5.1 shows {*y*} with θ = 0.5, giving a S/N of −17dB. The background random noise is created by repeating 20 random data generated by MATLAB. One can observe that the signal is very weak when compared to the noise even up to a time step of 2000. The conventional test of variance in engineering is obviously not capable of detecting the initiation of the growing wavy signal, but can only indicate such change when the wave has grown to a magnitude comparable to that of the stationary noise.

The growing wave results in a sudden change in the D^* at the instant $i = i_s$ at $\theta =$ 0.02 (S/N = –3dB) with n = 20 as shown in Figure 5.2. This is the instant when the moving time frame starts to include the wave. The initial value of the D^* depends on the stationary data and thus can vary substantially. However, the abrupt sharp change in D^* remains very clear. The effect of θ on the recovery of i_s is illustrated in Figure 5.3. The corresponding data for θ = 0.02 can be found in Figure 5.2. Though the relative magnitude of the D^* change is reduced as θ increases in general, it is still observable up to θ = 0.5. It is observed that D^* approaches 0.05 when the exponentially growing signal dominates the data set $\{y\}$. One can observe the instant i_s more easily when the rate of change of D^* is adopted as the indicator as shown in Figure 5.4. The change is clearly recovered even at θ = 0.5 (Figure 5.4b). Perhaps it is not very surprising that a change in D^* can be found, but the D test appears to be very sensitive to the introduction of an alien signal as the error in i_s recovery is negligible even at a S/N of –17dB if one uses the instant of the first change in D^* as the indicator. The corresponding error is about 3 time steps (15% of n) if one chooses the maximum local absolute rate of change of D^* for the purpose. A jump of the rate of change of D^* can also be observed at θ = 10 (S/N = –30dB) as shown in Figure 5.4c, but the corresponding jump cannot be clearly seen in the D^* time variation. The corresponding error in the i_s recovery is less than 13 time step

(~65% of n). However, these errors depends on the background noise and will be discussed further later.

One can observe that the JB test produces similar results, but the JB jump is distinctively observable here for $\theta = 0.5$ as shown in Figure 5.5a. It is noted that JB approaches 1.88 when the wave dominates the data set $\{y\}$. The corresponding rate of change of JB is illustrated in Figure 5.5b.

It is found that both JB and D^* provide an accurate location of the wave initialization if one adopts the first abrupt change of their rates of change from zero as the indicator. This phenomenon is independent of the stationary noise used. However, the stationary background noise will play a role in the recovery process if one chooses to use the maximum local rate of change to confirm the existence of the additional wavy signal together with the first abrupt change. To study this effect, 10 different sets of stationary noises are used and the average results with simple statistics are tabulated in Table 5.1. The average errors and their standard deviations are rounded up to the nearest integers (time steps). The numbers in parenthesis are the results obtained by using the first large change of the rate of change of JB or D^* after their first abrupt change from zero. It is observed that the maximum rate of change of the two statistical parameters is not a good indicator of i_s. The latter, that is the large change in the rate of change of D^* or JB, appears to be a better choice. Also, the results become steady when the S/N drops below −20dB. Since the errors are not normally distributed, the corresponding skewness and kurtosis are presented so as to indicate the error distributions.

One can observe from Table 5.1 that on average JB performs better that D^* if one adopts the abovementioned large change of the rate of change of JB and D^* as the wave initiation indicator, especially when θ increases beyond 0.5 (S/N < −17dB).

The more positive skewness also suggests that the mode of the error distribution is smaller than the mean error. The larger kurtosis indicates that the error values are more concentrated around the mode.

Results in Table 5.1 show that the recovery of i_s using JB and D^* depends on the background noises, even when the noises are of the same strengths. In order to minimize the background noise effect, a new parameter CT is proposed :

$$CT = \frac{D^*}{D^*_o} - \frac{JB}{JB_o}, \tag{5.7}$$

where the suffix o denotes the quantity before the introduction of the wave (that is, for $i < i_s$). One can find in Table 5.1 that CT gives a more stable and a slightly better prediction of i_s than JB on average.

Figures 5.6a to 5.6c show the variations of the rate of change of CT, D^* and JB in the presence of the same background noise with $\theta = 10$ respectively. The ordinates are normalized by the corresponding maximum magnitude of the rate of change for $900 < i < 1100$. In this example, D^* is doing better than JB and it is found that the rate of change CT follows that of D^* very closely. Another example for $\theta = 10$ is illustrated in Figure 5.7. This time JB gives better recovery of i_s, while D^* results in ambiguous information around $i \sim i_s$. First large change of the rate of change of CT (that is, the second time derivative of CT) appears to take place closer to i_s than that predicted by D^*, though time variation of the rate of change of CT still follows roughly that of D^*. It appears that CT has inherited the advantages of both JB and D^* and thus its better performance on average.

5.4 Conclusion

Two statistical tests, namely the Jarque-Bera test and D'Agostino test, which have been adopted in the past for checking data normality, are used for recovering the instant of introduction of an exponentially growing sinusoidal wave in the presence of a stationary background noise in this chapter. This type of unsteady signal is of prime importance in engineering. The effects of the signal-to-noise ratio on the recovery are also examined. A moving time frame of constant width is adopted in the calculation of the time variations of the statistical test parameters.

It is found that there is a jump when the calculation starts to include the wave. However, the magnitude of the jump depends on the background noise. The second time derivative of the statistical test parameter is found to be a more appropriate choice for the recovery of the wave initialization instant. Though it may not be very surprising that there is a change in the statistical parameter when an alien signal is introduced to the ordinary noise, it is the sensitivity of the change that is impressive. The Jarque-Bera test parameter gives better prediction. The error in the recovery is 3 time steps on average even at a signal-to-noise ratio of –30dB.

A new parameter derived from the parameters from the two tests is introduced. This parameter appears to have inherited the advantages of the two abovementioned statistical test parameters. The corresponding average errors are slightly smaller than those from the Jarque-Bera test parameter. However, the errors from the newly proposed parameter show insignificant dependence on the background noise.

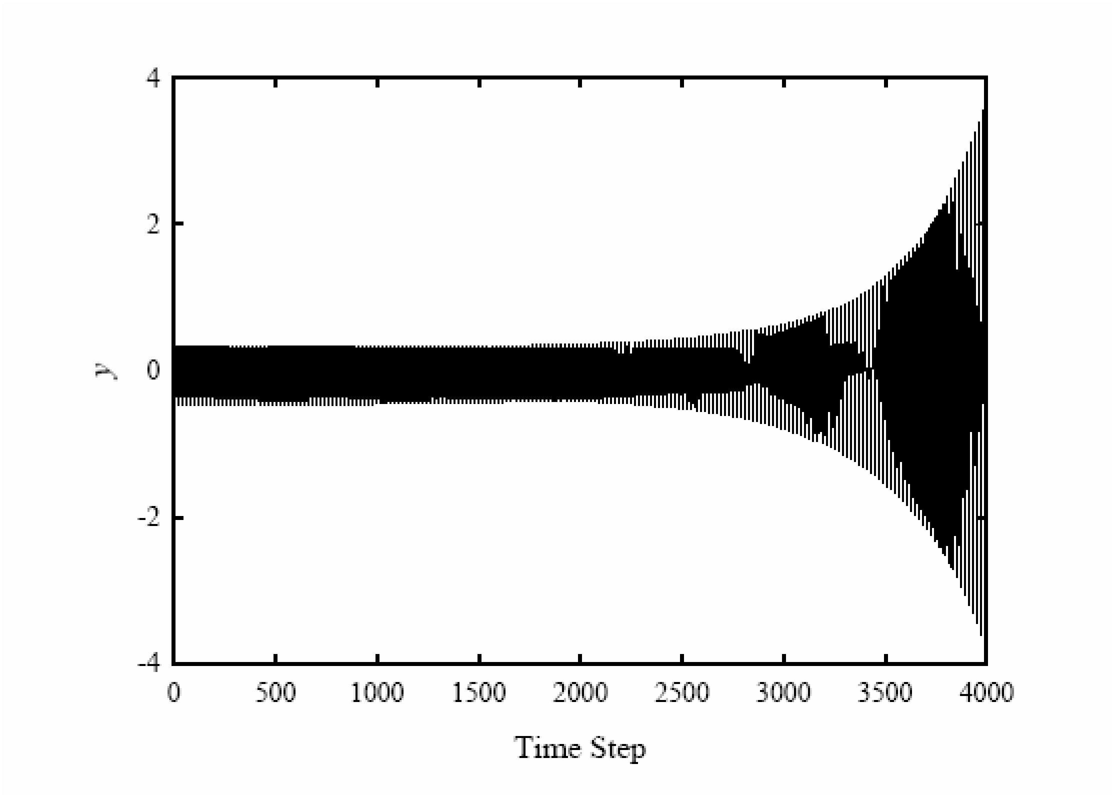

Figure 5.1 Data set $\{y\}$ with $\theta = 0.5$.

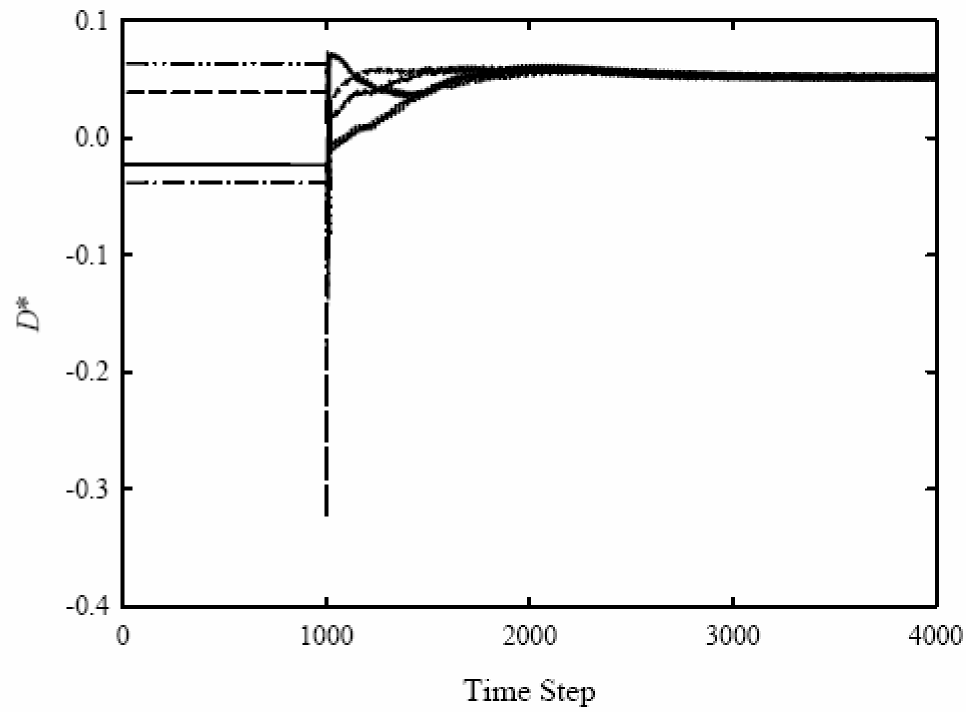

Figure 5.2 Examples of D^* variations for $\theta = 0.02$.

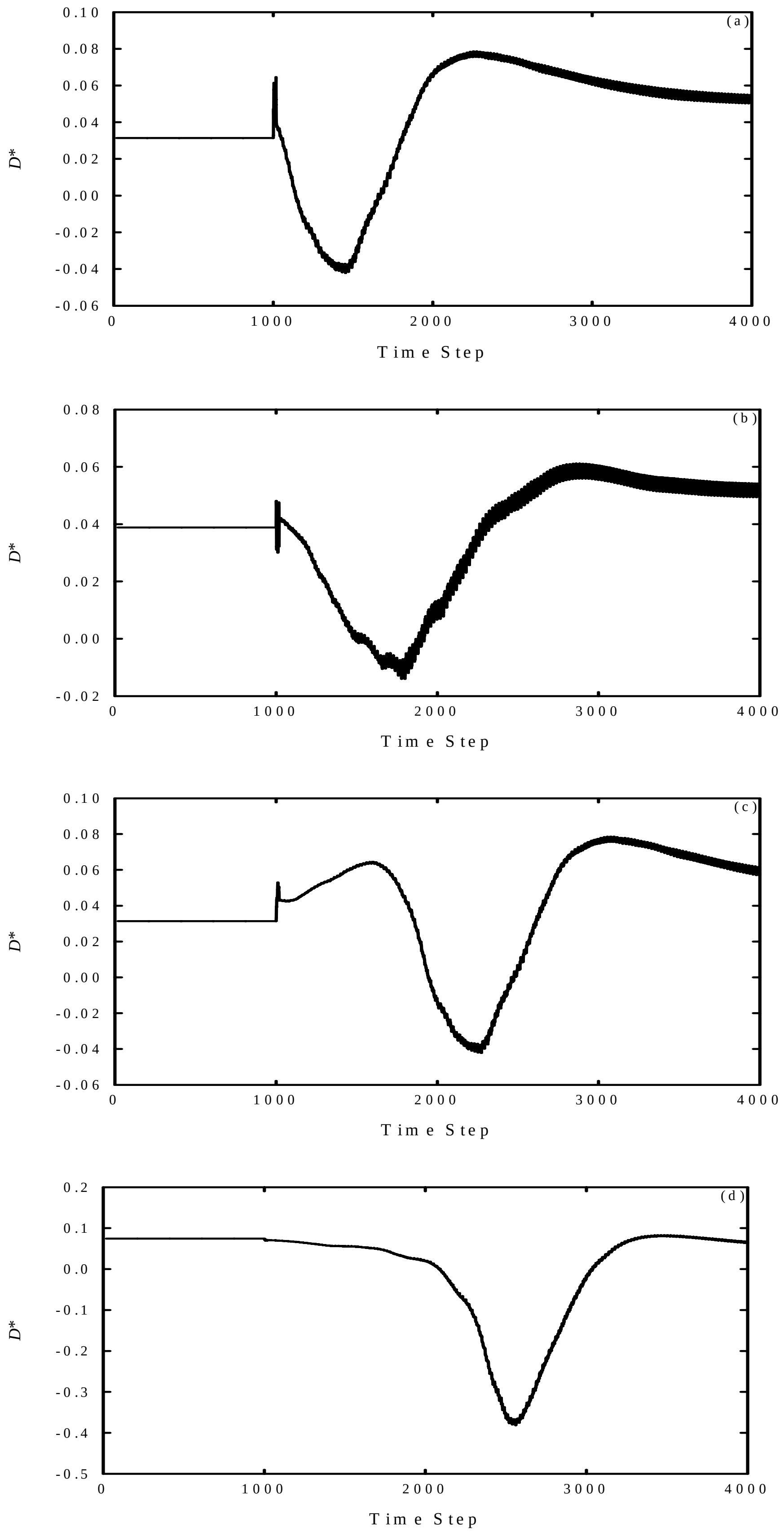

Figure 5.3 Effects of background stationary noise magnitude on D^* variations.

(a) $\theta = 0.05$; (b) $\theta = 0.1$; (c) $\theta = 0.25$; (d) $\theta = 0.5$.

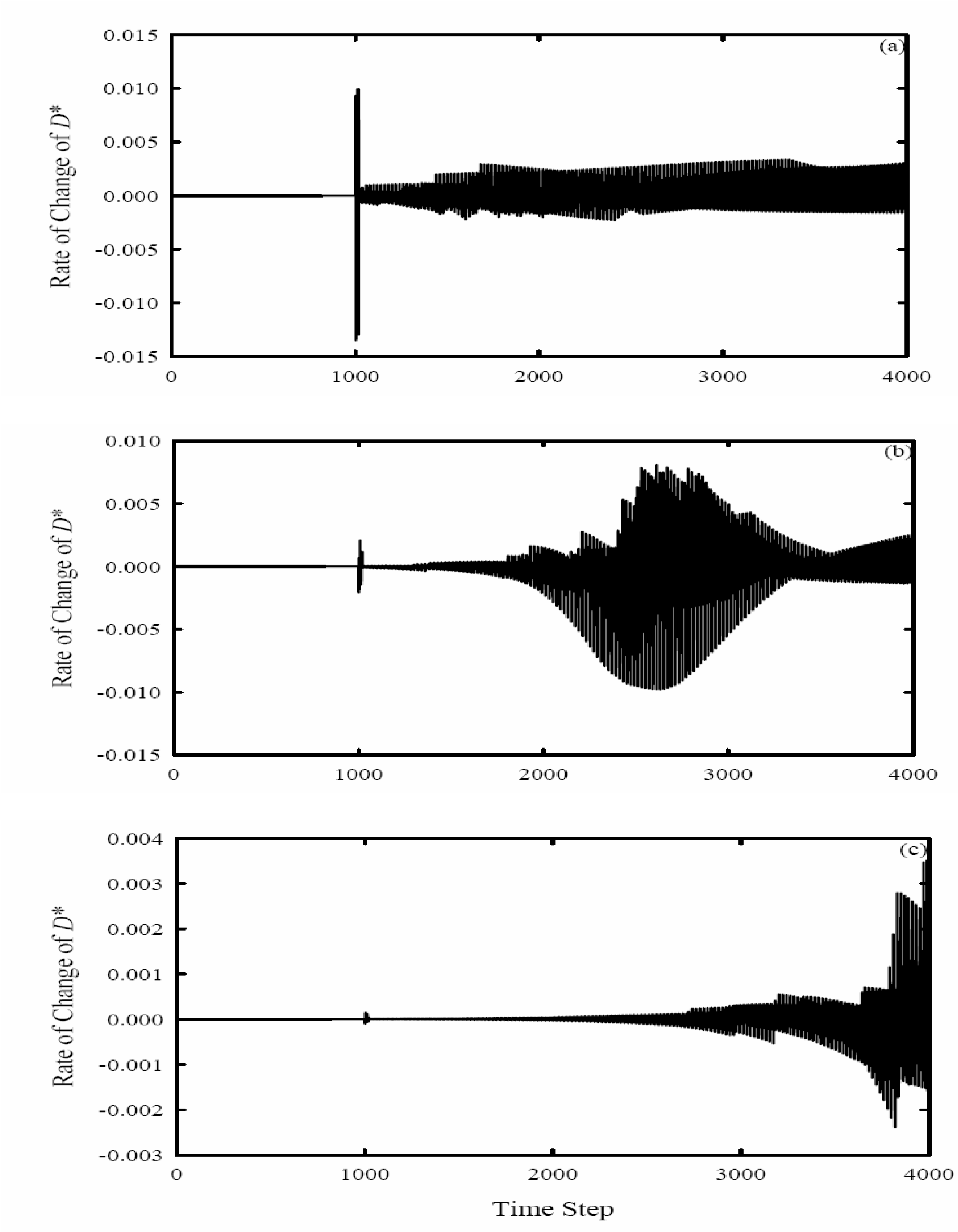

Figure 5.4 Rate of change of D^*.

(a) $\theta = 0.1$; (b) $\theta = 0.5$; (c) $\theta = 10$.

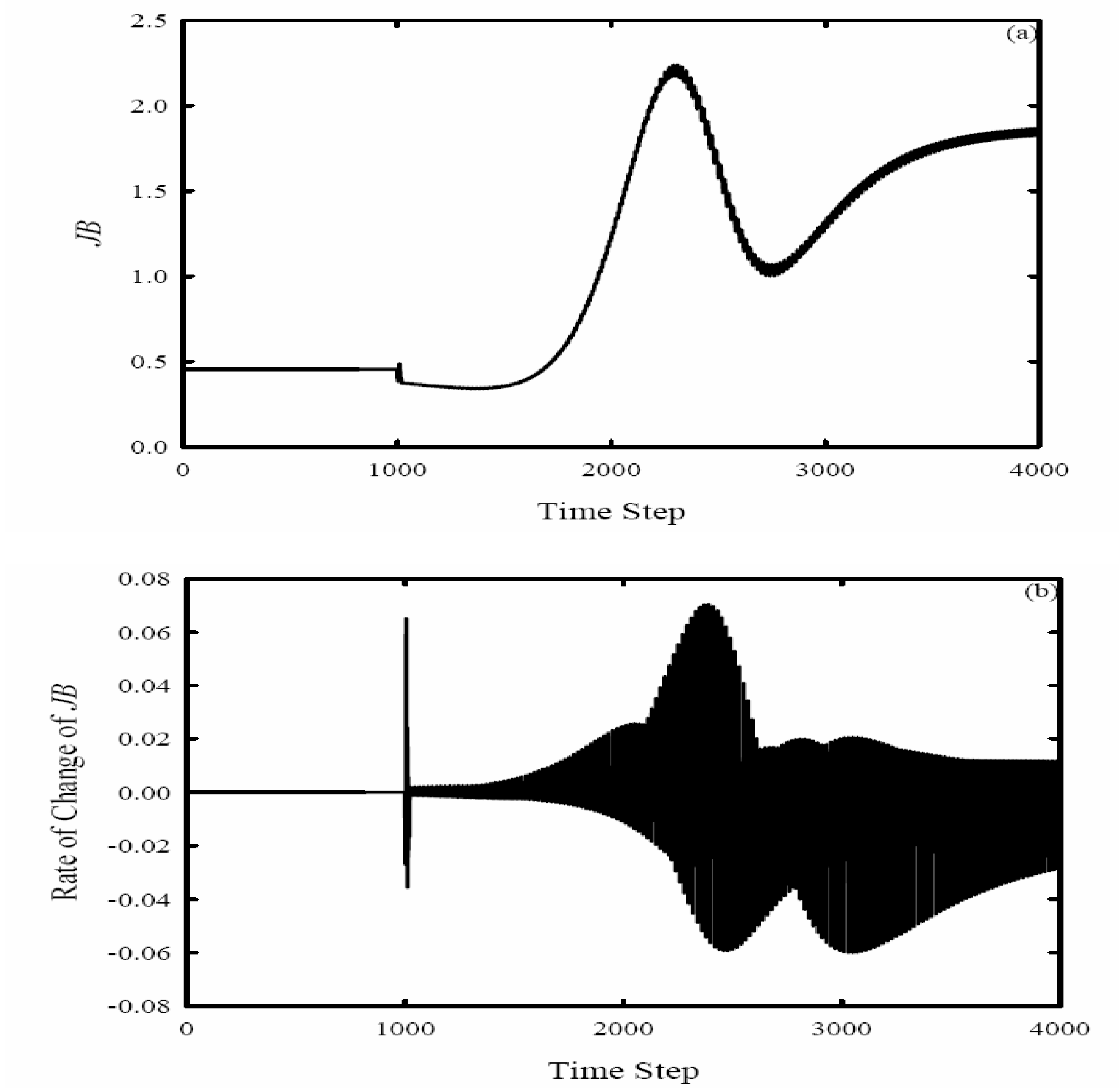

Figure 5.5 Performance of *JB* test.

(a) *JB* variation; (b) variation of the rate of change of *JB*.

$\theta = 0.5$.

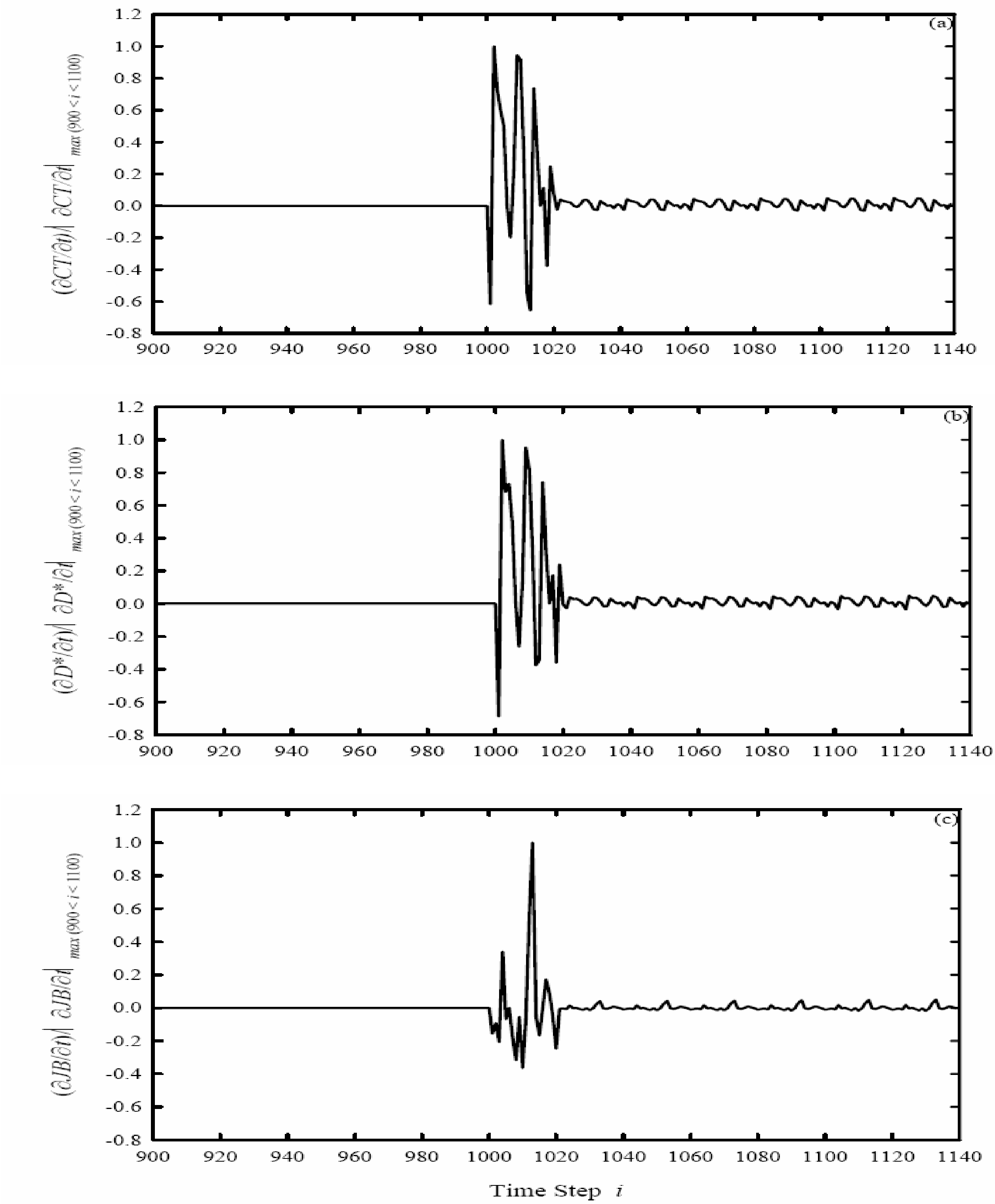

Figure 5.6 Example of the time variation of the rate of change of indicator around i_s where D^* gives better performance than JB.

(a) CT; (b) D^*; (c) JB. $\theta = 10$.

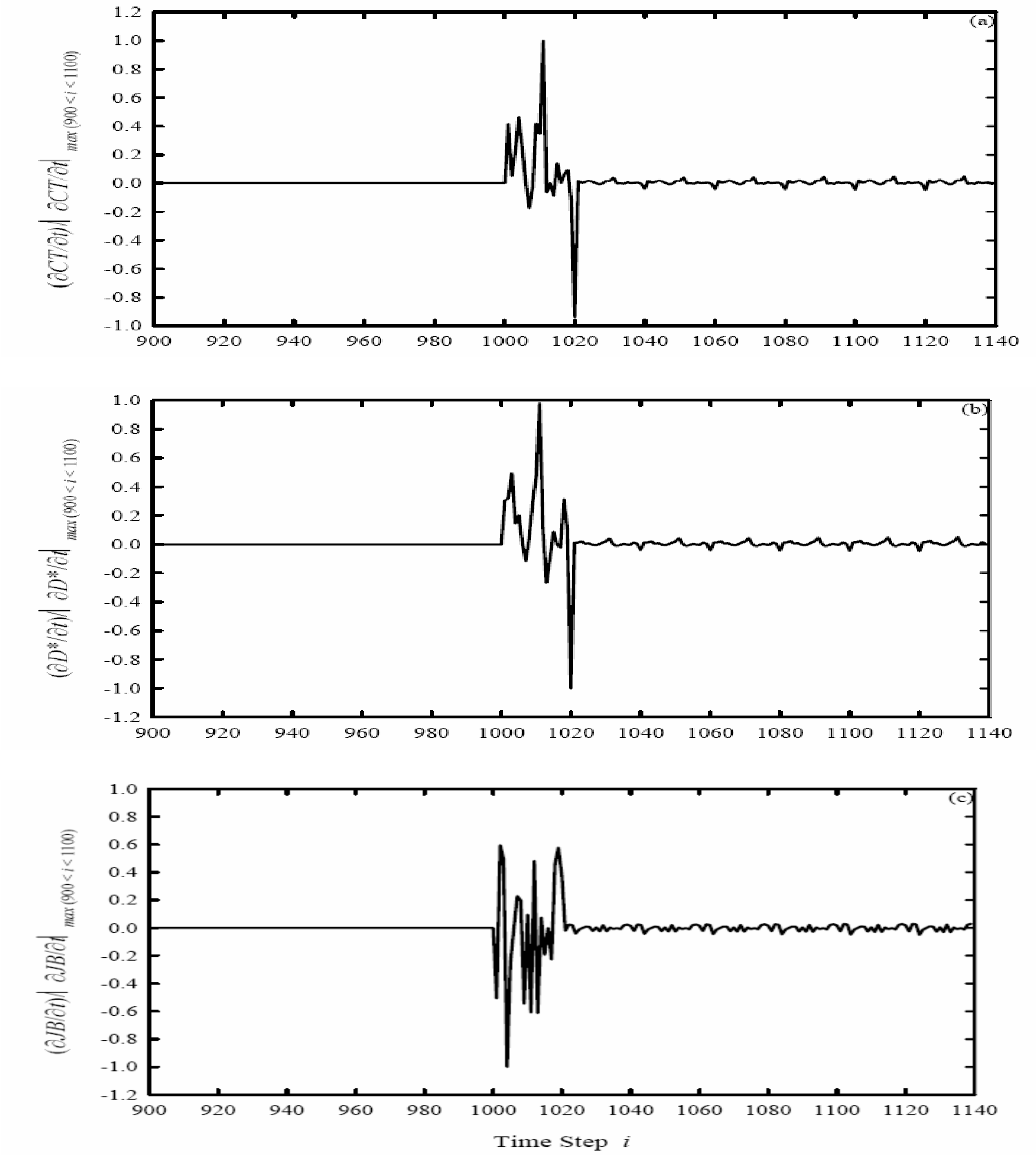

Figure 5.7 Example of the time variation of the rate of change of indicator around i_s where *JB* gives better performance than *D**.

(a) *CT*; (b) *D**; (c) *JB*.

$\theta = 10$.

θ	S/N (dB)	Statistical Parameter	Error in i_s		
			*D**	*JB*	*CT*
0.01	0	Mean	2 (1)	2 (2)	1 (2)
		S.D.	3 (1)	3 (1)	2 (1)
		s	2.9 (0.7)	1.6 (0.9)	2.7 (0.2)
		k	8.6 (-1.0)	1.2 (-0.5)	7.9 (-0.7)
0.02	–3	Mean	8 (4)	8 (3)	9 (3)
		S.D.	6 (3)	6 (2)	6 (2)
		s	0.1 (1.0)	0.6 (1.3)	-0.2 (1.7)
		k	-0.3 (-0.3)	0.6 (1.2)	0.7 (3.4)
0.05	–7	Mean	11 (4)	8 (3)	12 (2)
		S.D.	7 (3)	6 (2)	7 (1)
		s	-0.4 (0.4)	0.6 (1.4)	-0.8 (0.2)
		k	-1.0 (-2.1)	0.8 (3.0)	-0.7 (-0.7)
0.10	–10	Mean	6 (4)	7 (4)	5 (3)
		S.D.	6 (3)	6 (3)	5 (2)
		s	0.7 (0.5)	0.7 (1.3)	-0.8 (0.2)
		k	-0.3 (0.2)	0.3 (0.9)	-2.2 (0.8)
0.25	–14	Mean	4 (5)	7 (3)	5 (3)
		S.D.	7 (4)	6 (3)	7 (2)
		s	1.5 (0.3)	0.9 (1.5)	1.1 (0.8)
		k	1.4 (-1.8)	0.2 (0.9)	0.1 (-0.1)
0.50	–17	Mean	6 (5)	9 (3)	6 (3)
		S.D.	6 (4)	6 (3)	7 (2)
		s	0.8 (0.2)	0.2 (1.5)	0.6 (1.2)
		k	-0.3 (-2.0)	-0.6 (0.9)	-0.8 (0.3)
1.00	–20	Mean	7 (5)	9 (3)	8 (2)
		S.D.	6 (4)	6 (2)	7 (2)
		s	0.5 (0.5)	0.2 (2.6)	0.4 (1.5)
		k	-0.3 (-2.0)	-0.6 (7.5)	-1.2 (1.9)
2.00	–23	Mean	9 (5)	9 (3)	7 (2)
		S.D.	5 (4)	6 (2)	6 (2)
		s	-0.2 (0.5)	0.2 (2.5)	0.4 (1.5)
		k	-2.0 (-0.3)	-0.6 (6.9)	-0.5 (1.9)
10.0	–30	Mean	9 (5)	9 (3)	7 (2)
		S.D.	5 (4)	6 (2)	6 (2)
		s	-0.1 (0.1)	0.2 (2.5)	0.4 (1.9)
		k	1.2 (-2.3)	-0.6 (6.9)	-0.4 (3.8)

† $CT = D^*/D_o - JB/JB_o$, s : skewness, k : kurtosis

Table 5.1 Error of i_s recovery†

Chapter 6

Detecting Weak Sinusoidal Signals Embedded in a Non-stationary Random Broadband Noise

6.1 Introduction

The focus of this chapter is on enhancing the detection of weak signals embedded in a stronger non-stationary signal / noise. A numerical treatment for a signal made up of the weak sinusoidal signal and the non-stationary noise, which can facilitate a better retrieval of the weak signal after the time-frequency procedure, is proposed. Its performance is examined through two illustrative examples.

6.2 Theoretical Considerations

The target of the present study is to establish a method to detect effectively a sinusoidal signal under a poor S/N ratio. This section describes the development of the method. The spectral technique will be used, but it is believed that the method also works for wavelet transforms. In the following analysis, all parameters are non-dimensional. Also, Ω denotes random noise with vanishing mean and magnitude bounded between ± 1 in the following discussions. These signals are generated using the software MATLAB.

6.2.1. Power Spectral Density of a Square Pulse Train

An infinite square pulse train is perhaps the most fundamental signal in building services engineering other than the sinusoidal wave. It is given by the expression

$$x(t) = A\sum_{i=-\infty}^{\infty}(-1)^i[U(t-it_o-(t_o-\Delta)/2)-U(t-it_o-(t_o+\Delta)/2)], \tag{6.1}$$

where U is the unit step function, and A and Δ are real constants which fix the magnitude and duration of each pulse in $x(t)$ respectively. The period is $2t_o$. Figure 6.1 illustrates an example of the train with $A = 1$, $t_o = 10$ and $\Delta = 6$. The power spectral density of x, $P(x,\omega)$, is [33],

$$P(x,\omega) = \underset{T\to\infty}{Lim}\frac{2}{T^2}\left|\int_0^T x(t)e^{-j\omega t}dt\right|^2, \tag{6.2}$$

where $j = \sqrt{-1}$ and ω the angular frequency. Suppose n pulses exist within the period T, then one can approximate that for very large T and n, $T \sim nt_o$ and

$$P(x,\omega) = \frac{8}{(nt_o)^2}\left[\frac{A}{\omega}\sin\left(\frac{\omega\Delta}{2}\right)\right]^2\left|\sum_{i=0}^{n-1}(-1)^i e^{-j\omega t_o(i+0.5)}\right|^2. \tag{6.3}$$

For $n\to\infty$, P is only significant when $\omega = \pi/t_o$ and

$$P(x,\pi/t_o) = \frac{8A^2}{\pi^2}\sin^2\left(\frac{\pi\Delta}{2t_o}\right). \tag{6.4}$$

If the amplitude of each pulse in the train is not constant, but is uniformly distributed between 0 to A, the average amplitude will be $A/2$ and P will be reduced by 4 times. It can therefore be concluded that the detection of a sinusoidal wave in the presence of a random background noise will be enhanced if a numerical procedure that can transform the wave into a pulse-like train of the appropriate width can be found as the spectral contents of the random frequency components will be attenuated more quickly than that of the sinusoidal wave after the transformation. However, this may not be true if the effective Δ is too small after the transformation.

6.2.2. The Proposed Transformation Function and Its Properties

A function which is simple enough for quick numerical application is required. In order to transform the sinusoidal wave into a pulse-like train, the crests and troughs of the wave must be preserved while the regions of large gradient must be down-weighted to small values. The function must also be an odd function. The following function

$$w(x) = 2x^{(2\alpha+1)} \big/ (1 + x^{2\beta}), \tag{6.5}$$

where α and β are non-zero integers appear to fit the present purpose. Subsequent transformations are hereinafter denoted by $w^{(1)} = w(x)$, $w^{(2)} = w(w(x))$, $w^{(3)} = w(w(w(x)))$ and so on. Figure 6.2 illustrates the $w^{(1)}$ of the cosine wave $x(t) = \cos(0.02\pi t)$ for various α with $\beta = 1$. One can notice that the effective Δ is reduced after repeated transformation, but the increase in the Δ reduction is much less pronounced for large α. Increasing β at a fixed α tends to flatten slightly the crests and troughs of the transformed wave. This is in fact true for other combinations of for α and β. Repeated transformation using the same combination of α and β results in more rapid reduction of the pulse width than that produced by increasing α. It should be noted that for sinusoidal wave of amplitude less than unity, each transformation will also result in a reduction of the pulse amplitude.

Figure 6.3 shows the associated power spectral densities of the transformed cosine wave $x(t) = 0.5\cos(0.02\pi t)$ for $\alpha = \beta = 1$. It is noted that harmonics appear after the transformation of the cosine wave. The first three transformation using w

results in faster attenuation of the average spectral energy levels of the random noise 0.5Ω compared to the peak spectral energy of the cosine wave as shown in Figure 6.4, which is expected from the deduction discussed in the previous section. Since the width of the pulse is reduced after each transformation, repeated transformation will eventually smooth out the spectral signatures of all discrete frequency components as suggested by Equation (6.4). This leads to the higher spectral attenuation of $x(t) = 0.5\cos(0.02\pi t)$ after the fourth transformation. Similar phenomenon is expected to happen for the wave with amplitude after the fifth or the sixth transformation. One can also observe that the first two transformations tend to attenuate more the spectral power density of the random noise for smaller magnitude signals. Since the ratio of the spectral power density is around 4000 for the un-transformed case, it can be concluded that the detection of the sinusoidal wave may still be possible if its amplitude is ~2% of that of the random noise without any transformation, but the isolation of its spectral peak will be extremely difficult because of the very spurious spectral peaks resulted from the random noise.

Figures 6.5a and 6.5b illustrates the $w^{(4)}$s of the wave $x(t) = 0.5\cos(0.02\pi t)$ and 0.5Ω respectively. The very periodical structure of the transformed cosine wave compared to that of the transformed random noise should enable a better detection of the wave in the presence of the random noise using the spectral method, even though the attenuation of the random noise is less serious compared to that of the cosine wave after this fourth transformation. This will be discussed further in the next section.

6.3. Illustrative Examples and Discussions

In this section, the effectiveness of the function w in resolving weak signals embedded inside a broadband random noise is illustrated. Two signals will be examined. The first one is a weak but steady sinusoidal signal and the other one is an abruptly generated sinusoidal signal that grows exponentially with time.

6.3.1. Weak Steady Sinusoidal Signal

The signal in this section consists of a weak sinusoidal signal embedded inside a random background noise and it takes the form of

$$x(t) = 0.5\Omega + A\sin(0.02\pi t). \tag{6.6}$$

The signal-to-noise ratio S/N is defined as $10\log_{10}(2A)$. Figure 6.6 shows the effects of repeated transformations (Equation 5) on $x(t)$ for $\alpha = \beta = 1$ and $A = 0.025$ (S/N ~ −13dB). The spectral power density calculations were done using MATLAB with data segment length of 8192 and 0% overlapping. Figures 6.7 and 6.8 illustrate the corresponding spectral power spectra for $A = 0.015$ (S/N ~ −15dB) and 0.007 (S/N ~ −19dB) respectively.

For $A = 0.025$, the spectral peak of the sinusoidal wave can be observed even without the transformation, but it is surrounded by the spurious spectral peaks from the noise. It is consistent with the observations made in Section 6.2. Subsequent transformations make this peak more distinctive from the background (Figure 6.6) until $w^{(5)}$. With a weaker sinusoidal wave of $A = 0.015$, the spectral peak due to this wave can be recovered distinctively after $w^{(2)}$ as shown in Figure 6.7. The performance starts to go worse after the fifth transformation. For $A = 0.007$, the spectral peak of the wave can only be recovered satisfactorily after $w^{(4)}$, but again the resolution of this spectral peak becomes poor after $w^{(5)}$.

6.3.2. Abruptly Generated Signal that Grows Exponentially with Time

The signal in this section consists of Ω and an exponentially growing sinusoidal wave initiated at time t_i:

$$x(t) = B\Omega + AH(t - t_i)\exp(\gamma(t - t_i))\cos(\omega(t - t_i)), \quad (6.7)$$

where γ denotes the growth rate of the sinusoidal wave. Without lost of generality, t_i, γ, A and ω are fixed at 1000, 0.002, 0.01 and 0.1π respectively. The signal-to-noise ratio here is defined as $10\log_{10}(A/B)$, which denotes the initial strength of the growing signal relative to the background noise. It is shown in the last section that the proposed transformation can enhance the recovery of a steady and continuous sinusoidal wave in the presence of a background noise down to a S/N of −19dB. In this section, the main focus is on the recovery of t_i – the instant of wave initiation. The two transformation parameters α and β are fixed at unity.

The moving time frame approach is adopted. The number of data in the time frame is denoted by n. The calculation is commenced by going through the transformations as in the last section using the first n data of $x(t)$. The time frame is then advanced by one time step and the calculations are repeated. Since the performance of the transformations depends on the magnitude of the data series, which can vary widely in practice, the data in the each time frame are normalized by the data range (the difference between the maximum and minimum values). The mean value of the data in each time frame is kept to be zero. The number of data put into the Fourier transform calculation is again n, giving a frequency resolution of $1/n$.

Figure 6.9 is an example of $x(t)$ with $B = 0.5$ such that the S/N = −17dB. Only random background noise exists for $t < 0$. One can observe that the sinusoidal wave cannot be seen even at $t = 2000$ under this S/N. Figure 6.10 illustrates the variation of the maximum spectral power density upon repeated transformations of the signal with $B = 0.01$ for $n = 200$. Since it is actually unable to know in advance the frequency of the sinusoidal growing signal in reality as far as machine monitoring process is concern, only maximum spectral power density $P_{\max}$ is considered in Figure 6.10 and in similar figures presented afterward. The frequency of the sinusoidal wave can be easily found out once it is detected. Certainly, one can anticipate that the maximum power density is equal to that of the growing sinusoidal wave after two transformations. One can notice from Figure 6.10 that the spectral power densities of the signals are the smallest near to the instant of the wave initiation. For n = 200 at this S/N of 0dB, the minimum $P_{\max}$ appears at t ~ 1010 after $w^{(3)}$ and a sharp drop in $P_{\max}$ precedes this instant. The largest time gradient of the maximum power density, $\left|\partial P_{\max}/\partial t\right|$, appears at t = 1001 (Figure 6.11) which is very close to t_i.

Figure 6.12 illustrates the effects of S/N on the recovery of t_i using the parameter $\left|\partial\left(P_{\max}(x,\omega)/P_{\max}(w^{(3)},\omega)\right)/\partial t\right|$. One can observe that a sharp rise in the parameter followed by closely packed spurious peaks of significant magnitudes can be found at t ~ t_i. The recovery of t_i is still satisfactory for a S/N of −17dB, although the associated error goes up to ~180 time steps. However, the sinusoidal wave is still not visually distinguishable from $x(t)$ at t = 300 (Figure 6.9). This parameter also works as good as $\left|\partial P_{\max}/\partial t\right|$ for the S/N = 0dB case. Figure 6.12 also indicates that the error of t_i recovery ranges from 0 to a maximum of 230 for a S/N above −17dB, but there appears no relationship between the S/N and the magnitude of the recovery error.

The number of data included in the moving average spectral calculation has crucial impacts on the recovery of t_i. It is expected that the increase in n will smooth out spectral irregularities especially within the random background noise and will increase the frequency resolution of the spectral calculation. It is found that a larger n also allows the use of more repeated transformations. The parameter

$\left|\partial\left(P_{\max}(x,\omega)/P_{\max}(w^{(5)},\omega)\right)/\partial t\right|$ is adopted in Figure 6.13, where the associated S/N equals −17dB. One can observe from Figure 6.13a that the associated error in the recovery is negligible for n = 4000, though there is a small spike at t ~ 550 due to the random background noise. The recovery is unambiguous when n is increased to 8000 (Figure 6.13b).

6.4. Conclusion

A novel but simple method for the detection of weak signals embedded in a non-stationary strong broadband background noise is derived in this chapter. A function, which tends to transform sinusoidal wave into a regular pulse train, is proposed to be use together with the Fourier transformation. Its performance is illustrated using two artificial numerical examples. The first one is a very weak but steady sinusoidal signal, while the other an exponentially growing sinusoidal wave generated abruptly within the background noise. For the latter, the recovery of the instant of wave initiation is the focus. The performance of repeated transformations on the signals is also investigated.

For the steady sinusoidal wave, the proposed transformation function enables its unambiguous detection even when the signal-to-noise ratio drops to −19dB after the fourth transformation. Further transformation is found to make thing worse and is not recommended. The recovery of the instant of the exponentially growing wave initiation is also enhanced after the application of the transformation function. The associated error is found to be within 300 time steps even for an initial signal-to-noise ratio of −17dB when 200 data are used in the spectral calculation. The recovery of the instant of the wave initiation is improved remarkably by increasing the number of data involved in the calculation. The recovery error becomes negligible when more than 4000 data are used to produce the moving spectral averages.

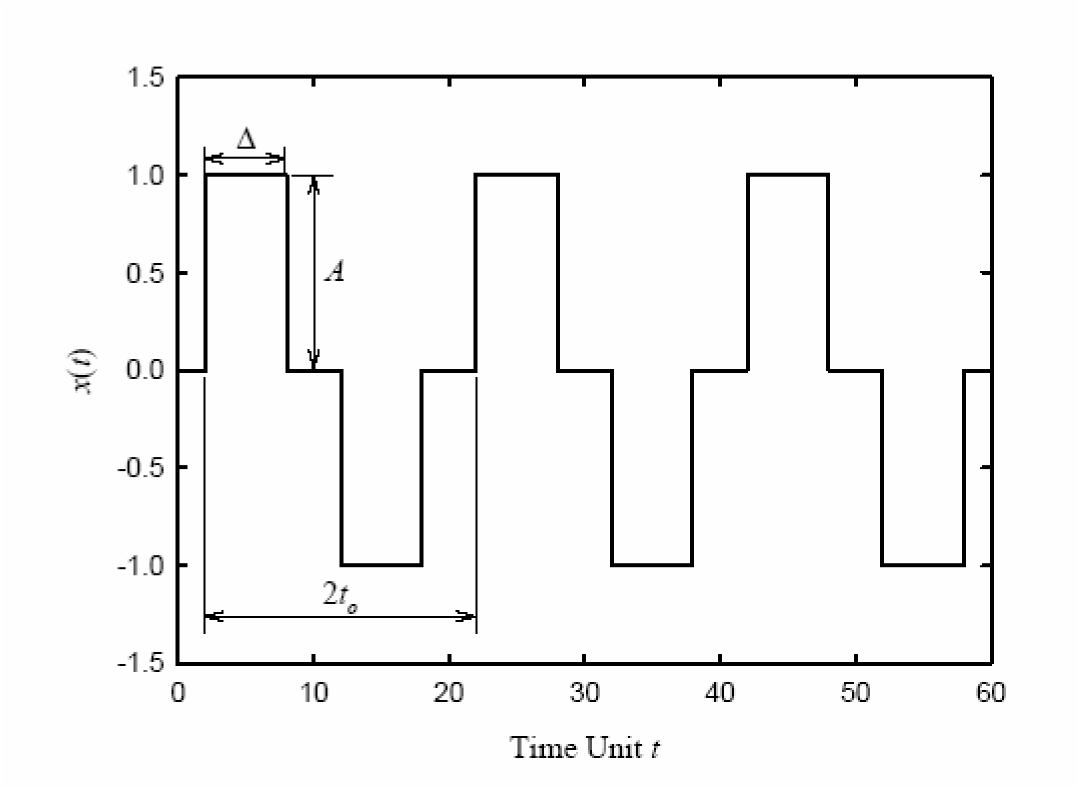

Figure 6.1 Example of a square pulse train and the nomenclatures.

$\Delta = 6$, $A = 1$, $t_o = 10$.

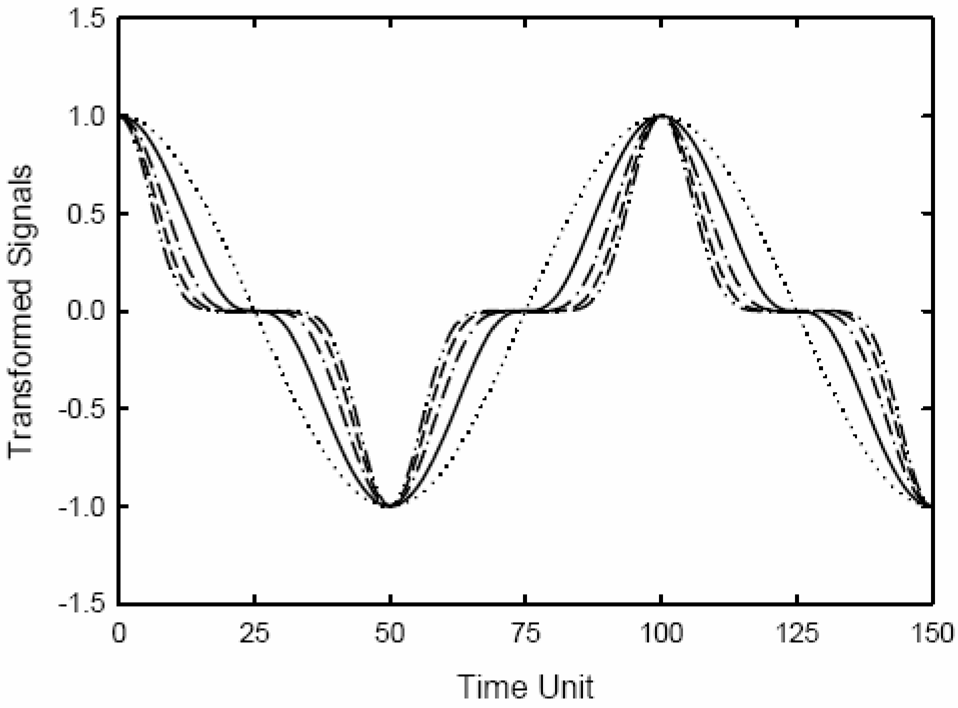

Figure 6.2 Examples of transformed cosine waves with $\beta = 1$.

············ : Original wave $\cos(0.02\pi t)$; ——— : $\alpha = 1$; — · — : $\alpha = 2$;

— — : $\alpha = 3$; — · · — : $\alpha = 4$.

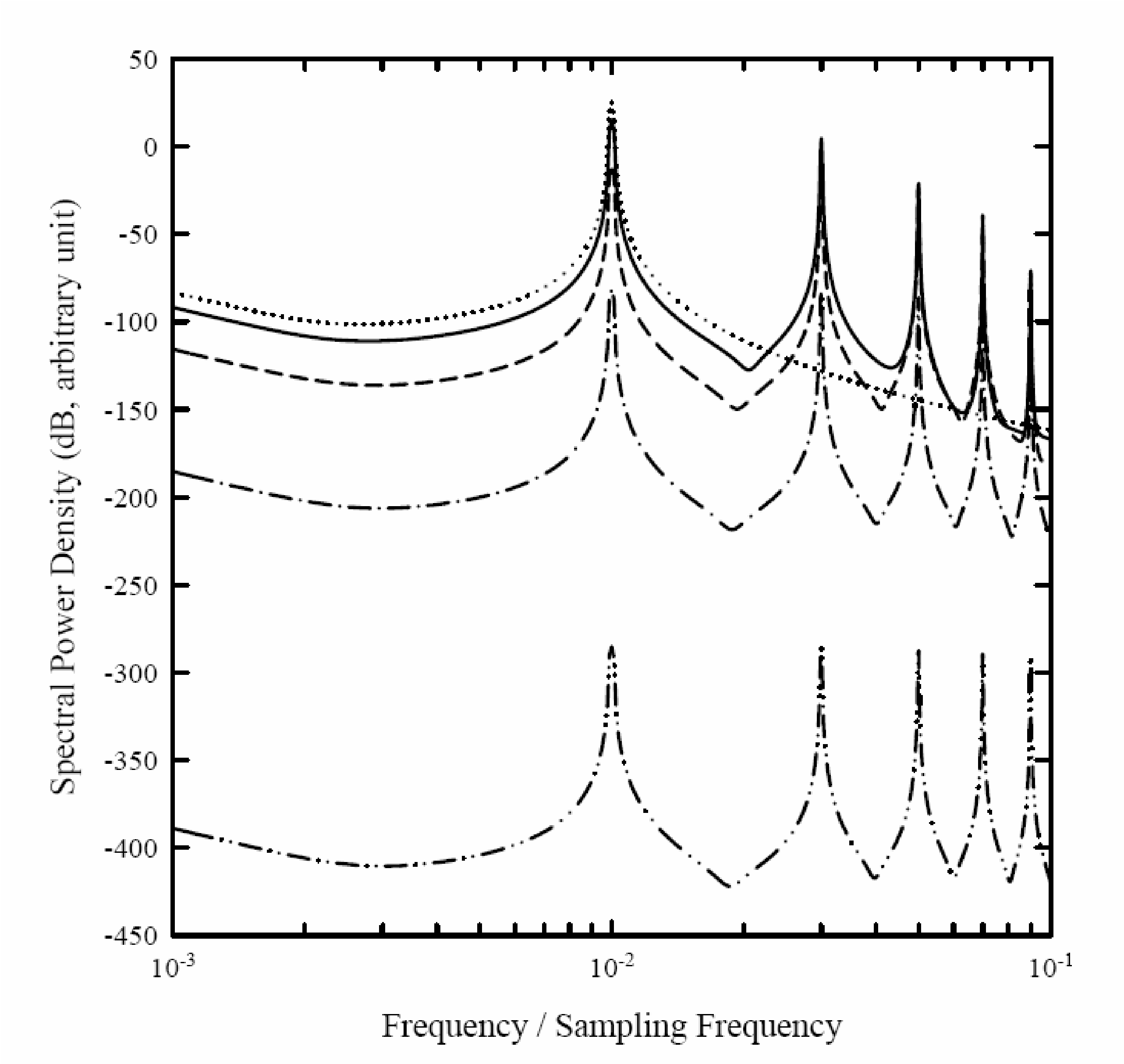

Figure 6.3 Spectral power densities of transformed cosine wave $x(t) = 0.5\cos(0.001\pi t)$;

.... : x; ——— : $w^{(1)}$; — · — : $w^{(2)}$; — — : $w^{(3)}$; — · · — : $w^{(4)}$.

$\alpha = \beta = 1$.

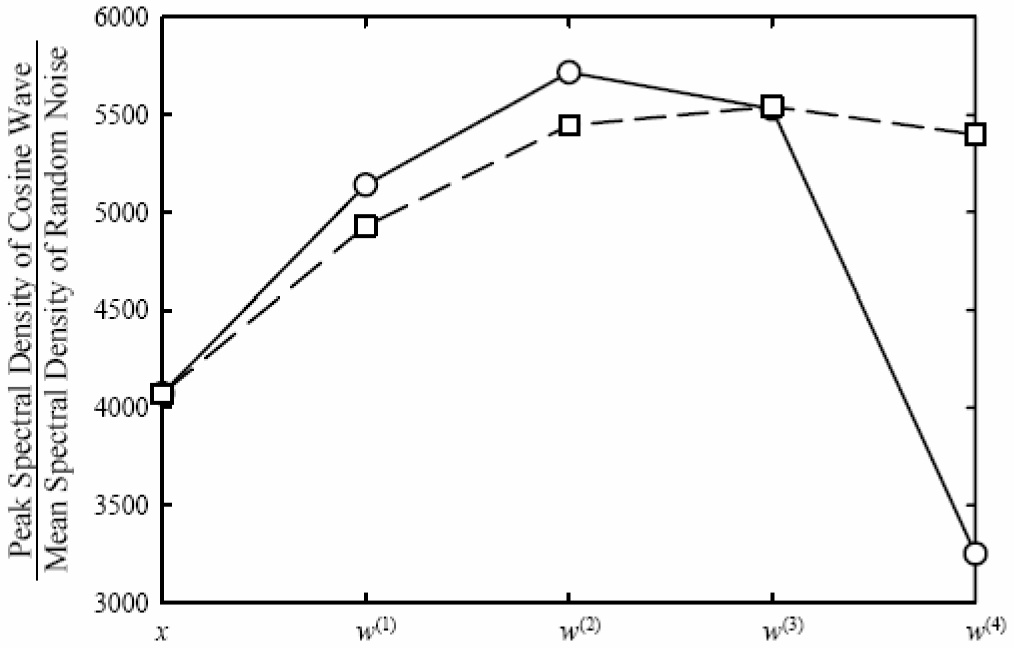

Figure 6.4 Spectral power density attenuations due to repeated transformations.

○ : $\cos(0.02\pi t)$ and Ω; □ : $0.5\cos(0.02\pi t)$ and 0.5Ω.

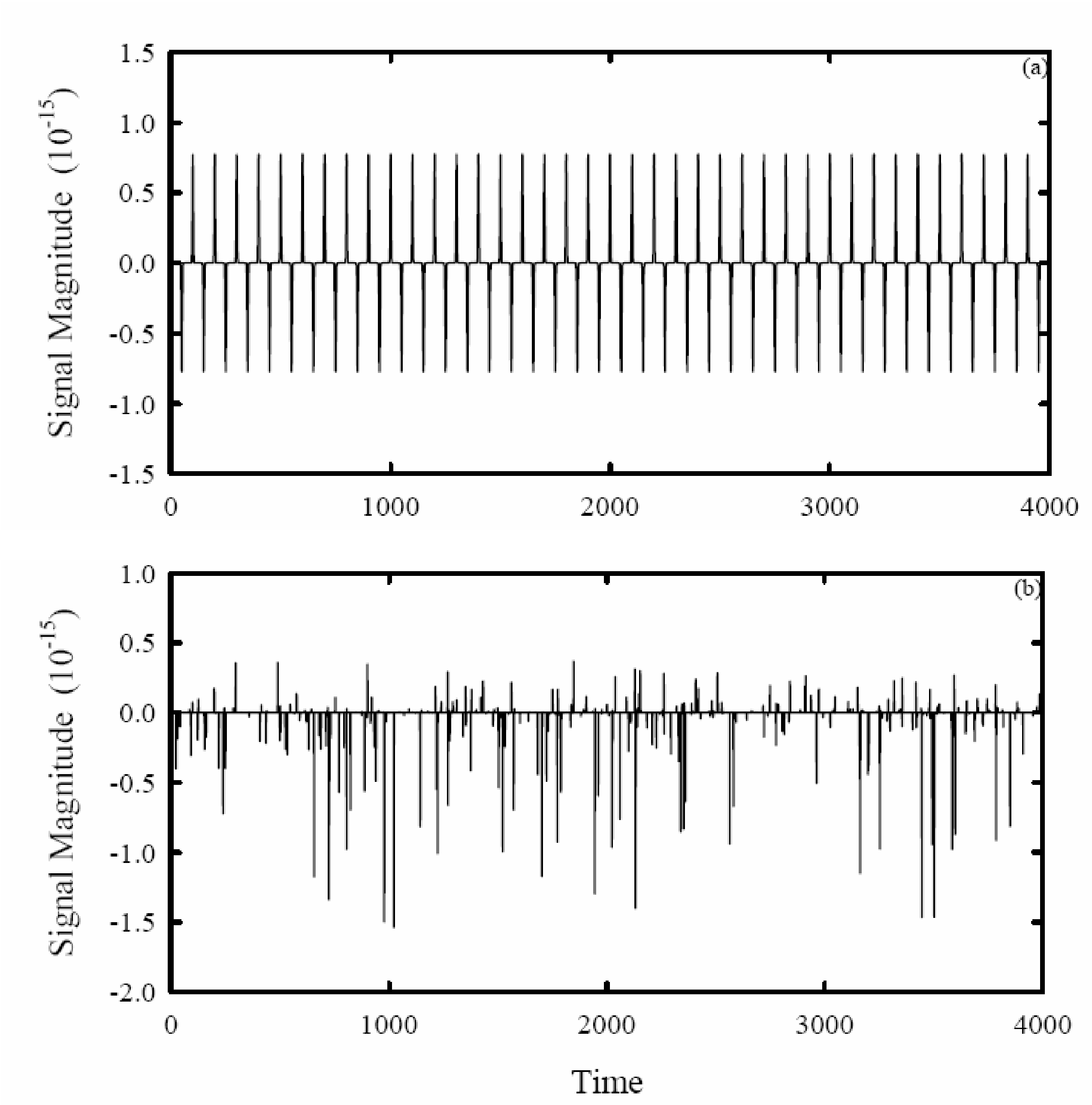

Figure 6.5 Time variations of $w^{(4)}$ for $\alpha = \beta = 1$.

(a) $0.5\cos(0.02\pi t)$; (b) 0.5Ω.

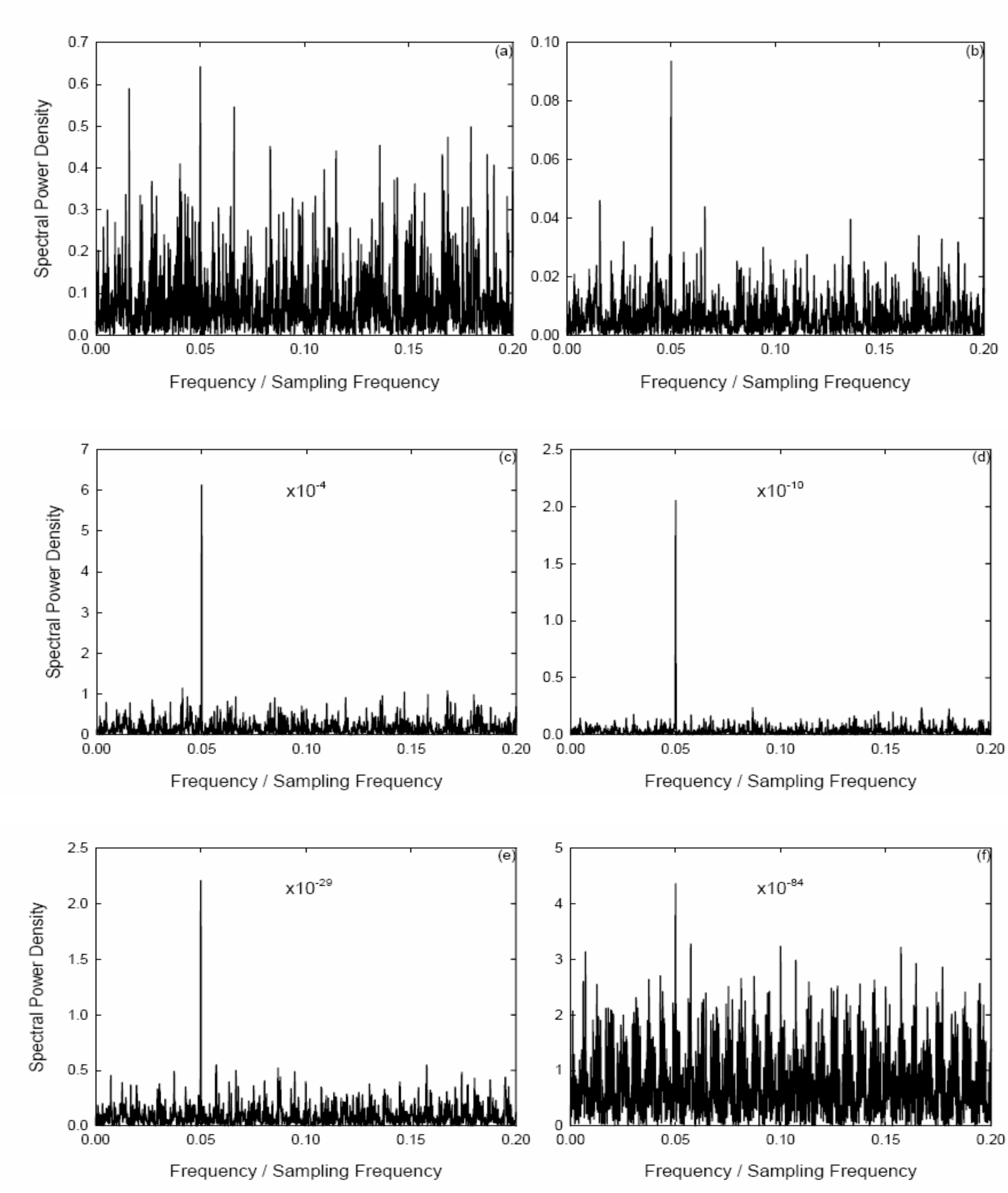

Figure 6.6 Effects of repeated transformations on spectral peak recovery in the presence of a background noise Ω for A = 0.025 (S/N = −13dB) and $\alpha = \beta = 1$.

(a) x; (b) $w^{(1)}$; (c) $w^{(2)}$; (d) $w^{(3)}$; (e) $w^{(4)}$; (f) $w^{(5)}$; (g) $w^{(6)}$.

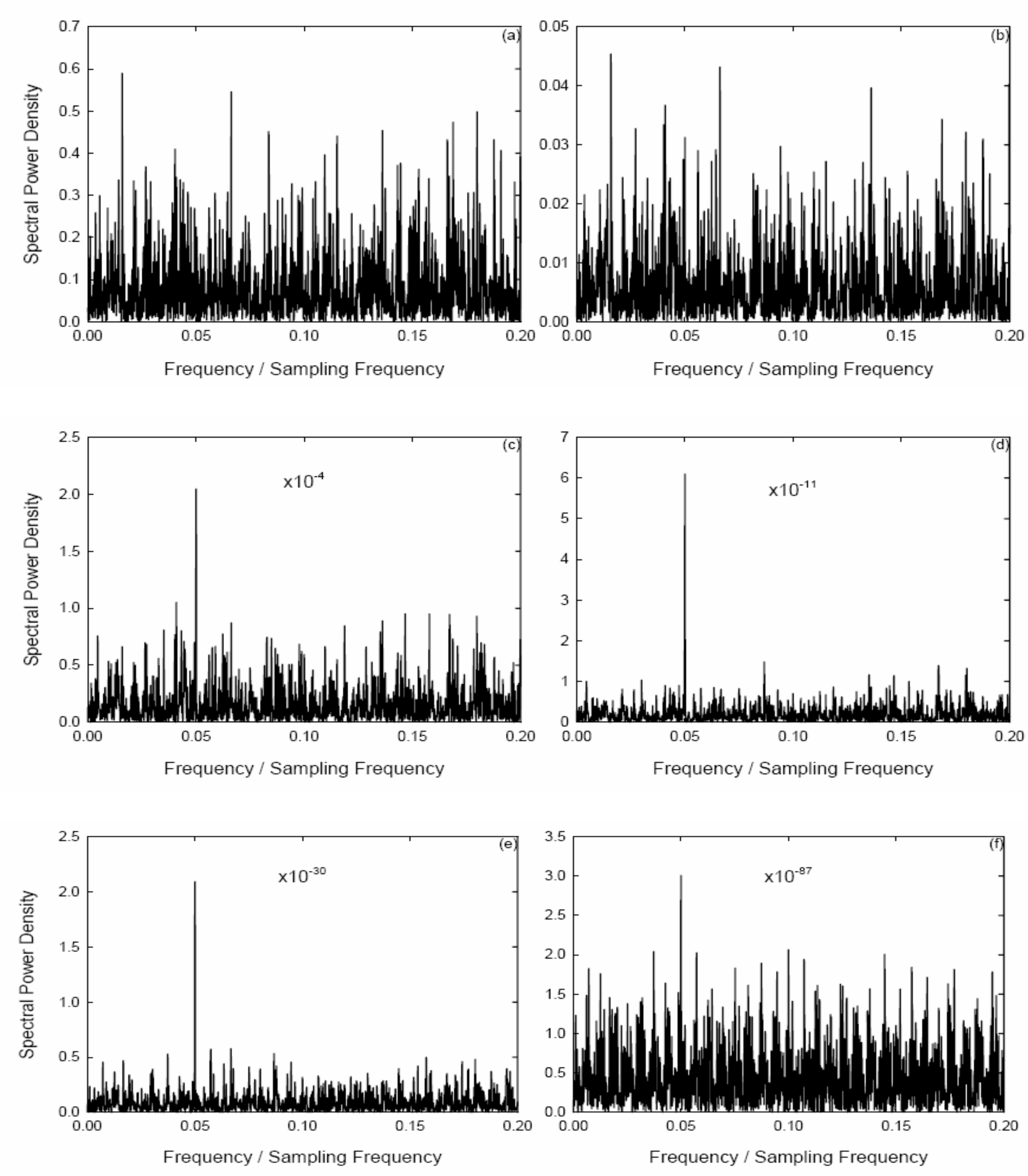

Figure 6.7 Effects of repeated transformations on spectral peak recovery in the presence of a background noise Ω for A = 0.015 (S/N = −15dB) and $\alpha = \beta = 1$.

(a) x; (b) $w^{(1)}$; (c) $w^{(2)}$; (d) $w^{(3)}$; (e) $w^{(4)}$; (f) $w^{(5)}$; (g) $w^{(6)}$.

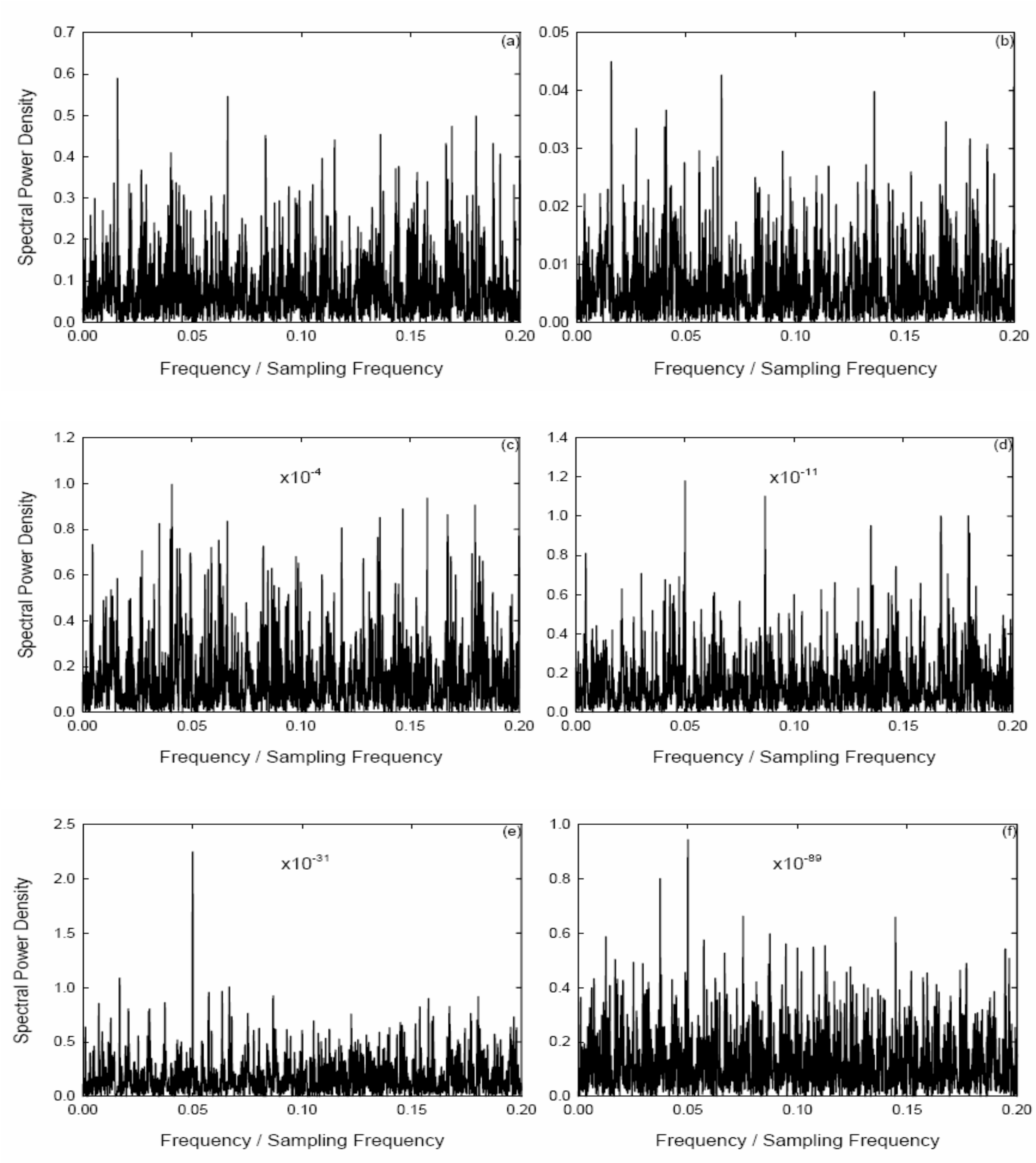

Figure 6.8 Effects of repeated transformations on spectral peak recovery in the presence of a background noise Ω for $A = 0.007$ (S/N = −19dB) and $\alpha = \beta = 1$.

(a) x; (b) $w^{(1)}$; (c) $w^{(2)}$; (d) $w^{(3)}$; (e) $w^{(4)}$; (f) $w^{(5)}$; (g) $w^{(6)}$.

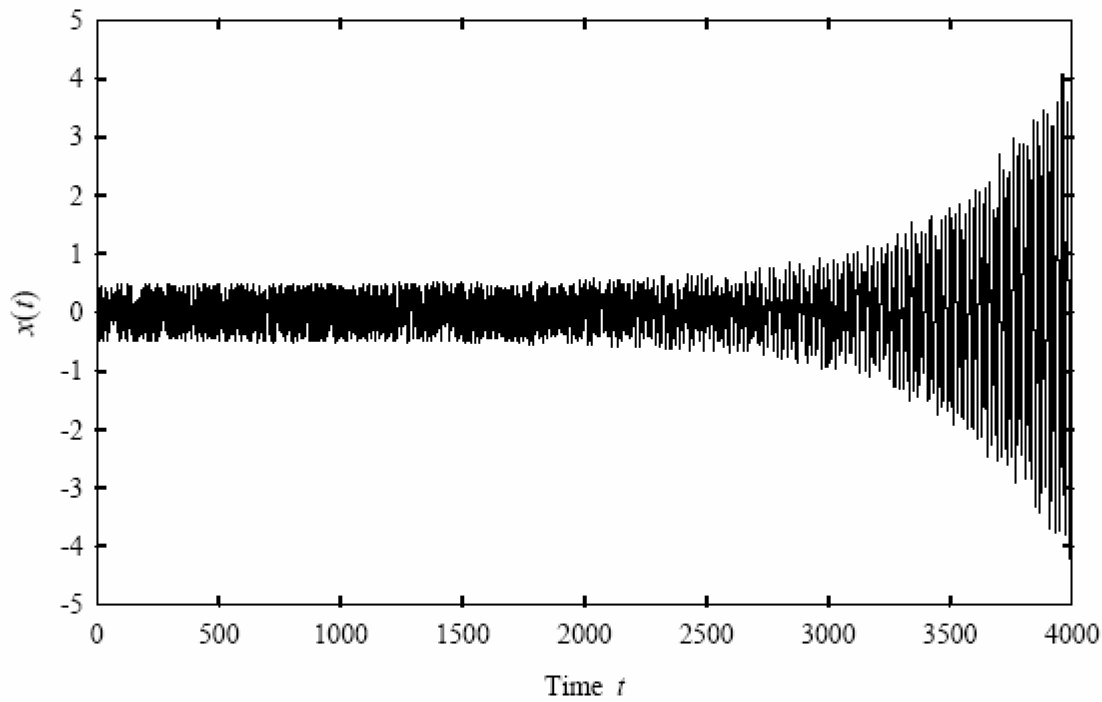

Figure 6.9 An example of exponentially growing signal embedded in a background noise.

B = 0.5 (S/N = –17dB).

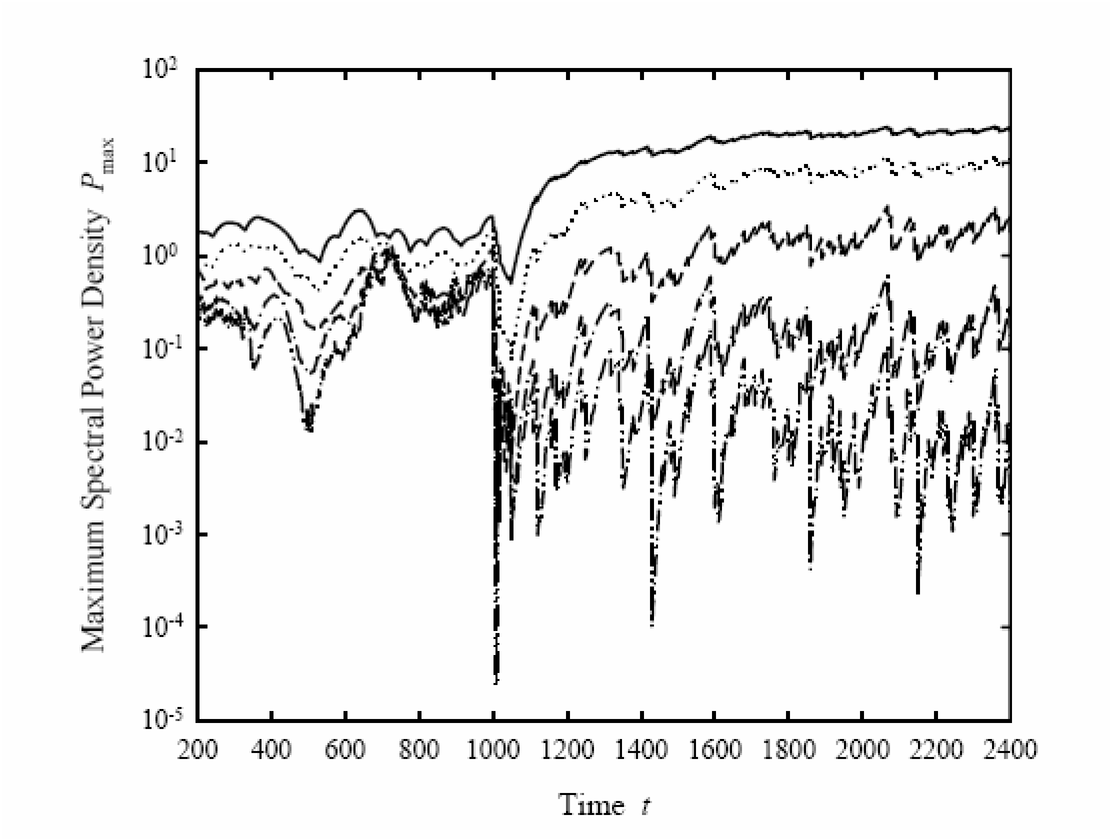

Figure 6.10 Time variation of the maximum spectral power density for transformed signals with exponentially growing components.

——— : x; ··········· : $w^{(1)}$; — — : $w^{(2)}$; — · — : $w^{(3)}$; — · · — : $w^{(4)}$.

$\alpha = \beta = 1$, $n = 200$.

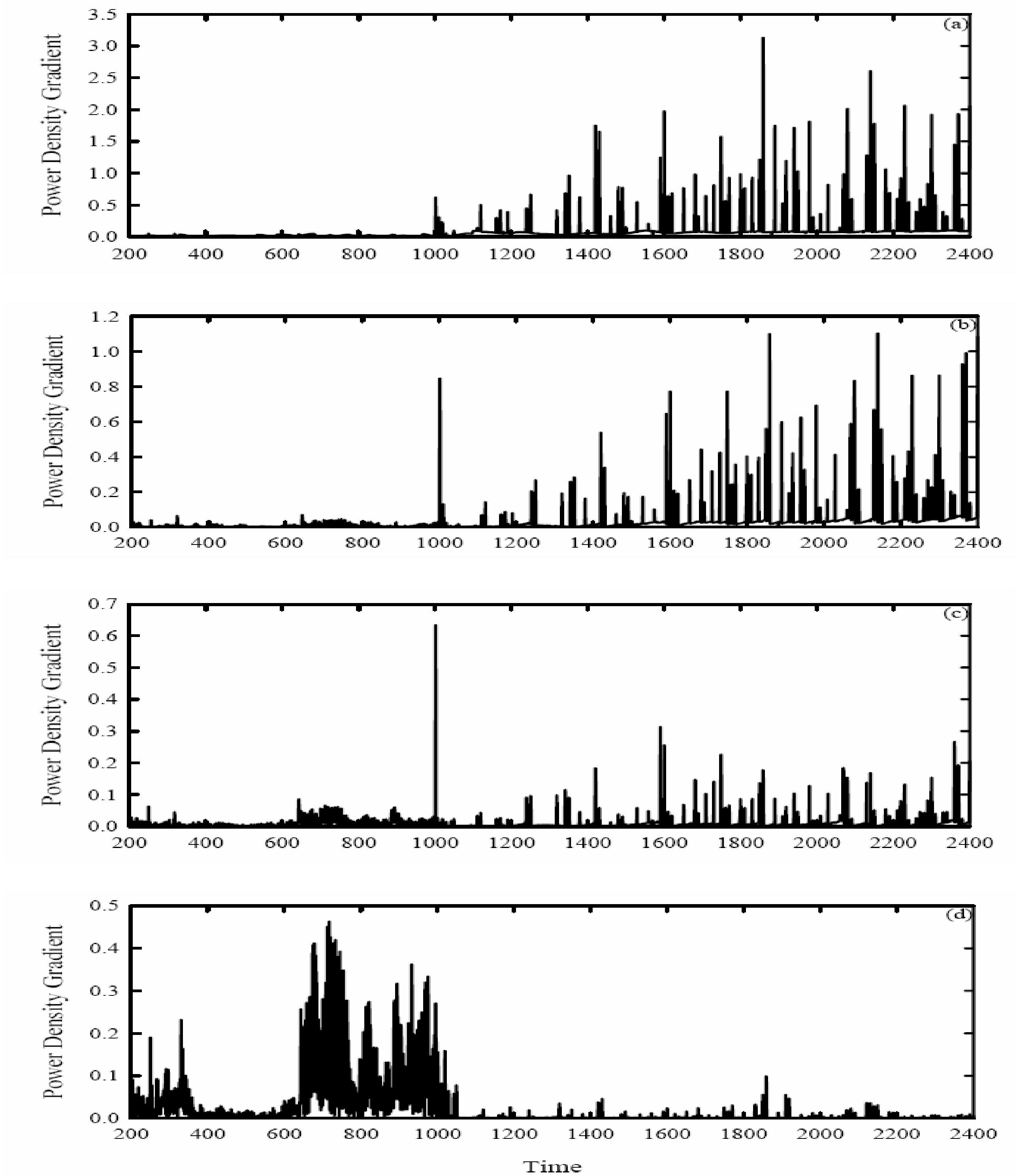

Figure 6.11 Time variation of the time gradient of the maximum spectral power density for transformed signals with exponentially growing components. (a) x; (b) $w^{(2)}$; (c) $w^{(3)}$; (d) $w^{(5)}$. $\alpha = \beta = 1$, $n = 200$.

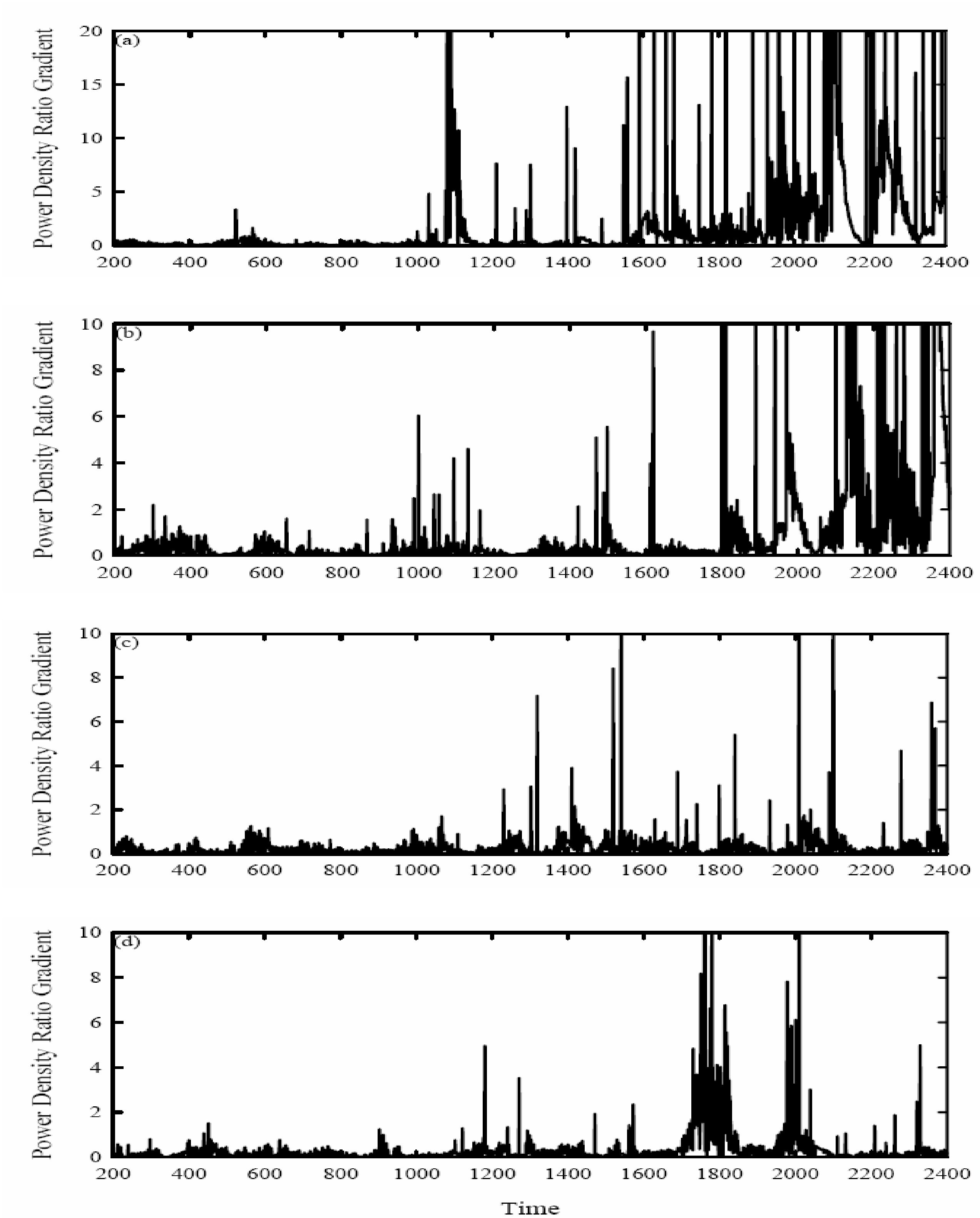

Figure 6.12 Time variation of $\left|\partial\left(P_{\max}(x,\omega)/P_{\max}(w^{(3)},\omega)\right)/\partial t\right|$.

(a) S/N = – 7dB; (b) S/N = –10dB; (c) S/N = –14dB; (d) S/N = –17dB.

$\alpha = \beta = 1$, $n = 200$.

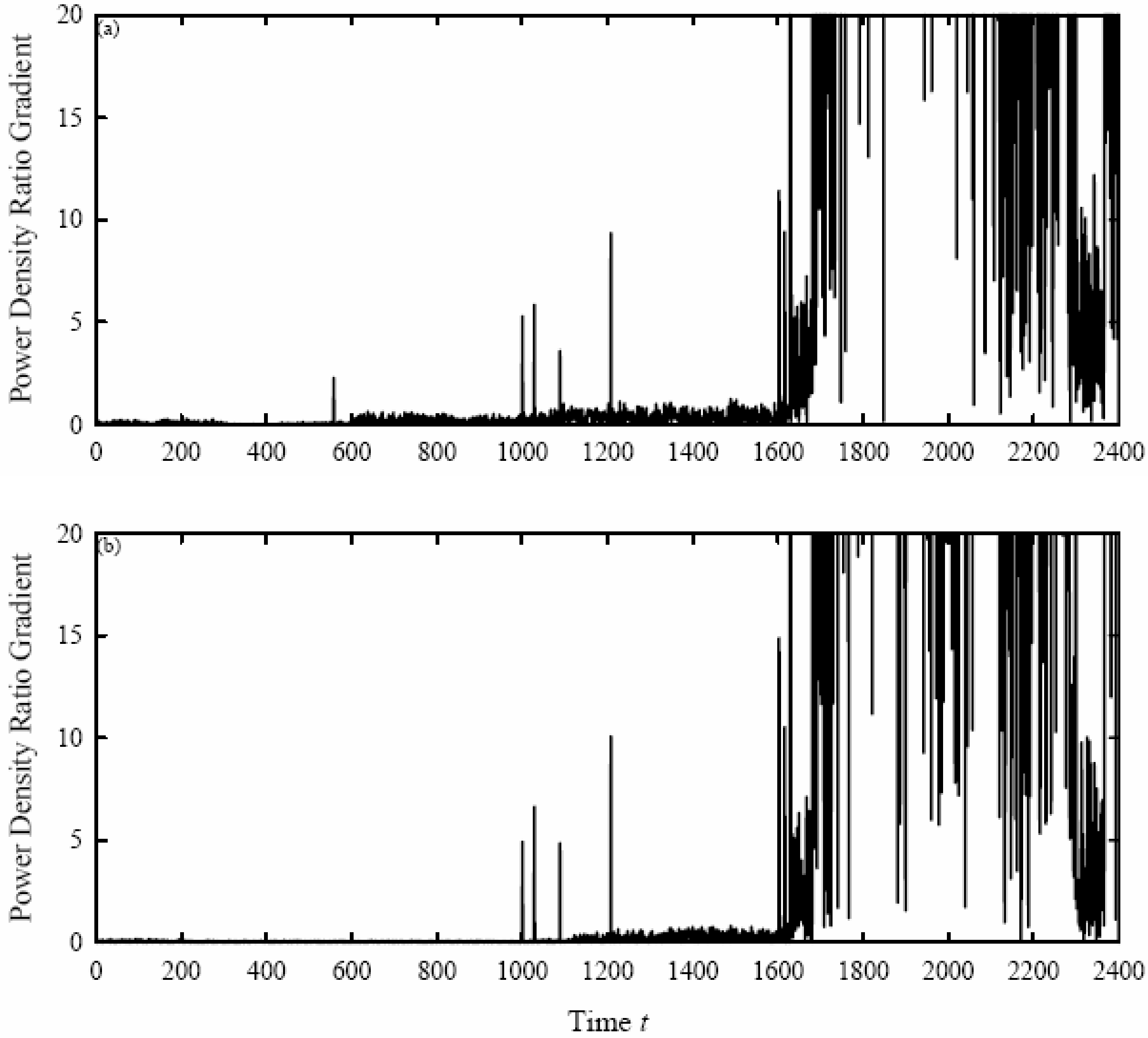

Figure 6.13 Effect of n on the recovery of t_i using $\left|\partial\left(P_{\max}(x,\omega)/P_{\max}(w^{(5)},\omega)\right)/\partial t\right|$.

(a) n = 4000; (b) n = 8000.

$\alpha = \beta = 1$, S/N = –17dB.

Chapter 7

Conclusions and Recommendations for Future Work

7.1 Conclusions

In this thesis, statistical methods have been derived to solve engineering problems related to machine condition monitoring and room acoustics. In particular, statistical functions and numerical functions are used to model the instant of the pulse initiation, detect very weak sinusoidal signals embedded in a non-stationary random broadband background noise and recover the decay constants of exponentially decaying pulses in the presence of strong background noises.

In Chapter 2, the use of stochastic volatility model incorporated with the exponential power distribution to capture the initiation and decay of exponentially decaying signals in the presence of random background noises has been proposed. The exponential power distribution is an alternative choice to the normal distribution as the tail of the exponential power distribution can be adjusted to both heavier and lighter than that of the normal distribution. Although the accuracy deteriorates as the signal-to-noise ratio decreases, the performance of the present volatility model is clearly better than that of the conventional short time Fourier transform even up to a signal-to-noise ratio of -7dB. Also, it is found that the introduction of a $\pm 10\%$ of time fluctuation in the signal frequency does not affect much the pattern of the time varying mixing parameter. The fluctuation in signal frequency does increase the log volatility when the background noise is strong relative to the signal. The detection of pulse will then become difficult when the signal-to-noise ratio drops below to 0dB.

In Chapter 3, the use of a heavy tail distribution has been demonstrated for the purpose mentioned in Chapter 2. The Student – t distribution is expressed into two stage scale mixture distributions so as to reduce the computational time. The performance is comparable to the exponential power distribution. The performance of the mixing parameter *u* in Student-t is basically the same as those obtained using EP distribution with a kurtosis of 0.75 in Chapter 2. The results of log-volatility *h* are also very close to that for the EP distribution.

In Chapter 4, a statistical modeling technique using full Bayesian approach is proposed for use in the detection of jumps and decaying pulses in the presence of background noises. Two heavily tailed distributions, namely, the exponential-power and the Student-t distributions, are adopted for the purpose. Results suggest that both of these distributions can give accurate recovery of the instants when the abrupt changes take place if the background noise level is lower than that of the changes by 3dB. They also indicate that the exponential-power distribution is more useful when the signal-to-noise ratio falls below 0dB.

In Chapter 5, two statistical tests, namely the Jarque-Bera test and D’Agostino tests, are adopted to recover the instant of initialization of an exponentially growing sinusoidal wave. It is found that the Jarque-Bera test gave better prediction than the D’Agostino test. A new parameter is derived from the two tests. This new parameter appears to have inherited the advantages of the two abovementioned statistical tests. It is also found that the errors from these parameters show insignificant dependence on the background noises.

In Chapter 6, a novel numerical method for the detection of weak signals embedded in a non-stationary strong broadband background noise has been derived and its capacity in detecting weak signals demonstrated. A function which tends to transform a sinusoidal wave into a regular pulse train is proposed to be used together with Fourier Transform in this new method. It is found that the present derived method can detect the presence of a sinusoidal signal even when the signal-to-noise ratio drops to around -19dB. It can also recover the instant of wave initiation with smaller amount of data under strong background noise situation.

7.2 Recommendations for Future Work

The analysis of exponentially decaying pulse signals and weak unsteady signals using statistical approach is focused in the present study. There are three areas which can be dealt with in the future.

7.2.1 Analysis of Exponentially Decaying Pulse Signals using Stochastic Volatility Model using Generalized-t Distribution on Multiple Jump Signals

In Chapter 2 and Chapter 3, the stochastic volatility model incorporating the exponential power distribution and Student-t distribution are shown to be able to retrieve the instant of the pulse initiation and the decay constant within engineering tolerance in the presence of background noises of an artificial signal. The recovery of the instants when the multiple jumps and pulses are created can be an extension of the present work. The introduction of generalized-*t* (GT) distribution is proposed to replace the abovementioned two distributions. The GT is proposed by McDonald and Newey [34] in regression analysis and is studied by Arslan and Genç [35] for robust estimation of location and scale parameters. Other than the location μ and scale σ, the GT density is governed by two shape parameters, p (> 0) and q (> 0) and is given by

$$f_X(x \mid \mu, \sigma, p, q) = \frac{p}{2q^{1/p}\sigma B(1/p, q)}\left(1 + \frac{1}{q}\left|\frac{x-\mu}{\sigma}\right|^p\right)^{-(q+1/p)} \tag{7.1}$$

The GT density in (7.1) can be expressed into

$$f_X(x \mid \mu, \sigma^2, p, q) = \int_0^\infty EP\left(u \mid \mu, \frac{2q^{\frac{2}{p}}c_0\sigma^2}{s}, \frac{2}{p}\right) GG\left(s \mid q, 1, \frac{p}{2}\right) ds \tag{7.2}$$

where $EP(u \mid \cdot,\cdot,\cdot)$ is the EP density function as in (7.1). Denoting $c_0 = \frac{\Gamma(3/p)}{\Gamma(1/p)}$ and GG is the generalized gamma distribution [35].

The performance of a SV model when incorporating the GT distribution on retrieving the properties of the multiple exponentially decaying pulses in the presence of random noises is expected to provide a significance result. It is hoped that the results will also provide useful information for the future enhancement of signal detection and machine diagnosis.

7.2.2 Experimentally Validations

In this thesis, statistical approach has been applied on the analysis of exponentially decaying pulse signals and weak unsteady signals. Numerical pure tone data are used. It will be worthwhile to consider the application of the present proposed statistical models to real data which are broadband in nature. Source of decaying pulse can be a room sound decay which can be obtained from reverberation time measurements, while background noise of various magnitudes can be adopted by loudspeakers.

7.2.3 Modeling Signal by Asymmetric Distribution

If a signal exhibits certain degrees of skewness, the use of the asymmetric distributions may be more effective than the symmetric distributions for modeling the signal. From statistical point of view, the asymmetric scale mixture distributions are still unavailable. There is a possibility to develop a class of heavy-tailed asymmetric distributions via scale mixtures from the existing symmetric scale mixture distributions, and to develop efficient simulation algorithms for Bayesian computation. This may be achieved by adding an auxiliary variable to the location of the existing scale mixture distributions in order to get the desired asymmetry. At the same time, the heaviness of the tail (either the left tail or the right tail) can be preserved.

References

[1] Hodges, C.H., Power, J. and J. Woodhouse, J., 1985, "The use of sonogram in structural acoustics and an application to the vibrations of cylindrical shells," J. Sound Vib. **101**, pp. 203 – 218.

[2] Daubechies, I., 1990, "The wavelet transform, time-frequency localization and signal analysis," IEEE T. Inform. Theory, **36**, pp. 961 – 1005.

[3] Forrester B.D., 1990, "Time-frequency analysis of helicopter transmission vibration," Aeronautical Research Laboratory, ARL Propulsion, Melbourne, Technical Report **180**.

[4] McFadden, P. D., and Wang, W. J., 1993, "Early detection of gear failure by vibration analysis – I. Calculation of the time-frequency distribution," Mechanical Systems and Signal Processing, **7**, pp. 193 – 203.

[5] Staszewski, W.J. and Tomlinson, G. R., 1994, "Ph.D. Thesis, Manchester School of Engineering, Mechanical Engineering Department, the University of Manchester." The application of time-variant analysis to Gearbox fault detection.

[6] Yesilyurt, I., 1997, "Gearbox fault detection and severity assessment using vibration analysis"

Ph.D. Thesis, Manchester School of Engineering, Mechanical Engineering Department. The University of Manchester.

[7] Vorländer, M. and Kob, M., 1997, "Practical aspects of MLS measurements in building acoustics," Appl. Acoust. **52**, pp.239 – 258.

[8] Tang, S. K., 2000, "On the time-frequency analysis of signals that decay exponentially with time," Journal of Sound and Vibration, **234**, pp. 241 – 258.

[9] Wong, H., Ip, W. C., and Li, Y., 2001, “Detection of jumps by wavelets in a heteroscedastic autogressive model,” Statistics and Probability Letters, **52**, pp. 365 – 372.

[10] Hastings, W. K., 1970, “Monte Carlo sampling methods using Markov chains and their applications, **57**, pp.97-109.

[11] Andrieu, C., Doucet, A., Fitzgerald, W.J. and Godsill, S.J., 1998, “An introduction to the theory and applications of simulation based computational methods in Bayesian signal processing,” in Proceedings of the International Conference on Acoustics, Speech, and Signal Processing, Settle, WA.

[12] Taylor, S. J., 1994, “Modelling stochastic volatility,” Math. Financ. **4**, pp. 183 – 204.

[13] Jacquier, E., Polson, N.G. and Rossi, P.E., 1994, “Bayesian analysis of stochastic volatility models (with discussion), Journal of Business and Economic Statistics, **12**, pp. 371-417.

[14] Choy, S. T. B. and Chan, C. M, 2003, “Scale mixtures distributions in insurance applications,” Austin Bull. **33**, pp. 93 – 104.

[15] Choy, S. T. B. and Chan, C. M., 2005, “Bayesian student-*t* stochastic volatility models via two-stage mixtures representation,” 20th International Workshop on Statistical Modelling. Sydney, Australia, 10-15, July,2005.

[16] Choy, S. T. B. and Smith, A. F. M., 1997, “Hierarchical models with scale mixtures of normal distribution,” TEST **6**, pp. 205 – 211.

[17] Box, G. E. P. and Tiao, G. C., 1992, Bayesian Inference in Statistical Analysis (Wiley, New York).

[18] Choy, S. T. B. and Walker, S. G., 2003, **“**The extended exponential power distribution and Bayesian robustness,” Stat. Probabil. Lett. **65**, pp. 227 – 232.

[19] Walker, S. G. and Gutierrez-Pena, E., 1999, in Bayesian Statistics 6, pp. 685 – 710 (Oxford University Press, Oxford, 1999).

[20] Gelfand, A. E., Hills, S. E., Racine-Poon, A. and Smith, A. F. M., 1990, “Illustration of Bayesian inference in normal modes using Gibbs sampling,” J. Am. Stat. Assoc. **85**, pp. 972 – 985.

[21] Michalopoulou, Z. H. and Picarelli, M., 2005, “Gibbs sampling for time-delay and amplitude estimation in underwater acoustics,” J. Acoust. Soc. Am. **117**, pp. 799 – 808.

[22] Daubechies, I., 1992, Ten Lectures on Wavelets. Philadelphia : SIAM.

[23] Choy, S. T. B. and Chan, C.M., 2005, “Bayesian Student-t stochastic volatility models.” Statistical Solutions to Modern Problem: Proceedings of the 20th International Workshop on Statistical Modelling, pp. 139-142.

[24] Robert, C. P., 1995, “Simulation of truncated normal variables,” Stat. Comput. **5**, pp. 121 – 125.

[25] Kuttruff, H., 2000, Room Acoustics 4th ed. London: E & FN Spon.

[26] Randall, R. B., 1987, Frequency Analysis. Nærum: Brüel & Kjær.

[27] Armstrong, J., 1986, Condition Monitoring – An Introduction to Its Application in Building Services, TN 1/86, BSRIA.

[28] Jarque, C., and Bera, A., 1980, "Efficient Tests for Normality, Heteroscedasticity, and Serial Independence of Regression Residuals," Econ. Lett., **6**, pp. 255 – 259.

[29] D'Agostino, R. B., 1971, "An Omnibus Test of Normality for Moderate and Large Size Samples," Biometrika, **58**, pp. 341 – 348.

[30] Shapiro, S. S., and Wilk, M. B., 1965, "An Analysis of Variance Test for Normality (Complete Samples)", Biometrika, **52**, pp.591 – 611.

[31] D'Agostino, R. B., 1972, "Small Sample Probability Points for the *D* Test of Normality," Biometrika, **59**, pp.219 – 221.

[32] Dong, L. B., and Giles, D. E. A., 2004, "An Empirical Likelihood Ratio Test for Normality", Econometrics Working Paper EW P 0401, Department of Economics, University of Victoria, Australia.

[33] Bendat, J.S. and Piersol, A.G., 1986, "Random Data : Analysis and Measurement Procedures", Wiley, New York.

[34] Mcdonalad, J.B. and Newey, W.K., 1988, Partially adaptive estimation of regression models via the generalized t distribution. Economic Theory 4, pp.428-457.

[35] Arslan, O and Genc, A.I.., 2003, Robust location and scale estimation based on the univariate generalized *t* (GT) distribution. *Communications in Statistics: Theory and Methods* **32**, pp.1505-1525.