MONASH UNIVERSITY
THESIS ACCEPTED IN SATISFACTION OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

ON........[redacted]...

..............[redacted]

~~for~~ Sec. Research Graduate School Committee

Addendum

Page 15, at last paragraph, add: "and its usefulness will be demonstrated throughout the thesis".
Page 20, first line after Eq. (3.2.2), delete: "... and j is a complex number, $j^2 = -1$".
Page 21, at the end of fourth sentence after heading 3.3, add: "as will be clearly shown and demonstrated in this chapter".
Page 23, second line, delete: "...and thus its real and "imaginary" parts are identical".
Page 24, sixth line, delete:"...and thus their real and "imaginary" parts are identical".
Page 68, Eq. (4.2.9) should be replaced by: $\psi_{Soliton}(\theta) = [sech(\beta\theta)]^2 \cdot cos(\pi\theta)$
Page 68, second line from bottom, delete "important" and read "... examine some properties..."
Page 70, Eq. (4.3.6) should be replaced by:

$$C_{constraint} = \int_{-\infty}^{+\infty} -\beta^2 \cdot sech\,(\beta\omega)\,\{1 - 2\cdot[sech\,(\beta\omega)]^2\}\,d\omega$$

Page 70, at the end of second last paragraph, add: "Eq. (4.3.6) could be analytically verified using Maple or Mathematica software. However, a simulation method is employed to evaluate integrals since the method is fast and yields further understanding on the integrand. Thus, simulation will be used throughout this chapter and subsequent chapters where applicable".
Page 95, first line after heading 4.4.5: "sometimes" for "sometime".
Page 116, at the end of second last paragraph, add: "The above range of $0.5 \leq \beta \leq 50$ could be considered "rule-of-thump" for the hyperbolic kernel. It is important to note that the above performance-analysis procedure for the hyperbolic kernel should be applied to other kernels so that their true performance can be obtained."

Time-Frequency Analyses of the Hyperbolic Kernel and Hyperbolic Wavelet

Lê, Nguyên Khoa

B. E. (Hons.)

A thesis submitted for the degree of

Doctor of Philosophy

in the

Department of Electrical and Computer Systems Engineering

Monash University

Clayton, VIC. 3168, Australia

8 March, 2002

TABLE OF CONTENTS

PREFACE

The research presented in this thesis is about time-frequency signal processing and parallel computing. Since Fourier transform method has been a dominant tool in signal processing for many decades, the role of time-frequency signal processing seems to be forgotten. In other words, there are many non-stationary signals, i.e. signals whose frequency spectrum varies with time, encountered in practice which requires the use of time-frequency signal processing. For example, for signals such as ECG, music, speech, underwater signal analysis, plasma physics, chaos, time-frequency signal analysis is required since their spectra vary with time. If the Fourier spectrum technique is employed, fine details of the spectrum will be lost and thus it does not clearly reveal the signal characteristics.

This thesis is written with a prime purpose to brighten up the time-frequency signal processing area by proposing a new kernel family. Unlike other kernels proposed in the literature, this kernel family is found by taking the advantage of summing two first-power exponential functions which are familiar in many aspects of electrical engineering. Moreover, properties of these functions have been well known and thus make the analysis of the new kernel family more effective and easier than using higher-power exponential functions. It should also be noted that time-frequency analysis is built upon the Fourier transform method except that the former can display power spectra in both time and frequency domains.

For all graphs in the thesis, the captions include important information about the graphs. The header of some graphs might have incomplete sentences or contains strings of numbers for programming purposes. In those cases, the captions should be solely consulted. Some captions may also contain a lot of text explaining about the graphs since it is believed that the graphs should be self-explainable and understandable at the first glance.

LIST OF SYMBOLS

$\Phi(\theta, \tau)$	Kernel function in terms of θ and τ
$P(t, \omega)$	Time-frequency power spectrum as a function of time t and frequency ω
τ	Lag parameter
$R(t, \tau)$	Local auto-correlation function as a function of time t and τ
$R(\tau)$	Auto-correlation function as a function of τ
$\cdot$	Multiplication operation
$\delta(\cdot)$	Impulse function
$F\{\cdot\}$	1-D Fourier transform operation
$\hat{F}(\cdot)$	Fourier transform of $(\cdot)$
WV	Wigner-Ville kernel function
CW	Choi-Williams kernel function
MTE	Multiform Tiltable Exponential kernel function
Hy	Hyperbolic kernel function
n	Order of the hyperbolic kernel family, $n = 1$ corresponds to first order.
W	Weighting function
β, σ	Control parameters of the hyperbolic and CW kernels respectively
α, r, β_{MTE}, γ, λ	Control parameters of the MTE
SNR	Signal-to-noise ratio
σ^2	Noise variance
$\psi(t)$	Wavelet function
$\varphi(t)$	Scaling function
C_ψ	Admissibility constant
t_0	First moment in time
ω_0	First moment in frequency
$\omega'_p = 2\pi f'_p$	Dimensional peak frequency of a wavelet function
a_{max}	Scale limit
j_{max}	Total number of possible calculated scales
ω_d	Scale resolution
T	Length of the time sampling interval of a wavelet or the window width of a wavelet

WT	Wavelet transform
FIR	Finite Impulse Response
N	Total number of samples for a wavelet function*
N_{wrap}	Number of wrap around points in the wavelet transform operation
M	Total number of input data samples
M_a	Vanishing moment order a
$h(t)$	Scaling filter of the scaling function $\varphi(t)$
η	Statistics of a binary hypothesis
A_0	Total energy of a discrete signal
B_0	Difference energy of a discrete signal
N_0	Noise variance or noise energy (analogous to A_0)
M_0	Difference energy of noise (analogous to B_0)
Q	Quality or improvement factor
CORR	Cross-correlator signal detector or matched-filter
X_1	Energy ratio of an input signal (ratio of B_0 to A_0)
X_2	Energy ratio of noise (ratio of M_0 to N_0)
HyD	Hyperbolic detector
CWWD	Choi-Williams detector
$WPS(t, \omega)$	Wavelet Power Spectrum as a function of time t and frequency ω
sym3	The *sym3* wavelet obtained from MATLAB software package
$B(f_1, f_2)$	The bispectrum as a function of two frequency variables f_1 and f_2
a	Scale index†
b	Time index
2-D $FFT\{\cdot\}$	2-D Fourier transform of $\{\cdot\}$
PCA	Power C Analyser
CPU	Central Processing Unit
HOTBLACK	Name of a 12-CPU super computer at Monash University
C	Computer programming language
MIMD	Multiple Instructions Multiple Data parallel machine
$Var(\cdot)$	Variance operation of the function $(\cdot)$
$O(\cdot)$	In the order of $(\cdot)$
$\forall(\cdot)$	for all values of $(\cdot)$

* N is used as "number of processors" in Chapter 8.
† a is used as "auto-term slope" in Chapter 3.

ACKNOWLEDGEMENT

I would like to thank my two wonderful supervisors to whom I owe great debt in completing this thesis, Dr. Kishor P. Dabke and Prof. Gregory K. Egan. I sincerely thank them for their useful guidance, kindness, patience and great sense of humor. They have taught me how to become an independent research student and what it takes to do a Ph.D. Dr. Kishor P. Dabke has also been one of my closest friends who will be there for me in bad and good times. I am grateful to Greg for his generous support on my expenditure when attending conferences. I am also grateful to Monash University for granting me the MGS and IPRS scholarships along with generous travel grants during my candidature.

I have been lucky to have so much great support from friends and family, thanks particularly to Philip Branch, Peter Freere and Jane Moodie for endless encouragement and for proof reading the early draft of the thesis. Ian Kaminskij has been generous for providing the music files. I am thankful to my Grand Mother for her great encouragement, my lovely aunts in Australia and in Vietnam for their constant support. Finally, this thesis is dedicated to my parents and my younger brother whose endless encouragement has motivated me to go beyond and further.

DECLARATION

The research presented in this thesis were performed in the Department of Electrical and Computer Systems Engineering, Monash University, Clayton, Victoria, Australia. This thesis contains no material which has been accepted for the award of any other degree or diploma in any university, or work published by another person, except where due reference is made in the text.

[redacted]

Khoa N. Le
Monash University, Clayton
Melbourne, Australia
March, 2002

LIST OF PUBLICATIONS

The following papers have been published, accepted, submitted or is under preparation for publication

Thesis Chapters	Corresponding Publications
Chapter 8	K. N. Le, G. K. Egan, and K. P. Dabke, "Parallel Computation of the Bispectrum," *ISSPA '99*, Brisbane, Australia, **1**, Aug. 1999, pp. 251-254.
Chapter 8	K. N. Le, G. K. Egan, and K. P. Dabke, "Parallel Computation of the Time-Frequency Power Spectrum," *ICII 2001 Proceedings – Conference C*, Beijing, China, Nov. 2001, pp. 322-28.
Chapter 4	K. N. Le, K. P. Dabke and G. K. Egan, "The Hyperbolic Wavelet Function," *SPIE AeroSense Proceedings: Wavelet Applications VIII*, April 2001, **4391**, Orlando, Florida, USA, pp. 411-22.
Chapter 5	K. N. Le, K. P. Dabke and G. K. Egan, "Signal Detection Using Non-Unity Kernel Time-Frequency Distribution," *Optical Engineering*, Dec. 2001, **40**, No. 12, pp. 2866-77.
Chapters 2 & 3	K. N. Le, K. P. Dabke and G. K. Egan, "The Hyperbolic Time-Frequency Power Spectrum," submitted for publication.
Chapter 6	K. N. Le, K. P. Dabke and G. K. Egan, "The Hyperbolic Wavelet Power Spectra of Typical Signals," to be submitted for publication
Chapter 7	K. N. Le, K. P. Dabke and G. K. Egan, "The Hyperbolic Wavelet Power Spectra of Music and Speech Signals," to be submitted for publication

ABSTRACT

This thesis proposes a new hyperbolic kernel family $[sech(\beta\theta\tau)]^n$, where $n = 1, 3, 5\ldots$, for time-frequency power spectrum analysis. An important relationship between time-frequency kernels and wavelet functions is found which leads to the discovery of the new hyperbolic wavelet function.

Theoretical background of the first-order hyperbolic kernel, which corresponds to $n = 1$, and its corresponding wavelet are examined in detail. The effectiveness of the first-order hyperbolic kernel is compared with previous kernels including Choi-Williams, Wigner-Ville and the multiform tiltable exponential. The hyperbolic, Morlet and Choi-Williams wavelets are examined so that appropriate applications of each wavelet can be identified.

There are two major applications of the hyperbolic kernel and hyperbolic wavelet presented in this thesis. The first application employs the hyperbolic kernel to establish a general non-unity kernel time-frequency detector and hyperbolic time-frequency detector for non-stationary signal detection. Detection of stationary signals such as sinusoids will also be investigated to validate the effectiveness of the detector. Performance of the general non-unity kernel time-frequency detector, hyperbolic detector, Choi-Williams detector, matched-filter detector and Wigner-Ville detector will be compared and discussed in detail by calculating their corresponding signal-to-noise ratios, *SNR*s.

The second application uses the proposed hyperbolic wavelet function for signal analysis, especially to differentiate among periodic, transient and chaotic properties of various signals, including Duffing oscillator, ECG, sinusoids, exponential, music and speech, by calculating their hyperbolic wavelet power spectra. The hyperbolic wavelet power spectrum and Fourier power spectrum techniques are compared. The merits of each technique will be clearly identified.

Parallel computing, as a useful tool, is employed to improve the efficiency in calculating the bispectrum and time-frequency power spectrum. A 12-processor parallel computer system is employed to run parallel programs whose speedup factors and efficiency are measured.

Chapter 1: INTRODUCTION

Signal analysis is important in understanding the behaviour of electrical systems such as power plants, control systems in telecommunication, chaotic systems such as the human electrocardiogram (ECG), plasma phenomena, oscillatory systems and other non-linear systems. This understanding enables electrical engineers to predict the behaviour of the system in the future, to measure the performance of the system, to study the system characteristics in detail and to effectively improve the system performance. Many tools for signal analysis have been developed over many years for stationary and non-stationary signals with constant improvements in efficiency, effectiveness, accuracy and cost.

The most important signal-processing tool is the Fourier transform which is very useful in spectral analysis. Spectral analysis is employed to establish important functions in higher-order statistics of signals such as the power spectrum (first-order statistics) and bispectrum (second-order statistics). The power, or energy, spectrum can be employed to study chaotic phenomena and turbulence [1-4]. However, it should be emphasised that for Fourier analysis to be effective and accurate, the signal is assumed to be stationary or wide-sense stationary. That is, its statistical properties do not vary with time.

The Fourier transform is formed by summing and then averaging a product of the input signal and a sum of two sinusoidal functions or the complex exponential function $exp(-j\omega t)$ whose time support is infinite. If the Fourier transform is used to examine non-stationary signals, fine-detailed information of the signal energy density will be suppressed since the Fourier transform yields the average signal energy density over an infinite time interval. In practice, non-stationary signals are often encountered and thus it is important to develop methods to study these signals. This is one of the main themes of the research presented in this thesis. Examples of non-stationary signals are voice and speech, underwater signals such as whale sounds, bat sounds, transient signals, ECG and other biological signals, chaotic signals and sun-spot intensity over a short- or long-time period [5-8]. It is important that the energy density or energy distribution of a non-stationary signal is viewed as a function of time so that it is known when a particular event happens.

To effectively investigate non-stationary signals, a method called the joint time-frequency analysis [8, 9] or time-varying spectrum analysis was proposed by Page, Rihaczek and others [10-12]. This method was later studied in detail and significantly modified by Cohen [8] so that time-frequency analysis could be generalised. This method uses a general transform function for a general expression of the time-frequency distribution. For each different transform function that satisfies a certain set of constraints, a new and unique time-frequency distribution is found.

To capture the fine-detailed information in the energy density of non-stationary signals, the time-frequency transform function must be a finite-time supported function. This transform function is called the kernel function and labelled $\Phi(\theta, \tau)$. The kernel function, $\Phi(\theta, \tau)$, has two arguments θ and τ, which are frequency and time respectively and are used to derive the joint time-frequency spectrum. It has been shown that Fourier analysis is a special case of the joint time-frequency analysis [8]. In the case of the Fourier transform, only one argument, ω, which corresponds to the frequency, is employed. The time variable is removed through averaging and so has no variation.

Different time-frequency spectra can be generated if there exist different kernel functions. The simplest kernel function that has been employed is the Wigner-Ville (WV) unity kernel, $\Phi_{WV}(\theta, \tau) = 1$. The WV time-frequency distribution has been extensively studied by many researchers [13-22], but suffers from a disadvantage that it generates cross terms in the time-frequency plane. These cross terms originate from instantaneous values of an auto-correlation function and hence make it difficult to interpret the WV time-frequency power spectrum. The main problem is that even when the signal is absent, the WV time-frequency power spectrum is not zero.

The existence of cross terms in the time-frequency domain motivated the search for new kernels that are more effective than the WV kernel with desirable properties such as effective cross-term suppression, fine auto-term resolution and high level of noise robustness. In 1989, the Choi-Williams (CW) kernel, $\Phi_{CW}(\theta, \tau) = e^{-\theta^2\tau^2/\sigma}$ where σ is the kernel control parameter, was proposed as one outstanding candidate to effectively suppress the cross terms. The main principle of the CW kernel is that the location of auto terms and cross terms is identified. That is, the auto terms are located in the vicinity of the origin and the cross terms away from the origin. The weighting function, which is the 1-D Fourier transform of the kernel, of the CW kernel peaks at the origin and decays to small values as the time and frequency variables τ and θ of the kernel increase. This means that only the auto terms are amplified while the cross terms are suppressed in the time-frequency plane, which means cross terms can be effectively avoided.

The CW kernel is a second-power exponential function which is sometimes difficult to integrate, especially when being combined with high-order power series. This complexity of the CW kernel has motivated the need to find simpler first-power kernels. However, the CW kernel has been applied in different fields and has found many applications in speech processing, image processing, wavelet theory, underwater signal analysis and other fields [6, 8, 9, 23-26]. Among the various applications, wavelet theory has been the most useful application of the CW kernel because of the presence of the Mexican-hat wavelet [51-67] function.

The wavelet transform and joint time-frequency power spectrum are closely related since they both employ time and frequency domains to describe the power spectrum of a non-stationary signal. The wavelet transform is a time-frequency transform which employs a wavelet function to display the signal energy density. The main difference between the joint time-frequency analysis and wavelet analysis is that the former employs the Fourier transform and the latter uses the wavelet transform which employs a wavelet function. Wavelet functions are usually finite time supportive so that fine-detailed information of the energy spectrum can be effectively monitored. If a new wavelet function is found, then a new wavelet transform, using that particular wavelet function, can be generated.

In 1995, the multiform tiltable exponential (MTE) kernel was proposed by Costa and Boudreaux-Bartels [27] which has been claimed to be more effective than the CW kernel in suppressing cross terms. The MTE kernel is given by the following equation

$$\Phi(\theta,\tau) = exp\left\{-\pi\left[\left[\frac{\tau}{\tau_0}\right]^2\left(\left[\frac{\theta}{\theta_0}\right]^2\right)^\alpha + \left(\left[\frac{\tau}{\tau_0}\right]^2\right)^\alpha\left[\frac{\theta}{\theta_0}\right]^2 + 2r\left(\left[\frac{\tau\theta}{\tau_0\theta_0}\right]^\beta\right)^\gamma\right]^{2\lambda}\right\} \tag{1.1.1}$$

where τ_0 and θ_0 are constants. The CW kernel is a special case of the MTE kernel when $r = 0$ and $\alpha = \lambda = 1$. The multiform tiltable exponential kernel has ten different forms whose shapes depend on values of its control parameters $\alpha, \beta, \gamma, \lambda$ and r.

Although the MTE kernel has been shown to be effective and diversified in terms of kernel variety and cross-term suppression [27], it is a very complicated kernel and it is not possible to generate its explicit time-frequency expressions. In addition, some of the MTE kernel types are not Fourier-transformable which makes the MTE kernel difficult to use. Consequently, a new kernel which is simpler and more effective than the CW and MTE kernels is urgently needed.

A kernel is said to be effective if it meets the following criteria: it can effectively suppress cross terms, support the auto terms by having a fine auto-term resolution and it is noise robust under noisy conditions or interference. These are the three criteria that are used to judge the performance of time-frequency kernels. Even if only two of the three criteria listed above are satisfied, the kernel can be said to be effective. In practice, all three criteria :e very difficult to satisfy simultaneously and there exists a trade-off among them.

Cross terms, generated because of coupling of various signals in the time-frequency plane, are undesirable in the time-frequency power spectrum. Cross-term suppression is crucial to prevent false interpretation of the signal characteristics as happens with the WV kernel. The ratio of the cross-term peak magnitude to auto-term peak magnitude measures the performance of a kernel. The lower this ratio is, the more effective the kernel.

Auto-term resolution is also important in time-frequency power spectrum analysis. The finer the auto-term resolution is, the better a particular kernel can support auto terms. Some kernels suppress the auto term and cross terms simultaneously since they cannot separate the auto terms from the cross terms. Auto terms are located near the origin of the time-frequency power spectrum and the cross terms are located further away from the origin. Thus, increasing the auto-term resolution of a time-frequency power spectrum makes its corresponding kernel more effective. In fact, the effectiveness of a kernel is measured by the peak-magnitude ratio of the cross terms to the auto terms. This is the key factor that must be taken into account when designing new kernels.

Noise robustness measures the ability of a time-frequency power spectrum to withstand external interference or noisy disturbances created by the surroundings such as transmission noise. Under noisy conditions, the time-frequency power spectrum of a particular kernel function should retain its original shape and the amount of deformation is expected to be minimal. The less the deformation is, the better the noise robustness of the kernel function.

Investigations on cross-term suppression [28], auto-term resolution and noise robustness of the CW kernel have been carried out by Stankovic and Amin [29, 30]. The MTE kernel has been studied in detail by Costa and Boudreaux-Bartels [27] and has been shown that its cross-term suppression is superior compared with the CW kernel. However, the MTE kernel has not been investigated and reported in the literature on auto-term resolution and noise robustness analyses. Even though the MTE kernel is the most effective kernel in suppressing cross terms in the time-frequency plane, its auto-term resolution and noise robustness ability raise unanswered questions about the kernel effectiveness. Although it has been shown that the CW kernel is effective in meeting all three criteria and the MTE is superior in cross-term suppression, it is still believed that a simpler and more effective kernel can be found.

Criteria on how to evaluate an effective kernel have been stated which enable one to propose an effective kernel and consequently derive an effective time-frequency power spectrum. Moreover, a signal-processing method should be efficiently calculated. One typical and common problem of the time-frequency power spectral analysis is that its computation is extensive because a large amount of instantaneous values of the auto-correlation function (for the time-frequency power spectrum) are required. Thus, there is a need to develop new tool(s) to improve the efficiency of time-frequency analysis.

1.1 Proposal of the New Kernel

This thesis proposes a new family of kernels, called the hyperbolic family kernels, $\Phi(\theta, \tau) = [sech(\beta\theta\tau)]^n$ where n is the order of the kernel family and $n = 1$ corresponds to the first-order hyperbolic kernel; β is the kernel control parameter. The corresponding hyperbolic time-frequency power spectrum is employed to study non-stationary signals. Properties of the hyperbolic kernel such as cross-term suppression, auto-term resolution and noise robustness will be investigated. In addition, a new wavelet function, hyperbolic wavelet, is generated and its properties are studied in detail. Applications of the hyperbolic kernel and hyperbolic wavelet are discussed in the fields of signal detection and studies of signal characteristics.

The first aim of the research reported in this thesis is to propose the new hyperbolic kernel for use with time-frequency power spectrum and compare it with other well-known kernels such as the CW and multiform tiltable exponential (MTE) which are used as comparison benchmarks. A trade-off among auto-term resolution, cross-term suppression and noise-robustness of the hyperbolic kernel, CW kernel and the MTE kernel will be stated.

The second aim of this research is to show that there is a strong link between time-frequency kernels and wavelet functions. It will be shown that if a new time-frequency kernel is proposed, it simultaneously gives rise to a new wavelet function (provided that the wavelet function satisfies admissibility constraint(s)) and vice versa. This concept extends the time-frequency and wavelet fields which means more useful time-frequency kernels can be generated from the corresponding wavelet functions and vice versa.

We explore the relationship between time-frequency kernels and wavelet functions in some detail. This is an important contribution of the thesis. We begin by carefully examining the Mexican-hat wavelet function and show that it is generated from the Gaussian pulse function, which is essentially the CW kernel. In other words, there is a strong relationship between time-frequency kernels and wavelet functions that has not been reported in the literature.

The third aim is to demonstrate possible applications of the time-frequency power spectrum, in particular, the hyperbolic time-frequency power spectrum, in areas such as signal detection and signal analysis, especially chaotic studies. Non-stationary signal detection has been identified as one such application of the time-frequency power spectrum because of the outstanding work by Kumar and Carroll [31, 32]. Detection of chaotic behaviour using the time-frequency power spectrum and wavelet transform, which has been inspired by the work of Milligen and Farge [33-35], is another typical application of the time-frequency power spectrum. Since the time-frequency power spectrum shows how the energy density of an input signal changes with time and frequency, it is possible to detect the transition from the periodic region to the chaotic region of the signal which is useful in determining when a non-linear signal is behaving chaotically.

The fourth aim of the research is to investigate the effectiveness of parallel computing in the calculation of the bispectrum and the hyperbolic time-frequency power spectrum as two typical signal-processing methods. Parallel computing has steadily developed over many decades, yet it has not been extensively used due to high cost and programming difficulties. It will be shown that the use of parallel computing is effective and parallel programming is a powerful tool for dealing with heavy-computation tasks of the hyperbolic time-frequency power spectrum and the bispectrum, that are common problems in the field of higher-order statistical signal processing.

1.2 Thesis overview

The thesis is organised as follows.

Chapter 2 reviews a number of already-published kernels in the literature so that an overview on time-frequency kernels is given and the proposal of the hyperbolic kernel is appropriate.

Chapter 3 proposes the new hyperbolic kernel with its family and studies its properties in detail. The advantages and disadvantages of this kernel family are discussed and the kernel is compared with the CW and the MTE kernels in terms of auto-term resolution, noise robustness and cross-term suppression.

Chapter 4 continues the analysis of the hyperbolic time-frequency kernel by investigating the new hyperbolic wavelet, which is generated from the hyperbolic kernel by taking its negative second-order derivative function. The research reported in this chapter was inspired by the fact that the popular Mexican-hat wavelet was generated from the CW kernel. This important fact, however, has not been reported in the literature. The hyperbolic, CW or Mexican-hat and Morlet wavelets are compared in this chapter from an engineering point of view in terms of aliasing effects, the number of sampling points for the wavelet function, the maximum scale and scale resolution. One typical example is provided to illustrate the advantages and disadvantages of the hyperbolic wavelet. From a mathematical point of view, the hyperbolic wavelet is briefly studied in the second part of the chapter.

The first application of the hyperbolic time-frequency power spectrum is in the field of signal detection in Chapter 5. In this chapter, performances of the hyperbolic, CW (non-unity kernels), WV (unity kernel) and *cross-correlator* signal detectors are compared in detail in terms of their signal-to-noise ratios (*SNR*s) and quality factors *Q*'s. This chapter shows that non-unity-kernel time-frequency signal detectors are better than the unity-kernel signal detector (WV signal detector) in terms of *SNR*. The hyperbolic signal detector will be shown to be better than the Choi-Williams signal detector by having a larger *SNR*. From this chapter, Chapters 3 and 4, the applicable range of β and trade-off among cross-term suppression, auto-term magnitude, noise robustness, scale resolution, signal-detection ability against auto-term resolution are stated.

Chapters 6 and 7 report the second application of the time-frequency power spectrum technique for signal analysis. In these chapters, the hyperbolic wavelet power spectra of signals such as sinusoids, exponentially decaying sinusoids, Duffing oscillator, the ECG, music and speech are calculated so that their instantaneous characteristics or any transitions from periodicity to chaos can be detected. Chapter 6 lays a foundation for Chapter 7 by forming a gallery of hyperbolic wavelet power spectra of various familiar signals including sinusoidal, exponential, exponentially decaying sinusoidal, Duffing oscillator and ECG. Music and speech signals are separately investigated in Chapter 7. The Fourier power spectrum and the wavelet power spectrum techniques are also compared in detail in Chapters 6 and 7.

Chapter 8 reports on the measured- and effective-speedup factors that can be achieved when a 12-processor parallel computer is used to estimate the bispectrum and time-frequency power spectrum. It will be shown that parallel computing can significantly improve the efficiency of the bispectrum and time-frequency power spectrum calculation processes.

Finally, Chapter 9 presents the conclusions of this research and summarises important contributions of the thesis. The chapter also outlines new directions for future research.

The first application of the hyperbolic time-frequency power spectrum is in the field of signal detection in Chapter 5. In this chapter, performances of the hyperbolic, CW (non-unity kernels), WV (unity kernel) and *cross-correlator* signal detectors are compared in detail in terms of their signal-to-noise ratios (*SNR*s) and quality factors *Q*'s. This chapter shows that non-unity-kernel time-frequency signal detectors are better than the unity-kernel signal detector (WV signal detector) in terms of *SNR*. The hyperbolic signal detector will be shown to be better than the Choi-Williams signal detector by having a larger *SNR*. From this chapter, Chapters 3 and 4, the applicable range of β and trade-off among cross-term suppression, auto-term magnitude, noise robustness, scale resolution, signal-detection ability against auto-term resolution are stated.

Chapters 6 and 7 report the second application of the time-frequency power spectrum technique for signal analysis. In these chapters, the hyperbolic wavelet power spectra of signals such as sinusoids, exponentially decaying sinusoids, Duffing oscillator, the ECG, music and speech are calculated so that their instantaneous characteristics or any transitions from periodicity to chaos can be detected. Chapter 6 lays a foundation for Chapter 7 by forming a gallery of hyperbolic wavelet power spectra of various familiar signals including sinusoidal, exponential, exponentially decaying sinusoidal, Duffing oscillator and ECG. Music and speech signals are separately investigated in Chapter 7. The Fourier power spectrum and the wavelet power spectrum techniques are also compared in detail in Chapters 6 and 7.

Chapter 8 reports on the measured- and effective-speedup factors that can be achieved when a 12-processor parallel computer is used to estimate the bispectrum and time-frequency power spectrum. It will be shown that parallel computing can significantly improve the efficiency of the bispectrum and time-frequency power spectrum calculation processes.

Finally, Chapter 9 presents the conclusions of this research and summarises important contributions of the thesis. The chapter also outlines new directions for future research.

Chapter 2: REVIEW OF TIME-FREQUENCY KERNELS

2.1 Introduction

This chapter gives details of the time-frequency kernels which have been published in the literature. The most notable kernels are the Choi-Williams and multiform tiltable exponential kernels which will be compared with the proposed kernel in Chapter 3. The general expression for the time-frequency power spectrum is [8]

$$P(t,\omega)=\frac{1}{4\pi^2}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-j\theta t-j\tau\omega+j\theta u}\cdot\Phi(\theta,\tau)\cdot R_{t,1}(t,\tau)\ dud\tau d\theta \tag{2.1.1}$$

The chapter is mainly based on the review papers by Cohen [8] and Janse and Kaiser [16]. Some of the already-published kernels in the literature are listed along with their distributions or time-frequency power spectra.

2.2 Revision of time-frequency kernels

2.2.1 The Wigner-Ville (WV) Unity Kernel

The WV kernel is a unity kernel which is given in Eq. (2.2.1)

$$\Phi(\theta,\tau)=1 \tag{2.2.1}$$

The corresponding Cohen distribution of this kernel is given by Eq. (2.2.2)

$$WCV(t,\omega)=\frac{1}{2\pi}\int e^{-j\tau\omega}\cdot\delta(t-u)\cdot x^*\left(u-\frac{\tau}{2}\right)\cdot x\left(u+\frac{\tau}{2}\right)d\tau \tag{2.2.2}$$

The function $\delta(t-u)$ is the Fourier transform of the unity function $\Phi(\theta, \tau) = 1$. For this function to exist, the temporary variable u must be equated to the time variable t which yields the Wigner-Ville time-frequency power spectrum as a function of the time t and frequency ω

$$WCV(t,\omega) = \frac{1}{2\pi}\int e^{-j\tau\omega}\cdot x^{*}\left(t-\frac{\tau}{2}\right)\cdot x\left(t+\frac{\tau}{2}\right)d\tau \qquad (2.2.3)$$

This is the simplest time-frequency kernel in Cohen's class. The characteristics of the WV distribution have been extensively studied in a series of three papers by Claasen and Mecklenbrauker [13-15]. The 1-D Fourier transform of the kernel is an impulse concentrated at every time instant t as can be seen in Eq. (2.2.2). Hence the term $\delta(t-u)$ can be ignored which gives the standard form of the distribution as given in Eq. (2.2.3) [8, 16]. It should be noted that Eq. (2.2.2) is basically the form of an ambiguity function as defined in [8, 27].

Consider first the input signal as a mono-component signal. The WV time-frequency power spectrum estimation of the input signal is purely based on the auto-correlation method. This means that if there exists a certain degree of similarity of the signal in the past and future (regardless the value of the signal) then the WV power spectrum will be non-zero as explained by Cohen [7] in terms of "overlapping of the signal". For the case of a multi-component input signal, where the input signal comprises of a number of mono-component signals, there will be interference among the power spectra of the mono-component signals which provide misleading information about the time-frequency power spectrum. Thus the search for new kernels that are effective at cross-term suppression is necessary.

2.2.2 Choi-Williams (CW) Kernel

The CW kernel was first proposed in 1989 and is given by Eq. (2.2.4)

$$\Phi(\theta,\tau) = exp\left(-\theta^{2}\tau^{2}/\sigma\right) \qquad (2.2.4)$$

where σ is the kernel control parameter.

The CW distribution (based on Cohen's class) is given by Eq. (2.2.5)

$$CWC(t,\omega)=\frac{1}{2\sqrt{\pi}}\iint\sqrt{\frac{\sigma}{\tau^2}}\cdot e^{-j\tau\omega}\cdot exp\left[-\frac{\sigma(t-u)}{4\tau^2}\right]\cdot x^*\left(u-\frac{\tau}{2}\right)\cdot x\left(u+\frac{\tau}{2}\right)du d\tau \tag{2.2.5}$$

where $\tau = 0, 1, \ldots, M-1$ and M is the length of the discrete input signal $x(\cdot)$.

The CW kernel was the first kernel proposed in the reduced interference kernel class [48] and is well recognised for its effectiveness in suppressing cross terms in the time-frequency plane.

2.2.3 The Generalised Choi-Williams Kernel

The generalised CW kernel is defined by Eq. (2.2.6)

$$\Phi(\theta,\tau)=exp\left[-\left(\frac{\theta}{\theta_1}\right)^{2N}\cdot\left(\frac{\tau}{\tau_1}\right)^{2M}\right] \tag{2.2.6}$$

where M, N, θ_1 and τ_1 are positive integer constants.

This kernel [25] attempts to reduce the transition region of the kernel so that the cross terms can be more effectively reduced when compared with the CW kernel. Derivations of the weighting function and hence the general formula of the generalised CW kernel are complicated. For example, for $N = M = 3$, the weighting function is complex and consists of products of Gamma and Lommel functions [44] which require further approximations. This kernel and the CW kernel are special cases of the multiform tiltable exponential kernel [27] which will be given in Section 2.2.6.

2.2.4 Butterworth Kernel

The kernel is defined by Eq. (2.2.7)

$$\Phi(\theta,\tau)=\frac{1}{1+\left(\frac{\theta}{\theta_1}\right)^{2N}\cdot\left(\frac{\tau}{\tau_1}\right)^{2M}}\text{, where } M \text{ and } N \text{ are arbitrary constants.} \tag{2.2.7}$$

This kernel [25] and the generalised CW kernel were studied by Papandreou and Boudreaux-Bartels. The main difficulty with the GED (Generalised Exponential Distribution) and the Butterworth kernels is that they are complicated (because of the high-power part of the kernels) and thus it is very difficult to derive their general distributions.

2.2.5 Cone-Shaped Kernel

This kernel was studied in [39, 46] and is given by Eq. (2.2.8)

$$\Phi(\theta,\tau)=\frac{sin\left(\frac{2\pi\theta|\tau|}{a}\right)}{\pi\theta}\cdot w(\tau) \tag{2.2.8}$$

where $w(\tau)$ is a function to be specified and $a \geq 2$ to ensure finite time support.

The general formula of the cone-shaped time-frequency power spectrum is given by [9]

$$ZAM(t,f)=\int_{-\infty}^{+\infty} w\tau\cdot e^{-j2\pi f}\int_{t-|\tau|/a}^{t+|\tau|/a} x^*\left(u-\tfrac{\tau}{2}\right)\cdot x\left(u+\tfrac{\tau}{2}\right)\,du\,d\tau \tag{2.2.9}$$

The cone-shaped kernel has been used to study speech and chirped signals and it was reported that the cone-shaped kernel produce good results in locating speech formants and pitch [47].

2.2.6 Multiform Tiltable Exponential (MTE) Kernel

This kernel was proposed in 1995 by Costa and Boudreaux-Bartels [27] and has the general form given by Eq. (2.2.10)

$$\Phi(\theta,\tau)=exp\left\{-\pi\left[\left[\frac{\tau}{\tau_0}\right]^2\left(\left[\frac{\theta}{\theta_0}\right]^2\right)^{\alpha}+\left(\left[\frac{\tau}{\tau_0}\right]^2\right)^{\alpha}\left[\frac{\theta}{\theta_0}\right]^2+2r\left(\left[\frac{\tau\theta}{\tau_0\theta_0}\right]^{\beta}\right)^{\gamma}\right]^{2\lambda}\right\} \tag{2.2.10}$$

where α, β, γ, λ and r are kernel control parameters which are independent of the signal parameters and subject to certain conditions given in [27]. The positive-valued parameters τ_0 and θ_0 can be designed to suit specific requirements.

The kernel given in Eq. (2.2.10) is a general formula for a number of kernels including the WV kernel, the CW, generalised CW kernel and the tilted Gaussian kernel. A detailed list of various types of the MTE kernel can be found in [27], which also includes the design procedures of their control parameters. The MTE kernel was shown to be effective in suppressing cross terms and to vary the stop band and pass band of the kernel depending on specific values of the kernel control parameters. This allows the MTE kernel flexibility in kernel design and a wider range of applications. However, the MTE kernel is a complicated kernel whose closed-form general time-frequency power spectrum cannot be obtained.

2.2.7 Reduced-Interference Kernel

Jeong and Williams [48] proposed a kernel design procedure to reduce cross terms in the power spectrum. They also compared a number of different kernels used in time-frequency distributions. Their kernel, the reduced interference kernel, can be used to derive other kernels, in particular, the Born-Jordan, G-Hamming, truncated-CW, truncated-*sinc* and triangular kernels. This kernel can be considered as one of the general kernels which is used to suppress cross terms by employing attenuation techniques in its time-frequency plane. Table 2.2.1 summarises some of the popular kernels that have been reported in the literature and their corresponding time-frequency distributions.

Table 2.2.1: Some popular kernels and their distributions [8]

Kernel Name	Kernel $\Phi(\theta, \tau)$	Distribution $P(t, \omega)$ (Eq. (2.1.1))
Wigner-Ville [16]	1	$\frac{1}{2\pi}\int e^{-j\omega\tau}x^*(t-\frac{\tau}{2})\cdot x(t+\frac{\tau}{2})\,d\tau$
Margenau And Hill [16]	$cos\left(\frac{\theta\tau}{2}\right)$	$\mathrm{Re}\left\{\frac{1}{\sqrt{2\pi}}x(t)\cdot e^{-j\omega t}\cdot \hat{X}^*(\omega)\right\}$
Kirkwood and Rihaczek [16]	$exp\left(\frac{j\theta\tau}{2}\right)$	$\frac{1}{\sqrt{2\pi}}x(t)\cdot e^{-j\omega t}\cdot \hat{X}^*(\omega)$
sinc [16]	$\frac{sin(a\theta\tau)}{a\theta\tau}$	$\frac{1}{4\pi a}\int\frac{1}{\tau}\cdot e^{-j\omega\tau}\int_{t-a\tau}^{t+a\tau}x^*(u-\frac{\tau}{2})\cdot x(u+\frac{\tau}{2})\,du d\tau$
Page [8]	$exp\left(\frac{j\theta\lvert\tau\rvert}{2}\right)$	$\frac{\partial}{\partial t}\left\lvert\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t}x(t_1)e^{-j\omega t_1}dt_1\right\rvert^2$
Choi-Williams [28]	$exp\left(-\frac{\theta^2\tau^2}{\sigma}\right)$	$\frac{1}{2\sqrt{\pi}}\iint\sqrt{\frac{\sigma}{\tau^2}}\cdot e^{-j\tau\omega}\cdot exp\left[-\frac{\sigma(t-u)}{4\tau^2}\right]\cdot x^*\left(u-\frac{\tau}{2}\right)\cdot x\left(u+\frac{\tau}{2}\right)du d\tau$
Product of Levin and Choi-Williams [11, 28]	$exp\left(-\frac{\theta^2\tau^2}{\sigma}\right)\cdot cos\left(\frac{\theta\tau}{2}\right)$	$\iint\sqrt{\frac{\sigma}{4\pi\tau^2}}\cdot e^{-j\tau\omega}\cdot exp\left[-\frac{\sigma(t-u)}{4\tau^2}\right]\cdot\frac{1}{2}\left[x^*(u-\tau)x(u)+x(u+\tau)x^*(u)\right]du d\tau$

2.3 Conclusion

This chapter has reviewed kernels that have been published in the literature. Second- and higher-power exponential kernels have been successfully employed in the field of time-frequency signal processing, especially the CW kernel and MTE kernel. However, first-power exponential kernels have not been effectively used. These kernels are simpler than the CW and MTE kernels but they do not provide effective cross-term suppression ability and noise robustness. Thus, there ought to be a new kernel, employing first-power exponential functions, which is simpler and also more effective than the CW and MTE kernels. This kernel will be presented in the next chapter.

Chapter 3: THE HYPERBOLIC KERNEL FOR TIME-FREQUENCY POWER SPECTRUM

This chapter proposes a new family of hyperbolic kernels $\Phi_{hyperbolic}(\theta, \tau) = [sech(\beta\theta\tau)]^n$, where $n = 1, 3, 5\ldots$. The first-order hyperbolic kernel $\Phi_{hyperbolic}(\theta, \tau) = sech(\beta\theta\tau)$ with $n = 1$ is mainly considered in this thesis. Theoretical aspects of the new hyperbolic kernel are examined and studied in detail. In the time-frequency context, the effectiveness of a kernel is determined by three factors: cross-term suppression, auto-term resolution and noise robustness. The effectiveness of the new kernel will be compared with other kernels including Choi-Williams, Wigner-Ville and multiform tiltable exponential (MTE) using two different signals: complex exponential and chirp. The results of this chapter form the foundation for the subsequent chapters in which the hyperbolic wavelet (Chapter 4), hyperbolic signal detector (Chapter 5) and hyperbolic time-frequency wavelet power spectrum technique (Chapters 6 and 7) will be examined.

3.1 Introduction

A cross-term effect in the power spectrum of multi-component signals represents interactions among the individual component signals. This effect, sometimes called the "artifact", is undesirable since the interactions among different mono-component signals in a multi-component signal provide no useful physical interpretation of the individual signals. For example, the artifact causes zero-valued regions of the original spectrum to be non-zero and complicates the interpretation of the time-frequency power spectrum as will be illustrated later. To eliminate artifacts, the modulus of cross terms in the time-frequency power spectrum must be reduced. However, cross terms cannot be completely reduced since a spectrum consists of both auto and cross terms [28]. One of the methods for reducing the effects of cross terms is to use an appropriate kernel for the computation of the power spectrum. A desirable property of a kernel is that it supports auto terms and suppresses cross terms in the time-frequency plane by multiplying them with its weighting function. A kernel is an arbitrary function which must satisfy a number of admissibility constraints. These constraints were studied in detail in [14] and are:

1. Kernel function, $\Phi(\theta, \tau)$, is independent of time t,
2. Kernel function is independent of frequency ω,
3. $\Phi(\theta, 0) = 1$ for all θ,
4. $\Phi(0, \tau) = 1$ for all τ,
5. Kernel function must be real, i.e. $\Phi(\theta, \tau) = \Phi^*(-\theta, -\tau)$, where "*" indicates the complex conjugate,
6. $\left.\frac{d}{d\tau}\Phi(\theta,\tau)\right|_{\tau=0} = 0, \quad \forall\theta$.
7. $\left.\frac{d}{d\theta}\Phi(\theta,\tau)\right|_{\theta=0} = 0, \quad \forall\tau$.

Detailed interpretations of the seven constraints above are given in Appendix A.

The main motivation in inventing new kernels is to more effectively suppress cross terms in the time-frequency power spectrum of multi-component signals. The Wigner-Ville (WV) distribution or Wigner-Ville time-frequency power spectrum, which employs a unity kernel, was first proposed by Wigner in 1932 to solve problems in quantum mechanics [8]. Since then, the WV time-frequency distribution has found many different applications including radar, speech recognition and loudspeaker design [16]. Further details on the WV distribution are provided in [8, 13-15, 36]. Since the WV kernel is unity, the cross terms in the time-frequency plane are not suppressed, i.e. they are scaled down by a unity factor, which is the main disadvantage of the WV distribution. These cross terms or "artifacts" provide misleading information about the WV time-frequency power spectrum. It should also be noted that the terms "time-frequency distribution", which was coined by Cohen [8, 37], and "time-frequency power spectrum", which was first used by Page and Rihaczek [10, 11], are identical. These terms have been extensively used by many different authors in the field of time-frequency signal processing. In this thesis, they will be interchangeably used without any difference in their meanings.

Currently, there are two kernels that have been shown to be useful and effective in the time-frequency power spectrum analysis which were introduced earlier. The first kernel was the Choi-Williams (CW) kernel which was proposed in 1989 by Choi and Williams [28]. The second kernel was the multiform tiltable exponential (MTE) kernel which was found in 1995 by Costa and Boudreaux-Bartels [27]. The CW kernel is a special case of the MTE kernel for some special values of the kernel parameters. The main problem of the MTE kernel is that some of the kernel types are not Fourier transformable which makes it difficult to use.

A number of alternative kernels have been proposed and studied in Chapter 2 such as $cos(0.5\theta\tau)$ by Margenau and Hill [8], $sinc(\beta\theta\tau)$ [10], exponential kernel $e^{j\theta\tau/2}$ [11], the compound kernel derived by taking a product of the Hill and CW kernels [38], cone-shaped kernel [39] and the generalised CW kernel [25]. These kernels, although easy to use, are not effective in cross-term suppression compared with the CW and MTE kernels.

The CW and MTE kernels are second- and higher-power exponential functions whose explicit expressions when integrating with power series do not exist. Thus, "first-power" exponential kernels are more suitable for a time-frequency distribution. The problem of first-power exponential kernels is that they are simple and thus some desirable properties such as effective cross-term suppression and noise robustness are missing. Therefore, finding the right kernel, which is easy to use and at the same time effective, is a difficult task. To overcome this difficulty is the main aim of this chapter.

The purpose of this chapter is to propose a new family of kernels, the family of hyperbolic functions, $[sech(\beta\theta\tau)]^n$ where n is the kernel family order, that can be used to suppress cross terms in the time-frequency power spectrum. These kernels provide better results than the CW kernel for well-chosen values of the kernel control parameter β. Since the MTE kernel is not Fourier transformable, it is not possible to compare its cross-term suppression and noise robustness with that of the hyperbolic and CW kernels. However, various forms of the MTE kernel are studied by estimating their auto-term widths, then comparisons are made among the MTE, hyperbolic and CW kernels. The hyperbolic and CW kernels are compared in detail in terms of cross-term suppression, auto-term resolution and noise robustness.

The chapter is organised as follows. The proposed hyperbolic kernel family is detailed in Section 3.3. Section 3.4 compares the weighting functions of the hyperbolic and CW kernels. Section 3.5 discusses cross-term suppression ability of the hyperbolic and CW kernels using multi-component chirped and complex exponential signals. Sections 3.6 and 3.7 compare the effectiveness of the CW, hyperbolic and multiform tiltable exponential (MTE) kernels in terms of their auto-term widths and noise robustness.

3.2 Background on Cohen's Time-Frequency Distribution

The general form of time-frequency power spectrum in Cohen's class for deterministic non-stationary signals is defined as [8]

$$P(t,\omega)=\frac{1}{4\pi^2}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-j\theta t-j\tau\omega+j\theta u}\cdot\Phi(\theta,\tau)\cdot R_{t,1}(t,\tau)\ dud\tau d\theta \qquad (3.2.1)$$

where $R_{t,1}(t,\tau)=x\left(u+\frac{\tau}{2}\right)\cdot x^*\left(u-\frac{\tau}{2}\right)$ is the local auto-correlation function, $\Phi(\theta,\tau)$ is the kernel function, $u=t+\frac{\tau}{2}$, τ is the lag parameter and t is the running time variable. The range of t is $0\le t\le t_0$, where t_0 is the signal window size over which the power spectrum of a non-stationary signal is estimated. From now on, the range of all integrals is from $-\infty$ to $+\infty$ unless otherwise stated.

The one-dimensional (1-D) Fourier transform of a function $x(t)$ is defined as [41]

$$\hat{F}(\omega)=\int_{-\infty}^{+\infty} x(t)\cdot e^{-j\omega t}\ dt \qquad (3.2.2)$$

where $\hat{F}(\omega)$ is the 1-D Fourier transform of $x(t)$ and j is a complex number , $j^2=-1$.

The formula for a time-frequency distribution is derived by first obtaining its weighing function. The weighting function [44, 45] is derived by taking the 1-D Fourier transform of the kernel, $\Phi(\theta,\tau)$. This weighting function, $W(t-u,\tau)$, determines how the cross terms of a time-frequency power spectrum are scaled down and thus reducing their effects in relation to the auto terms.

Eq. (3.2.1) can be rewritten in the form of the weighting function $W(t-u,\tau)$ as given by

$$P(t,\omega)=\frac{1}{4\pi^2}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\underbrace{\left[e^{-j\theta(t-u)}\cdot\Phi(\theta,\tau)\right]}_{W(t-u,\tau)}\cdot e^{-j\tau\omega}\cdot R_{t,1}(t,\tau)\ dud\tau d\theta \qquad (3.2.3)$$

In the case of the time-frequency power spectrum, the local auto-correlation function is defined [8, 19] as $R_{t,1}(t,\tau) = x\left(u+\frac{\tau}{2}\right)\cdot x^*\left(u-\frac{\tau}{2}\right)$. It should be noted that the auto terms are located over the small-valued region of the lag parameter τ [40] and the cross terms in the high-valued region as the auto-correlation function is a measure of the similarity of the signal with itself as a function of the lag parameter τ [41]. Higher-order time-frequency spectra have been studied in [20, 21, 42, 43] by defining a new form of the local auto-correlation function. However, this chapter is devoted to the time-frequency power spectrum and a hyperbolic kernel which is studied next.

3.3 The Hyperbolic *sech*($\beta\theta\tau$) Kernel and Its Family

A number of kernels have been reviewed in the previous chapter. In this section, we propose and analyse a new kernel family. The main motivation for finding the hyperbolic kernel (first-power exponential functions) is that previous kernels are high-power exponential functions such as the CW kernel (second-power exponential function) and the MTE kernel (second- or higher-power exponential functions). First-power exponential functions are mathematically much simpler, easier to interpret and more effective than second- or higher-power exponential functions. Consider the following function given by Eq. (3.3.1)

$$\Phi(\theta,\tau) = cosh\,(\beta\theta\tau) = \frac{exp\,(\beta\theta\tau) + exp\,(-\beta\theta\tau)}{2} \tag{3.3.1}$$

The function can be rewritten for the general case as

$$\Phi(\theta,\tau_1,\tau_2,\ldots,\tau_k) = cosh\big(\beta\theta(\tau_1+\tau_2+\ldots+\tau_k)\big) \tag{3.3.2}$$

The function given in Eq. (3.3.1) satisfies the seven constraints given in Section 3.1, however, it is not bounded which results in an infinite volume under the surface of its Fourier transform function or the weighting function. It should be noted that the boundedness of the weighting function of a kernel function is crucial for its cross-term suppression. If the weighting function is not bounded, i.e. the cross terms will be very large since they are multiplied with a large weighting-function factor, then the corresponding kernel, even though satisfies the seven constraints, is not effective in suppressing cross terms in the time-frequency plane of the spectrum. Thus, the additional constraint on boundedness on a kernel function should be included with thos constraints already stated in Section 3.1.

From Eqs. (3.3.1) and (3.3.2), a new kernel function can be derived by taking the reciprocal of the hyperbolic function

$$\Phi(\theta,\tau_1,\tau_2,...,\tau_k)=\frac{1}{cosh\ [\beta\theta(\tau_1+\tau_2+...+\tau_k)]}=sech\ [\beta\theta(\tau_1+\tau_2+...+\tau_k)] \tag{3.3.3}$$

where β is a parameter to control the exponential terms of the hyperbolic function.

The use of the control parameter β is important. As β tends to infinity in Eq. (3.3.3), the kernel will approach zero. If $\beta = 0$, the hyperbolic distribution will become the WV distribution. Thus, the chosen values of β should not be too large or too small. Depending on a particular application, β should be accordingly chosen so that satisfactory performance in terms of cross-term suppression, auto-term resolution and noise robustness can be achieved. It is also important to note that the hyperbolic $sech(\cdot)$ kernel, given in Eq. (3.3.3), is not the MTE kernel given in Chapter 2, Eq. (2.2.10), even though the CW kernel is a special case of the latter kernel. This makes the hyperbolic kernel unique and thus it hopefully might provide some improvements to the CW and MTE kernels as investigations on the hyperbolic kernel unfold till the end of the thesis.

As the new first-order hyperbolic kernel is proposed, it is necessary to justify that it satisfies the seven constraints listed in the Introduction. First, it is clear that the kernel satisfies the first two constraints since it is independent of the time t and the frequency ω. For the third and fourth constraints, we always have

$$sech(\beta\theta\tau)\big|_{\tau=0} = sech(\beta\theta\tau)\big|_{\theta=0} = 1 \tag{3.3.4}$$

since the hyperbolic kernel is an even function of θ and τ. From Eq. (3.3.4), constraints number 3 and 4 are satisfied by the hyperbolic first-order kernel.

Constraint number 5 is satisfied by the new kernel since the $sech(\cdot)$ function is a real function and thus its real and "imaginary" parts are identical. Constraint number 6 for the kernel is examined as follows

$$\frac{d}{d\tau}\left[sech(\beta\theta\tau)\right]\bigg|_{\tau=0} = -2\cdot\frac{\beta\theta\cdot exp(\beta\theta\tau)-\beta\theta\cdot exp(-\beta\theta\tau)}{\left[exp(\beta\theta\tau)+exp(-\beta\theta\tau)\right]^2}\Bigg|_{\tau=0} = 0\ ,\ \forall\theta \tag{3.3.5}$$

From Eq. (3.3.5), it is clear that the first-order hyperbolic kernel satisfies constraint number 6. Similarly, since the kernel is an even function of θ and τ, constraint number 7 is also satisfied. Thus, the first-order hyperbolic kernel satisfies the seven admissibility constraints and it can be used as a valid time-frequency kernel for further time-frequency signal processing analysis.

The time-frequency power spectrum using the hyperbolic $sech(\beta\tau\theta)$ kernel can be derived by substituting $\Phi(\theta,\ \tau)\ =\ sech(\beta\tau\theta)$ into Eq. (3.2.3) with $R_{t,1}(t,\tau) = x\left(u+\frac{\tau}{2}\right)\cdot x^*\left(u-\frac{\tau}{2}\right)$ as follows

$$P(t,\omega)=\frac{1}{4\pi^2}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-j\tau\omega}\cdot\left(sech(\beta\theta\tau)\cdot e^{-j\theta(t-u)}\right)\cdot x^*\left(u-\frac{\tau}{2}\right)\cdot x\left(u+\frac{\tau}{2}\right)du\,d\tau\,d\theta \tag{3.3.6}$$

Hence, the general time-frequency power spectrum of the hyperbolic kernel is obtained as

$$P(t,\omega)=\int_{\tau} e^{-j\tau\omega}\int_{u}\frac{1}{2\beta\tau}\cdot sech\left(\frac{\pi(t-u)}{2\beta\tau}\right)\cdot x^*\left(u-\frac{\tau}{2}\right)\cdot x\left(u+\frac{\tau}{2}\right)du\,d\tau \tag{3.3.7}$$

The general expression of the higher-order hyperbolic family kernel is given in Eq. (3.3.8)

$$\Phi(\theta,\tau)=\left[sech(\beta\theta\tau)\right]^n,\ n>1 \text{ and } n \text{ is a positive integer.} \tag{3.3.8}$$

For the hyperbolic kernel family, it is important to show that it satisfies the admissibility constraints as does the first-order hyperbolic kernel. First, it is clear that the family kernel satisfies constraints number 1 and 2 since the hyperbolic family kernel is independent of the time t and frequency ω. From Eq. (3.3.4), it is clear that constraints number 3 and 4 are satisfied by the family kernel. Constraint number 5 is similarly satisfied by the family kernel since hyperbolic functions are real and thus their real and "imaginary" parts are identical. Constraint number 6 can be examined based on Eq. (3.3.5),

$$\frac{d}{d\tau}\left[sech(\beta\theta\tau)\right]^{n}\Big|_{\tau=0} = -2n\cdot\frac{\beta\theta\cdot exp(\beta\theta\tau)-\beta\theta\cdot exp(-\beta\theta\tau)}{\left[exp(\beta\theta\tau)+exp(-\beta\theta\tau)\right]^{2}}\cdot\left[sech(\beta\theta\tau)\right]^{n-1}\Big|_{\tau=0} = 0$$

for all values of θ.

(3.3.9)

From Eq. (3.3.9), constraint number 6 and likewise constraint number 7 are satisfied by the hyperbolic kernel family. Thus, all higher-order hyperbolic kernels are also valid time-frequency kernels and they can be employed for time-frequency signal processing analysis. This shows the generality and diversification of the hyperbolic family kernel. This diversification of these kernels has the ability to generate different wavelets as will be shown in Chapter 4.

There are two separate cases of even and odd values of the parameter n. For n = 2, 4, 6, ..., the weighting factors provided by these kernels are not effective as compared with the case of n = 1. The main reason for this is that their weighting functions have infinite volumes under the surface in the $((t-u), \tau)$ plane as will be shown later. Thus, they are not as cross-term effective as the first-order hyperbolic and CW kernels. For n = 3, 5, 7, ..., initial investigations show that the weighting functions of these kernels have smaller side lobes which can suppress cross terms more effectively than the first-order hyperbolic $sech(\beta\theta\tau)$ kernel. Thus, the odd-set of hyperbolic kernels might find useful applications in signal analysis which requires more research work in this direction. However, this thesis is devoted to the first-order hyperbolic kernel and thus the discussion on higher-order hyperbolic kernels is stopped here. Further studies on these kernels can be found in [127]. The weighting functions of the hyperbolic and CW kernels are compared in Section 3.4.

3.4 Comparison of the Hyperbolic and Choi-Williams Weighting Functions

Mathematically, the hyperbolic kernel is easier to integrate than the CW kernel. The weighting functions (the 1-D Fourier transform of the kernel with the "frequency" variable ($t - u$)) of the two kernels are given by Eqs. (3.4.1) and (3.4.2) respectively

$$W_{hyperbolic} = \frac{\pi}{\beta\tau} \cdot sech\left[\frac{\pi(t-u)}{2\beta\tau}\right] = \frac{\pi}{\beta\tau} \cdot \frac{2}{exp\left[\frac{\pi(t-u)}{2\beta\tau}\right] + exp\left[-\frac{\pi(t-u)}{2\beta\tau}\right]} \tag{3.4.1}$$

$$W_{CW} = \frac{\sigma\sqrt{\pi}}{\tau} \cdot exp\left[-\frac{\sigma(t-u)^2}{4\tau^2}\right] \tag{3.4.2}$$

where β and σ are the kernel parameters of the hyperbolic and CW kernels respectively, τ is the lag parameter used to calculate the auto-correlation function.

The 2-D contour plots of the weighting functions of the CW and hyperbolic kernels are shown in Figure 3.4.1. This shows that the hyperbolic kernel is more "local" in the ($t - u$)-direction than in the τ-direction. The CW kernel extends wider in the ($t - u$) direction and therefore it can be said not to be "local" in that direction [28]. Thus, the hyperbolic kernel is more concentrated in the vicinity of the origin in the ($t - u$) direction than the CW kernel. The peak of the hyperbolic weighting function is also $\sqrt{\pi}$ times larger than that of the CW kernel at the origin which might suggest that the hyperbolic auto terms are $\sqrt{\pi}$ times larger than the CW auto terms. Detailed simulations in Sections 3.5.2 and 3.5.3 will justify this fact.

The 3-D plots of the weighting functions of the CW and hyperbolic kernels, which correspond to the contour plots displayed in Figure 3.4.1, are shown in Figure 3.4.2. The hyperbolic weighting function has larger non-zero values in the vicinity of the origin as seen in Figure 3.4.2. If the auto terms are mostly distributed along the horizontal straight line ($t - u$) = 0 in Figure 3.4.1, then the hyperbolic kernel is most suitable in amplifying these auto terms since it is localised in the ($t - u$) axis.

Since the CW weighting function is localised in the direction of the τ axis, it is most suitable for auto terms that are located along the $(t-u)$ axis. Overall, the CW kernel is more localised around the origin than the hyperbolic kernel, i.e. the lobes of the CW contours around the origin are wider than those of the hyperbolic weighting function (as seen in Figure 3.4.1). This suggests that the CW kernel can support the auto terms more effectively than the hyperbolic kernel, in other words, the CW kernel is expected to have a finer auto-term resolution than that of the hyperbolic kernel. From Figure 3.4.1, it can be seen that the hyperbolic kernel has large main lobes that extend in the direction of the τ axis, i.e. its weighting function has a fast decaying rate, which suggests that it is more effective than the CW kernel in terms of cross-term suppression. The cross-term suppression ability of these kernels will be discussed in Sections 3.5.2 and 3.5.3 using a sum of two complex exponential and chirp signals respectively. The auto-term resolution of the CW, hyperbolic and some MTE kernels will be investigated in Section 3.6 so that a trade-off between cross-term suppression and auto-term resolution can be established.

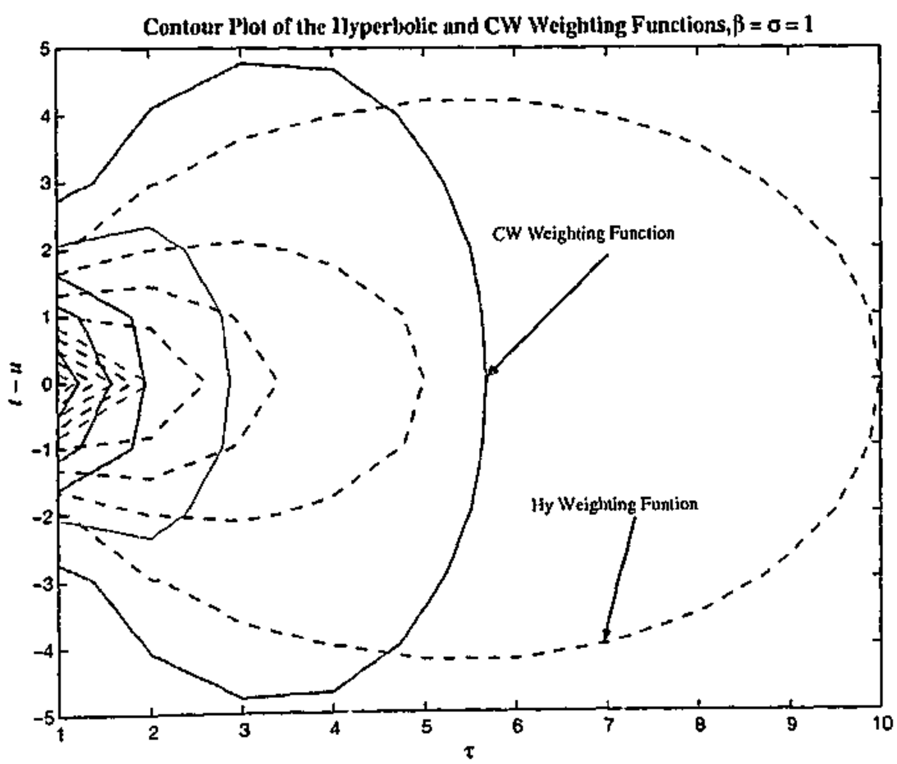

Figure 3.4.1: Contour plots of the CW and hyperbolic weighting functions

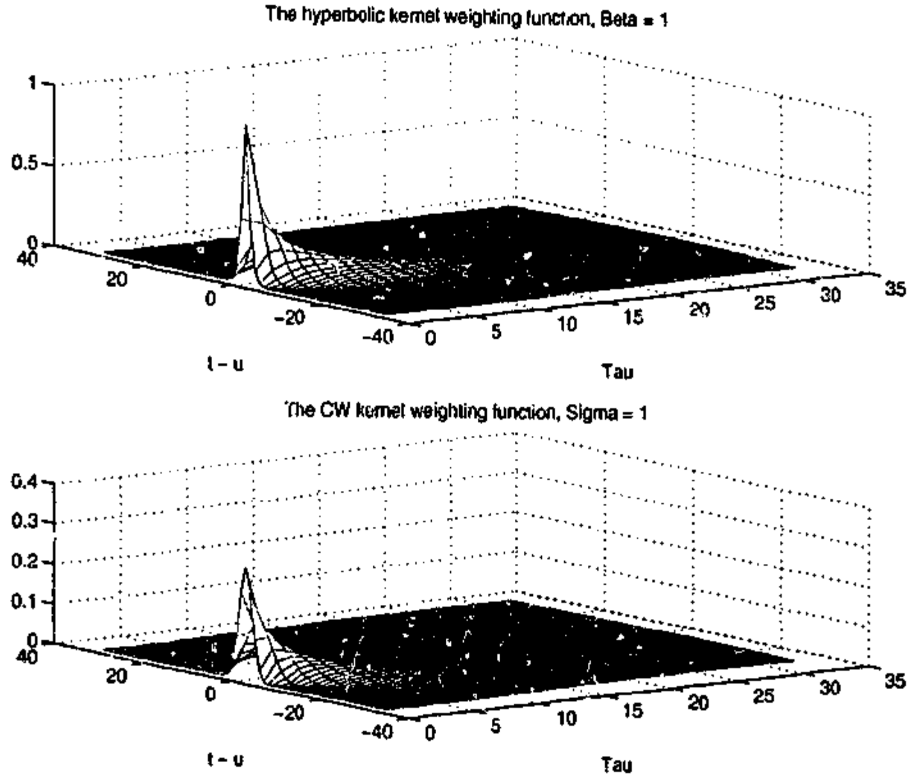

Figure 3.4.2: 3-D plots of the hyperbolic and CW weighting functions

The 3-D plots in Figure 3.4.2 of the weighting functions of the hyperbolic and CW kernels validate the effects of kernel functions on time-frequency power spectra. They show auto-term supportive regions (around the vicinity of the origin in both the $(t-u)$ and τ axes) of the kernels and therefore it is possible to choose the application to the appropriate kernels with the minimum amount of cross terms and maximum amount of auto terms.

The contour plot of the second-order hyperbolic kernel, $[sech(\beta\theta\tau)]^n$ for $n = 2$, is given in Figure 3.4.3. The main and side lobes of the weighting function of the $[sech(\beta\theta\tau)]^2$ kernel are unbounded at the centre frequency suggesting that its volume under the surface is infinite. Thus, it is not suitable for suppressing cross terms in the time-frequency plane and therefore the even-order of the hyperbolic family kernels will not be investigated further in this thesis. Although the even-order hyperbolic family kernels do not provide effective cross-term suppression, they still satisfy the seven constraints. In the following section, performance of the CW and hyperbolic kernels are compared in terms of cross-term suppression through simulation.

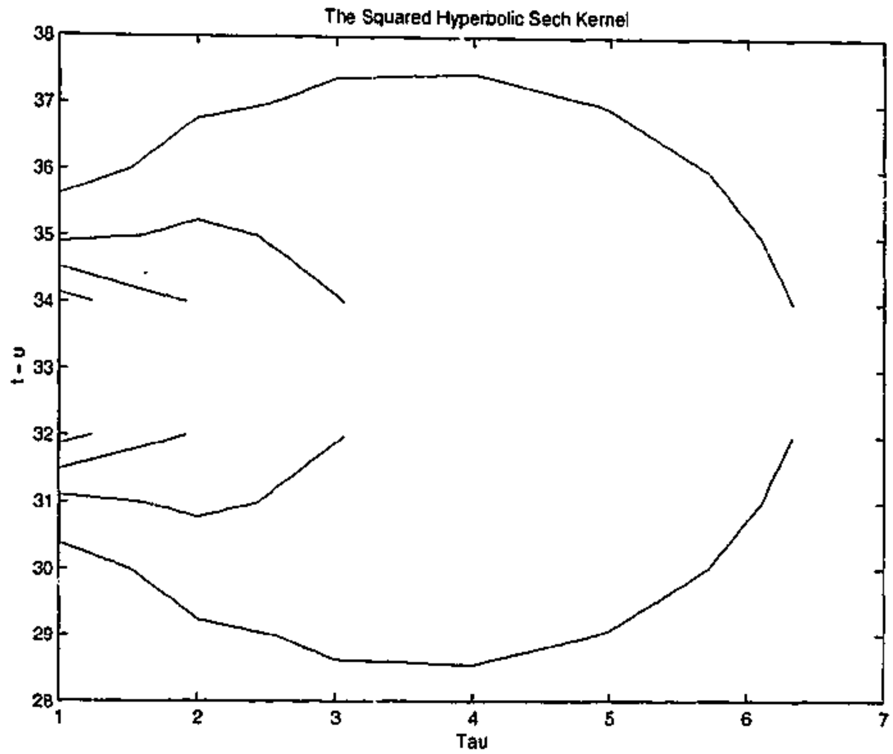

Figure 3.4.3: The weighting function of the second-order hyperbolic kernel $\Phi(\theta, \tau) = [sech(\beta\theta\tau)]^2$ with $\beta = 1$

3.5 Cross-Term Suppression Comparison

The effectiveness of the CW and hyperbolic kernels in suppressing cross terms will be compared with two types of multi-component signals: a sum of two complex exponential and two chirp signals. Performance of the WV kernel is also compared with the CW and hyperbolic kernels. One of the key factors that can be used to judge the performance of a particular kernel is to estimate the normalised peak-magnitude ratio of the cross terms to auto terms. The lower this ratio, the more effective cross-term suppression the kernel is.

Firstly, the WV, CW and hyperbolic time-frequency power spectra are compared so that the disadvantages of the WV unity kernel are shown and the advantages of the CW and hyperbolic time-frequency power spectra are demonstrated. In the rest of this chapter, the CW, MTE and the first-order hyperbolic $sech(\beta\theta\tau)$ kernels are studied and compared (where appropriate) in terms of normalised cross-term magnitude ratio (Section 3.5.2), normalised peak-magnitude ratio of the cross terms to auto terms (Section 3.5.3), auto-term resolution or auto-term width (Section 3.6) and noise robustness (Section 3.7). The normalisation, that has been used throughout this chapter, unless otherwise specified, is done by dividing the particular values by their maximum value. For example, the normalised ratio of the cross terms to auto terms is obtained by dividing all of the ratios by their maximum value.

3.5.1 A Typical Example

In this section, the effects of cross terms and artifacts are demonstrated so that the effectiveness of the CW and hyperbolic kernels can be clearly identified and understood. The MTE time-frequency distribution does not have a general expression and thus it is not included here. Since the well-known CW kernel is a special case of the MTE kernel, it can be chosen as a representative kernel for the MTE kernel.

A simulated speech signal, shown in Figure 3.5.1, is used as an input signal to obtain the WV, CW and hyperbolic time-frequency power spectra. A silent period [5-9] is present since it is unavoidable in normal conversations.

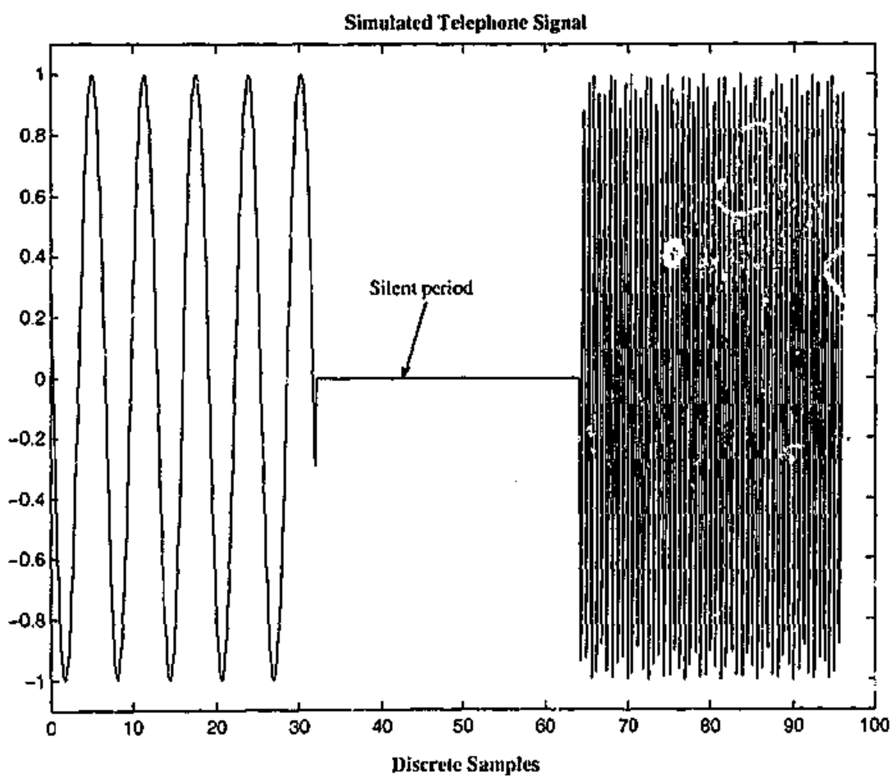

Figure 3.5.1: A simulated speech signal with a silent period

The WV, CW and hyperbolic time-frequency power spectra are displayed in Figure 3.5.2 to Figure 3.5.4 respectively. The most important thing that determines the effectiveness of a kernel is that during the silent period of the speech signal, its time-frequency power spectrum must be "silent" or there is effectively no energy-smearing.

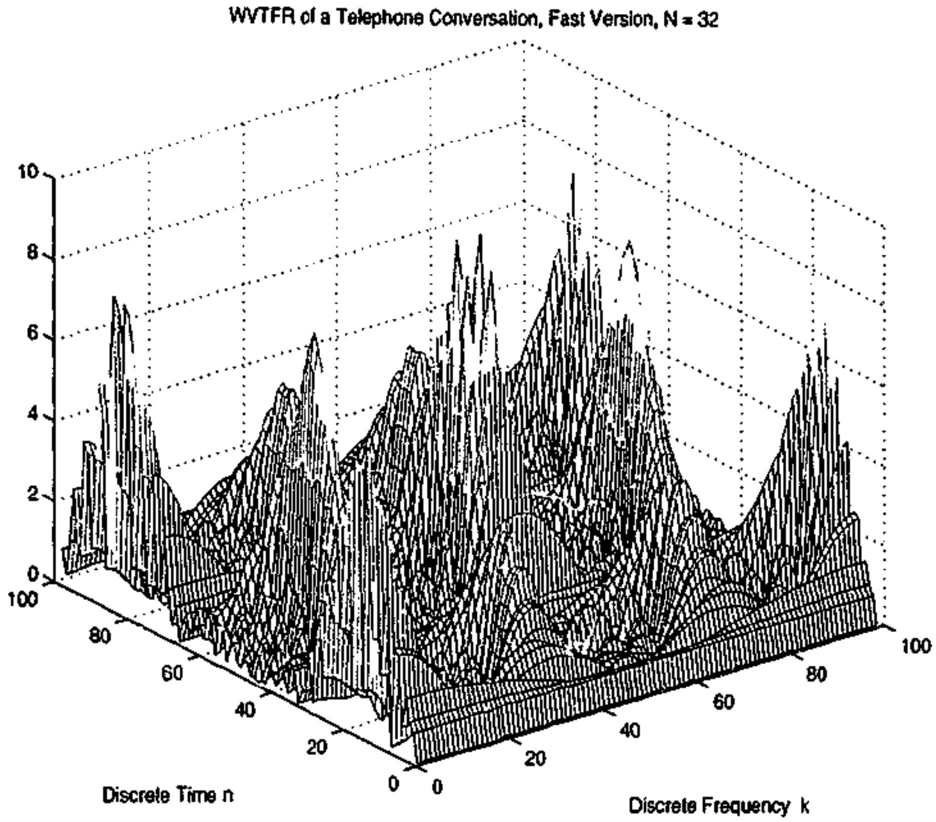

Figure 3.5.2: The WV time-frequency power spectrum of a speech signal displayed in Figure 3.5.1

As can be seen from Figure 3.5.2, the WV time-frequency power spectrum is non-zeroed over the silent period (from discrete times of 32 to 64) of the conversation. There are many humps and considerable energy smearing over the silent period. This creates misleading information about the nature of the input signal and thus it shows that the WV unity kernel is not effective in suppressing cross terms in the time-frequency plane.

Figure 3.5.3 displays the CW time-frequency power spectrum which shows zeroed spectrum over the silent period. This is a major improvement over the WV time-frequency power spectrum. However, there is still energy smearing over the silent period. The "humps" are cleaner and smaller but they should be ideally removed from the spectrum.

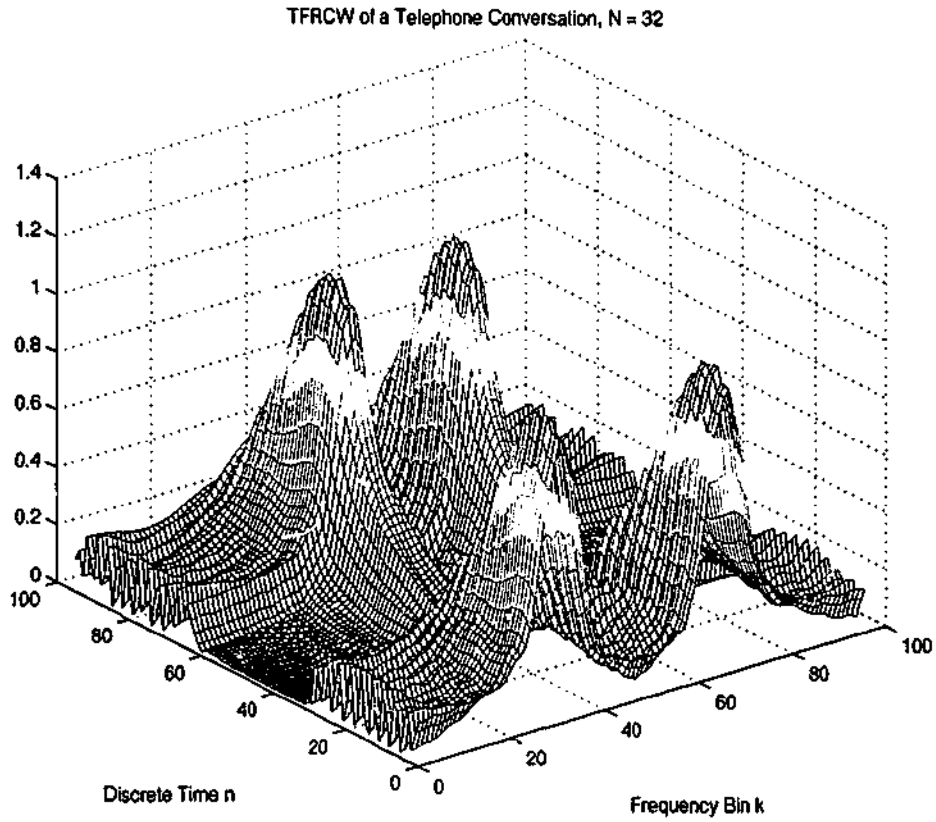

Figure 3.5.3: The CW time-frequency power spectrum with $\sigma = 1$ of a speech signal displayed in Figure 3.5.1. The "Frequency Bin k" axis should read "Discrete Frequency".

Figure 3.5.4 shows the hyperbolic time-frequency power spectrum with a clear display of the silent period. The edges are sharp and the amount of energy smearing is considerably reduced. There are still small "humps" over the silent period but these humps are much smaller and cleaner than those in the WV and CW time-frequency spectra. This suggests that the hyperbolic kernel can perform better cross-term suppression than the CW and WV kernels. The subsequent sections compare cross-term suppression ability of the hyperbolic, CW and WV kernels by using a sum of two complex exponential and two chirp signals. Auto-term resolution and noise robustness of the hyperbolic, CW and MTE kernels are examined and compared in Sections 3.6 and 3.7 respectively.

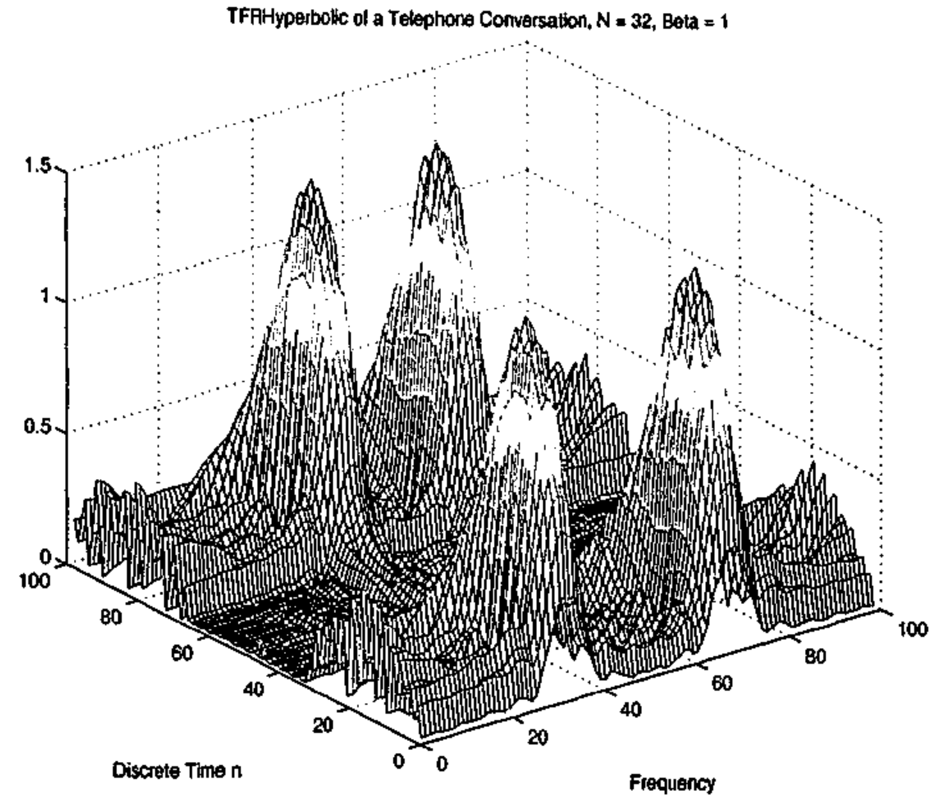

Figure 3.5.4: The hyperbolic time-frequency power spectrum with $\beta = 1$ of a speech signal displayed in Figure 3.5.1. The "Frequency" axis should read "Discrete Frequency".

3.5.2 A Sum of Two Complex-Exponential Signals

Given the input signal

$f(t) = A_1 exp\left[j(\omega_1 t + \theta_1)\right] + A_2 exp\left[j(\omega_2 t + \theta_2)\right]$

where A_1, A_2 are arbitrary real constants and θ_1, θ_2 are the phases of the exponential terms, $\omega_1 = 30$ rad/s† and $\omega_2 = 34$ rad/s. The CW time-frequency power spectrum of $f(t)$ is given by [28, 102]

$$\begin{aligned} ED_{CW}(t,\omega) = {} & 2\pi A_1^2 \delta(\omega - \omega_1) + 2\pi A_2^2 \delta(\omega - \omega_2) \\ & + 2A_1 A_2 \cos\left[(\omega_1 - \omega_2)\cdot t + \theta_1 - \theta_2\right] \cdot WEIGHT_{CW} \end{aligned} \tag{3.5.1}$$

where

† For comparison purposes, values of ω_1 and ω_2 are taken from the paper by Choi and Williams [28].

$$WEIGHT_{CW} = \sqrt{\frac{\pi\sigma}{(\omega_1 - \omega_2)^2}} \cdot exp\left[-\frac{\sigma}{4(\omega_1 - \omega_2)^2} \cdot \left(\omega - \frac{\omega_1 + \omega_2}{2} \right)^2 \right] \tag{3.5.2}$$

The WV time-frequency power spectrum is given by

$$\begin{aligned} WV(t,\omega) &= 2\pi A_1^2 \delta(\omega - \omega_1) + 2\pi A_2^2 \delta(\omega - \omega_2) \\ &+ 2A_1 A_2 \cos\left[(\omega_1 - \omega_2)\cdot t + \theta_1 - \theta_2\right] \cdot WEIGHT_{WV} \end{aligned} \tag{3.5.3}$$

where $WEIGHT_{WV} = 1$.

The auto terms and cross terms of the hyperbolic time-frequency power spectrum are identical to those of the CW time-frequency power spectrum (as seen in Eq. (3.5.1)). However, the hyperbolic weighting factor is different from that of the CW kernel and is given by

$$WEIGHT_{hyper} = \frac{\pi}{\beta(\omega_1 - \omega_2)} \cdot sech\left[\frac{\pi}{2\beta(\omega_1 - \omega_2)} \cdot \left(\omega - \frac{\omega_1 + \omega_2}{2} \right) \right] \tag{3.5.4}$$

From Eqs. (3.5.1)–(3.5.4), it is clear that the WV kernel does not effectively suppress cross terms, i.e. its weighting factor is unity as seen in Eq. (3.5.3). The weighting factors of both the CW and hyperbolic kernels are much less than unity and thus they are more effective in suppressing cross terms than the WV kernel. Figure 3.5.5 shows the 3-D plot of the normalised ratio of the hyperbolic weighting factor (Eq. (3.5.4)) to that of the CW kernel (Eq. (3.5.2)) as a function of ω and β. This ratio is very small except for small values of β. This means that for approximately $\beta \leq 1.5$ and for frequencies less than 5 rad/s, the hyperbolic weighting function is much larger than that of the CW kernel and thus the former is not effective in suppressing cross terms. However, for $\beta > 1.5$, the hyperbolic weighting factor appears to be smaller than that of the CW kernel and therefore the former kernel is more effective at cross-term suppression than the latter.

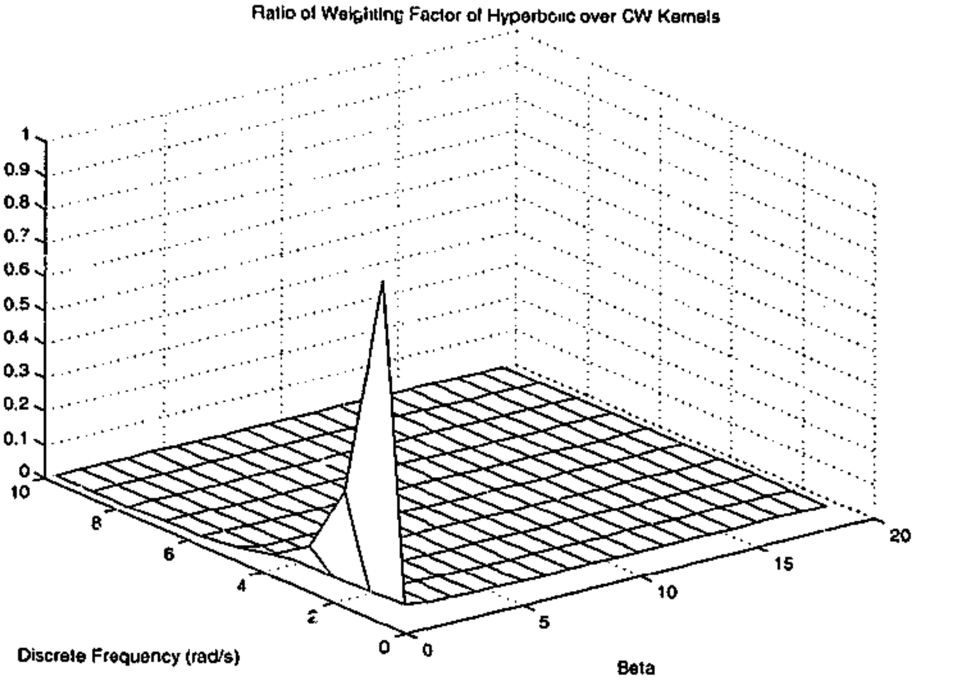

Figure 3.5.5: Normalised ratio of the hyperbolic weighting factor (Eq. (3.5.4)) to that of the CW kernel (Eq. (3.5.2)) for a sum of two complex-exponential signals

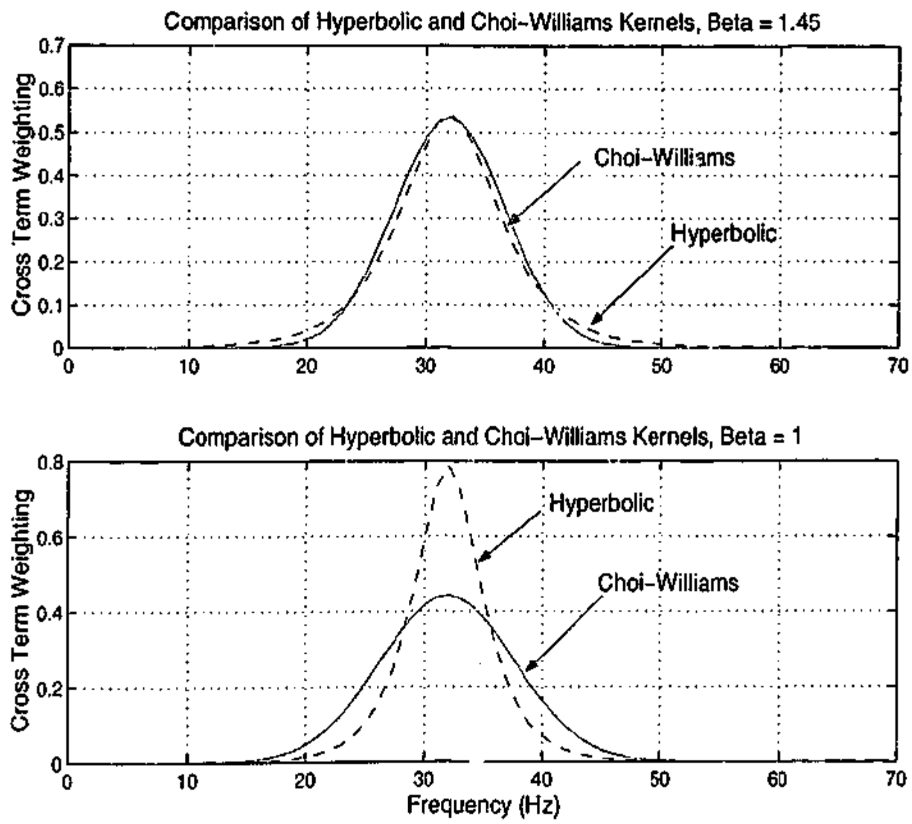

Figure 3.5.6: Comparison of the CW and hyperbolic kernels using a sum of two complex-exponential signals

Figure 3.5.6 displays the cross-term suppression ability of the CW and hyperbolic kernels for $\beta = 1$ and $\beta = 1.45$. As explained in Section 3.3 and from Figure 3.5.5, when β increases, better performance in terms of cross-term suppression is obtained since the main-lobe magnitude of the hyperbolic weighting function is reduced. For small values of $\beta \leq 1.5$ ($\sigma \geq 0.67$), as explained earlier, the CW distribution gives better results since its main lobe is smaller in magnitude than that of the hyperbolic distribution. However, it should be noted that β should not be chosen too small or too large or accordingly σ should not be chosen too large or too small (as explained in Section 3.3) since extreme values of β or σ can make the kernel become the WV kernel which does not have effective cross-term suppression.

The hyperbolic, CW and WV auto-term magnitude remains constant at $2\pi A_1^2$ and $2\pi A_2^2$ as shown in Eqs. (3.5.1) and (3.5.3). Since the auto terms remain constant, their ratios to the corresponding cross terms are not shown. Instead, the normalised cross terms of the CW and hyperbolic kernels (given also in Eq. (3.5.1)) for a sum of two complex-exponential signals is shown in Figure 3.5.7.

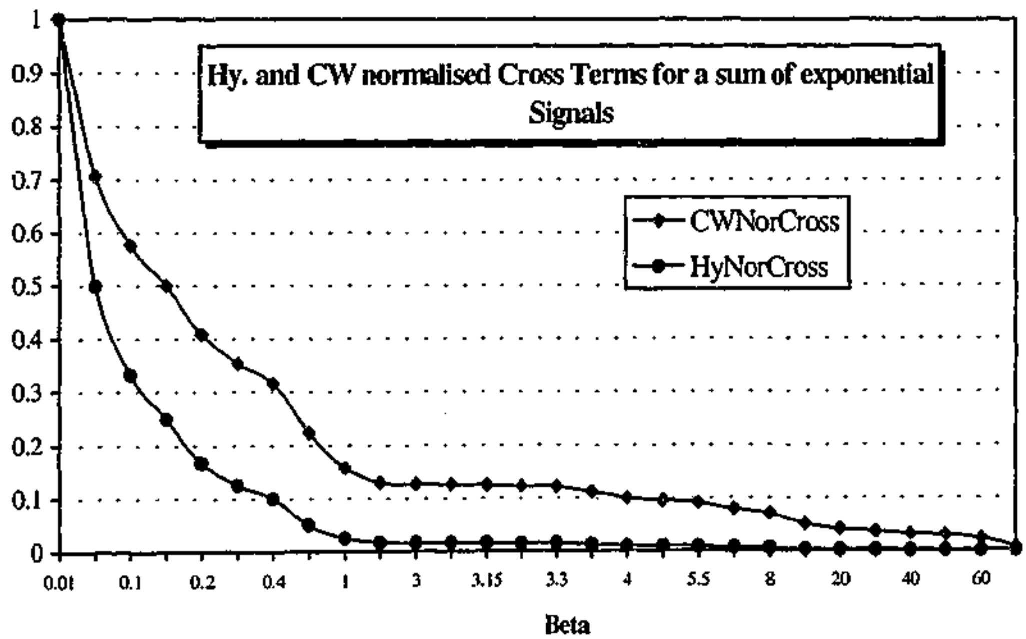

Figure 3.5.7: Normalised CW and hyperbolic cross terms for a sum of two complex-exponential signals. The lower this value is, the better the cross-term suppression ability of the kernel. In this case, the hyperbolic kernel is better than the CW kernel for $\beta \geq 1.45$.

From Figure 3.5.6, for β near 1.45 ($\sigma \approx 0.7$), the hyperbolic cross terms have identical peaks with those of the CW. When $\beta \geq 1.45$, the hyperbolic kernel starts outperforming the CW kernel by having a smaller cross-term peak magnitude. The normalised cross-term magnitude ratio, which is shown in Figure 3.5.7, decreases as β increases. From Figure 3.5.7, the faster decaying rate of the hyperbolic normalised cross terms compared with that of the CW cross terms suggests that the hyperbolic kernel is more effective in suppressing cross terms than the CW kernel as predicted in Section 3.4. In fact, from Figure 3.5.7, the CW normalised cross terms are always larger than those of the hyperbolic kernel for values of β not large, i.e. typically, $\beta \leq 50$.

If β is large, the hyperbolic kernel approaches a "zeroed" kernel (which is not very useful) and the CW kernel becomes the WV kernel which does not provide effective cross-term suppression. In addition, under this extreme condition of β, the normalised cross terms of the two kernels will be getting closer in value and it is expected that they will be identical for very large values of β. Thus, the value of β and σ should be carefully chosen with the specific application to avoid the above limitations of the hyperbolic and CW kernels. For a sum of two complex exponential signals, from Figure 3.5.5 to Figure 3.5.7, the useful range of β for effective cross-term suppression is $\beta \geq 1.45$ with β is not to be chosen very large. Another frequently encountered non-stationary signal in practice is the chirp signal. A sum of two chirp signals is examined in the following section.

3.5.3 A Sum of Two Chirp Signals

Let the input signal, $f(t)$, be a sum of two chirp signals of the form

$f(t) = A_1 exp\left(\frac{j\alpha_1 t^2}{2}\right) + A_2 exp\left(\frac{j\alpha_2 t^2}{2}\right)$, where $\alpha_1 = 1$, $\alpha_2 = 3$ and $A_1 = A_2 = 1$ (for simplicity).

For a sum of two chirp signals, the integration cannot be analytically calculated, thus, approximation methods (by means of simulation) using discrete techniques are used to estimate the integrals. The general form of the time-frequency power spectrum can be written as $P(t, \omega) = AUTO + CROSS$. The WV time-frequency power spectrum is given by

$$AUTO_{WV} = A_1^2\delta(\omega - \alpha_1 t) + A_2^2\delta(\omega - \alpha_2 t) \tag{3.5.5}$$

and

$$CROSS_{WV} = 2A_1A_2\int_\tau exp\left\{-j\left(\omega - \tfrac{\alpha_1+\alpha_2}{2}t\right)\tau\right\}\cdot cos\left[\tfrac{1}{2}(\alpha_1 - \alpha_2)(\tfrac{\tau^2}{4} + t^2)\right] d\tau \tag{3.5.6}$$

The time-frequency power spectrum given by the CW distribution is [28, 102]

$$\begin{aligned} AUTO_{CW} &= A_1^2\int_\tau \left\{exp\left(-\frac{\alpha_1^2\tau^4}{\sigma}\right)\right\}\cdot\left(e^{-j(\omega-\alpha_1 t)\tau}\right) d\tau \\ &+ A_2^2\int_\tau \left\{exp\left(-\frac{\alpha_2^2\tau^4}{\sigma}\right)\right\}\cdot\left(e^{-j(\omega-\alpha_2 t)\tau}\right) d\tau \end{aligned} \tag{3.5.7}$$

$$\begin{aligned} &CROSS_{CW} = A_1A_2\int_\tau exp\left[-j\left(\omega - \tfrac{\alpha_1+\alpha_2}{2}t\right)\tau\right]\cdot \\ &\int_u e^{j(\alpha_1+\alpha_2)u\tau/2}\cdot\left\{\frac{\sigma}{2\tau\sqrt{\pi}}\cdot exp\left(-\frac{\sigma u^2}{4\tau^2}\right)\right\} \\ &\cdot cos\left[(\alpha_1 - \alpha_2)(u+t)^2/2 + (\alpha_1 - \alpha_2)\tau^2/8\right] du d\tau \end{aligned} \tag{3.5.8}$$

The auto terms and cross terms of the hyperbolic time-frequency power spectrum of a sum of two chirp signals are given by

$$AUTO_{Hy} = A_1^2 \int_\tau \left\{ sech\left(\beta\alpha_1\tau^2\right)\right\} \cdot \left(e^{-j(\omega-\alpha_1 t)\tau}\right) d\tau + A_2^2 \int_\tau \left\{ sech\left(\beta\alpha_2\tau^2\right)\right\} \cdot \left(e^{-j(\omega-\alpha_2 t)\tau}\right) d\tau \tag{3.5.9}$$

$$CROSS_{Hy} = A_1 A_2 \int_\tau exp\left[-j\left(\omega - \tfrac{\alpha_1+\alpha_2}{2} t\right)\tau\right] \cdot \int_u e^{j(\alpha_1+\alpha_2)u\tau/2} \cdot \left\{\frac{1}{2\beta\tau} \cdot sech\left(\frac{\pi u}{2\beta\tau}\right)\right\} \cdot cos\left[(\alpha_1-\alpha_2)(u+t)^2/2 + (\alpha_1-\alpha_2)\tau^2/8\right] du\, d\tau \tag{3.5.10}$$

From Eqs. (3.5.6), (3.5.8) and (3.5.10), the cross-term weighting factors of the WV, CW and hyperbolic distributions are

$$CROSS\text{-}WEIGHT_{WV} = cos\left[\frac{1}{2}(\alpha_1-\alpha_2)\left(\tfrac{\tau^2}{4}+t^2\right)\right] \tag{3.5.11}$$

$$CROSS\text{-}WEIGHT_{CW} = \left\{\frac{\sigma}{2\tau\sqrt{\pi}} \cdot exp\left(-\frac{\sigma u^2}{4\tau^2}\right)\right\} \cdot cos\left[(\alpha_1-\alpha_2)(u+t)^2/2 + (\alpha_1-\alpha_2)\tau^2/8\right] \tag{3.5.12}$$

$$CROSS\text{-}WEIGHT_{Hy} = \left\{\frac{1}{2\beta\tau} \cdot sech\left(\frac{\pi u}{2\beta\tau}\right)\right\} \cdot cos\left[(\alpha_1-\alpha_2)(u+t)^2/2 + (\alpha_1-\alpha_2)\tau^2/8\right] \tag{3.5.13}$$

From Eqs. (3.5.11)–(3.5.13), it is evident that the WV kernel has a larger weighting factor than those of the hyperbolic and CW kernels. Thus, the WV unity kernel is not effective in cross-term suppression. The normalised ratio of the hyperbolic weighting factor (Eq. (3.5.13)) to the CW weighting factor (Eq. (3.5.12)) is given in Figure 3.5.8.

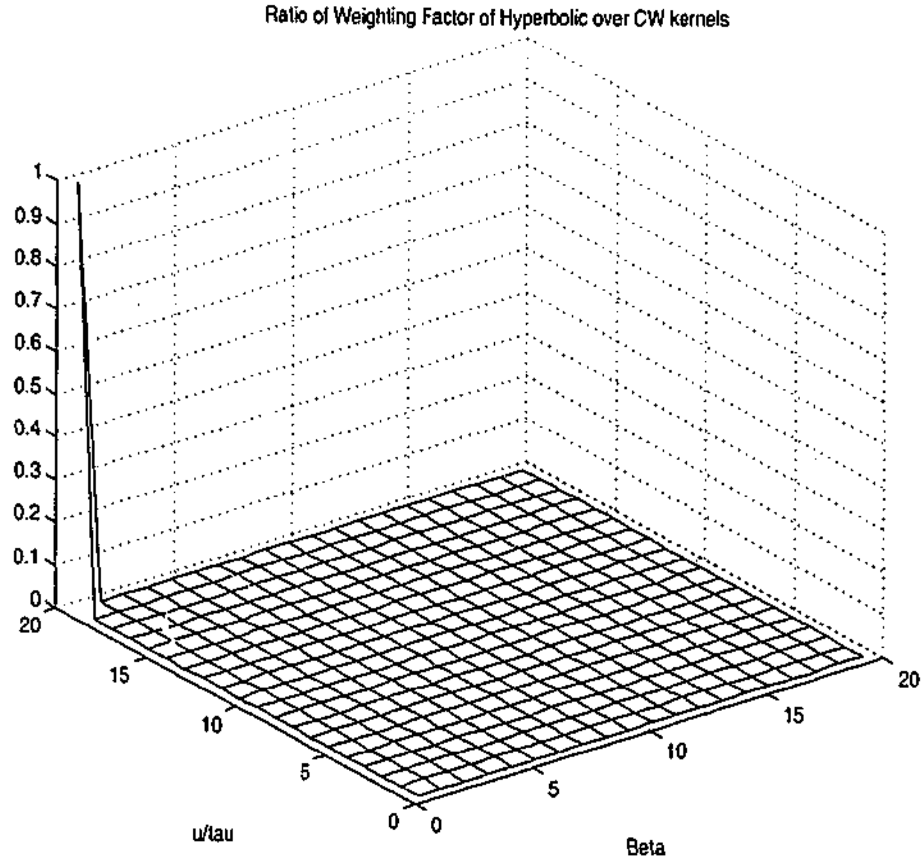

Figure 3.5.8: Normalised ratio of the hyperbolic weighting factor to that of the CW kernel for a sum of two chirp signals as a function of $\frac{u}{\tau}$ and β. The approximate useful range of β is $\beta \geq 0.5$.

Figure 3.5.8 displays the 3-D plot of the normalised ratio of the hyperbolic weighting factor (Eq. (3.5.13)) to that of the CW kernel (Eq. (3.5.12)) as a function of $\frac{u}{\tau}$ and β in which the ratio is small except for large values of $\frac{u}{\tau}$ and small values of β. This is similar to the case of a sum of two complex-exponential signals investigated earlier. As stated in Section 3.5.2, as β increases, better performance of the hyperbolic kernel compared with the CW kernel will be obtained. Increasing β will reduce the volume under the surface of the weighting functions of the hyperbolic and CW kernels. The faster the reduction rate of this volume with respect to β, the larger the peak-magnitude ratio of the auto terms over the cross terms.

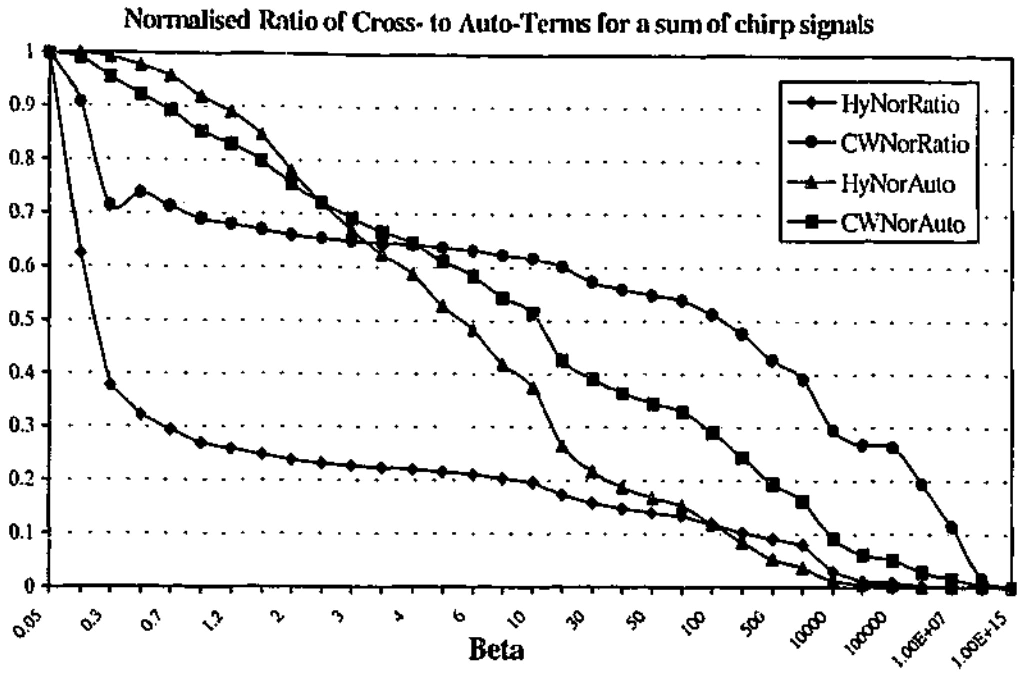

Figure 3.5.9: Normalised peak-magnitude ratio of the cross terms to auto terms of the CW kernel (ratio of Eq. (3.5.8) to Eq. (3.5.7), *CWNorRatio*) and hyperbolic kernel (ratio of Eq. (3.5.10) to Eq. (3.5.9), *HyNorRatio*) for a sum of two chirped signals. The lower this ratio is, the better the cross-term suppression ability of the kernel. The normalised values of the auto terms of the two kernels (*HyNorAuto* and *CWNorAuto*) are also shown in the same graph. The higher this value, the better the auto term magnitude. At $\beta \approx 2.5$, the normalised values of the auto terms of the two kernels are approximately equal. The hyperbolic normalised ratio of the cross terms to auto terms is much lower than that of the CW kernel for all values of β except at $\beta = 0.05$. The useful range of β is therefore $0.3 \leq \beta \leq 2.5$.

Figure 3.5.9 displays the peak-magnitude ratio of the cross terms to auto terms and the normalised auto terms of the two kernels as β varies. This ratio is more important than the individual magnitude of the cross terms and auto terms since it reflects the effectiveness of the kernel in supporting auto terms and suppressing cross terms. If the cross terms are small in magnitude, say 0.1, and the auto terms under the same conditions are much smaller than the cross terms, say 0.000001, then the kernel is not effective even though the cross terms are small. This explains why the ratio of cross terms to auto terms of a kernel is considered to be the most important factor and therefore it is used as a benchmark to compare the effectiveness of different kernels. It is clear that the smaller this ratio, the more effective the kernel. From Figure 3.5.9, theoretically, for $\beta \geq 0.05$, i.e. $\sigma \leq 20$, the hyperbolic kernel will perform better than the CW kernel by having a small cross-term to auto-term magnitude ratio. The worst performance occurs when the hyperbolic auto terms have lower magnitude than those of the CW kernel which corresponds to $\beta \geq 100$, i.e. $\sigma \leq 0.01$.

It should be noted that the decaying rate of the hyperbolic cross terms is faster than that of the CW cross terms as discussed earlier in Section 3.4 which yields better cross-term suppression as can be seen in Figure 3.5.9. This effect has also been observed by Boudreaux-Bartels and Papandreou [25]. From Figure 3.5.9, the useful range of β is approximately $0.3 \leq \beta \leq 2.5$ to ensure that the hyperbolic kernel is more effective than the CW kernel by having better auto-term magnitude and cross-term suppression ability. Using the observed range of β from Figure 3.5.8 of $\beta \geq 0.5$, the range of β now becomes $0.5 \leq \beta \leq 2.5$. It should be noted that the lower limits of β obtained from Figure 3.5.8 ($\beta \geq 0.5$) and from Figure 3.5.9 ($\beta \geq 0.3$) are in the same order of magnitude which suggests that both methods of calculating the ratio of the kernels weighting factors or magnitude ratio of auto terms and cross terms are valid.

From Section 3.5.2, the useful range of β for a sum of two complex exponential signals is $\beta \geq 1.45$. Thus, to enable the hyperbolic kernel to perform better than the CW kernel, practically, β should be in the range of $1.45 \leq \beta \leq 2.5$. From Figure 3.5.9, it should be noted that for $20 \geq \beta \geq 2.5$, the hyperbolic kernel still performs well, but with a slightly smaller auto-term magnitude compared to that of the CW kernel. If only the cross-term suppression ability is considered, then the larger β is, the better the cross-term suppression. However, if β is very large (about 10^7), detailed simulation shows that the auto-term peak magnitude becomes saturated at about 0.001 for a sum of two chirp signals.

The auto terms of the two kernels are plotted in Figure 3.5.10 for $t = 0$, $\sigma = 1$ and Figure 3.5.11 shows the cross terms of the CW and hyperbolic time-frequency power spectra for $\beta = 3.5$ and $t = 0$ to give further understanding on the effectiveness of the hyperbolic and CW kernels.

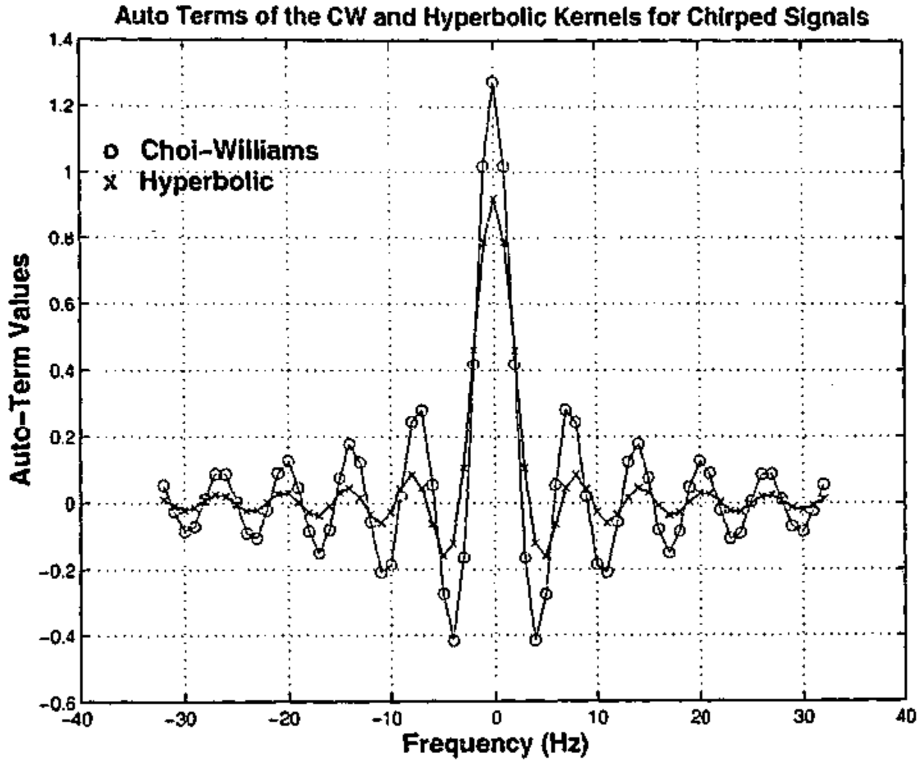

Figure 3.5.10: Auto-term magnitude of the CW and hyperbolic time-frequency power spectra for $t = 0$ and $\beta = 1$

From Figure 3.5.10, it can be seen that the hyperbolic auto-term peak magnitude is less than that of the CW kernel for $\beta = 1$. From Figure 3.5.11, the hyperbolic cross-term peak magnitude is equal to that of the CW kernel for $\beta = 3.5$. From Figure 3.5.9, it was shown that for $20 \geq \beta \geq 2.5$, the hyperbolic normalised auto-term peak magnitude is less than that of the CW kernel and for $\beta \geq 0.3$, the magnitude ratio of cross terms to auto terms of the hyperbolic kernel is less than that of the CW kernel. As explained earlier, the above ratio truly reflects the effectiveness of the kernel rather than the cross-term peak magnitude. Therefore, the useful range of β is not going to be chosen as $\beta \geq 3.5$, as it was the method used to obtain the range of β for the case of a sum of two complex exponential signals in Section 3.5.2.

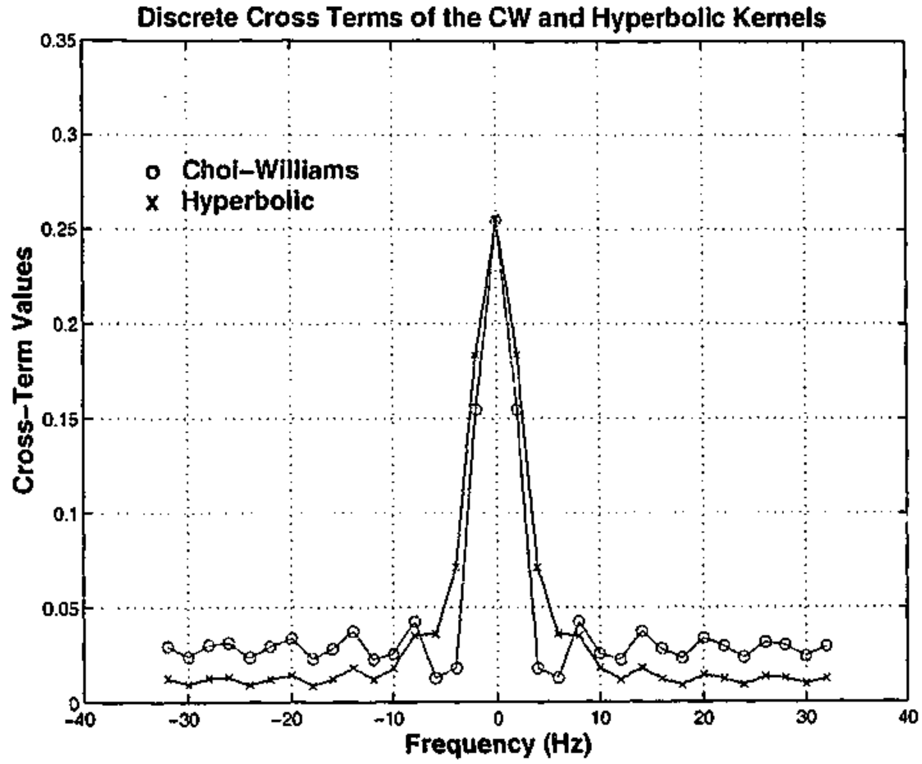

Figure 3.5.11: Cross-term magnitude of the CW and hyperbolic time-frequency power spectra for $t = 0$ and $\beta = 3.5$ which suggest that for $\beta \geq 3.5$, better cross-term suppression can be achieved by using the hyperbolic kernel rather than the CW kernel.

However, from Figure 3.5.9, at $\beta = 3.5$, the hyperbolic normalised auto terms are only slightly less than the CW normalised auto terms (about 5 %) and thus it can be accepted as a useful value of β. Thus, the most useful range of β, which yields optimum performance for the hyperbolic kernel in cross-term suppression and auto-term magnitude compared with the CW kernel, can be expanded to $1.45 \leq \beta \leq 3.5$. The applicable range of β for a satisfactory performance of effective cross-term suppression and acceptable auto-term magnitude is therefore $0.5 \leq \beta \leq 20$. This range of β will be discussed further along with other trade-offs in Chapters 4 and 5 so that the most applicable range of β can be clearly identified.

Although the hyperbolic kernel can suppress cross terms more effectively than the CW kernel for well-chosen values of β, increasing β to a very large value will saturate the auto-term peak magnitude as discussed earlier and as observed by Choi and Williams [28]. Making β too large does not provide useful information since the hyperbolic kernel approaches a "zeroed" kernel as explained in Section 3.4. If β is too large then the peak-magnitude ratio of the cross terms to auto terms decreases as shown in Figure 3.5.7 and Figure 3.5.9. In addition, the normalised auto-term magnitude of the CW and hyperbolic kernels also decreases. Thus, it can be suggested that increasing β (or decreasing σ) enhances cross-term suppression but decreases the auto-term magnitude. A question arises at this point — *Are there any other trade-off(s) associated with increasing β, such as auto-term resolution and noise robustness?* — Sections 3.6 and 3.7 examine the auto-term resolution or auto-term width and noise robustness (as β varies) of the CW, hyperbolic and some of the MTE kernels in some detail so that the relationships and trade-off(s) among the above mentioned quantities can be established.

3.6 Auto-Term Functions and Auto-Term Widths

Sections 3.5.2 and 3.5.3 examined the effectiveness of the hyperbolic and CW kernels by estimating the peak-magnitude ratio of their auto terms to cross terms. The effectiveness of a kernel can also be measured based on its auto-term width or auto-term resolution which can be estimated from its auto-term function. The auto-term function is a function of the lag parameter τ but with the substitution of $\theta = -a\tau$, where a is the slope of the auto-term line in the kernel time-frequency plane.

The auto-term width is defined as the frequency at which the auto-term magnitude decreases by e = 2.718 times its peak magnitude [49]. The larger the auto-term width, the finer the auto-term resolution. Previous work by Stankovic [49] calculated the auto-term functions and auto-term widths of a number of kernels including the Born-Jordan kernel, the pseudo WV kernel, the optimal kernel, the CW kernel and *sinc* kernel [49]. This section is devoted to compare the hyperbolic *sech($\beta\theta\tau$)* kernel with the CW and MTE kernels as the kernel control parameters $\beta = \frac{1}{\sigma}$ (for hyperbolic kernel), σ (for CW kernel) α, r, β_{MTE}, γ and λ (for MTE kernel) vary. The auto-term function is given in general by

$$Auto\text{-}term\ Function = \int_{-\infty}^{+\infty} \Phi(\theta,\tau)\big|_{\theta=-a\tau} \cdot e^{-j\omega\tau}\, d\tau \qquad (3.6.1)$$

The auto-term functions of the CW and hyperbolic kernels are given in Eqs. (3.6.2) and (3.6.3) respectively

$$AUTO_{CW} = \int_{-\infty}^{+\infty} exp\left[-\frac{a^2\tau^4}{\sigma}\right] \cdot e^{-j\omega\tau}\, d\tau \tag{3.6.2}$$

$$AUTO_{Hy} = \int_{-\infty}^{+\infty} sech\left[-a\beta\tau^2\right] \cdot e^{-j\omega\tau}\, d\tau \tag{3.6.3}$$

Eqs. (3.6.2) and (3.6.3) cannot be further reduced to their closed forms although the integrands are well-behaved functions. To estimate the auto-term widths of the hyperbolic and CW kernels, the discrete Fourier transform versions of Eqs. (3.6.2) and (3.6.3) were used based on simulations in MATLAB. The normalised auto-term widths of the hyperbolic and CW kernels plotted against β are shown in Figure 3.6.1 in which the maximum auto-term width of each series is used as the normalisation factor. The auto-term functions and auto-term widths of the MTE *diamond case 1* and *2* forms along with those of the hyperbolic and CW kernels are plotted in Figure 3.6.2 and Figure 3.6.3 respectively. Figure 3.6.4–Figure 3.6.6 show the auto-term functions as a function of the frequency ω of various forms of the MTE kernel. Table 3.6.1 lists the auto-term widths of various types of the MTE kernels for $a = 1$ and compares them with those of the hyperbolic kernel and CW kernel.

As explained in Section 3.4, the auto terms are located around the origin and the hyperbolic kernel supports auto terms in the direction of the τ axis and the Choi-Williams kernel does so in the direction of the $(t - u)$ axis. It has also been shown that the CW kernel is more effective than the hyperbolic kernel since it is more concentrated around the origin whereas the hyperbolic kernel has large main lobes that extend in the direction of the τ axis. From Figure 3.6.1, the above remark can be validated. It is clear that the CW kernel is more auto-term supportive than the hyperbolic kernel by having a finer auto-term resolution. Thus, it can be drawn that auto terms are mainly located in the direction of the $(t - u)$ axis (vertically) rather than in the direction of the τ axis (horizontally).

From Figure 3.6.1, the hyperbolic auto-term resolution approaches that of the CW kernel when β is very small (σ is very large). For other values of β, the CW kernel outperforms the hyperbolic kernel which is a trade-off of having more effective cross-term suppression of the hyperbolic kernel at the expense of having a poorer auto-term resolution.

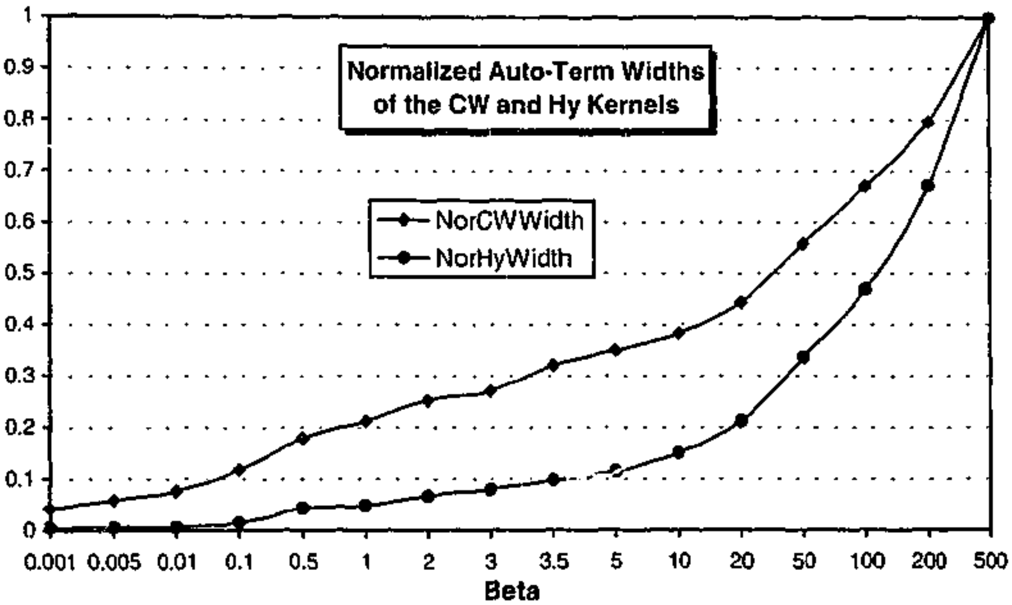

Figure 3.6.1: Normalised auto-term width of the hyperbolic and CW kernels

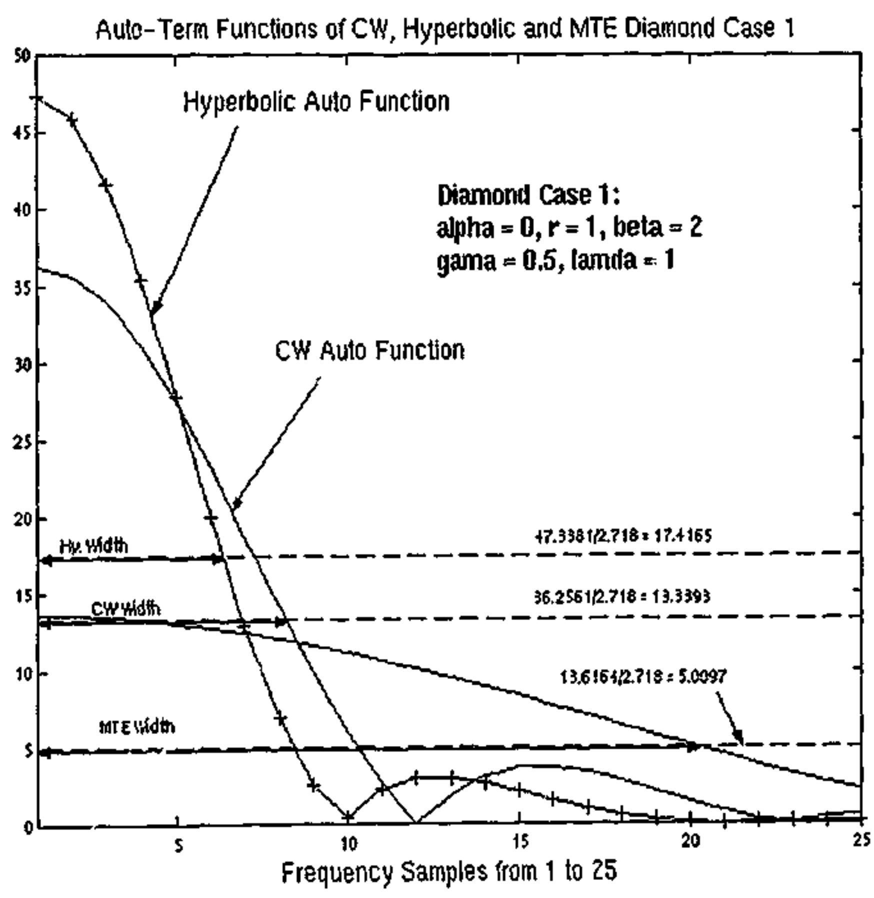

Figure 3.6.2: Auto-term functions and auto-term widths of the MTE *diamond case 1* kernel, hyperbolic and CW kernels

In Table 3.6.1, the following parameter values are chosen: a = 1, τ_0 = θ_0 = 1 and $\beta = \frac{1}{\sigma} = 1$ for simplicity which will not affect the generality of comparison. The auto-term width of the CW and hyperbolic kernels for other values of β is displayed in Figure 3.6.1 in which the hyperbolic kernel has a smaller auto-term width than that of the CW kernel. This clearly indicates the trade-off between cross-term suppression ability and auto-term resolution. Increasing β increases the auto-term resolution (seen in Figure 3.6.1) but also decreases the auto-term magnitude as seen in Figure 3.5.9. It should be noted that the MTE kernel becomes the WV kernel when $\lambda = 0$. In that case, the MTE kernel can be rewritten as $e^{-\pi\times 1}$, which is essentially the WV kernel multiplied by a constant $e^{-\pi} \approx 0.0432$. From Table 3.6.1, it should also be noted that the MTE kernel has 5 parameters which can generate up to (5!) = 120 different MTE kernels with different sets of parameters. The main aim of this work is not going to analyse the MTE kernel in detail but to show that there is still room for improvements even though the MTE has been shown to be an effective kernel [27]. Thus, only some popular forms of the MTE kernel are studied in this thesis. Further studies of the MTE kernel can also be found in [27].

Table 3.6.1: Auto-term widths (in frequency samples) for a = 1 of various forms of the MTE kernel. The auto-term widths and auto-term functions of the CW, hyperbolic kernels and MTE *Diamond Case 1* and *Case 2* forms are given in Figure 3.6.2 and Figure 3.6.3 respectively. Figure 3.6.4 to Figure 3.6.6 display the auto-term functions of various forms of the MTE kernel for a = 0.5, 1 and 2 respectively.

MTE Kernel	Parameter Value					Auto-term Width for a = 1		
	α	r	β	γ	λ	MTE	Hyperbolic‡	CW
Parallel Strip	0	1	1	1	1	0.5	5.5	7.0
Cross	0	−1	2	0.5	1	0.5	5.5	7.0
Snowflake§	0	$r = -2$	2	0.5	1	14.5	5.5	7.0
Untilted Elliptical	0	0	1	1	1	14.5	5.5	7.0
Tilted Elliptical	0	0.5	1	1	2	9.2	5.5	7.0
Diamond Case1	0	1	2	0.5	1	18.5	5.5	7.0
Diamond Case 2	0.1	0	1	1	1	13.0	5.5	7.0
Hyperbolic	1	0	1	1	1	10.5	5.5	7.0
Rectangular	10^{10}	0	1	1	1	6.5	5.5	7.0

‡ The parameters of the hyperbolic and CW kernels are $\beta = 1/\sigma = 1$ throughout the table.

§ For this set of parameters, the MTE *snowflake* and *untilted-elliptical* forms have identical auto-term functions. The auto-term functions of the MTE *snowflake* forms with $\gamma = 1$ and $\gamma = 10$ are shown in Figure 3.6.4–Figure 3.6.6 along with other MTE forms.

Table 3.6.1 shows the advantages and disadvantages of various forms of the MTE kernel over the CW and hyperbolic kernels in terms of auto-term width. From Table 3.6.1, it can be suggested that the MTE kernel can produce better auto-term quality than the hyperbolic and CW kernels (for $\beta = 1/\sigma = 1$) as larger auto-term widths are obtained from various types of the MTE kernel. Except in cases of the *parallel* and *cross* MTE kernels where the MTE auto-term widths are 0.5 ($\alpha = 0$, $r = \beta = \gamma = \lambda = 1$) compared with 5.5 and 7.0 of the hyperbolic and CW kernels respectively.

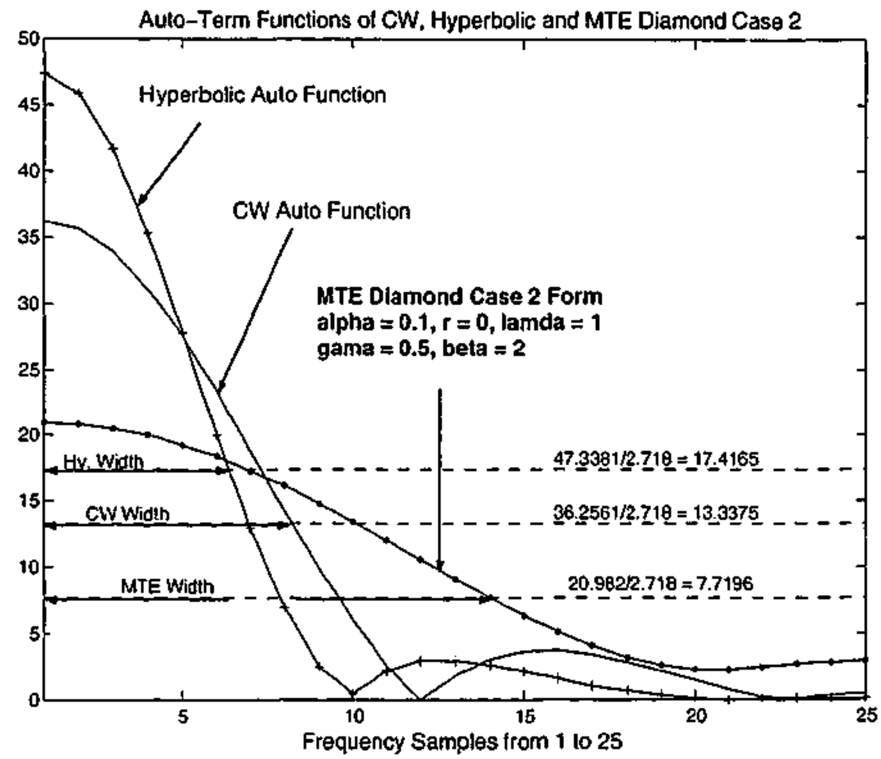

Figure 3.6.3: Auto-term functions and auto-term widths of the MTE *diamond case 2* kernel, hyperbolic kernel and CW kernel

Figure 3.6.2 and Figure 3.6.3 respectively display and compare the auto-term widths of the MTE *Diamond Case 1* and *Case 2* kernels with those of the hyperbolic and CW kernels for $a = 1$. The auto-term widths of other remaining MTE kernels are displayed in Figure 3.6.4 to Figure 3.6.6 for $a = 0.5$, 1 and 2 respectively. The shape of different types of the MTE kernel is clearly displayed for visualisation purposes only and their auto-term widths can be roughly estimated. All auto-term widths of various types of the MTE kernel were given in Table 3.6.1 for $a = 1$. It should be noted that the "hyperbolic" labelled in these figures is the *hyperbolic* MTE kernel (in *italic* font), not our proposed hyperbolic $sech(\beta\theta\tau)$ kernel.

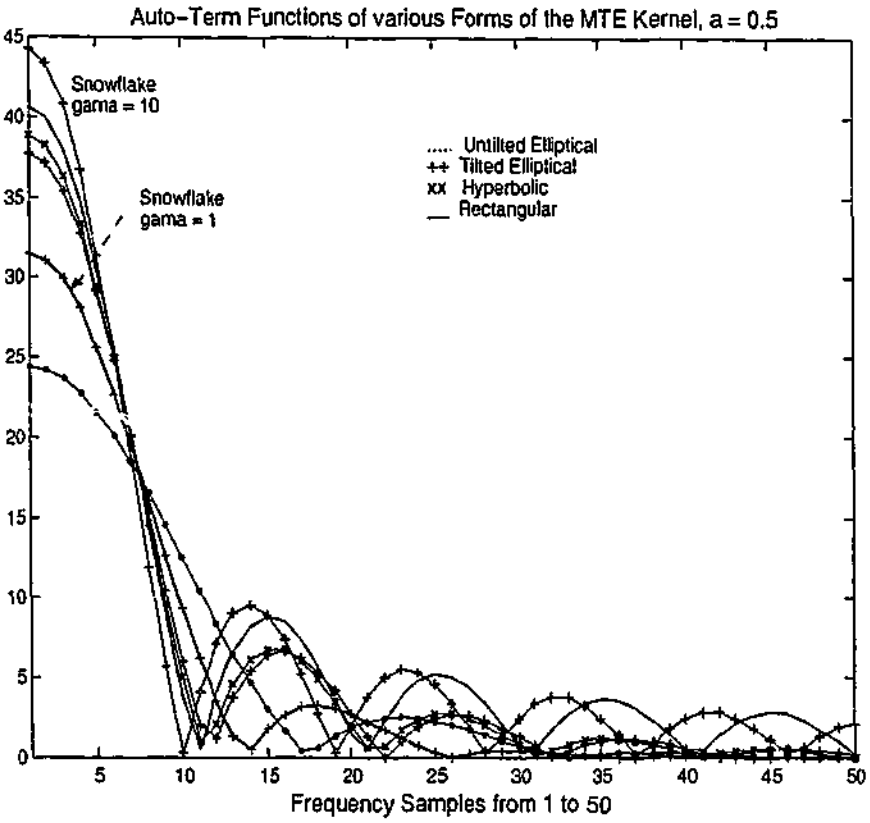

Figure 3.6.4: Auto-term functions for $a = 0.5$ of the *rectangular*, *snowflake*, *untilted elliptical*, *tilted elliptical* and *hyperbolic* MTE forms whose parameter values are shown in Table 3.6.1

The following conclusions on the MTE kernel are drawn after observing Figure 3.6.4, Figure 3.6.5 and Figure 3.6.6. The larger the auto-term slope a in the (θ, τ) plane of the kernel function $\Phi(\theta, \tau)$, the finer the auto-term resolution. It also appears that the *untilted elliptical* MTE kernel has the finest auto-term resolution and is most sensitive to the auto-term slope compared to other types of the MTE kernel, the hyperbolic and CW kernels. The *tilted elliptical* MTE kernel appears to have the coarsest auto-term width. The auto-term functions of the remaining MTE kernels (except the *untilted elliptical* MTE kernel) are almost identical (and so are their auto-term widths) for a small value of $a = 0.5$ as seen in Figure 3.6.4. From Figure 3.6.5 and Figure 3.6.6, the MTE *hyperbolic* and MTE *tilted elliptical* kernels have identical auto-term functions and hence equal auto-term widths. From Figure 3.6.5, the auto-term functions of the *snow flake* MTE kernel ($\gamma = 1$) and the *untilted* *elliptical* MTE kernel are identical when $a = 1$. This might suggest that at some specific values of a, the auto-term functions of various types of the MTE kernel are identical, yielding convergence of various forms of the MTE kernel, which reduces its uniqueness.

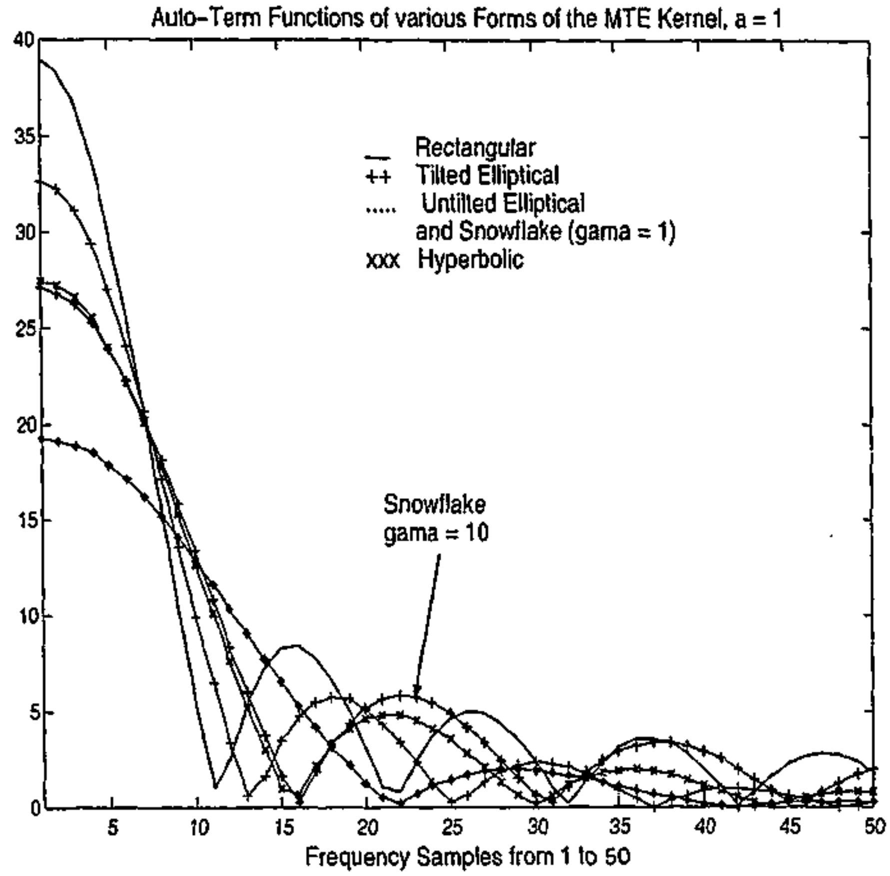

Figure 3.6.5: Auto-term functions for $a = 1$ of various forms of the MTE kernel, except the *diamond case 1* and 2 forms

Depending upon the kernel control parameter(s), specific requirements can be met. The MTE kernel is flexible, since it can generate various types of different kernels, but one of its disadvantages is that the *parallel* and *cross* forms have coarse auto-term resolutions in which their auto-term functions are identical triangular pulses with very large peaks. Further, the auto-term resolutions of the MTE *snowflake* and *untilted elliptical* forms are equal in value as seen in Table 3.6.1 for identical auto-term functions as observed earlier. For larger values of λ, the MTE *snowflake* auto-term function departs from that of the MTE *untilted-elliptical* kernel which suggests that these kernels can only be effectively used when λ is large.

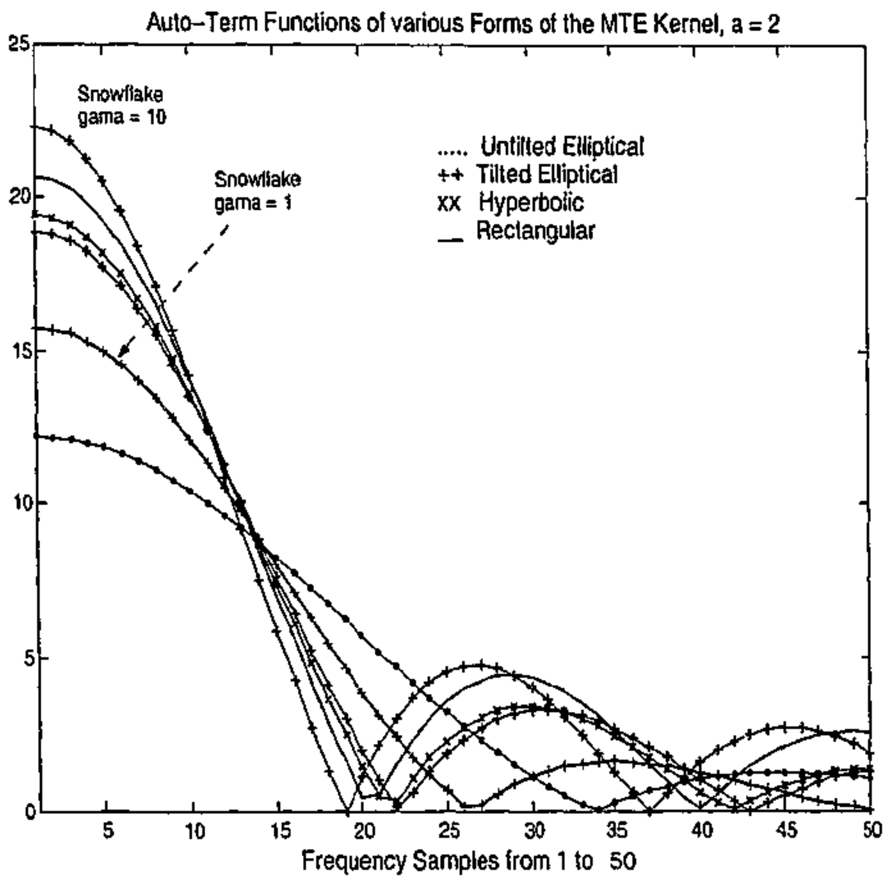

Figure 3.6.6: Auto-term functions for $a = 2$ of various forms of the MTE kernel except the *diamond case 1* and 2 forms

In this section, the relationship between the auto-term resolution, auto-term magnitude and β has been established. The larger the control parameter β is, the higher the auto-term resolution but the smaller the auto-term magnitude. There is also a trade-off between the auto-term resolution and cross-term suppression ability of a kernel. The finer the auto-term resolution, the less effective the kernel is in cross-term suppression. From this, it might be suggested that the MTE kernel is less cross-term suppression effective compared with the hyperbolic kernel and CW kernel since most MTE kernels have finer auto-term resolutions than those of the former two kernels as was shown earlier.

Section 3.7 examines the noise variance of the hyperbolic and CW time-frequency power spectra so that further conclusion(s) on the trade-off among auto-term resolution, cross-term suppression and noise robustness can be established.

3.7 Noise Variance Calculation

In practice, one usually deals with a complex noise and real white Gaussian noise. Previous work done by Stankovic and Ivanovic [29], Hearon and Amin [30, 50] found that given an input complex white Gaussian noise with variance σ_{in}^2, the noise variance σ^2 produced by the input noise in time-frequency power spectrum is given by

$$\sigma^2 = \sigma_{in}^4 \sum_{\tau=-\infty}^{+\infty} \sum_{(t-u)=-\infty}^{+\infty} \left| W(\tau,(t-u)) \right|^2 \tag{3.7.1}$$

where $W(\tau, t-u)$ is the weighting function of the kernel function $\Phi(\theta, \tau)$.

Eq. (3.7.1) can be clearly interpreted as the volume under the surface of the squared weighting function which is independent of the frequency ω. Since it is almost impossible to estimate the Fourier transform of the MTE kernel in closed forms, its noise variance could not be performed in this work.

The normalised noise variance of the CW and hyperbolic kernels, as a function of β, is plotted in Figure 3.7.1, from which it can be suggested that the hyperbolic kernel is more noise robust than the CW kernel for $\beta \geq 3$. For detailed analysis of the noise variance of other kernels, see [29, 50]. Hence, it can be concluded that kernels that can effectively suppress cross terms tend to be more noise robust (the hyperbolic kernel) than kernels that are less cross-term effective but have a finer auto-term resolution (in this case, the CW and the MTE kernels). This important relationship agrees with what was reported in [13-15, 28].

For the case of real noise, the noise variance is given by [29, 50]

$$\sigma^2(\omega) = \sigma_{in}^4 \sum_{\tau=-\infty}^{+\infty} \sum_{(t-u)=-\infty}^{+\infty} \left| W(\tau,(t-u)) \right|^2 + W(\tau,(t-u)) \cdot W^*(\tau,(t-u)) \cdot e^{-j4\omega\tau} \tag{3.7.2}$$

where all notations have the same meaning as in Eq. (3.7.1) and "*" indicates complex conjugate operation.

From Eq. (3.7.2), it can be seen that the noise variance is a function of ω and gains its maximum value when $\omega = 0$, thus the maximum real noise variance in the time-frequency power spectrum is given by

$$\sigma_{\max}^2 = 2\sigma_{in}^4 \sum_{\tau=-\infty}^{+\infty} \sum_{(t-u)=-\infty}^{+\infty} |W(\tau,(t-u))|^2 \tag{3.7.3}$$

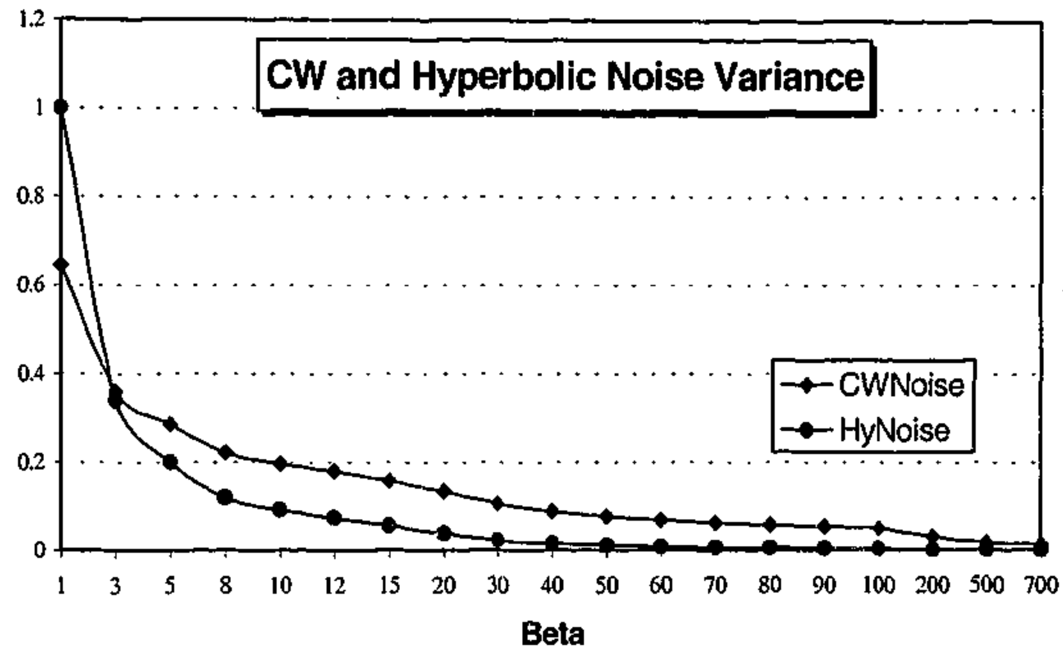

Figure 3.7.1: Normalised noise variance of the CW and hyperbolic kernels as a function of β

Eq. (3.7.3) is evidently a function of the volume under the squared weighting function. Thus, it is important to note that to ensure robustness in the time-frequency power spectrum, the volume under the squared weighting function should be minimised which means effective cross-term suppression. Eqs. (3.7.1) and (3.7.3) have a similar form except that in the case of real noise, the noise variance is a function of the frequency which peaks at $\omega = 0$ and has a magnitude of twice as large compared with that of the complex noise given by Eq. (3.7.1).

Figure 3.7.2 and Figure 3.7.3 display contour plots of the hyperbolic and CW time-frequency power spectra respectively for a sum of two chirp signals without noise interference. Figure 3.7.4 and Figure 3.7.5 display contour plots of the CW and hyperbolic time-frequency power spectra respectively for a sum of two chirped signals embedded in a 3-dB Gaussian noise. The corresponding 3-D plots of the hyperbolic and CW time-frequency power spectra without noise interference, whose contours plots are displayed in Figure 3.7.2 and Figure 3.7.3, are given in Figure 3.7.6 and Figure 3.7.7 respectively.

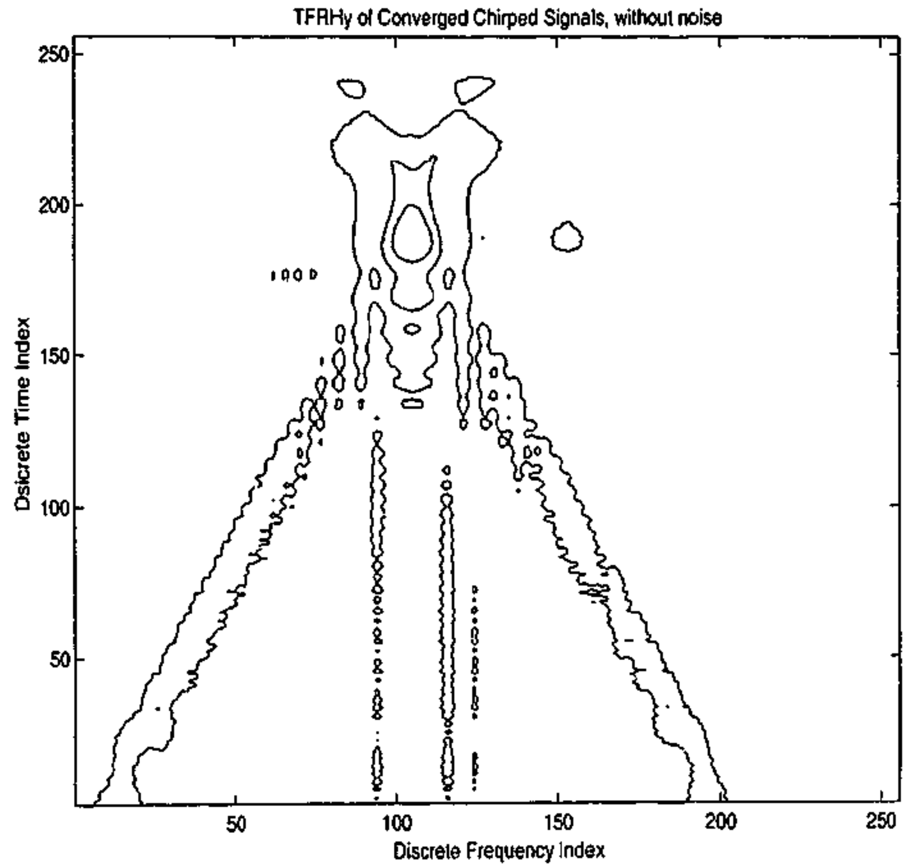

Figure 3.7.2: Contour plot of the hyperbolic time-frequency power spectrum of two chirped signals when no noise is added, $\beta = 10$. The cross-term region is approximately from discrete frequencies 95 to 125. The x-axis and y-axis should read "Discrete Frequency" and "Discrete Time" respectively. This convention is also applied to the remaining graphs in this chapter.

As expected, by comparing Figure 3.7.2 and Figure 3.7.3, it might be suggested that the hyperbolic time-frequency power spectrum is clearer than the CW time-frequency power spectrum due to a smaller amount of cross terms in the region between the two auto-term arms. In addition, at the intersection of the two arms, there are less interference from the auto terms themselves than in the case of the CW time-frequency power spectrum as displayed in Figure 3.7.3, which is another advantage of the hyperbolic kernel over the CW kernel.

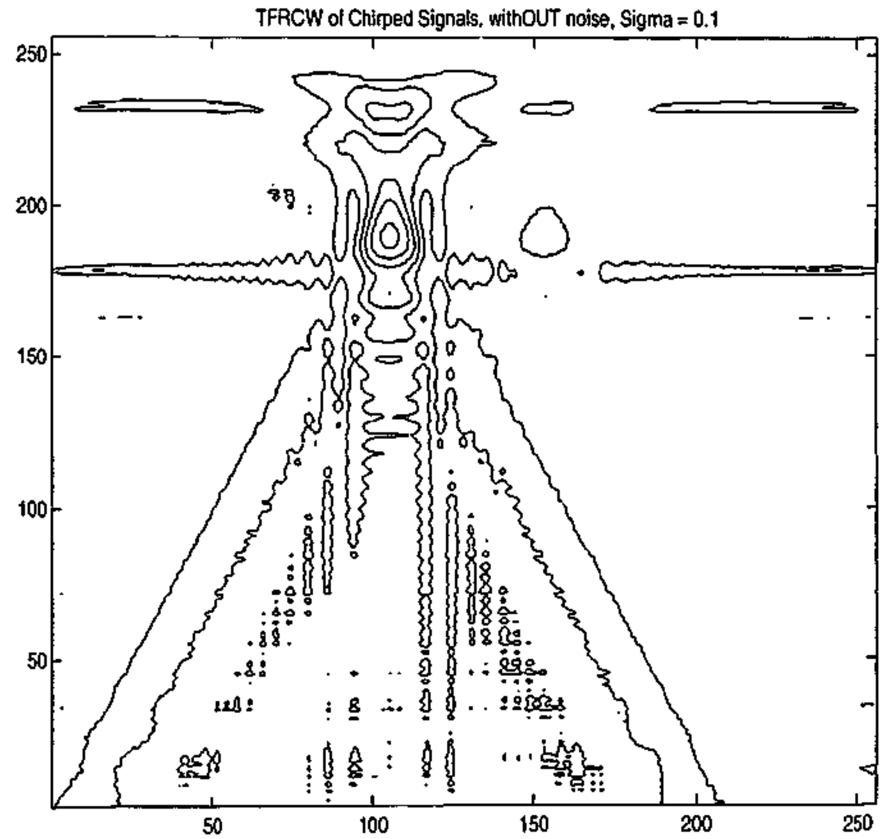

Figure 3.7.3: Contour plot of the CW time-frequency power spectrum of two chirped signals when no noise is added, $\sigma = 0.1$. The x-axis is "Discrete Frequency" and the y-axis is "Discrete Time". The cross-term region is approximately from discrete frequencies 45 to 175.

As stated earlier, the CW time-frequency power spectrum has more cross terms in the region between the two auto-term arms and in the directions along the arms as seen in Figure 3.7.3 which is a disadvantage of the CW kernel compared with the hyperbolic kernel. However, the CW kernel, due to its finer auto-term resolution, has stronger auto-term arms in the time-frequency power spectrum as shown in Figure 3.7.3 compared with those of the hyperbolic time-frequency power spectrum in Figure 3.7.2. This advantage establishes an important trade-off between auto-term resolution and cross-term suppression of the two kernels as discussed throughout this chapter. However, there appears one more important parameter in this trade-off (as stated earlier in this section) which is the noise robustness, which will be graphically shown in Figure 3.7.4 and Figure 3.7.5.

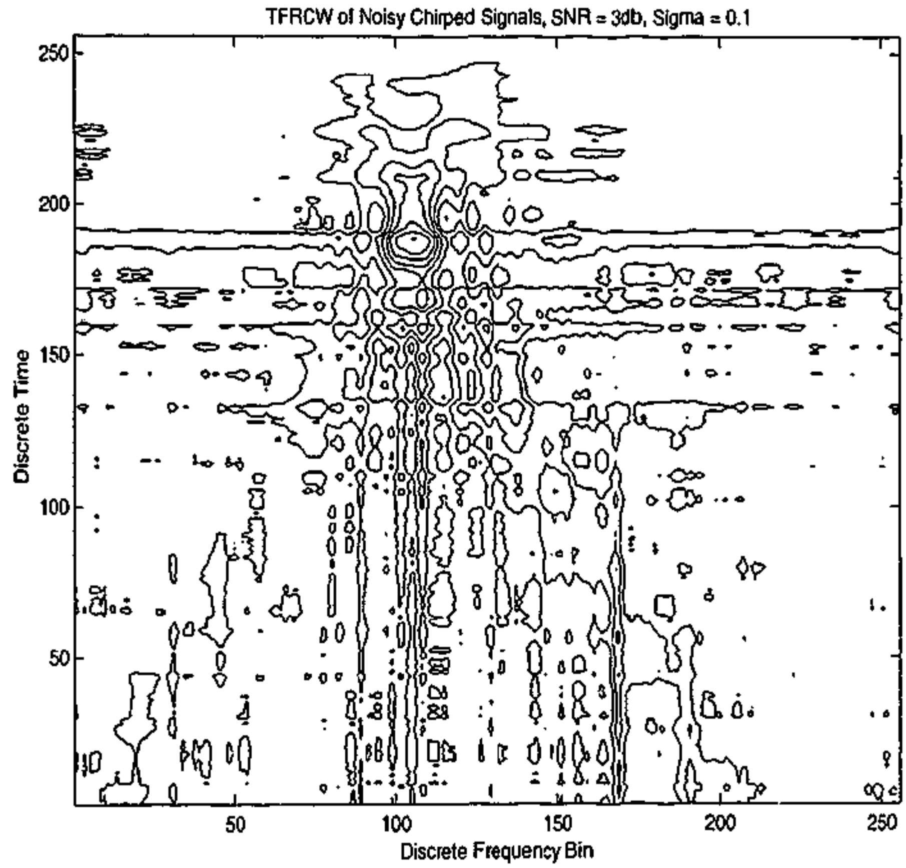

Figure 3.7.4: Contour plot of the CW time-frequency power spectrum of chirp signals embedded in a 3-dB noise, $\sigma = 0.1$

From Figure 3.7.4, it is seen that the CW time-frequency power spectrum is significantly distorted under the effects of a 3-dB noise source. It is very hard to distinguish the two main auto-term arms of the spectrum and therefore it might be said that the CW time-frequency power spectrum is not robust. The cross terms appear to remain almost unchanged under the effects of a noise source even though they are slightly degraded.

From Figure 3.7.5, the hyperbolic time-frequency power spectrum, although is better than the CW spectrum, still suffers from noise interference. The left auto-term arm of the power spectrum is distorted, however, the right auto-term arm can still be recognisable as it was not the case for the CW time-frequency power spectrum displayed in Figure 3.7.4. The hyperbolic cross terms are also degraded (as were the CW cross terms) as compared with the case in which no noise was added in Figure 3.7.2. However, the amount of cross terms appears to remain unchanged. This might suggest that noise sources do not considerably affect cross terms in time-frequency power spectra, however, the auto terms are significantly reduced.

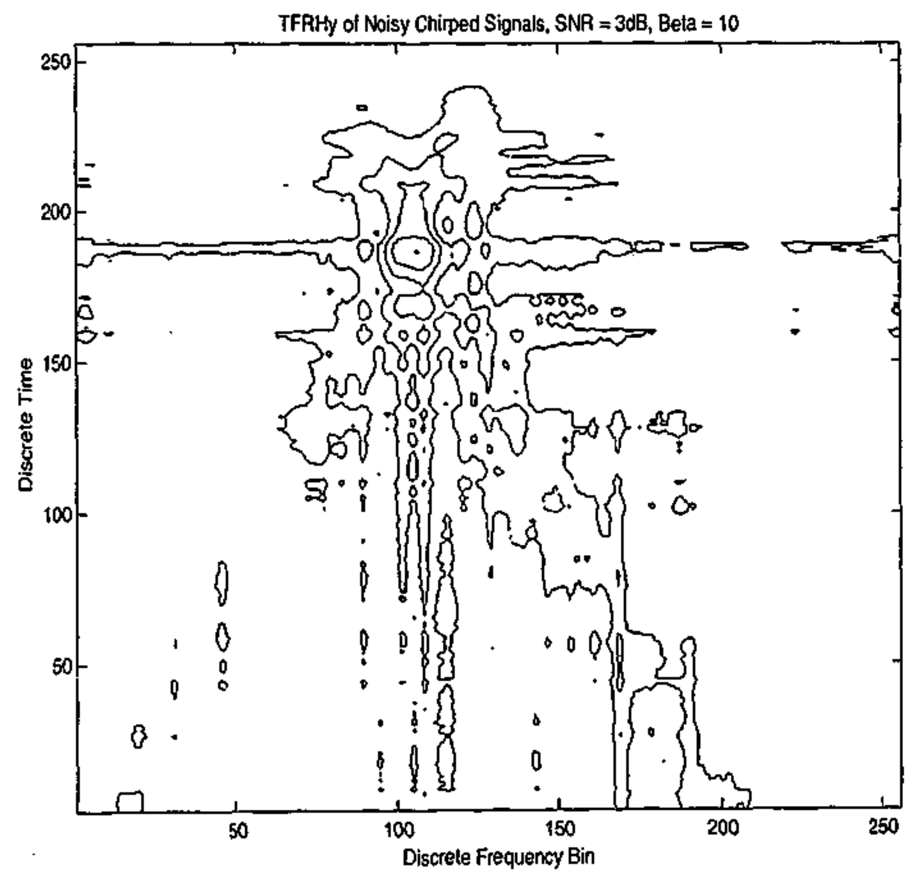

Figure 3.7.5: Contour plot of the hyperbolic time-frequency power spectrum of a sum of two chirped signals embedded in a 3-dB noise, $\beta = 10$

As can be seen in Figure 3.7.4, the CW time-frequency power spectrum, by having a finer auto-term resolution, considerably suffers under the effects of noise interference compared to the hyperbolic time-frequency spectrum (Figure 3.7.5). Obviously, the latter can withstand tougher conditions than the former. This suggests that the more effective the kernel is at cross-term suppression, auto-term magnitude and noise robustness, the poorer its auto-term resolution. This is the prime result that this chapter aims to achieve. The next few chapters further develop this trade-off so that more understanding on time-frequency kernels can be gained.

The 3-D mesh plots of the hyperbolic and CW time-frequency power spectra are provided in Figure 3.7.6 and Figure 3.7.7 to give further understanding on the effects of a noise source on the spectrum. Mesh plots of the CW and hyperbolic time-frequency power spectra embedded in a 3-*dB* noise source are given in Figure 3.7.8 and Figure 3.7.9 respectively.

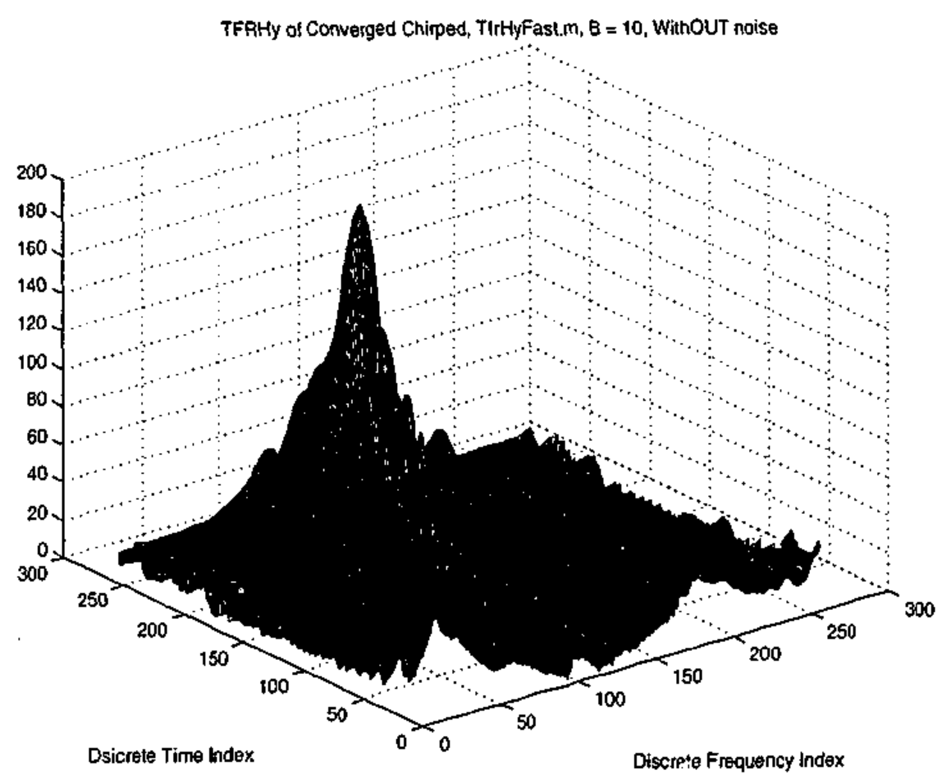

Figure 3.7.6: 3-D plot of the hyperbolic time-frequency power spectrum of a sum of two chirp signals, $\beta = 10$, no additional noise is added

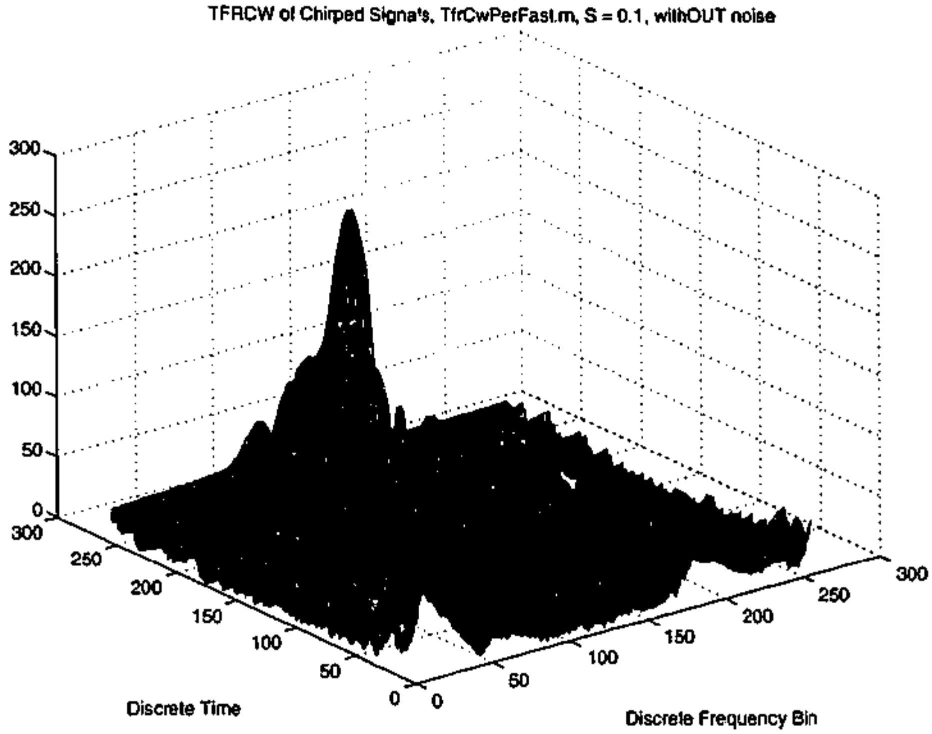

Figure 3.7.7: 3-D plot of the CW time-frequency power spectrum of a sum of two chirp signals, $\sigma = 0.1$, no additional noise is added

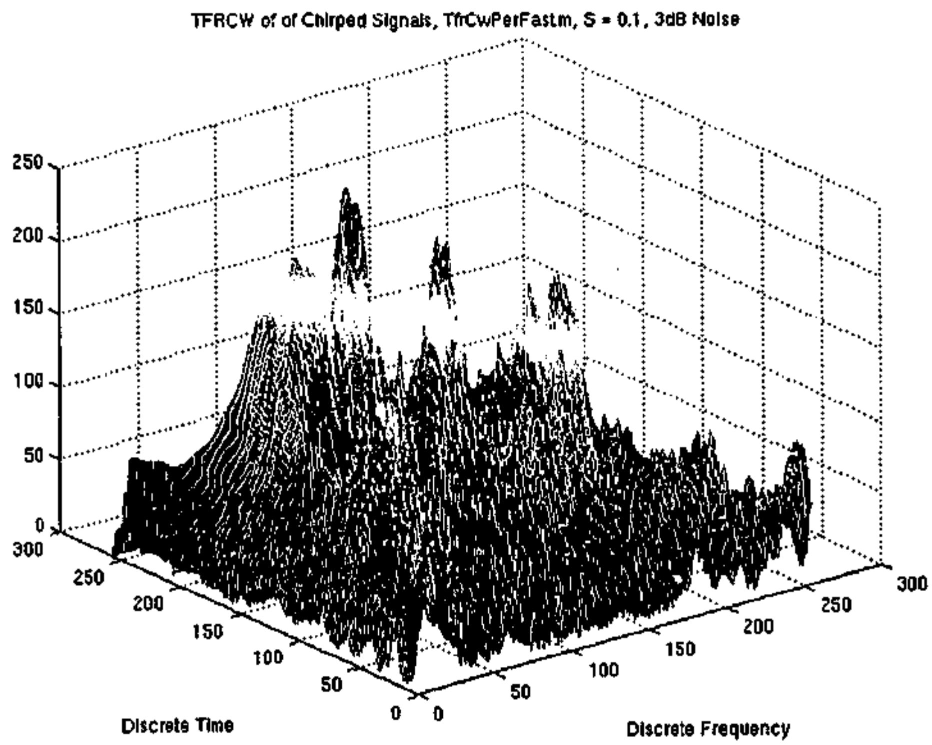

Figure 3.7.8: Mesh plot of the Choi-Williams time-frequency power spectrum embedded in a 3-*dB* noise, $\sigma = 0.1$

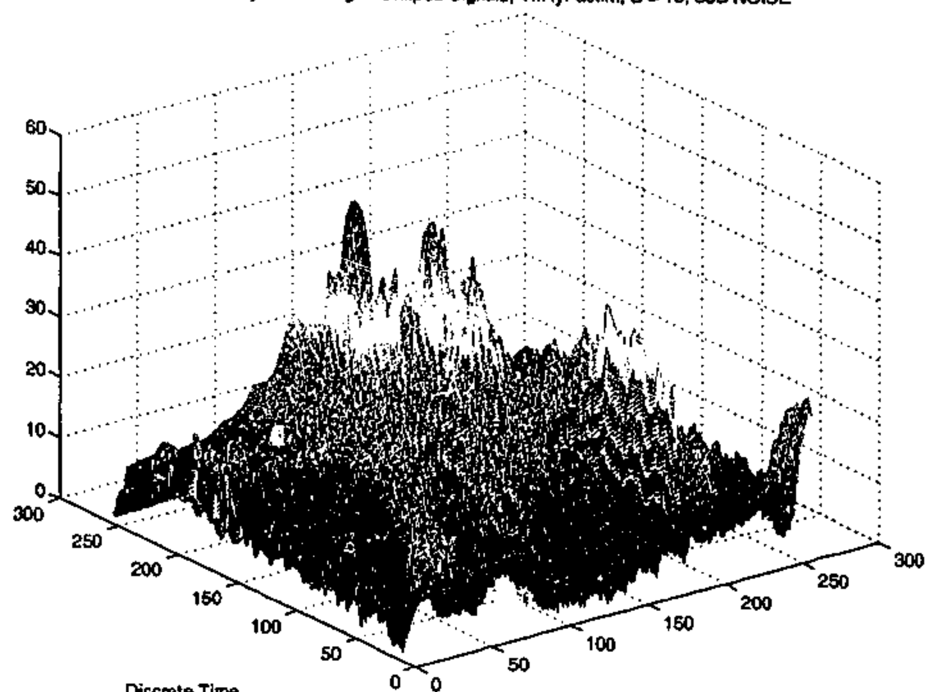

Figure 3.7.9: Mesh plot of the hyperbolic time-frequency power spectrum embedded in a 3-*dB* noise, $\beta = 10$

3.8 Conclusion

The hyperbolic $[sech(\beta\theta\tau)]^n$ (with $n = 1$) kernel has been shown to be effective in cross-term suppression. In particular, we have shown its effectiveness for a sum of two complex-exponential signals, for $\beta \geq 1.45$ and in the case of a sum of two chirp signals, for $20 \geq \beta \geq 0.5$. The hyperbolic kernel has also been shown to be better than the CW kernel in terms of cross-term suppression ability and lower noise variance for well-chosen values of $\beta \geq 3$. Thus, the applicable range of β is $20 \geq \beta \geq 3$.

However, the hyperbolic kernel has a smaller auto-term resolution than that of the CW kernel and most types of the MTE kernels, except in the case of the MTE *rectangular* form where the auto-term widths of the three kernels are approximately equal. There appears to be a trade-off among auto-term resolution, auto-term magnitude, cross-term suppression ability and noise robustness. The more effective the kernel is at cross-term suppression, auto-term magnitude and noise robustness, the poorer its auto-term resolution. This is an important trade-off that should be considered in choosing the appropriate kernel for a particular application. Further research needs to be carried out to investigate other members of the hyperbolic kernel family, such as the $[sech(\beta\theta\tau)]^3$ kernel or higher-order kernels, for further improvements on auto-term resolution and noise robustness. An additional constraint on the boundedness of a kernel weighting function has also been stated.

The first contribution of this chapter is to propose the new hyperbolic kernel and its family of kernels for the time-frequency power spectrum. The second contribution is that the hyperbolic kernel has been shown to be more effective than the well-known Choi-Williams in terms of cross-term suppression and noise robustness but less effective in terms of auto-term resolution. It has also been shown that the hyperbolic kernel is simpler than the MTE kernel but its auto-term resolution is poorer than that of most types of the MTE kernel. The third contribution is that the important trade-off among auto-term resolution against cross-term suppression and noise robustness is established. This relationship has been reported in the literature by a number of researchers but the effects of noise and noise robustness have not been previously identified. The next chapter introduces the hyperbolic wavelet and studies its properties in detail.

Chapter 4: THE HYPERBOLIC WAVELET FUNCTION

This chapter continues the study of the hyperbolic kernel in Chapter 3 by exploring the new hyperbolic wavelet. The primary aim of this chapter is to further explore the differences between the hyperbolic and Choi-Williams (CW) kernels by comparing the hyperbolic wavelet and the Choi-Williams or Mexican-hat wavelet. More importantly, the main aim is to show that there exists a strong link between time-frequency kernels and wavelets. This relationship helps to expand the time-frequency and wavelet areas in the field of signal processing. The contents of this chapter are necessary background for Chapters 6 and 7 in which the hyperbolic wavelet power spectra of a number of stationary and non-stationary signals are examined.

The next chapter introduces the first application of the hyperbolic kernel as a time-frequency signal detector. The WV detector is used as a benchmark for signal detection comparison. The CW and hyperbolic detectors are also compared in detail.

4.1 Introduction

Studies of wavelet functions and wavelet transform have been done over many years, starting with the simplest wavelet, the Haar wavelet [51]. There is a strong connection between the wavelet transform technique and time-frequency power spectrum technique since both of these techniques view the energy density of a non-stationary signal in both time and frequency domains. One of the most popular wavelets is the Mexican-hat wavelet, which is the negative second derivative function [52] of the Gaussian pulse, $\Phi(\theta, \tau) = e^{-\theta^2\tau^2/\sigma}$, where σ is the pulse control parameter.

In 1989, the Gaussian pulse was used by Choi and Williams [28] as a time-frequency kernel to suppress cross terms in the time-frequency power spectrum which means the Gaussian pulse and the CW kernel are identical except that they have been used in different areas of signal processing. In other words, there exists a strong link between wavelet theory context and time-frequency power spectrum context in which the CW kernel and Mexican-hat wavelet pair is one typical example. However, there have not been investigations on the relationship between the Mexican-hat wavelet and the CW time-frequency kernel in the literature.

Based on the relationship between the CW kernel and the Mexican-hat wavelet, there ought to be a strong connection between the hyperbolic kernel and the hyperbolic wavelet (by taking the negative second derivative of the hyperbolic kernel) if this derivative function satisfies an admissibility constraint which is given by

$$\int_{-\infty}^{+\infty} \psi(t)dt = 0 \tag{4.1.1}$$

Over the years, a large number of wavelets have been proposed and extensively studied, starting with the Haar wavelet [51, 52] proposed in 1910. In the 1980s, a number of excellent wavelets were proposed such as the Daubechies wavelet [52], the Meyer wavelet [53] and the Mallat wavelet [54-59]. These wavelets provide excellent features such as orthogonality, bi-orthogonality, vanishing moments, existence of the scaling function, continuity and discrete transform of the wavelet function, which have received considerable attention from mathematicians. One common feature of these wavelets is that their mother wavelets are not symmetrical and do not possess explicit expressions.

It has been claimed that wavelets which do not have explicit expressions tend to have more useful properties than those that have explicit expressions. For example, the Daubechies wavelet [52, 60-62] can only be represented by recursive relations since it does not have an explicit expression, however, it has many desirable properties including orthogonality, FIR filtering and vanishing moments. Among the wavelets, those that do not have explicit expressions are more common than those that do have such explicit expressions.

Along with wavelets that are non-symmetrical and do not have e[illegible]it expressions, there exist a small number of wavelets that are symmetrical and ha[illegible]pressions. The CW and Morlet wavelets both belong to this particular cl[illegible] named "crude" wavelet class [67]. Since the hyperbolic and CW [illegible]ntial functions and similar in shape as was shown in Chapter 3, it i[illegible] strong possibility that the hyperbolic wavelet might exist. In addi[illegible]c and CW wavelets might have similar characteristics. The scaling fu[illegible]mbers of the "crude" wavelet class do not exist which make them un-orthog[illegible]hich is the main disadvantage of this class. However, these wavelets are symmetrical and have explicit expressions which make their investigations and understanding much easier.

There are a large number of papers written about wavelets from a mathematical point of view as listed throughout the chapter. However, this chapter is written from an engineering point of view in which intensive mathematics is avoided. Instead, the physical interpretation and exemplification will be discussed and explored in detail. Properties such as orthogonality, bi-orthogonality, scaling property and regularity will not be discussed in detail.

The purpose of this chapter is to investigate the hyperbolic wavelet function if it satisfies the admissibility constraint imposed by Eq. (4.1.1). In a sense, this chapter emphasises that there exists a strong relationship between time-frequency kernels and wavelet functions. This relationship is important in terms of diversifying the wavelet and time-frequency areas in which new kernels can be generated from corresponding wavelets and vice versa. The hyperbolic, CW (Mexican-hat) and Morlet wavelets are compared in terms of scale resolution, scale limit and aliasing effects. From that, particular applications for each wavelet can be found. Some useful overview papers that summarise the developments in the field of wavelet theory can be found in [55, 57, 59, 63, 64].

This chapter is organised as follows. Section 4.2 provides the literature survey of some popular and well-known wavelets. Section 4.3 investigates important properties of the hyperbolic, Morlet and CW wavelets by calculating their fundamental parameters (Section 4.3.1) and band-peak frequency (Section 4.3.2), examining the aliasing effects (Section 4.3.3), estimating the maximum scale (Section 4.3.4) and the scale resolution in Section 4.3.5. Section 4.4 investigates other properties of the hyperbolic wavelet from a mathematical point of view including symmetry (Section 4.4.1), orthogonality and bi-orthogonality (Section 4.4.2), compactly supported orthogonality and bi-orthogonality (Section 4.4.3), an arbitrary number of vanishing moments (Section 4.4.4), existence of the scaling function $\varphi(t)$ (Section 4.4.5) and the *FIR* (Finite Impulse Response) filtering property (Section 4.4.6).

4.2 The "Crude" Wavelet Group and the Hyperbolic Wavelet Function

There are many interesting wavelets that have been proposed and studied by many researchers and mathematicians such as Daubechies, Mallat, Meyer, Morlet whose wavelets are named after them. These wavelets have been extensively studied and their many interesting and useful properties can be found in [51-53, 55, 56, 58-61, 63-66]. The Daubechies wavelet is probably the most popular wavelet due to its desirable properties such as orthogonality, bi-orthogonality, vanishing moments and existence of the scaling function. The Daubechies wavelet family is recursive, which means that the formula of any wavelet cannot be explicitly expressed but as a function of another wavelet in the family. This fact, although not convenient, provides the Daubechies wavelet family with many useful properties [52]. Other existing wavelets such as Cauchy, Poisson, chirp can be found in [65].

Wavelet functions have been classified into four classes [67]

1. Type 1 (orthogonal with *FIR* filtering): the wavelet is orthogonal and its *FIR* filter exists. This class includes the Daubechies, Coiflets and Symlets wavelets.

2. Type 2 (bi-orthogonal with *FIR* filtering): the wavelet is bi-orthogonal and its *FIR* filter exists. The Bior-Splines wavelet belongs to this class.

3. Type 3 (orthogonal with scale function): the wavelet and its scale function exist, its *FIR* filter does not exist, however. The Meyer wavelet is a typical member of this class.

4. Type 4 (*FIR* filter and scaling function do not exist): this class has been considered as a "crude" wavelet class since the wavelet scaling function and its *FIR* filter do not exist. However, the support range of the wavelets in this class can be identified by the time-base interval (T in Section 4.3). The wavelets in this class are usually symmetrical and have explicit expressions. As already noted, this chapter only deals with this particular class of wavelets. The hyperbolic, CW and Morlet wavelets belong to this class.

Unlike the Daubechies wavelet family, the Mexican-hat (CW) and Morlet wavelets have explicit expressions and are odd symmetrical about the origin. By having explicit expressions, the Morlet and CW wavelets are considered as "crude" wavelets in which their scaling functions have been shown to be non-existent [67]. The CW wavelet, given in Eq. (4.2.1), is found by taking the negative second derivative of the CW kernel [28] (discussed in Chapter 3) as given by

$$\psi_{CW}(t) = \frac{2}{\sigma} \cdot exp\left(-\tau^2/\sigma\right) \cdot \left(-1 + 2\tau^2/\sigma\right) \tag{4.2.1}$$

The Morlet wavelet [66] is given by

$$\psi_{Morlet}(t) = exp(j\omega_\psi t) \cdot exp\left(-\frac{|t|^2}{\sigma}\right), \text{ where } \sigma \text{ is the wavelet control parameter.} \tag{4.2.2}$$

The frequency representations of the CW and Morlet wavelets are give by

$$F\{\psi_{CW}(t)\} = \hat{\psi}_{CW}(\omega) = \sqrt{\pi\sigma} \cdot \omega^2 \cdot exp\left(-\sigma\omega^2/4\right), \text{ and} \tag{4.2.3}$$

$$F\{\psi_{Morlet}(t)\} = \hat{\psi}_{Morlet}(\omega) = \sqrt{\pi\sigma} \cdot exp\left(-\sigma(\omega-\omega_\psi)^2/4\right) \tag{4.2.4}$$

where the symbol $F\{\cdot\}$ denotes the Fourier transform operation of the function $\{\cdot\}$.

The CW and Morlet wavelets belong to Type-4 wavelet group which is the "crude" wavelet group since they are symmetrical and have explicit expressions. It should be stressed that this class of wavelets currently consists of only very few members. Thus, it is important to further explore their characteristics and potential applications.

The hyperbolic wavelet function is generated by taking negative second derivative of the hyperbolic kernel which was proposed in Chapter 3. It is recalled that the hyperbolic kernel is given by $\Phi(\theta) = [sech(\beta\theta)]^n$. The first derivative function of the hyperbolic kernel of order n is given by

$$\frac{d\Phi}{d\theta} = -n\beta\cdot[tanh(\beta\theta)]\cdot[sech(\beta\theta)]^n \tag{4.2.5}$$

The second derivative function of the hyperbolic kernel of order n or the hyperbolic wavelet function is therefore given by

$$\frac{d^2\Phi}{d\theta^2} = n\beta^2\cdot[sech(\beta\theta)]^n\cdot\{n-(n+1)\cdot[sech(\beta\theta)]^2\} \tag{4.2.6}$$

The hyperbolic wavelet function, $\psi_{Hy}(\theta)$, can be formed by taking the negative second derivative function given by Eq. (4.2.6) as

$$\psi_{Hy}(\theta) = (-1)\cdot n\beta^2\cdot[sech(\beta\theta)]^n\cdot\{n-(n+1)\cdot[sech(\beta\theta)]^2\} \tag{4.2.7}$$

For $n = 1$, taking the Fourier transform of Eq. (4.2.7), the frequency domain representation of the hyperbolic wavelet function is given by

$$F\{\psi_{Hy}(\theta)\} = \hat{\psi}_{Hy}(\omega) = \frac{\pi\omega^2}{\beta}\cdot sech\left(\pi\omega/2\beta\right) \tag{4.2.8}$$

By substituting $n = 1$ into Eq. (4.2.7), the first-order hyperbolic wavelet function is obtained. The hyperbolic wavelet will be examined later by simulation in Section 4.3.1 to determine whether it satisfies the odd-symmetry condition imposed by Eq. (4.1.1), i.e. the area under the curve of the second derivative of the hyperbolic kernel is zero.

In 1992, Szu [125, 126] proposed the soliton wavelet to study nonlinearity dynamics in sonar, ocean waves and so on. The soliton wavelet can be given as

$$\psi_{H_y}(\theta) = [sech(\beta\theta)]^2 \cdot cos(\pi\theta) \tag{4.2.9}$$

By comparing Eqs. (4.2.7) and (4.2.9), it should be noted that the hyperbolic wavelet is not the soliton wavelet proposed by Szu. In fact, the soliton wavelet employs one of the members of the hyperbolic kernel, the second-order hyperbolic kernel $[sech(\beta\theta)]^2$ studied in Chapter 3. This shows that the hyperbolic kernel family can find useful applications in the field of non-linear signal processing. From Eq. (4.2.9), it can be seen that the soliton wavelet is symmetrical and thus it belongs to the "crude" wavelet group. However, as stated earlier, since the soliton wavelet employs one of the hyperbolic kernels in its expression, the investigation of this wavelet will be not be studied in this chapter but only the hyperbolic wavelet.

Other wavelets in the same family with different scales can be obtained by using a translation and dilation relationship or multi-resolution relationship [51]

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \cdot \psi_{a,b}\left(\frac{t-b}{a}\right) \tag{4.2.10}$$

where a is the scale index and b is the translation or time index of the wavelet. The mother wavelet corresponds to $a = 1$ and $b = 0$.

For each value of the scale index a, there is one unique corresponding wavelet function which can be considered as a band-pass filter. In the frequency domain, the multi-resolution relationship becomes [66]

$$\hat{\psi}_{a,b}(\omega) = \sqrt{a} \cdot \hat{\psi}_{a,b}(a\omega) \cdot exp(-j\omega b) \tag{4.2.11}$$

where $\hat{\psi}_{a,b}(\omega)$ is the Fourier transform of the wavelet function $\psi_{a,b}(t)$.

In this section, explicit expressions of the CW wavelet (Eqs. (4.2.1)–(4.2.3)), Morlet wavelet (Eq. (4.2.4)) and the hyperbolic wavelet (Eqs. (4.2.7) and (4.2.8)) have been given in both time and frequency domains. It is important to examine some important properties of these wavelets by estimating their numbers of sampling points, aliasing effects, the maximum possible scales and their scale resolutions. These properties will be studied in detail in Section 4.3.

4.3 Properties of the Choi-Williams, Morlet and Hyperbolic Wavelets

From an engineering point of view, to study properties of a wavelet function, it is necessary to investigate its scale resolution, maximum scale used, number of sampling points, relation to input sampling interval and aliasing effects. The main reason that sampling of a wavelet function is of concern is that digital signal processing is practical and important. In addition, the input waveform is usually a discrete set of samples from a continuous process. This section examines the above mentioned properties in detail. Firstly, fundamental parameters of the CW, Morlet and hyperbolic wavelets are estimated.

4.3.1 Fundamental Parameters

The Morlet wavelet was studied and used to study the transition to turbulence in [66] by Jordan, Miksad and Powers in which the following useful parameters are numerically estimated

The admissibility constant $C_\psi = 2\pi \int_{-\infty}^{+\infty} \frac{|\hat{\psi}(\omega)|^2}{\omega} \, d\omega$ (4.3.1)

The first moment in time domain $t_0 = \frac{\int_{-\infty}^{+\infty} t \cdot |\psi(t)|^2 \, dt}{\int_{-\infty}^{+\infty} |\psi(t)|^2 \, dt}$ (4.3.2)

The time variance or time width from the mean of the wavelet function [68] is given by Eq. (4.3.3).

$$\sigma_t = \sqrt{\left(\int_{-\infty}^{+\infty} (t - t_0)^2 \cdot |\psi(t)|^2 \, dt \right)} \qquad (4.3.3)$$

The larger the width around the mean, the less energy-concentrated the wavelet is. Wavelet functions that have narrow widths about their means have high energy density. The quantity σ_t, given in Eq. (4.3.3), is also the standard deviation from the mean or the second-moment of the wavelet. The smaller the frequency domain standard deviation is, the more energy-concentrated the wavelet is in frequency domain. This feature will be useful for comparison purposes.

The first moment ω_0 in frequency domain is given by

$$\omega_0 = \frac{\int_{-\infty}^{+\infty} \omega \cdot |\psi(\omega)|^2 d\omega}{\int_{-\infty}^{+\infty} |\psi(\omega)|^2 d\omega} \tag{4.3.4}$$

The frequency variance σ_ω of the wavelet function is given by Eq. (4.3.5)

$$\sigma_\omega = \sqrt{\left(\int_{-\infty}^{+\infty} (\omega - \omega_0)^2 \cdot |\psi(\omega)|^2 d\omega \right)} \tag{4.3.5}$$

At this point, it is necessary to check whether the first-order hyperbolic wavelet given by Eq. (4.2.7) satisfies the admissibility constraint given by Eq. (4.1.1). By substituting Eq. (4.2.7) into Eq. (4.1.1) we obtain

$$C_{constraint} = \int_{-\infty}^{+\infty} -\beta^2 \cdot sech\,(\beta\theta)\,\{1 - 2\cdot[sech\,(\beta\theta)]^2\}\; d\omega \tag{4.3.6}$$

Simulation results showed that $C_{constraint}$ was very small which proved that the hyperbolic wavelet is a valid wavelet. As will be shown later in Figure 4.3.1, the hyperbolic wavelet is a symmetrical wavelet which means it belongs to the symmetrical Type 4 "crude" wavelet group.

Numerical values of fundamental parameters of the CW, Morlet and hyperbolic wavelets for $\beta = 1$ are given in Table 4.3.1.

Table 4.3.1: Fundamental parameters of the CW, Morlet and hyperbolic wavelets for $\beta = 1$

Wavelet	Parameter values, $\beta = 1$				
	C_ψ	t_0	σ_t	ω_0	σ_ω
Choi-Williams	3.98	0.0	1.48	2.12	2.386
Morlet	1.352	0.0	0.4	5.0	2.806
Hyperbolic for $n = 1$	1.25	0.0	0.87	1.6	1.23

The admissibility constant C_ψ represents the inverse-transform ability of a wavelet [66], i.e. when a function is transformed using a particular wavelet then it can be successfully recovered by using the inverse wavelet transform. As can be seen from the above table, all three wavelets have finite admissibility constants which means that they are valid wavelets.

From Table 4.3.1, the hyperbolic wavelet has higher energy density than the CW and Morlet wavelets as shown by having a smaller σ_ω. This property is useful when analysing signals that have broad power spectra. The value of σ_t is used to obtain the length of the wavelet time-base sampling interval T. Typically, $T \geq 4\sigma_t$ to ensure that the mother wavelet is completely sampled. However, to make sure that the time-base sampling interval is long enough, graphical display of the mother wavelet is required. Fundamental parameters of the three wavelets are re-calculated for $\beta = 0.5$ and are given in Table 4.3.2. For various values of the control parameters σ and β, from simulation, it can be concluded that the hyperbolic wavelet satisfies the admissibility constraint imposed by Eq. (4.1.1) by having a very small area under the curve. The error in this case is always $O(10^{-6})$ or less from simulation results.

Table 4.3.2: Fundamental parameters of the CW, Morlet and hyperbolic wavelets for $\beta = 0.5$ using Eqs. (4.3.1)–(4.3.5)

Wavelet	Parameter values, $\beta = 0.5$				
	C_ψ	t_0	σ_t	ω_0	σ_ω
Choi-Williams	1.785	0.0	1.245	1.47	1.08
Morlet	1.58	0.0	0.6656	5.0	2.36
Hyperbolic	0.15	0.0	0.62	0.817	0.213

The larger the value of σ_t and σ_ω, the less time and frequency support the corresponding wavelet has respectively. The smaller σ_ω is, the less the number of frequencies that are required to express the spectrum of an input signal. In other words, this feature is closely related to the compression effects of a wavelet which are of practically significant. The more effective the compression effect of the wavelet is, the less the number of required scales of the wavelet are for an input signal which reduces the required computational time of the wavelet power spectrum and increases the efficiency of the calculation process.

Sections 4.3.2–4.3.5 estimate the band-peak frequencies, examine the aliasing effects, the maximum scales and scale resolutions as well as the total numbers of required scales of the CW, Morlet and hyperbolic wa... s.

4.3.2 Dimensional Expressions and Band-Peak Frequency

It is assumed that a dimensional sampling interval of the input data series of length M is $(\Delta t')$ and a non-dimensional sampling interval of a wavelet, whose time base interval is from $-T$ to T, is (Δt), where the symbol " ′ " indicates a dimensional quantity [66]. Let N be the number of samples that should be taken for the wavelet function. To calculate the non-dimensional time base of the wavelet function, we have to map the sampling interval of the input waveform to that of the wavelet, i.e. $[-T, T] \leftrightarrow [0, N(\Delta t')]$. The wavelet time base is therefore given by

$$t = \frac{2T}{N(\Delta t')} \cdot t' \tag{4.3.7}$$

The expression for the non-dimensional frequency f is obtained by taking the inverse of Eq. (4.3.7) yielding

$$f = \frac{N(\Delta t')}{2T} \cdot f' \text{ or } \omega = \frac{\pi \cdot N(\Delta t')}{T} \cdot \omega' \tag{4.3.8}$$

The wavelet functions of the CW, Morlet and hyperbolic in terms of the dimensional quantities are therefore given by Eqs. (4.3.9), (4.3.10) and (4.3.11) respectively

$$\psi_{a,b'}^{CW}(t') = \frac{2}{\sigma\sqrt{a}} \cdot \left(1 - \frac{2}{\sigma} \cdot \left(\frac{2T}{aN(\Delta t')}(t'-b')\right)^2\right) \cdot exp\left[\frac{1}{\sigma} \cdot \left(\frac{2T}{aN(\Delta t')}(t'-b')\right)^2\right] \tag{4.3.9}$$

$$\psi_{a,b'}^{Morlet}(t') = \frac{1}{\sqrt{a}} \cdot exp\left(\frac{j2T\omega'_\psi}{aN(\Delta t')}(t'-b')\right) \cdot exp\left[-\frac{1}{\sigma} \cdot \left(\frac{2T}{aN(\Delta t')}(t'-b')\right)^2\right] \tag{4.3.10}$$

$$\psi_{a,b'}^{Hy}(t')$$
$$= \frac{-n\beta^2}{\sqrt{a}}\left[sech\left(\frac{2T\beta}{aN(\Delta t')}(t'-b')\right)\right]^n \left\{n-(n+1)\left[sech\left(\frac{2T\beta}{aN(\Delta t')}(t'-b')\right)\right]^2\right\} \tag{4.3.11}$$

where n is the order of the hyperbolic kernel and also the hyperbolic wavelet, $n = 1$ corresponds to the first-order hyperbolic wavelet. The dimensional quantity b' is similarly defined by Eqs. (4.3.7) and (4.3.8).

The corresponding dimensional frequency expressions of the CW, Morlet and hyperbolic wavelets are given by Eqs. (4.3.12), (4.3.13) and (4.3.14)

$$\hat{\psi}_{a,b'}^{CW}(f') =$$
$$\sqrt{a\pi\sigma} \cdot exp\left(-\frac{jN\pi f'(\Delta t')b'}{T}\right) \cdot \left(\frac{N\pi f'(\Delta t')}{T}\right)^2 \cdot exp\left[-\frac{\sigma}{4} \cdot \left(\frac{a\pi f' N(\Delta t')}{T}\right)^2\right] \tag{4.3.12}$$

$$\hat{\psi}_{a,b'}^{Morlet}(f') = \sqrt{\frac{a}{2\pi}} \cdot exp\left(-\frac{jN\pi f'(\Delta t')b'}{T}\right) \cdot exp\left[-\frac{1}{\sigma} \cdot \left(\frac{a\pi f' N(\Delta t')}{T} - \omega'_\psi\right)^2\right] \tag{4.3.13}$$

$$\hat{\psi}_{a,b'}^{Hy}(f') = \frac{\pi\sqrt{a}}{\beta} \cdot exp\left(-\frac{jN\pi f'(\Delta t')b'}{T}\right) \cdot \left(\frac{\pi f' N(\Delta t')}{T}\right)^2 \cdot sech\left(\frac{\pi \cdot a \cdot \pi f' N(\Delta t')}{2\beta \cdot T}\right) \tag{4.3.14}$$

where typically, $5.0 \le \omega'_\psi \le 6.0$ rad/s is used to ensure that the constraint imposed by Eq. (4.1.1) is met. Throughout this chapter, $\omega'_\psi = 5.0$ rad/s is employed for the Morlet wavelet.

The band-peak frequency, f'_p, is the frequency at which the wavelet filter has the maximum value. To estimate the band-peak frequency, the first derivatives of the real parts of the dimensional frequency expressions of the wavelets should be firstly obtained. Since the real parts of the first-derivative functions are exponential functions, the second-derivative functions are not required. For the Morlet wavelet, to maximise $\hat{\psi}_{a,b'}^{Morlet}(f')$ (given by Eq. (4.3.13)), the exponent of the exponential term is made to be zero which yields [66]

$$f'_{p(Morlet)} = \frac{T_{Morlet}\omega_\psi}{\pi a N_{Morlet}(\Delta t')} = \frac{\omega'_\psi}{a C_{Morlet}}, \text{ where } C_{Morlet} = N_{Morlet}\pi(\Delta t')/T_{Morlet}. \quad (4.3.15)$$

The band-peak frequency of the CW wavelet is obtained by equating the first derivative of Eq. (4.3.12) to zero which yields

$$f'_{p(CW)} = \frac{2}{aC_{CW}\sqrt{\sigma}} = \frac{2T_{CW}}{aN_{CW}\pi(\Delta t')\sqrt{\sigma}} \quad (4.3.16)$$

where σ is the kernel control parameter of the CW kernel.

Varying σ yields different values of the band-peak frequency and affects the effectiveness in cross-term suppression [28] of the kernel as was shown in Chapter 3. For the hyperbolic and CW wavelets, there is a point at which the frequency expressions of the wavelets attain the minimum value, which is at the origin. For other non-zero values of the frequency, the band-peak frequency of the hyperbolic wavelet can be similarly obtained

$$f'_{p(Hy)} \approx \frac{4\beta}{a\pi C_{Hy}} = \frac{4\beta T_{Hy}}{aN_{Hy}\pi^2(\Delta t')}, \text{ where } C_{Hy} = \frac{N_{Hy}\pi(\Delta t')}{T_{Hy}} \quad (4.3.17)$$

Eq. (4.3.17) is an approximate expression of the band-peak frequency of the hyperbolic wavelet, $f'_{p(Hy)}$. The assumption that has been used to obtain it is that C_{Hy} is a small number since N is typically not very large compared to the wavelet time base interval T_{Hy} and $(\Delta t')$ is very small. It should be noted that N is the number of sampling points for the mother wavelet which is of the same order of magnitude of T.

4.3.3 Aliasing Effects

In this section, the number of sampling points of the hyperbolic, CW and Morlet wavelets are calculated.

To avoid aliasing effects in sampling the wavelet non-dimensionally and in sampling the input time series dimensionally, the Nyquist criterion must be satisfied. The Nyquist frequency of the input time series with the sampling interval $(\Delta t')$ can be given by

$$f'_{Ny} = \frac{1}{2(\Delta t')} \tag{4.3.18}$$

where $(\Delta t')$ is the dimensional sampling interval of the input series.

To avoid aliasing in the mother wavelet itself, the overlapping fraction α of two adjacent wavelet filters at different scales must be prescribed so that it is less than a threshold value. This fraction can be defined as an absolute value of the ratio of the wavelet at the frequency $f'_{overlapp}$ at which α is sufficiently small to the magnitude of the wavelet at the band-peak frequency f'_p (Eq. (4.3.19)). At the time that two adjacent wavelet filters overlap, to recover the input signal and to avoid aliasing of the wavelet filters, the overlapping frequency must be at least equal to the Nyquist frequency f'_{Ny}, i.e. $f'_{overlapp}$ = f'_{Ny}. The mathematical expression of the ratio α is therefore given by Eq. (4.3.19) and graphical representation of α is seen in Figure 4.3.1.

$$\alpha = \left| \frac{\hat{\psi}_{a=1,b'}(f'_{Ny})}{\hat{\psi}_{a=1,b'}(f'_p)} \right| \tag{4.3.19}$$

If α is known beforehand, then it is possible to estimate the number of sampling points for the wavelet function. Jordan, Miksad and Powers [66] calculated the required number of sampling points N_{Morlet} for the Morlet wavelet for a typical case of $\sigma = 2$. The number of sampling points of the Morlet wavelet function N_{Morlet} for a general value of σ is given by

$$N_{Morlet} = \frac{2T_{Morlet}}{\pi} \cdot \left(\omega_\psi + \sqrt{-\sigma \ln \alpha}\right), \text{ where } \omega_\psi = 5.0 \text{ rad/s and } \alpha \leq 1. \tag{4.3.20}$$

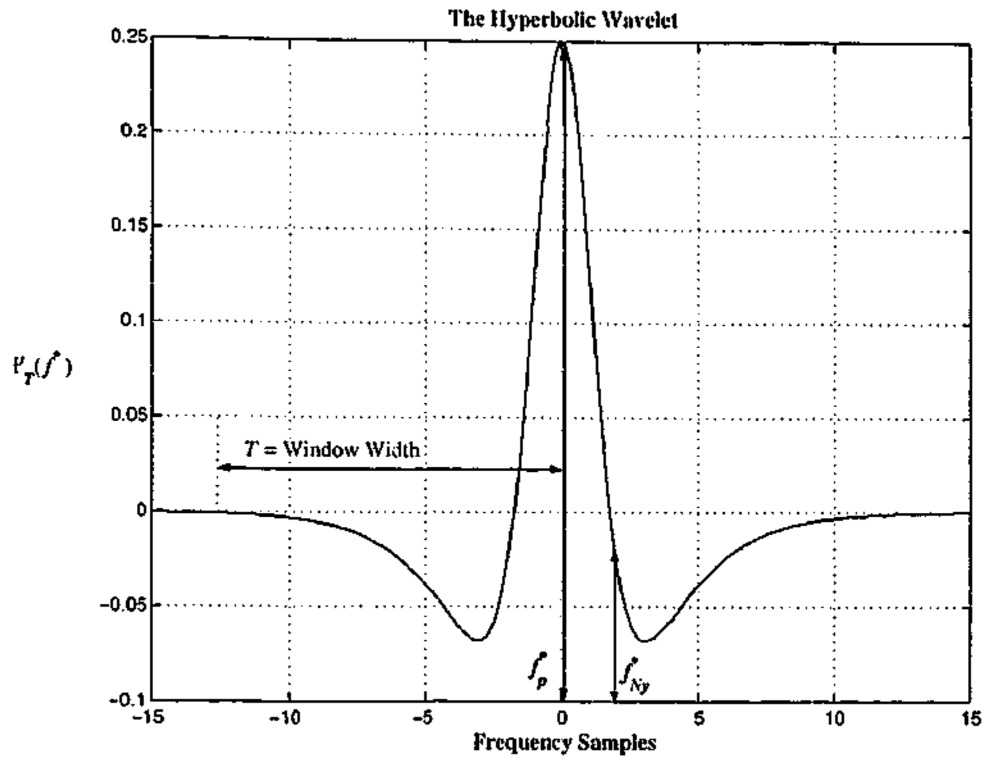

Figure 4.3.1: Aliasing effect of the hyperbolic wavelet for $\beta = 0.5$, f'_P and f'_{Ny} are dimensional band-peak and overlapped frequencies respectively. The ratio α (Eq. (4.3.19)), which can be viewed as the ratio of f_p^* and f_{Ny}^* in the graph, should be small so that aliasing problems can be effectively avoided.

The following calculations attempt to estimate the minimum number of sampling points for the CW and hyperbolic wavelets. For these wavelets, the estimation process is more difficult and the solutions are found by employing a graphical method. In using this method, functions on the left- and right-and sides of an equation are plotted on one co-ordinate system and the intersection(s) of the graphs of two functions are approximately the roots of the equation. In this case, the next odd value of the number of sampling points is chosen since only $N - 1$ points are employed to sample the wavelets as will be seen in subsequent sections.

By using Eq. (4.3.18) for f'_{Ny}, Eq. (4.3.17) for f'_p and Eq. (4.3.14) for the expression of $\hat{\psi}^{Hy}_{a=1,b'}(f')$, we obtain the minimum number of sampling data points N_{Hy} for the hyperbolic wavelet. If $N_{Hy} \geq 16$, then N_{Hy} can be approximated by

$$N_{Hy} \approx \frac{4\beta T_{Hy}}{a\pi^2} \cdot \ln\left(\frac{11.46a^2}{\alpha\beta^2 T_{Hy}^2}\right) \tag{4.3.21}$$

where $a = 1$ for the mother wavelet and it is assumed that $N_{Hy} \geq 16$.

If $N_{Hy} < 16$, then the number of sampling points N_{Hy} is found by the graphical method by plotting the graphs of two functions f_1 and f_2 given by the following equation

$f_1 = f_2$ or

$$\ln\left[\frac{\alpha\beta^2 T_{Hy}^2}{5.73a^2} \cdot \left(exp\left(N_{Hy} \cdot \frac{a\pi^2}{2\beta T_{Hy}}\right) + 1\right)\right] = \ln(2) + N_{Hy} \cdot \frac{a\pi^2}{4\beta T_{Hy}} \tag{4.3.22}$$

Eq. (4.3.22) yields a good estimate of N_{Hy} and is therefore used throughout this chapter. Eq. (4.3.21) provides the approximate value of N_{Hy} only in the case of $N_{Hy} \geq 16$. Similarly, by using Eq. (4.3.12) for the expression of $\hat{\psi}_{a=1,b'}^{CW}(f')$, Eq. (4.3.18) for f'_{Ny}, Eq. (4.3.17) for f'_p and after some mathematical manipulations, the sufficient number of sampling points N_{CW} for the CW wavelet is given by

$$0.617\sigma \cdot \left(\frac{N_{CW}}{T_{CW}}\right)^2 = \ln\left[\frac{1.67\sigma}{\alpha} \cdot \left(\frac{N_{CW}}{T_{CW}}\right)^2\right] \tag{4.3.23}$$

The approximate minimum number of sampling points for the Morlet, hyperbolic and CW wavelets have been estimated and are given by Eqs. (4.3.20), (4.3.22) and (4.3.23) respectively. Eqs. (4.3.22) and (4.3.23) give the approximate values of N_{Hy} and N_{CW} by employing the graphical method. Eq. (4.3.20) yields the exact expression of N_{Morlet}. The maximum scale that can be used for each wavelet is examined in the next section.

4.3.4 Scale Limit

Scales, in wavelet theory context, are inversely proportional to frequencies in the frequency domain. For each wavelet function, there exists the maximum number scale number that the wavelet function can display. The larger the scale limit, the better the wavelet in terms of representing broad-spectrum signals.

The maximum scale number used for a wavelet is determined based on the number of wrapped-around points or end-points of the input time series since these points do not provide useful information. It has been observed that the number of end points is proportional to the scale a [66]. That means, if the scale increases to a certain value, the number of end points will dominate the estimated wavelet transform coefficients and not much information can be gained about the signal if the scale increases further.

From [66], the number of wrap-around points at one end is a function of the scale a and is approximately given by

$$N_{wrap}(a) \approx \frac{a(N-1)}{2} \tag{4.3.24}$$

To estimate the maximum scale number of a wavelet, let us introduce η as a fraction of the number of wrap-around points N_{wrap} and $M = 2^m$, the number of input data points into the wavelet, then we obtain

$$\frac{a_{\max}(N-1)/2}{2^m} \le \frac{\eta}{2} \text{ and } \eta = \frac{N_{wrap}}{M}, \text{ where } N_{wrap} \text{ is given in Eq. (4.3.24);} \tag{4.3.25}$$

where N is the number of sampling points of the wavelets. For the Morlet, hyperbolic and CW wavelets, their numbers of sampling points are given by Eqs. (4.3.20), (4.3.22) and (4.3.23) respectively.

To speed up the calculation process of the wavelet transform coefficients, M should be a power of 2. The fraction $\eta = \frac{1}{3}$, which corresponds to about 30% of the wavelet transform coefficients being overlapped by the wrap around points, was used by Jordan and Miksad [66] for the estimation of the maximum scale number. In this chapter, we leave η to be arbitrary so that the general expression of $a_{\max}$ can be obtained. For both ends and from Eq. (4.3.25), the upper limit of the maximum scale number $a_{\max}$ is given by

$$a_{\max} \le \frac{2^m \cdot \eta}{N-1} \tag{4.3.26}$$

The number of input sampling points M can be estimated from the maximum scale number $a_{\max}$ using Eq. (4.3.27)

$$M = 2^m \geq \frac{(N-1)\cdot a_{\max}}{\eta} \text{ or } m \geq 1.443\cdot\ln\left(\frac{(N-1)\cdot a_{\max}}{\eta}\right) \tag{4.3.27}$$

where η is a ratio of the number of wrap around points at the maximum scale number to the total number of points in the time series.

The next section calculates the scale resolution of the three wavelet functions.

4.3.5 Scale Resolution

The scale resolution constant ω_d is defined as the distance between two band-peak frequencies of two adjacent wavelet filters [66]. The finer the scale resolution ω_d is, the smaller the resolution constant. The distance between two adjacent band-peak frequencies can be determined by specifying a variable λ which has the mathematical form given by Eq. (4.3.28)

$$\lambda = \frac{\hat{\psi}(a\omega'_p)}{\hat{\psi}(a\omega'_p + \omega_d)} \tag{4.3.28}$$

where ω'_p is the dimensional peak frequency, ω_d is the scale resolution constant and $\hat{\psi}(\omega)$ is the frequency expression of the wavelet function given by Eqs. (4.2.3), (4.2.4) and (4.2.8) for the Morlet, CW and hyperbolic wavelets respectively. In most practical wavelets, the scale resolution constant must be small to capture rapid changes in the energy density of the input waveform, which is usually non-stationary such as turbulence and chaos [66], ECG [69], music signal [70-72] or random processes [73]. It is important to note that in Eq. (4.3.28), the frequency quantities are non-dimensional, thus appropriate conversion of the variables must be used to obtain the correct answer.

As the scale a increases, the scale resolution constant will decrease since the frequency of a wavelet is inversely proportional to the scale [51, 52]. If j is the index of an instantaneous scale that is going to be used, then we have the following relationship

$$\omega'_{p(j+1)} - \omega'_{p(j)} = -\frac{\omega_d}{a_j} \tag{4.3.29}$$

where a_j is the j^{th} scale of the wavelet and $\omega'_{p(j+1)}$ is the band-peak frequency at the $(j + 1)^{th}$ scale.

The scale resolution of the Morlet wavelet can be analytically solved. For the CW and hyperbolic wavelets, the scale resolutions are approximately obtained by eliminating the third- and higher-power terms in the time series of $\ln(1 - x)$ [45], where x is substituted as the scale resolution ω_d. The main reason that the third-power terms are ignored is that the scale resolution ω_d is expected to be less than 1. In addition, for these two particular wavelets, the third-power constants are quite small, thus, they can be safely ignored without making large differences in value of the final answer. For the Morlet wavelet function, the exact scale resolution constant ω_d^{Morlet} is found to be

$$\omega_d^{Morlet} = \sqrt{-\sigma \cdot \ln\lambda}\text{, where } \lambda \le 1 \text{ and } \ln\lambda \le 0. \tag{4.3.30}$$

The approximate scale resolution constant of the hyperbolic wavelet ω_d^{Hy} is given by Eq. (4.3.31)

$$\omega_d^{Hy} \approx \frac{4\beta \cdot \sqrt{-\ln\lambda}}{\pi} = \frac{4\sqrt{-\ln\lambda}}{\pi\sigma} \tag{4.3.31}$$

The approximate scale resolution constant of the CW wavelet ω_d^{CW} is given by

$$\omega_d^{CW} \approx \sqrt{-\frac{2\cdot\ln\lambda}{\sigma}} = \sqrt{-2\beta\cdot\ln\lambda} \tag{4.3.32}$$

Since the scale resolution is always less than unity, the applicable ranges of β and σ can be worked out using the chosen value of λ. From Eqs. (4.3.30)–(4.3.32), by making the scale resolutions less than unity, the ranges of β are obtained as $\beta \geq 0.1$, $\beta \leq 2.4$ and $\beta \leq 4.7$. Thus, the applicable range of β is $0.1 \leq \beta \leq 2.4$, i.e. $10 \geq \sigma \geq 0.42$, with $\lambda = 0.9$. For other values of λ, different ranges of β and σ can be obtained. The closer λ is to unity, the larger the value of β and the wider the range of β becomes. When $\lambda = 0.99$, from Eqs. (4.3.30)–(4.3.32), the ranges of β are $\beta \geq 0.01$, $\beta \leq 7.8$ and $\beta \leq 49$ which results in the final range of β to be $0.01 \leq \beta \leq 7.8$, i.e. $100 \geq \sigma \geq 0.13$. It should be noted that the closer the value of ω_d to unity, the finer is the resolution of a wavelet.

From Eqs. (4.3.30)–(4.3.32), it is evident that the scale resolutions ω_d of the three wavelets are independent of the sampling interval $(\Delta t)'$ which makes them unique. The above equations are analytically obtained or with practical approximations. The following table lists values of the scale resolution of the three wavelets for $\beta = 0.5, 1, 2, 2.4$ and λ is chosen to be 0.9.

Table 4.3.3: The approximate scale resolution constants ω_d of the hyperbolic wavelet (Eq. (4.3.31)), Morlet wavelet (Eq. (4.3.30)) and CW wavelet (Eq. (4.3.32)) for $\beta = 0.5, 1, 2, 2.4$ and $\lambda = 0.9$

β	$\omega_d^{Hyperbolic}$	ω_d^{Morlet}	ω_d^{CW}
0.5	0.20	0.46	0.33
1	0.41	0.33	0.46
2	0.83	0.23	0.65
2.4	1.0	0.21	0.71

From Table 4.3.3 and for the range of the hyperbolic control parameter $2.4 \geq \beta \geq 0.1$, the hyperbolic wavelet appears to have a finer scale resolution constant compared with those of the Morlet and CW wavelets. When $\beta = 2$, the scale resolution $\omega_d^{Hyperbolic} = 0.83$ which approaches unity. From a kernel point of view, for $20 \geq \beta \geq 0.5$, the hyperbolic kernel outperforms the CW kernel in terms of cross-term suppression and noise robustness but being outperformed by the CW kernel with respect to auto-term resolution as was shown in Chapter 3. This is a trade-off of achieving more effective cross-term suppression at the expense of having a poorer auto-term resolution. In this chapter, from a wavelet point of view, it has been shown that for $2.4 \geq \beta \geq 0.1$ and $\lambda = 0.9$, the hyperbolic wavelet outperforms the CW and Morlet wavelets by having a finer scale resolution constant ω_d. For other values of λ, the range of β will be changed which yields different values of the scale resolution. In the next chapter, it will be shown that another trade-off in signal detection performance between the hyperbolic and CW kernels exists. The following table lists all values of λ from 0.9 to 0.99 (increasing in steps of 0.01) and the approximate corresponding applicable range of β.

Table 4.3.4: The approximate corresponding applicable range of β for λ varies from 0.9 to 0.99 in increasing steps of 0.01

λ	Minimum value of β (for the Morlet wavelet, Eq.(4.3.30))	Maximum value of β (for the hyperbolic wavelet, Eq. (4.3.31))	Maximum value of β (for the CW wavelet, Eq. (4.3.32))	Approximate applicable range of β
0.9	0.10	2.4	4.7	$0.1 \le \beta \le 2.4$
0.91	0.09	2.5	5.3	$0.09 \le \beta \le 2.5$
0.92	0.08	2.7	6.0	$0.08 \le \beta \le 2.7$
0.93	0.07	2.9	6.9	$0.07 \le \beta \le 2.9$
0.94	0.06	3.1	8.0	$0.06 \le \beta \le 3.1$
0.95	0.05	3.5	9.7	$0.05 \le \beta \le 3.5$
0.96	0.04	3.9	12.2	$0.04 \le \beta \le 3.9$
0.97	0.03	4.5	16.4	$0.03 \le \beta \le 4.5$
0.98	0.02	5.5	24.7	$0.02 \le \beta \le 5.5$
0.99	0.01	7.8	49.7	$0.01 \le \beta \le 7.8$

From Table 4.3.4, it is clear that for values of $0.9 \le \lambda \le 0.99$, the widest applicable range of β is $0.01 \le \beta \le 7.8$. From Chapter 3, for effective cross-term suppression of the hyperbolic kernel compared with the CW kernel, β needs to be in the range of $\beta \ge 1.45$ (Section 3.5.2 for a sum of two complex exponential signals) and $0.5 \le \beta \le 20$ (Section 3.5.3 for a sum of two chirp signals). For larger normalised auto-term magnitude (Figure 3.5.9), the range of β is $\beta \le 100$ and to improve noise robustness, the range of β is $\beta \ge 3$ (Figure 3.7.1). Thus the applicable range of β of $0.01 \le \beta \le 7.8$ (for the hyperbolic wavelet having a fine scale resolution* obtained in this chapter) lies well within the above ranges of β for the effectiveness of the hyperbolic kernel.

* The right-hand side of all ranges of β in Table 4.3.4 is from the maximum allowable value of β of the hyperbolic wavelet. This yields the hyperbolic scale resolution a unity value and of other kernels values of less than unity.

It should be noted that, for all values of β, the hyperbolic normalised auto-term resolution is always less than that of the CW kernel (Figure 3.6.1) which is a trade-off as explained earlier in Chapter 3. After taking into consideration all of the above factors, the most effective range of β is $0.5 \le \beta \le 20$ for satisfactory performance in time-frequency power spectrum and simultaneously having a fine wavelet scale resolution. Thus, another trade-off among cross-term suppression, auto-term magnitude, noise robustness and scale resolution constant against auto-term resolution has been established.

To obtain the total number of scales that can be utilised in a wavelet (provided that the scale resolution constant is known), it is convenient to take the first band-peak frequency to be the reference frequency. The subsequent band-peak frequencies are obtained by dividing the reference band-peak frequency by the scale that corresponds to the particular band-peak frequency, i.e. $\omega_{(p)j} = \omega_{(p)1}/a_j$. Using this relation and Eq. (4.3.29) one can obtain [66]

$$\frac{\omega_{(p)1}}{a_{j+1}} - \frac{\omega_{(p)1}}{a_j} = -\frac{\omega_d}{a_j} \tag{4.3.33}$$

The minus sign on the right-hand side of Eq. (4.3.33) is employed to ensure that the total number of scales j_{max} is a positive number (Eq. (4.3.36) without affecting the correctness of the equation.

The recursive relationship for the scale a is then given by

$$a_{j+1} = \left(\frac{\omega_{(p)1}}{\omega_{(p)1} + \omega_d}\right) a_j = \kappa\, a_j, \text{ where } \kappa = \frac{\omega_{(p)1}}{\omega_{(p)1} - \omega_d}. \tag{4.3.34}$$

The first band-peak frequencies (which corresponds to $a = 1$ for the mother wavelet) of the Morlet, CW and hyperbolic wavelets can be estimated by using Eqs. (4.3.15)–(4.3.17) respectively, which are given in Section 4.3.2.

From Eq. (4.3.34), it is evident that the present scales are dependent on the previous scales. This relationship can be understood via the constant κ, which is a function of the peak frequency of the first-scale wavelet (mother wavelet) and the scale resolution constant ω_d. As the scale a becomes larger, the width of the corresponding wavelet becomes smaller since it is inversely proportional to the scale a as can be seen in Eq. (4.2.10). Assuming that

$a_1 = 1$, i.e. choosing $j = 1$ as the starting point, Eq. (4.3.34) can be rewritten to find the total number of scales of a wavelet under certain conditions [66]

$$a_j = \kappa^{j-1}, \text{ where } \kappa \text{ was defined by Eq. (4.3.34).} \tag{4.3.35}$$

From Eqs. (4.3.34) and (4.3.35), one can obtain an expression for the required total number of scales j_{max}

$$j_{max} = \left(\ln(a_{max}) \big/ \ln\kappa \right) + 1 \tag{4.3.36}$$

By using the maximum value of a_{max} given by Eq. (4.3.26), the required total number of scales j_{max} for a wavelet can be obtained. For each wavelet, the number of sampling points of the mother wavelet is different and so are the band-peak frequency, scale resolution, a_{max} and the total number of required scales. To gain more practical insight into the three wavelets, the following section calculates their parameters which have been discussed in Sections 4.3.1–4.3.5. Wavelet power spectra of the English speech vowel "e" signal are also given to demonstrate compression ability of the hyperbolic wavelet.

4.3.6 Parameter calculations and wavelet power spectra of a speech signal

One practical example was used in [66] in which the transition to turbulence in a subsonic wake was investigated using the Morlet wavelet transform. Major conclusions about the behaviour of the subsonic wake were made in [66] and will not be repeated here. This section compares the Morlet, CW and hyperbolic wavelets by calculating the following parameters: band-peak frequency, maximum scales, aliasing, scale resolution and the total number of scales used in this particular application. For these wavelets, the value of $\beta = 1/\sigma = 0.5$ is used.

The sampling interval of the input time series was $(\Delta t') = 0.2$ *ms*. The aliasing parameter is chosen to be $\alpha = 0.01$ (1%) so that only 1% of the mother wavelet is aliased. From Table 4.3.2, Figure 4.3.1 and Figure 4.3.2, the one-sided length of the hyperbolic, CW and Morlet mother wavelets are $T_{Hy} \approx 10$, $T_{CW} \approx 5$ and $T_{Morlet} \approx 3$ respectively. The values of the required number of sampling points of the mother wavelets are hence found by direct calculation using values of Ts and Eq. (4.3.20) and by the graphical method using Eqs. (4.3.22) and (4.3.23). From Eqs. (4.3.22) and (4.3.23), the approximate number of sampling points of the hyperbolic and CW wavelets are given as $N_{Hy} = 9$ and $N_{CW} = 13$ respectively.

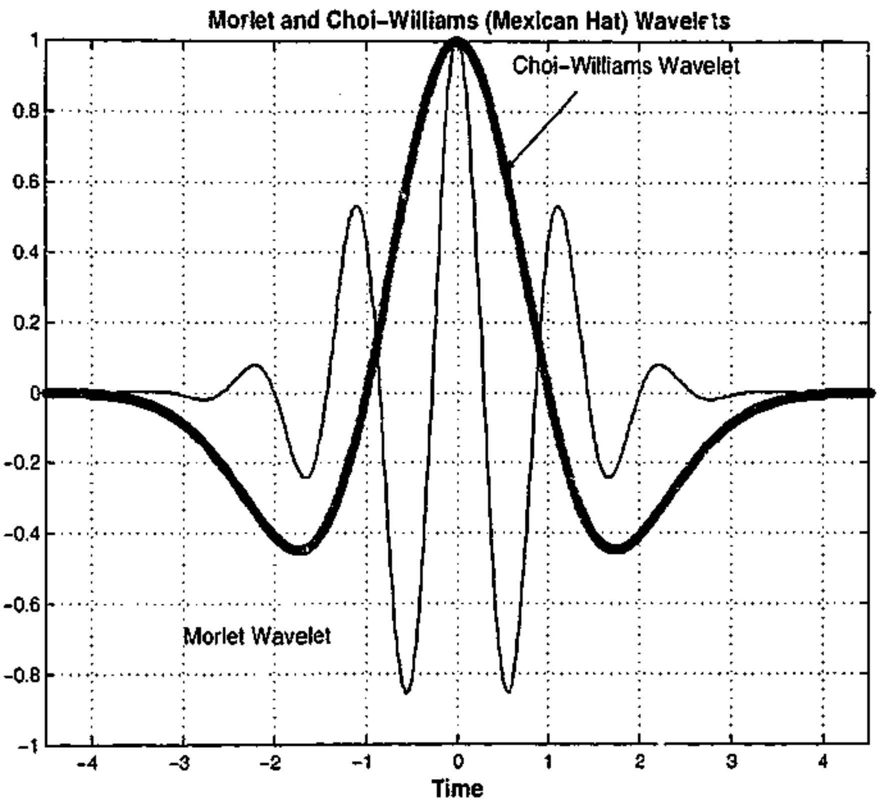

Figure 4.3.2: Time base interval T_{CW} and T_{Morlet} of the CW and Morlet wavelets for $\sigma = 2$, i.e. $\beta = 0.5$. T_{Morlet} and T_{CW} are used to estimate N_{Morlet} and N_{CW} by the graphical method using Eqs. (4.3.20) and (4.3.23) respectively.

The band-peak frequencies are given by

$$f'_{Hy} = \frac{1125.8}{a},\ f'_{CW} = \frac{865.7}{a} \text{ and } f'_{Morlet} = \frac{1452}{a}. \tag{4.3.37}$$

Further, the band-peak frequency can be scaled down to about 30 Hz, which allows us to estimate the maximum scale number a_{max}. The maximum scale numbers of each wavelet are given by

$$a_{max}^{Hy} = 37.53,\ a_{max}^{CW} = 28.8 \text{ and } a_{max}^{Morlet} = 48.4. \tag{4.3.38}$$

From Eqs. (4.3.27) and (4.3.38), the required numbers of data points for each wavelet with $\eta = \frac{1}{3}$ are given by

$$M_{Hy} = 1024, M_{CW} = 2048 \text{ and } M_{Morlet} = 4096. \tag{4.3.39}$$

It should be noted that the number of input sampling points could be varied by changing the value of η to provide satisfactory solutions to a particular application or problem. However, the value of η should be kept small so that aliasing can be effectively avoided. For $\beta = 0.5$, i.e. $\sigma = 2$, the scale resolution of each wavelet is estimated next using Eqs. (4.3.30)–(4.3.32) which yields

$$\omega_d^{Hy} = 0.2066,\ \omega_d^{CW} = 0.3246 \text{ and } \omega_d^{Morlet} = 0.459. \tag{4.3.40}$$

The dimensional peak frequencies of the Morlet, hyperbolic and CW mother wavelets are $\omega'^{1}_{p(Mrolet)} = 5.0$ rad/s, $\omega'^{1}_{p(Hy)} = 0.6366$ rad/s and $\omega'^{1}_{p(CW)} = 1.4142$ rad/s respectively. The total number of scales that can be computed is directly proportional to the scale resolution ω_d. By employing Eq. (4.3.36), the total number of scales of each wavelet can be approximately worked out as $j_{max(Hy)} \approx 11$, $j_{max(CW)} \approx 14$ and $j_{max(Morlet)} \approx 42$.

The Morlet wavelet, as expected, has the largest number of computed scales since it has the coarsest scale resolution as calculated by Eq. (4.3.40) compared with the hyperbolic and CW wavelets. Table 4.3.5 summarises values of important parameters and highlights the significant values of the hyperbolic, CW and Morlet wavelets that have been estimated in this section.

Table 4.3.5: Summary of important parameters of the hyperbolic, CW and Morlet wavelets for the case of $\beta = 1/\sigma = 0.5$. The shaded cells indicate important parameters of the wavelets.

Wavelet	One sided-length T	Wavelet sampling points N	Maximum scale number a_{max}	Scale resolution ω_d	Total number of scales j_{max}
Morlet	3	17	49	0.459	42
Choi-Williams	5	13	29	0.3246	14
Hyperbolic	10	9	38	0.2066	11

From Table 4.3.5, the hyperbolic wavelet appears to have the finest scale resolution compared with the Morlet and CW wavelets. However, the total number of scales of the hyperbolic wavelet is smaller than those of the CW and Morlet wavelets which suggests that it is not suitable for signals that have energy distributed over a wide frequency range. With regard to this aspect, the Morlet wavelet can be considered most suitable for wide-range frequency signals compared to the hyperbolic and CW wavelets. The hyperbolic wavelet is most suitable for transient signals, which do not have a wide frequency range to resolve. Clearly, the most appropriate wavelet depends on the application and the nature of the problem. It should be noted that the hyperbolic and CW wavelets do not have a wide scale range and their total numbers of scales are in the same order of magnitude. This might suggest that the hyperbolic wavelet is more advantageous than the CW wavelet since the former has a finer scale resolution.

Since the wavelet power spectrum is going to be used later in this section, it is appropriate to define the wavelet transform and wavelet power spectrum at this point. The wavelet transform $WT(a, b)$ of a function $x(t)$ is given by [51, 52]

$$WT(a,b) = \int_{-\infty}^{+\infty} x(t) \cdot \psi\left(\frac{t-b}{a}\right) dt \tag{4.3.41}$$

where $\psi\left(\frac{t-b}{a}\right)$ is the mother wavelet, a and b are the scale and time indices respectively.

The wavelet power spectrum of $x(t)$, as analogous to its Fourier power spectrum counterpart $P(\omega) = \hat{X}(\omega) \cdot \hat{X}^*(\omega) = \left|\hat{X}(\omega)\right|^2$, is given as

$$WPS(t,\omega) = WT(t,\omega) \cdot WT^*(t,\omega) = \left|WT(t,\omega)\right|^2 \tag{4.3.42}$$

Eqs. (4.3.41) and (4.3.42) form the background for plotting the Morlet, CW and hyperbolic wavelet power spectra of the English vowel "e". The wavelet power spectrum, in particular the hyperbolic wavelet power spectrum, will be investigated in detail in Chapter 6 in which detailed comparisons between the Fourier power spectrum and wavelet power spectrum are made.

It should also be noted that the scale resolution constant ω_H is independent of the external dimensional parameter $(\Delta t)'$, thus each wavelet has its own scale resolution constant which is determined by the nature of the mother wavelet. For this particular example, the hyperbolic wavelet has the finest scale resolution for $\beta = 0.5$ compared with the Morlet and CW wavelets. Depending on the application, the appropriate values of β can be chosen to yield the most suitable wavelet.

To demonstrate the effects of having a small total number of required scales, the Morlet, CW and hyperbolic wavelets are used to examine a speech signal of the English vowel "e" in which their wavelet power spectra are displayed in Figure 4.3.3 and Figure 4.3.4 respectively. It should be noted that the contour scale is not quantitatively included in these graphs as it is not the main emphasis of this chapter.

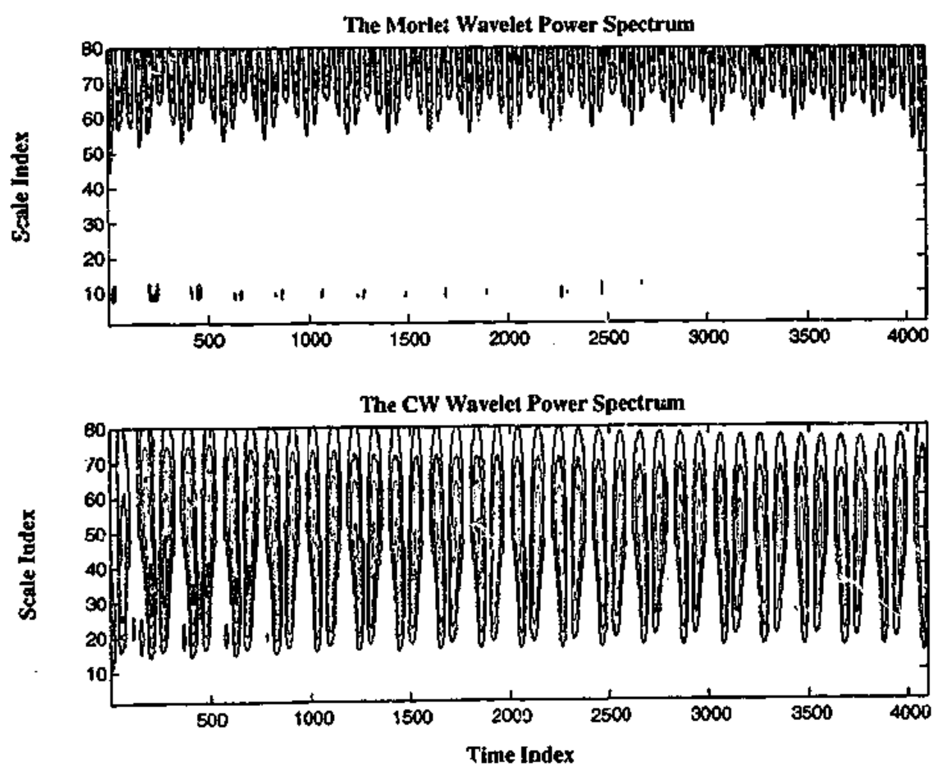

Figure 4.3.3: Contour plots of the Morlet and CW wavelet power spectra of the English vowel "e" speech signal

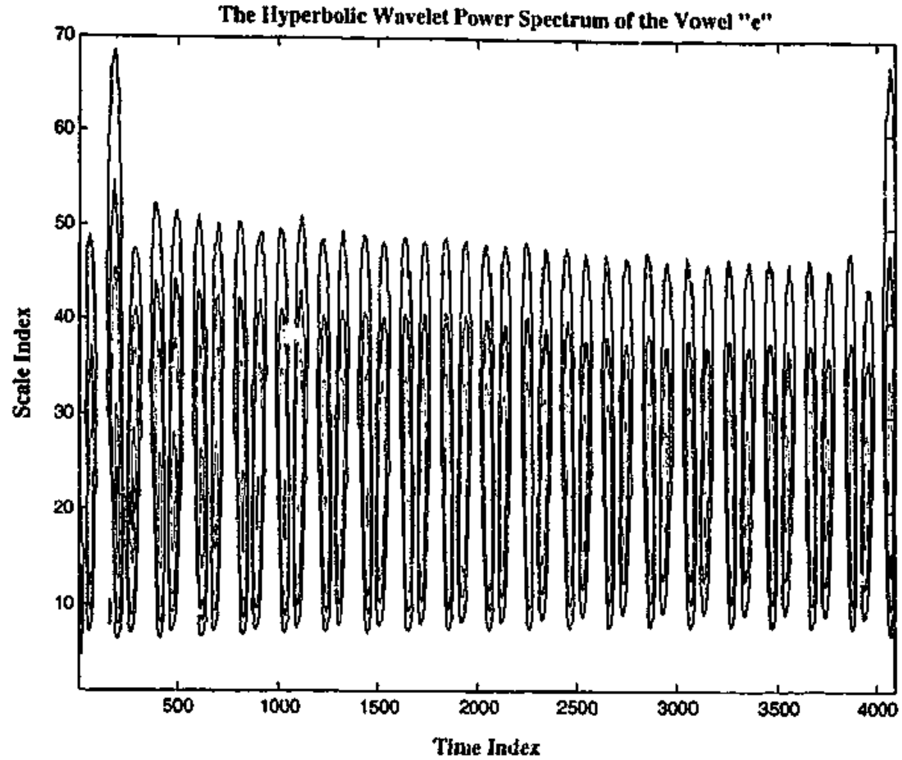

Figure 4.3.4: Contour plot of the hyperbolic wavelet power spectrum of the English vowel "e" speech signal

As can be seen, the Morlet wavelet does not reveal energy components in the approximate scales of 20 to 45, whereas the hyperbolic and CW wavelets do. Moreover, the hyperbolic wavelet can display components at very high frequencies which correspond to scales smaller than 10. The CW and hyperbolic wavelet power spectra are similar, except the latter has a finer scale resolution by clearly showing all harmonics and sub-harmonics as can be seen from Figure 4.3.3 and Figure 4.3.4. This suggests that the hyperbolic wavelet power spectrum cannot be used to examine signals that have broad power spectra as stated earlier in this chapter.

The main advantage of the hyperbolic wavelet over the Morlet and CW wavelets is that it has a smaller total number of required scales which considerably reduces the calculation time. In other words, with the same number of scales of 70 (Figure 4.3.4), the wavelet power spectrum of the vowel "e" speech signal can be successfully shown using the hyperbolic wavelet. Whilst, the CW and Morlet wavelets cannot display the wavelet power spectrum of the input signal within the above scale range due to having a coarser scale resolution and a larger total number of scales. This shows that the hyperbolic wavelet power spectrum can be successfully compressed whereas the Morlet and CW wavelet power spectra require a larger number of scales to display the energy distribution of the speech signal. Thus, it can be suggested that the hyperbolic wavelet is more effective and efficient than the Morlet and CW wavelets.

The effectiveness of the hyperbolic wavelet will be examined further by calculating the hyperbolic wavelet power spectra of different signals in Chapters 6 and 7 of the thesis. Table 4.3.6 summarises the advantages and disadvantages of the hyperbolic, Morlet and CW wavelets.

Table 4.3.6: Detailed qualitative comparisons of the Morlet, CW and hyperbolic wavelets

	Morlet Wavelet	CW Wavelet	Hyperbolic Wavelet
Common features	Easy to generate and analyse; symmetrical and all have explicit expressions; un-orthogonal, un-biorthogonal and their scaling functions do *not* exist; classified as "crude" wavelets.		
Aliasing effects	Calculation for the mother wavelet can be analytically done.	The graphical method must be employed to find the number of sampling points for the mother wavelet.	Similar to the CW wavelet
Scale limit a_{max}	High	Low	Higher than CW wavelet's but smaller than Morlet wavelet's
Total number of scales j_{max}	High	Moderate	Low
Scale resolution ω_d	Coarse	Moderate, depending on the control parameter σ	Fine, depending on values of the control parameter $\beta = 1/\sigma$
Suitable input signals	Wide-frequency range signals including chaos, music and speech	In between the Morlet and hyperbolic wavelets, depending upon the particular application	Narrow-frequency range signals including transients, and ECG

In this section, the hyperbolic, CW and Morlet wavelets have been studied in detail from an engineering point of view by comparing some crucial parameters and properties. The next section briefly investigates properties of the hyperbolic wavelet from a mathematical point of view including symmetry, orthogonality, bi-orthogonality, existence of scaling function, *FIR* and vanishing moments. It should be emphasised that these properties have been extensively studied in the past decade by many mathematicians and they are worth investigating.

4.4 Other Properties of the Hyperbolic Wavelet Function

Section 4.3 investigated some important properties of the hyperbolic, CW and Morlet wavelets from an engineering point of view. These properties include aliasing effects, scale limit, scale resolution, total number of required scales and some preliminary parameters of the hyperbolic, CW and Morlet wavelets. In this section, other properties of the hyperbolic wavelet are briefly investigated from a mathematical point of view in the following order

1. Explicit expression — The wavelet function is clearly defined in the time and frequency domains with unique expressions.
2. Symmetry — The wavelet function is symmetrical about the vertical axis. This property is desirable for avoiding de-phasing in image processing [67]. The wavelets investigated in this chapter, i.e. hyperbolic, CW and Morlet wavelets, are symmetrical.
3. Orthogonal and bi-orthogonal analyses — This property and the regularity of a wavelet allow fast algorithmic calculation of the wavelet transform.
4. Compactly supported orthogonal and bi-orthogonal.
5. Existence of the scaling function $\varphi(t)$ — This affects the existence of the *FIR* filtering property
6. Vanishing moments for $\varphi(t)$ — This property is desirable for compression purposes of polynomial power series up to a certain order.
7. *FIR* (Finite Impulse Response) filter — This filter represents the connection between the wavelet function $\psi(t)$ and the scaling function $\varphi(t)$.

The above properties have been discussed and reported in detail in [67]. This section mainly focuses on the hyperbolic wavelet, which has not been previously studied in the literature. Some desirable properties of the wavelet are studied to find out more about the hyperbolic wavelet and to compare it with other wavelets such as CW, Morlet, Daubechies, Mallat, Meyer and Cauchy. Although it might seem that the hyperbolic wavelet has similar properties to those of the CW and Morlet wavelets, it is necessary to examine its properties in some detail so that detailed conclusions on the wavelet can be drawn.

4.4.1 Explicit Expression and Symmetry

The expression of the hyperbolic wavelet function was given in Eq. (4.2.7). It is evident that the hyperbolic wavelet is a symmetrical function and has finite time support. The graphical representations of the hyperbolic and CW wavelet functions are displayed in Figure 4.4.1.

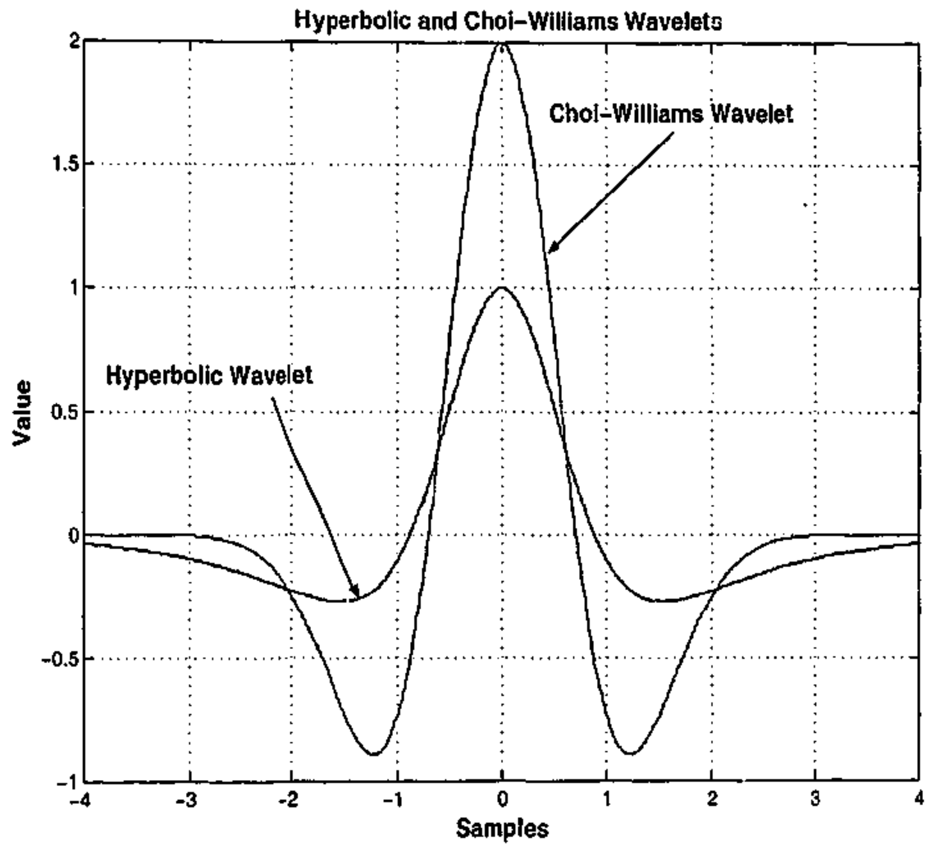

Figure 4.4.1: The CW (Eq. (4.2.1) and hyperbolic (Eq. (4.2.7)) wavelets for $\beta = 2$, i.e. $\sigma = 0.5$

It should be emphasised that wavelet functions that have explicit expressions are difficult to find. For example, Daubachies wavelets are excellent wavelets but their values are only obtained by recursive numerical calculation. Some wavelets that have explicit expressions include the Morlet, Haar (Daubachies order 1) and Mexican-hat or CW. The hyperbolic wavelet function is a continuous and symmetrical function with no singularity in the time domain as seen in Figure 4.4.1.

4.4.2 Orthogonal and Bi-Orthogonal Analyses

Let us define $f_k(t)$ as a set of functions or vectors that span the space F. The set $f_k(t)$ is orthogonal if every function or vector within the set is independent of each other, i.e. each function or vector is unique and no relationship can be established between them [51]. Consider a wavelet function as a set of wavelet functions. If the wavelet is orthogonal, the following condition must be met [51, 52]

$$\int_{-\infty}^{+\infty} \psi_m(t)\cdot\psi_l(t)dt = 0\text{, for all } m \neq l \tag{4.4.1}$$

Eq. (4.4.1) represents the area under the curve of the inner product of two wavelet functions $\psi_m(t)$ and $\psi_l(t)$. If this area is zero, the wavelet is said to be orthogonal. For the hyperbolic wavelet, we have to examine the following inner product of two hyperbolic wavelet functions

$$\int_{-\infty}^{+\infty} m^2[sech(m\tau)]\{-[sech(m\tau)]^2+[tanh(m\tau)]^2 \tag{4.4.2}$$

$$\cdot\, l^2[sech(l\tau)]\cdot\{-[sech(l\tau)]^2+[tanh(l\tau)]^2\}\, d\tau$$

for all $m \neq l$.

To examine the final value of Eq. (4.4.2), firstly, some familiar values of m and l are used. If Eq. (4.4.2) appears to satisfy the condition of orthogonality stated in Eq. (4.4.1), then the general case of m and l will be examined. For simplicity, by putting $m = 1$ and $l = 2$, we obtain

$$\int_{-\infty}^{+\infty} [sech(\tau)]\cdot\{-[sech(\tau)]^2+[tanh(\tau)]^2\} \tag{4.4.3}$$

$$\cdot\, 2^2\cdot[sech(2\tau)]\cdot\{-[sech(2\tau)]^2+[tanh(2\tau)]^2\}\, d\tau$$

The inner product of Eq. (4.4.3) is plotted against τ in Figure 4.4.2. It is evident that the area under the curve is non-zero (the area under the curve is ≈ 0.48 units) for $m = 1$ and $l = 2$, which means the hyperbolic wavelet is not orthogonal.

Since only symmetrical wavelets are considered in this thesis, bi-orthogonality characteristics are not important for these wavelets. However, bi-orthogonality is still briefly discussed for completeness.

If a wavelet (expansion set), $\psi_m(t)$, is not orthogonal within itself, then if there exists a dual set $\psi_l'(t)$, which is orthogonal to the expansion set $\psi_m(t)$, then the wavelet is said to be bi-orthogonal since it requires two sets of vectors or functions to make it orthogonal. Mathematically, the definition of bi-orthogonality can be given by [51]

$$\int_{-\infty}^{+\infty} \psi_m(t) \cdot \psi_l'(t)dt = \delta(m-l) \tag{4.4.4}$$

It is evident from Eq. (4.4.4) that the hyperbolic wavelet is not bi-orthogonal since it has a finite area under the curve as seen in Figure 4.4.2.

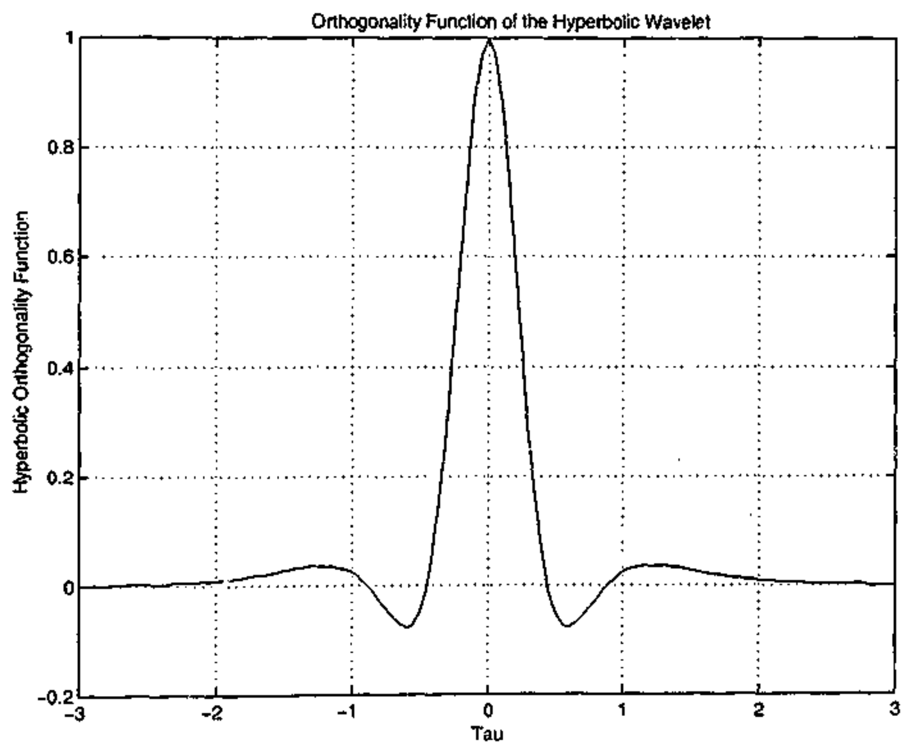

Figure 4.4.2: The orthogonal inner product function of the hyperbolic wavelet (Eq. (4.4.1)). The area under the curve is evidently not zero which means that the hyperbolic wavelet is not orthogonal or bi-orthogonal.

4.4.3 Compactly Supported Orthogonal and Bi-Orthogonal Analyses

If the scaling function of a wavelet does not exist, the wavelet is said to be un-orthogonal since it requires both the wavelet and scaling functions to be orthogonal or bi-orthogonal. From Section 4.4.2, it has been shown that the hyperbolic wavelet is not orthogonal and bi-orthogonal, thus it is not compactly supported orthogonal and bi-orthogonal. These properties are closely related to orthogonality and bi-orthogonality investigated in Section 4.4.2 and they were briefly mentioned here for completeness.

4.4.4 An Arbitrary Number of Vanishing Moments

In the context of wavelet analysis, the moment function of order a, M_a, is defined by Eq. (4.4.5)

$$M_a = \int_{-\infty}^{+\infty} t^a \cdot \psi(t)\, dt \tag{4.4.5}$$

where a is in the range of 0 to $(a_0 - 1)$ and a_0 is the maximum moment number.

If $M_a = 0$ for $a \geq a_{max}$, then the moment function is said to vanish at and beyond the order a_{max}. A wavelet that has vanishing moments of order a_{max} can suppress polynomial signals of the same order [67], i.e. polynomial signals of order a_{max} will have zero wavelet coefficients as seen in Eq. (4.4.5).

4.4.5 Existence of the Scaling Function $\varphi(t)$

The existence of the scaling function $\varphi(t)$ is sometime not easy to determine. The scaling and translating relationship of the wavelet function $\psi(t)$ and scaling function $\varphi(t)$ can be stated as [51]

$$\psi(t) = 2^{j/2} \cdot \psi(2^j t - k) \text{ and } \varphi(t) = 2^{j/2} \cdot \varphi(2^j t - k) \tag{4.4.6}$$

where k and j are the time translation and scale indices respectively.

The scaling function $\varphi(t)$ and the wavelet function $\psi(t)$ both can be recursively expanded. This relationship is called the dilation or multi-resolution relationship

$$\varphi(t) = \sum_{n=0}^{N-1} h(n) \cdot \sqrt{2} \cdot \varphi(2t - n) \text{ and } \psi(t) = \sum_{n=0}^{N-1} g(n) \cdot \sqrt{2} \cdot \psi(2t - n) \tag{4.4.7}$$

where N is the order of the FIR filter $h(n)$.

The scaling function is related to the wavelet function $\psi(t)$ by the following relation [51]

$$\psi(t) = \sum_{n=0}^{N-1} h_1(n) \cdot \sqrt{2} \cdot \varphi(2t - n) \tag{4.4.8}$$

where $h_1(n)$ is a set of finite coefficients and can be related to $h(n)$ by

$$h_1(n) = (-1)^n \cdot h(1 - n) \tag{4.4.9}$$

where $h(n)$ is the scaling filter of the scaling function $\varphi(t)$.

If the wavelet function is known beforehand such as in the case of the hyperbolic wavelet function, then the expression for the FIR filter $h(n)$ can be found. From Eq. (4.4.7), the first dilation equation for the scaling function $\varphi(t)$ may or may not exist which determines the status of $\varphi(t)$. If the dilation equation does not exist (which is the case for the Morlet and CW wavelets), the FIR filter $h(n)$ cannot be found.

The existence of the scaling function $\varphi(t)$ is determined by the existence of its scaling filter $h(n)$. If $h(n)$ does not exist, the scaling function $\varphi(t)$ cannot recursively expand itself (Eqs. (4.4.6) and (4.4.7)) through the given space.

4.4.6 FIR Filter

The finite impulse response (FIR) filter is the link between the wavelet function $\psi(t)$ and its scaling function $\varphi(t)$. If the scaling function does not exist, then the FIR filter will not exist as concluded in Section 4.4.5.

4.5 Conclusion

The new hyperbolic wavelet has been generated from the new time-frequency hyperbolic kernel, proposed in Chapter 3. The important link between time-frequency kernels and wavelets has been established. The hyperbolic, Choi-Williams and Morlet wavelets have been compared in terms of aliasing effects, scale resolution, total number of required scales and of maximum required-scale number. It can be concluded that the first-order hyperbolic wavelet is not orthogonal, nor bi-orthogonal, its scaling function does not exist and therefore no *FIR* filter can be generated for this wavelet. Instead, the hyperbolic wavelet is symmetrical about the vertical axis and can be identified with a unique explicit expression. The hyperbolic wavelet is a member of the "crude" wavelet group, which includes the Morlet and CW wavelets. It has been shown that the hyperbolic wavelet has a finer scale resolution but it has a smaller total number of required scales compared to those of the CW and Morlet wavelets. In other words, by having a smaller scale limit, the hyperbolic wavelet power spectrum can be calculated more efficiently than the CW and Morlet wavelets.

The first contribution of this chapter is to propose the new hyperbolic wavelet function which is generated from the hyperbolic kernel, proposed in the previous chapter. Detailed comparisons of the hyperbolic wavelet with the CW and Morlet wavelets are carried out in this chapter from an engineering point of view. The second contribution of the chapter is to establish an important link between wavelet functions and time-frequency kernels so that a new wavelet can be found if a new kernel exists.

The next chapter presents the first application of time-frequency power spectrum analysis in detecting non-stationary signals using the hyperbolic, CW, Wigner-Ville and *cross-correlator* signal detectors. Chapters 6 and 7 calculate the hyperbolic wavelet power spectra of signals including ECG, music and speech.

Chapter 5: SIGNAL DETECTION USING NON-UNITY KERNEL TIME-FREQUENCY DISTRIBUTIONS

This chapter examines the first application of the hyperbolic kernel in detecting non-stationary and stationary signals in the presence of noise. Moyal's formula for non-unity kernel time-frequency detectors is derived based on Moyal's formula for a unity-kernel time-frequency signal detector (Wigner-Ville detector). Performance comparisons of the hyperbolic detector, Choi-Williams detector, general non-unity kernel detector, Wigner-Ville detector and matched-filter correlator detector in terms of signal-to-noise ratio (*SNR*) are made. The second application of the hyperbolic wavelet function on wavelet power spectrum analysis is presented in Chapter 6 and Chapter 7.

5.1 Introduction

Detection of known and deterministic signals in the presence of noise is a classical problem which has been extensively studied in the literature [68, 74]. To solve this problem, the signals and the additive noise are assumed to be stationary or wide-sense stationary and zero-mean processes. A matched-filter technique has been shown to be the most effective method to detect signals in this case. However, if the signal is non-stationary, i.e. its power spectrum varies with time, or the signal is not known beforehand, then the classical method using the matched-filter technique is limited. Non-stationary signals include radar, sonar, chaotic, ECG, speech, music signals and image matching [75, 76]. For such non-stationary signals, time-frequency signal detectors need to be employed so that the signals can be effectively detected.

One typical time-frequency detector is the Wigner-Ville unity-kernel detector which can be used to solve a simple binary detection problem [31, 32]. There are two reasons that the Wigner-Ville time-frequency detector is popular. First, the Wigner-Ville distribution is simple and easy to implement and it provides perfect frequency concentration in the time-frequency plane [77]. Second, originally used in quantum mechanics [78], Moyal's formula, which is required for calculation of the *SNR*, is readily available for the Wigner-Ville distribution. Noise sources, which are assumed to be complex, wide-sense stationary, can be of two common types that are usually encountered in practice, namely white and coloured noise. Using the Wigner-Ville unity kernel detector, detection of non-stationary signals in white noise was done by Flandrin [77] and in coloured noise by Marinovich [75]. Both researchers used a method to detect signals by estimating a statistical function η which is then compared with a threshold value [18, 24, 31, 32, 77, 79]. If η is greater than the threshold, then the signal is said to be present; otherwise, the signal is not present.

The non-unity kernel time-frequency signal detectors form a class of detectors of which the Wigner-Ville unity-kernel signal detector is a special case. This class of detectors employs Cohen's time-frequency distributions with different kernel functions. Each kernel corresponds to a unique distribution and hence to a unique signal detector. The kernel function strongly influences the performance of the detector in terms of *SNR* and the higher the *SNR*, the better the performance of the signal detector. The simplest non-unity kernel of the Cohen time-frequency class is the Rihaczek kernel, $\Phi_{Rihaczek}(\theta, \tau) = e^{j\theta\tau/2}$. The Choi-Williams kernel signal detector (*CWWD*) can be considered as the most useful and popular detector due to the effectiveness of the Choi-Williams kernel in suppressing cross terms and its robustness in noisy conditions. A different class of signal detectors is the bilinear signal detectors in which the non-stationary structure of the signal is exploited to ensure the best match of the signal to the detector filter [76]. Another class is the quadratic class of time-frequency power spectrum, called the hyperbolic class, which was first proposed by Papandreou and Bartels [80]. Signal detection using this particular class is examined in [26, 81, 82] using the same method of estimating the statistical function η.

Non-unity kernel signal detectors have been studied in some detail in the literature, in particular, detectors using the Rihaczek and Choi-Williams kernels. A comparison of the Wigner-Ville and Rihaczek distributions has been done in [77] in which the Wigner-Ville distribution was found to be more suitable than the Rihaczek distribution in terms of signal detection and preservation of the inner product of Moyal's formula. The Wigner-Ville detector was compared with the Choi-Williams detector [24] for the case of the Doppler target-return signal using the same method presented in [77]. In [24], the reverberation ratio SRR was estimated instead of the SNR due to the specific requirements of calculating the target return.

Although the Choi-Williams time-frequency distribution has been used to detect the Doppler signal, the statistical function η of the general non-unity kernel signal detector ($GNKD$) has not been derived. It should also be noted that to estimate the SNR of a time-frequency detector, Moyal's formula for the corresponding time-frequency distribution of the detector is required. While Moyal's formula has been derived for the case of the Wigner-Ville time-frequency distribution of unity kernel only, this formula has not been derived for a non-unity kernel time-frequency distribution. We derive this formula for non-unity kernel time-frequency distributions and then apply it to the statistical function η to calculate the SNR of the detector. Thus, deriving Moyal's formula for a non-unity kernel time-frequency distribution is an important step before any performance calculation of a non-unity kernel time-frequency detector is carried out. Furthermore, using Moyal's formula is the only method currently available to estimate the SNR of time-frequency detectors. If Moyal's formula for a particular class of time-frequency detectors, i.e. Moyal's formula for the corresponding time-frequency distributions, does not exist, then it is not possible to estimate performance of detectors employing these distributions.

This chapter aims to achieve three goals. First, to derive the prerequisite Moyal formula for non-unity kernel detectors. This formula can be used for any non-unity kernel detector if a new kernel function and hence its corresponding time-frequency distribution are available. Secondly, the hyperbolic detector (*HyD*) and Choi-Williams detector (*CWWD*) are compared so that the effectiveness of the hyperbolic kernel in comparison with the Choi-Williams kernel can be clearly identified. Thirdly, the ability of non-unity kernel detectors in detecting practical signals such as ECG, music and speech is examined in detail.

The chapter is organised as follows. Section 5.2 briefly defines the binary signal detection problem and outlines the general expression of the *SNR*. Section 5.3.2 derives Moyal's formula in detail for the non-unity kernel time-frequency distribution. In Section 5.3.3, the general detailed expression of the *SNR* of the *GNKD* is given using Moyal's formula. The relative performance of the *HyD* and *CWWD* is compared by using the geometrical features of the hyperbolic and Choi-Williams weighting functions. Section 5.4 calculates the *SNR* by using Moyal's formula from Section 5.3.2 and compares the loss factor Q of three signal detectors, namely *GNKD*, *CWD* and correlated signal detector (*CORR*). Values of an energy ratio X_1, which plays an important role in determining the performance of a signal detector, are estimated in Section 5.4.4 for a number of signals including a sinusoid $sin(t)$, an exponential transient $exp(-t)$, an exponentially decaying sinusoid $sin(t)\cdot exp(-t)$, a chirped signal $cos(Ct^2)$, the ECG and speech signals including all the English vowels and the "sh"-sound signals.

5.2 The Binary Detection Problem

The binary detection problem can be understood as a problem of determining the presence of a non-stationary signal $s(t)$ in the presence of a stationary, white, zero mean and complex noise $w(t)$, given the received noisy signal $f(t)$. The signal energy and the noise variance are assumed to be A_0 and N_0 respectively. These parameters are used in this chapter to estimate a detector performance by estimating its *SNR*.

Since the signal is non-stationary, the classical method employed for stationary signals cannot be used. Instead, time-frequency signal detectors have to be employed to detect the presence of non-stationary and unknown signals which are corrupted by channel noise and other noise sources. It is assumed that it is not possible to separate frequency power spectra of the wanted signal and the noisy received signal $f(t)$ and also that the signal is completely masked by the noise $w(t)$. The two hypotheses for detecting the signal that need to be considered are given in Eq. (5.2.1)

$$H_0 : f(t) = w(t) \text{ and } H_1 : f(t) = s(t) + w(t) \tag{5.2.1}$$

in which H_0 means that the signal $s(t)$ is not present and H_1 means the signal is present. The reference signal $s(t)$ is assumed to be unknown, non-stationary and could be of random type.

The hypotheses are then examined and the main goal is to decide which one of them is likely to hold. This is done by forming a statistics η using the received noisy signal $f(t)$ and the reference signal $s(t)$. The hypotheses are then decided by comparing the statistics η with a threshold value. If η is greater than the threshold, the signal is said to be present. Otherwise, the signal is not present [75-77]. The performance of a particular statistics η is determined by estimating its *SNR*. The *SNR* of the statistical function η for random variables, which is equivalent to the likelihood ratio, is given by [31, 32]

$$SNR = \frac{\left| E\{\eta|_{H_1}\} - E\{\eta|_{H_0}\} \right|}{\sqrt{\frac{1}{2}\left(Var\{\eta|_{H_1}\} + Var\{\eta|_{H_0}\}\right)}} \tag{5.2.2}$$

where $E(\cdot)$ and $Var(\cdot)$ denote the expectation and variance operations respectively on the statistical function η under the hypotheses H_0 and H_1. The *SNR* of the matched filter or *cross-correlator* detector (*CORR*) can be found by using the general formula (Eq. (5.2.2)) which will be shown in detail in Section 5.4.1. The next section derives Moyal's formula for the general non-unity kernel time-frequency distributions based on Moyal's formula for the unity-kernel Wigner-Ville time-frequency distribution.

5.3 Derivation of the Discrete Moyal Formula for a General Time-Frequency Distribution

5.3.1 The Discrete Moyal Formula for the Wigner-Ville Time-Frequency Distribution

To successfully estimate the *SNR* of a time-frequency detector, Moyal's formula of a particular time-frequency distribution must be known. The discrete Moyal formula for the Wigner-Ville time-frequency distribution has been derived by Moyal and forms the basis for deriving Moyal's formula for a general time-frequency distribution which is vital in estimating its *SNR*.

The general time-frequency distribution is denoted as $TFR(\omega, t)$ in continuous form or $TFR(m, n)$ in discrete form with ω and t the frequency and time variables respectively, and m, n the discrete frequency and time variables respectively. From Chapter 3, the general time-frequency distribution [6, 8] is given in Eq. (5.3.1) as

$$TFR(t,\omega)=\frac{1}{4\pi^2}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\underbrace{\left[e^{-j\theta(t-u)}\cdot\Phi(\theta,\tau)\right]}_{F(t-u,\tau)}\cdot e^{-j\tau\omega}\cdot R_{t,1}(t,\tau)\,du\,d\tau\,d\theta \tag{5.3.1}$$

where $F(t-u, \tau)$ is the weighting function which is the 1-D Fourier transform of the kernel function $\Phi(\theta, \tau)$, $u=t+\frac{\tau}{2}$ and the local auto-correlation function $R_{t,1}(t,\tau)=x\left(u+\frac{\tau}{2}\right)\cdot x^*\left(u-\frac{\tau}{2}\right)$. The Choi-Williams kernel is given by $\Phi_{CW}(\theta, \tau) = e^{-\theta^2\tau^2/\sigma}$ [28] and the first-order hyperbolic kernel is given by $\Phi_{Hy}(\theta, \tau) = [sech(\beta\theta\tau)]^n$, where $n = 1$. Two other kernels that can be used for signal detection are the third-order hyperbolic kernel $\Phi_{CubHy}(\theta, \tau) = [sech(\beta\theta\tau)]^n$ (where $n = 3$) and the Choi-Williams-Butterworth (CWB) kernel $\Phi_{CWB}(\theta, \tau) = \frac{e^{-\theta^2\tau^2/\sigma}}{\theta^2\tau^2+1}$. It should be noted that the $\Phi_{CWB}(\theta, \tau)$ kernel satisfies the admissibility constraints discussed in Chapter 3 [8, 13, 14] and has not been reported in the literature. The CWB and cubic hyperbolic kernels are briefly mentioned in this chapter but they will not be investigated further in the research presented in this thesis.

The weighting functions F_{CW}, F_{Hy}, F_{CubHy} and F_{CWB} of the Choi-Williams, hyperbolic, cubic hyperbolic and Choi-Williams-Butterworth kernels are given by Eqs. (5.3.2)–(5.3.5) respectively

$$F_{CW} = \frac{\sigma\sqrt{\pi}}{\tau} exp\left(-\frac{\sigma(t-u)^2}{4\tau^2}\right) \tag{5.3.2}$$

$$F_{Hy} = \frac{\pi}{\beta\tau} sech\left(\frac{\pi(t-u)}{2\beta\tau}\right) \tag{5.3.3}$$

$$F_{CubHy} = \frac{\pi(\beta^2\tau^2+u^2)}{2\beta^3\tau^3} sech\left(\frac{\pi(t-u)}{2\beta\tau}\right) \tag{5.3.4}$$

$$F_{CWB} = \frac{\pi \cdot exp\left(1/\sigma\right)}{2\tau} \cdot \left\{exp\left(-\frac{t-u}{\tau}\right) \cdot Erfc\left(\frac{1}{\sigma} - \sqrt{\sigma} \cdot \frac{t-u}{2\tau}\right) + exp\left(\frac{t-u}{\tau}\right) \cdot Erfc\left(\frac{1}{\sigma} + \sqrt{\sigma} \cdot \frac{t-u}{2\tau}\right)\right\} \tag{5.3.5}$$

The continuous Wigner-Ville time-frequency distribution is given by substituting $\Phi(\theta, \tau) = 1$, which is a unity kernel [6, 8, 13-15], into Eq. (5.3.1) to obtain

$$W(t,\omega) = \int x\left(t + \tfrac{\tau}{2}\right) \cdot x^*\left(t - \tfrac{\tau}{2}\right) \cdot e^{-j\omega\tau} d\tau \tag{5.3.6}$$

where the range of integration is from $-\infty$ to $+\infty$ for the rest of the chapter unless otherwise stated.

In discrete form, the Wigner-Ville distribution of two signals $f(\cdot)$ and $s(\cdot)$ is given by Eq. (5.3.7)

$$W_{fs}(n,m) = 2 \cdot \sum_{\tau_k=-(M/2-|n|)}^{M/2-|n|} \left[f(n+\tau_k) \cdot s^*(n-\tau_k) \cdot exp(-j2\pi km/(M+1)) \right] \tag{5.3.7}$$

where τ_k is the lag parameter and M is the number of data samples. More detailed background on the Wigner-Ville time-frequency distribution can be found in [13-15].

The continuous Moyal formula for the Wigner-Ville time-frequency distribution derived by Moyal in 1949 [78] is given by Eq. (5.3.8) [18, 31, 32, 83]

$$\beta_{WV} = \int_\omega \int_t W_{fg}(t,\omega) \cdot W_{hs}^*(t,\omega) dt d\omega = \left\{ \int f(t) \cdot g^*(t) dt \right\} \cdot \left\{ \int h^*(t) \cdot s(t) dt \right\} \tag{5.3.8}$$

which is a product of two energy terms of four functions $f(t)$, $g(t)$, $h(t)$ and $s(t)$, i.e. the inner product has been reserved for the Wigner-Ville time-frequency distribution [77]. As will be seen later, the discrete Moyal formula for the non-unity kernel time-frequency distribution is complicated with the involvement of the odd and even samples of the signal in the time and frequency domains.

The discrete Moyal formula for the Wigner-Ville distribution is given by Eq. (5.3.9)

$$\beta_{WV} = \sum_{n=-M/2}^{M/2} \sum_{m=0}^{M-1} W_{fg}(n,m) \cdot W_{hs}^*(n,m) \tag{5.3.9}$$

The derivation of the discrete Moyal formula for the Wigner-Ville time-frequency distribution (Eq. (5.3.10)) is given in detail in [31] and is given here as

$$\begin{aligned}\beta_{WV} &= 2M \cdot \left[\sum_{u=-M/2}^{M/2} f(u) \cdot h^*(u)\right] \cdot \left[\sum_{v=-M/2}^{M/2} g^*(v) \cdot s(v)\right] \\ &+ 2M \cdot \left[\sum_{u=-M/2}^{M/2} (-1)^u f(u) \cdot h^*(u)\right] \cdot \left[\sum_{v=-M/2}^{M/2} (-1)^v g^*(v) \cdot s(v)\right]\end{aligned} \tag{5.3.10}$$

To apply the discrete Moyal formula to find the *SNR* of the *GNKD*, the following identities are applied to Eq. (5.3.10): $g(\cdot) \equiv h(\cdot) \equiv s(\cdot)$. The following section derives the discrete Moyal formula for the *GNKD*.

5.3.2 Derivation of the Discrete Moyal Formula for the General Non-Unity Kernel Signal Detector

The discrete Moyal formula for the Wigner-Ville distribution was given in the previous section. This section extends Moyal's formula for the general Cohen non-unity kernel time-frequency distribution. Given the reference signal $s(t)$ with energy A_0 and the white noise, zero mean process $w(t)$ of variance N_0, the problem we have to solve is to determine the existence of the reference signal in noisy conditions. The signal energy A_0 and the absolute energy difference B_0 between the even and odd samples of the signal $s(t)$ are defined by

$$A_0 = \sum_{k=-M/2}^{M/2} |s(k)|^2 \quad \text{and} \quad B_0 = \left| \sum_{k_{even}=-M/2}^{M/2} |s(k_{even})|^2 - \sum_{k_{odd}=-M/2}^{M/2} |s(k_{odd})|^2 \right| \tag{5.3.11}$$

The energy and energy difference of the noise $w(t)$ are similarly defined by Eq. (5.3.12)

$$N_0 = \sum_{k=-M/2}^{M/2} |w(k)|^2 \quad \text{and} \quad M_0 = \left| \sum_{k_{even}=-M/2}^{M/2} |w(k_{even})|^2 - \sum_{k_{odd}=-M/2}^{M/2} |w(k_{odd})|^2 \right| \tag{5.3.12}$$

From Eqs. (5.3.11) and (5.3.12), the dimensionless energy ratios of the signal $s(t)$ and the noise $w(t)$ are defined as $X_1 = \frac{B_0}{A_0}$ and $X_2 = \frac{M_0}{N_0}$ respectively. It is evident that the ratios X_1 and X_2 are positive and less than unity since $A_0 \geq B_0 \geq 0$ and $N_0 \geq M_0 \geq 0$. Generally, values of B_0 could be in the range of $-A_0 \leq B_0 \leq A_0$, however, in this chapter, only the positive half of B_0 is considered due to its usefulness and convenience in practical situations. The same convention is applied to M_0 of the noise. The physical meanings of X_1 and X_2 will be discussed in detail in Sections 5.3.3 and 5.4.3.

The discrete form of the general time-frequency distribution is also given by Eq. (5.3.9) but with $W(n, m)$ replaced by $TFR(n, m)$ as shown in Eq. (5.3.13)

$$\beta_{GNKD} = \sum_{n=-M/2}^{M/2} \sum_{m=0}^{M-1} TFR_{fg}(n,m) \cdot TFR^*_{hs}(n,m) \tag{5.3.13}$$

The discrete form of the general time-frequency distribution with a non-unity kernel is given by

$$TFR(n,m) = 2\sum_{\tau=-L}^{L} \sum_{u=-M/2}^{M/2} f(u+\tau) \cdot g^*(u-\tau) \cdot F(n-u,\tau) \cdot exp(-j2\pi m\tau/M) \tag{5.3.14}$$

where $L = \frac{M}{2} - |n|$, $F(n-u, \tau)$ is the 1-D Fourier transform of the kernel functions $\Phi(\theta, \tau)$, M is the length of the input discrete signal, m and n are the discrete time and frequency variables respectively.

The discrete Moyal formula for a non-unity kernel distribution is obtained by taking a product of two discrete $TFR(n, m)$'s and is given in Eq. (5.3.15)

$$\beta_{GNKD} = 4\sum_{n=-M/2}^{M/2} \left[\sum_{m=0}^{M-1} exp(j2\pi \cdot (\tau_l - \tau_k) \cdot m/M) \right] \cdot$$
$$\left\{ \left[\sum_{\tau_k=-L}^{L} \sum_{u_k=-M/2}^{M/2} F \cdot f(u_k+\tau_k) \cdot g^*(u_k-\tau_k) \right] \cdot \left[\sum_{\tau_l=-L}^{L} \sum_{u_l=-M/2}^{M/2} F \cdot h^*(u_l+\tau_l) \cdot s(u_l-\tau_l) \right] \right\} \tag{5.3.15}$$

where F is the weighting function of the kernel.

The summation with respect to m in Eq. (5.3.15) can be replaced by $M\cdot\delta(\tau_l - \tau_k)$ [31, 84] which results in $\tau_l = \tau_k = \tau$ so that the impulse function exists. After putting $p_k = u_k + \tau$ and $q_k = u_k - \tau$ and similarly $p_l = u_l + \tau$ and $q_l = u_l - \tau$, Eq. (5.3.15) can be rewritten as

$$\beta_{GNKD} = 4M \sum_{n=-M/2}^{M/2} \{A \cdot B\}$$

$$= 4M \sum_{n=-M/2}^{M/2} \left\{ \left[\sum_{q_k=-|n|}^{|n|} \sum_{p_k=-M+|n|}^{M-|n|} F_{p_k} \cdot f(p_k) \cdot g^*(q_k) \right] \cdot \left[\sum_{q_l=|n|}^{|n|+M} \sum_{p_l=-M+|n|}^{M+|n|} F_{p_l} \cdot h^*(p_l) \cdot s(q_l) \right] \right\}$$

(5.3.16)

where A and B correspond appropriately to the square-bracketed terms in Eq. (5.3.16).

From the above, we also obtain $P_k = p_k + q_k = 2u_k$ and $P_l = p_l + q_l = 2u_l$ which are even numbers. Thus, to allow the summation over the specified range given in Eq. (5.3.16), the factors $\frac{1}{2}\left(1 + (-1)^{p_k+q_k}\right)$ and $\frac{1}{2}\left(1 + (-1)^{p_l+q_l}\right)$ are inserted into the expressions A and B in Eq. (5.3.16) respectively without affecting the correctness of the expressions since the inserted factors are unity in value. After multiplying, separating and rearranging the variables appropriately, we obtain

$$\beta_{GNKD} = 4M \sum_{n=-M/2}^{M/2} [A \cdot B]$$

$$= 4M \sum_{n=-M/2}^{M/2} \left[\sum_{q_k=-|n|}^{|n|} g^*(q_k) \sum_{p_k} F_{p_k} \cdot f(p_k) + \sum_{q_k=-|n|}^{|n|} g^*(q_k) \cdot (-1)^{q_k} \sum_{p_k} F_{p_k} \cdot f(p_k) \cdot (-1)^{p_k} \right] \cdot$$
$$\left[\sum_{q_l=-|n|}^{|n|} h^*(q_l) \sum_{p_l} F_{p_l} \cdot s(p_l) + \sum_{q_l=-|n|}^{|n|} h^*(q_l) \cdot (-1)^{q_l} \sum_{p_l} F_{p_l} \cdot s(p_l) \cdot (-1)^{p_l} \right]$$

(5.3.17)

where A and B correspond to the square-bracketed items and the running ranges of p_k and q_l are similar to those given in Eq. (5.3.16). Eq. (5.3.17) is the final form of the discrete Moyal formula for the general time-frequency power spectrum with a non-unity kernel. The next section gives the calculations of the SNR of the statistics of the hypotheses H_0 and H_1 for the non-unity kernel general case by using Eq. (5.3.17).

5.3.3 *SNR* Calculation of the General Non-Unity Kernel Detector and Performance Comparison of Different Non-Unity Kernel Detectors

Having obtained Moyal's formula for the general time-frequency distribution, detailed derivation of the *SNR* of the *GNKD* can be made by employing Eq. (5.2.2). The mean and variance of the statistical function η are given by

$E\{\eta_{TFR}|_{H_0}\} = 0$ since $E\{w(t)\} = 0$

$E\{\eta_{TFR}|_{H_1}\} = 2MC_C$, where M is the length of the input data samples and

$$C_C = \left(\sum_{q_k} s^*(q_k)\sum_{p_k} F_k \cdot s(p_k)\right)\cdot\left(\sum_{q_l} s(q_l)\sum_{p_l} F_l \cdot s^*(p_l)\right)$$
$$+\left(\sum_{q_k} s^*(q_k)\sum_{p_k} F_k \cdot s(p_k)\right)\cdot\left(\sum_{q_l} s^*(q_l)\cdot(-1)^{q_l}\sum_{p_l} F_l \cdot s(p_l)\cdot(-1)^{p_l}\right) \quad (5.3.18)$$

Under the special conditions $p_k = p_l$, $q_l = q_{l1}$ and $q_k = q_l$, C_C (given by Eq. (5.3.18)) becomes a constant $C_C = C = (A_{0F})^2 + (B_{0F})^2$, where

$A_{0F} = \sum_{q_k} s^*(q_k)F_k \cdot s(q_k)$ and $B_{0F} = \sum_{q_l} s^*(q_l)\cdot F_l \cdot s(q_l)\cdot(-1)^{q_l}$.

Since $\eta_{TFR}|_{H_1} = C + \eta_{TFR}|_{H_0}$, the variance of $\eta_{TFR}|_{H_1}$ is equal to that of $\eta_{TFR}|_{H_0}$. Detailed derivations of the variances of the statistics and its *SNR* are given in Appendix B. The variance of the statistical function $\eta_{TFR}|_{H_1}$ is given by

$$Var\{\eta_{TFR}|_{H_0}\} = Var\{\eta_{TFR}|_{H_1}\} = E\{[\eta_{TFR}|_{H_1}]^2\}$$
$$= M^2 \cdot \sum_{n=-M/2}^{M/2} \sum_{j=-M/2}^{M/2} (A_0 \cdot A_{01FF} \cdot N_{0FF} \cdot A_{01}) + (A_0 \cdot B_{01} \cdot N_{0FF} \cdot B_{01FF})$$
$$+ (A_0 \cdot A_{01} \cdot N_{0FF} \cdot A_{01FF}) + (A_0 \cdot N_{0FF} \cdot B_{01} \cdot B_{01FF}) + (2A_0 \cdot N_{0FF} \cdot B_0 \cdot B_{0FF})$$
$$+ (2A_0 \cdot B_0 \cdot N_{0FF} \cdot B_{0FF}) + (2B_0 \cdot M_{0FF} \cdot A_0 \cdot A_{0FF}) + (2B_0 \cdot M_{0FF} \cdot B_0 \cdot B_{0FF})$$
$$+ (2B_0 \cdot M_{0FF} \cdot B_0 \cdot B_{0FF}) + (2B_0 \cdot M_{0FF} \cdot A_0 \cdot A_{0FF})$$
$$= M^2 \cdot \sum_{n=-M/2}^{M/2} \sum_{j=-M/2}^{M/2} 2[(A_0 \cdot A_{01FF} \cdot N_{0FF} \cdot A_{01}) + (A_0 \cdot B_{01} \cdot N_{0FF} \cdot B_{01FF})]$$
$$+ 2[2(A_0 \cdot N_{0FF} \cdot B_0 \cdot B_{0FF}) + 2(B_0 \cdot M_{0FF} \cdot A_0 \cdot A_{0FF}) + 2(B_0 \cdot M_{0FF} \cdot B_0 \cdot B_{0FF})] \tag{5.3.19}$$

where

$$A_{01FF} = \sum_{p_l} F_l \cdot s(p_l) \cdot \sum_{p_{l1}} F_{jl1} \cdot s^*(p_{l1}), \quad A_{0FF} = \sum_{p_l} F_l \cdot s(p_l) \cdot F_{jl} \cdot s^*(p_l),$$
$$B_{0FF} = \sum_{p_l} F_l \cdot s(p_l) F_l \cdot s(p_l) \cdot (-1)^{p_l}, \quad B_{01} = \sum_{q_l} s^*(q_l) \cdot (-1)^{q_l} \cdot \sum_{q_{l1}} s(q_{l1}) \cdot (-1)^{q_{l1}},$$
$$B_{01FF} = \sum_{p_l} F_l \cdot s(p_l) \cdot (-1)^{p_l} \cdot \sum_{p_{l1}} F_{l1} \cdot s(p_{l1}) \cdot (-1)^{p_{l1}}, \quad A_{01} = \sum_{q_l} s^*(q_l) \cdot \sum_{q_{l1}} s(q_{l1}),$$
$$M_{0FF} = \sum_{p_k} F_k \cdot w(p_k) \cdot F_k \cdot w(p_k) \cdot (-1)^{p_k}, \quad N_{0FF} = \sum_{p_k} F_k \cdot w(p_k) \cdot F_k \cdot w(p_k) \tag{5.3.20}$$

The *SNR* of the *GNKD* is given by

$$SNR_{GNKD} = \frac{E\{\eta_{TFR}|_{H_1}\}}{\sqrt{Var\{\eta_{TFR}|_{H_1}\}}} = \frac{2MC}{\sqrt{Var\{\eta_{TFR}|_{H_1}\}}} \tag{5.3.21}$$

The SNR_{GNKD} for the special case of $p_k = p_l$, $q_k = q_l$ and $q_l = q_{l1}$ is given by

$$SNR_{GNKD} = \frac{\sqrt{2} \cdot (A_0^2 + B_0^2)}{A_0 \sqrt{N_0 \cdot A_0 \cdot \left\{1 + \frac{B_0}{A_0} + 2\left[\left(\frac{B_0}{A_0}\right)^2 + \frac{B_0}{A_0} \cdot \frac{M_0}{N_0} + \frac{M_0}{N_0} \cdot \left(\frac{B_0}{A_0}\right)^3\right]\right\}}} \tag{5.3.22}$$

It is worth repeating that $X_1 = \frac{B_0}{A_0}$ and $X_2 = \frac{M_0}{N_0}$, which were defined by Eq. (5.3.11) and Eq. (5.3.12) in Section 5.3.2, are ratios of the absolute energy difference between the even and odd samples of a signal and noise to its total energy of the input signal and noise respectively. The ratio X_1 can be estimated by using simulation at different sampling rates. It will be shown later in Section 5.4.4 that the sampling rate can affect the value of X_1, which in turn will affect the performance of the signal detector.

The physical meanings of the energy ratio X_1 can be understood as the ratio of the bandwidth (the Difference_Energy given by Eq. (5.3.23)) to the total energy of the input signal. As it will be shown later in Figure 5.4.1 to Figure 5.4.3 that the smaller the value of X_1, the higher the signal detector performance in detecting a particular signal. In addition, satisfactory performance can be achieved by having the value of X_1 close to 1.0 provided that X_2 is small (Section 5.4.3). However, the latter scenario is not applicable to situations in which the X_2 ratio of the noise is large. The energy of the input signal can be expressed in terms of the even and odd energy of the input signal

$$\text{Total\_Energy} = \text{Even\_Energy} + \text{Odd\_Energy} = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \left|\hat{F}(\omega)\right|^2 d\omega = \int_{-\infty}^{+\infty} f^2(t)\,dt$$

$$\text{Difference\_Energy} = \left|\text{Even\_Energy} - \text{Odd\_Energy}\right| \sim \text{Signal\_Bandwidth}$$

$$\text{Absolute\_Energy\_Ratio} = \frac{\text{Difference\_Energy}}{\text{Total\_Energy}} \tag{5.3.23}$$

where $\hat{F}(\omega)$ is the Fourier transform of the input signal $f(t)$.

Theoretically, the constant signal ($-\infty \le t \le +\infty$), which according to Eq. (5.3.23) has a zero bandwidth, is most effectively detected since there is no energy difference between the even and odd samples of the signal. The Fourier transform or the energy density of the constant signal is a single impulse $\delta(\omega)$ located at the origin. This impulse is regarded as a perfect way to concentrate the energy in the frequency domain since there is no energy smearing. The bandwidth of a constant signal is zero since there is no "width" in the frequency domain for an impulse. In the case of periodic sinusoidal signals, the Fourier transforms of the functions $sin(\omega_0 t)$ and $cos(\omega_0 t)$ are impulses located at frequencies $\pm\omega_0$. These impulses perfectly concentrate the energy of the input signal in the frequency domain, as a result, their bandwidths are effectively zero. Thus, it can be concluded that signals that have a zero Fourier-frequency-domain bandwidth such as the constant and periodic sinusoid signals are effectively detected by using a time-frequency signal detector. Simulation results in Section 5.4.4 shows that the periodic sinusoid has a zero-valued X_1 which is consistent with the theoretical prediction of Eq. (5.3.23).

The two cases mentioned above validate Eq. (5.3.23) in which the absolute energy difference between the even and odd samples of an input digital signal is directly proportional to its Fourier-frequency-domain bandwidth. Other types of signals including exponential transient $exp(-t)$, chirped $cos(C \cdot t^2)$, exponentially decaying sinusoidal $sin(\omega_0) \cdot exp(-t)$ have non-zero bandwidths which result in a larger energy ratio X_1. Hence, detecting signals having wider bandwidths is more difficult than detecting those with narrower bandwidths. Real signals such as the ECG and speech are effectively detected using a time-frequency signal detector as will be seen in Section 5.4.4.

The fact that wide-band signals are more difficult to detect than narrow-band signals can also be explained by looking at the problem from a filter point of view. If the signal is wide-band, it is more likely to be contaminated by other signals such as noise disturbances that have been sent at the same time in a pass band of a filter. The role of the filter is to extract detected signal(s) in its pass band. If the pass band contains not just the wanted signal but a mixture of two or more signals, then the signal is more difficult to detect.

From Eqs. (5.3.18) and (5.3.19), it is evident that the *SNR* of a signal detector is proportional to the volume under the surface of the weighted signal, i.e. a product of the signal $s(t)$ and the weighting function of the kernel, and inversely proportional to the volume of the weighted-signal variance. Thus, if the volume under the surface of the weighted signal is larger than that of the weighted-signal variance, *SNR* of the corresponding detector is high. Furthermore, it has been found that the hyperbolic kernel is more robust [29, 30] than the Choi-Williams kernel. Thus, the *HyD* provides a smaller noise variance than that of the *CWWD* for well chosen values of $\beta \geq 3$ as can be seen in Figure 5.3.1 in which the volume under the surface of the weighted-signal variance is displayed. As a result, the *SNR* of the *HyD* is better than that of the *CWWD*. Table 5.3.1 gives the volumes under the surface of the weighting functions of four kernels: hyperbolic, cubic hyperbolic (the third power of the hyperbolic kernel), Choi-Williams and Choi-Williams-Butterworth kernel (a product of the Choi-Williams and Butterworth kernels [25]). The volumes under the surface of the weighted variance of the four kernels for some typical values of β are also listed in Table 5.3.2. It is evident from Table 5.3.2 and Figure 5.3.1 that for different values of the control parameter β, a different volume under the surface is obtained. Thus the control parameter of a kernel plays an important role in determining the performance of the corresponding signal detector and of time-frequency kernels as it was shown in Chapter 3.

Table 5.3.1: Volumes under the surface of the Choi-Williams, hyperbolic, cubic hyperbolic and CW-Butterworth (Eqs. (5.3.2)–(5.3.7) respectively) weighting functions

β	Volume under the surface of the weighting function			
	Hyperbolic kernel	Cubic Hyperbolic kernel	Choi-Williams (CW) kernel	CW-Butterworth kernel
0.1	12.08	12.04	12.014	11.91
1	11.997	11.88	11.98	11.78
5	9.9	7.58	11.02	10.78
10	7.39	4.98	9.953	9.787
20	4.884	2.98	8.63	8.53
50	2.44	1.33	6.78	6.74
100	1.3173	0.68	5.47	5.45

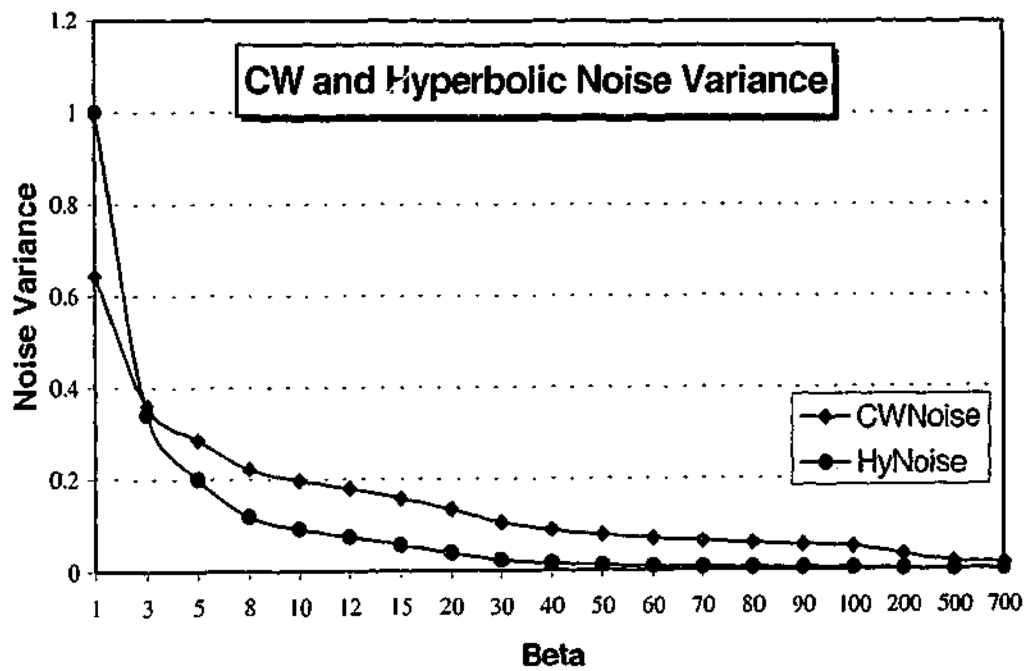

Figure 5.3.1: Noise variance of the *HyD* and *CWWD* (this figure is taken from Chapter 3, Figure 3.7.1)

Table 5.3.2: Volumes under the surface of the weighted variance of the four kernels

β	Volume under the surface of the variance of the weighting function			
	Hyperbolic kernel	Cubic Hyperbolic kernel	Choi-Williams (CW) kernel	CW-Butterworth kernel
0.1	5.25	2.766	1.02	0.36
1	0.51	0.274	0.322	0.234
5	0.1	0.05	0.1425	0.13
10	0.046	0.02	0.0985	0.094
20	0.0195	0.008	0.067	0.065
50	0.0054	0.0018	0.0387	0.0038
100	0.00175	0.000494	0.025	0.025

The volume under the surface of the weighted-signal variance is directly proportional to the variance of a time-frequency signal detector. The smaller this volume, the better is the performance of a particular time-frequency signal detector. The performance of the *GNKD* in terms of *SNR* is dependent on the volume under the surface of the weighted signal and its variance. The loss factor Q of the *GNKD* over the Wigner-Ville unity-kernel signal detector, i.e. a ratio of SNR_{GNKD} to SNR_{CWR}, is given by

$$Q(GNKD/CWD) = \frac{(V_{GNKD})^2}{\sqrt{4(V_{GNKD})^2 \cdot SV_{GNKD} + 12(SV_{GNKD})^2}} \tag{5.3.24}$$

where V_{GNKD} and (SV_{GNKD}) are the volumes under the surface of the weighted signal and its variance respectively. Eq. (5.3.24) can be used to estimate the improvement factor for different non-unity kernel time-frequency signal detectors.

To measure the relative performance of the *HyD* and *CWWD*, their loss factor Q is formed as [32]

$$Q(HyD/CWWD) = \frac{SNR_{HyD}}{SNR_{CWWD}} = \frac{C_{HyD}}{C_{CWWD}} \cdot \sqrt{\frac{Var\{\eta_{CWWD}|_{H_0}\}}{Var\{\eta_{HyD}|_{H_0}\}}} \approx \left(\frac{V_{HyD}}{V_{CWWD}}\right)^2 \cdot \frac{SV_{CWWD}}{SV_{HyD}}$$

(5.3.25)

where V_{HyD} and V_{CWWD} are the volumes under the surface of the weighted signal and SV_{HyD}, SV_{CWWD} are the volumes under the surface of the weighted-signal variance of the hyperbolic and Choi-Williams kernels respectively.

Using the data provided by Table 5.3.1 and Table 5.3.2, the improvement factor Q of the *HyD* and *CWWD* are calculated and given in Table 5.3.3.

Table 5.3.3: Improvement factors Q of the *HyD* and *CWWD*, $3 \le \beta \le 10$

Signal Detector	Improvement Factor Q
HyD	$22.5 \le Q_{HyD} \le 24.8\ dB$
CWWD	$22.554 \le Q_{CWWD} \le 23.9\ dB$

The relative performance of the *HyD* to that of the *CWWD* is graphically displayed in Figure 5.3.2. From Eq. (5.3.25) and using the appropriate values from Table 5.3.1 and Table 5.3.2, the *CWWD* is more effective than the *HyD* by 63% for $\beta = 1$. From Figure 5.3.1 and Figure 5.3.2, it is evident that for $\beta \le 3$, the performance of the *HyD* is worse than that of the *CWWD* due to a larger noise variance or larger volume under the surface of the weighted-signal variance. For $\beta > 3$, and typically $\beta = 5$, the *HyD* provides a larger *SNR* by a factor of 1.15 (15%) than that of the *CWWD*. For β = 10, the performance of the *HyD* is approximately 1.18 (18%) times better than the *CWWD* in terms of *SNR*. As β further increases, the performance of the *HyD* gradually degrades even though at $\beta \ge 500$ the performance is slightly improved. This is due to an unequal rate of change of the volume under the surface of the weighted signal and that of the weighted-signal variance.

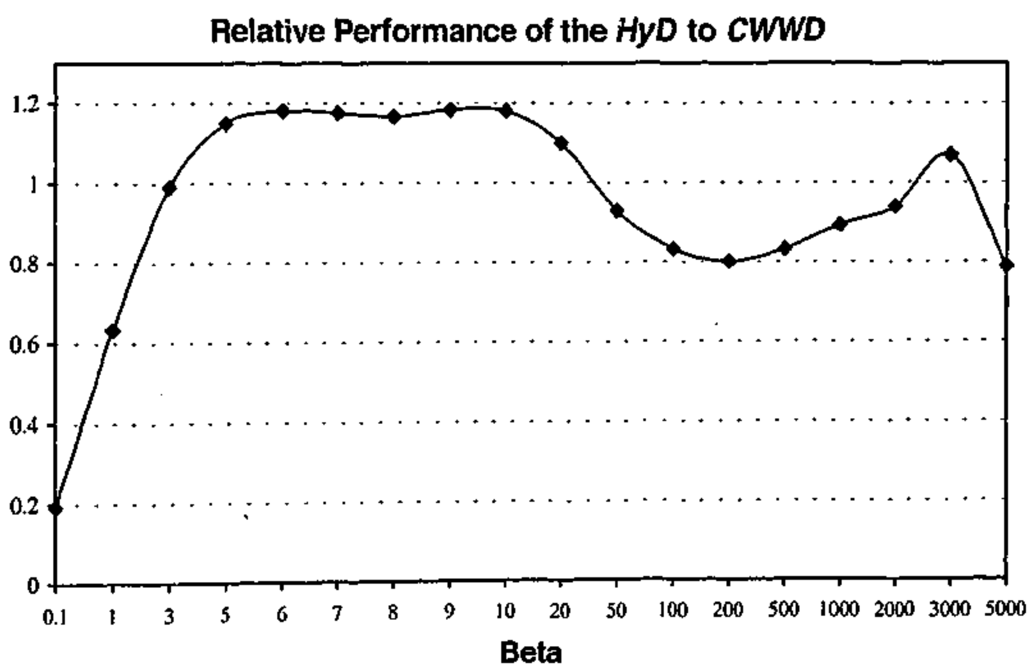

Figure 5.3.2: Loss factor Q of the *HyD* and *CWWD* (Eq. (5.3.25)) as a function of the kernel control parameter $\beta = 1/\sigma$. The useful range of β is evidently from 3 to 10.

From Figure 5.3.2, it can be suggested that the *HyD* is better than the *CWWD* in terms of *SNR* over the typical range of the control parameter β of $3 \leq \beta \leq 10$. Outside this range, the *CWWD* outperforms the *HyD*. For large values of β ($\beta \geq 500$, i.e. $\sigma \leq 0.002$) the *HyD* might provide a large *SNR* which is mainly due to a relatively large volume under the surface of the weighted-signal variance. It should be noted that large values of β are not applicable in practice since the hyperbolic weighting function collapses (in shape) into a near-flat function with a very small volume under the surface. This shape of the weighting function indicates that the kernel is not stable under these specific conditions of large β (small σ for the Choi-Williams kernel) and should not be employed as a time-frequency kernel. This fact was explained in detail in Chapter 3. In contrast, the Choi-Williams weighting function retains its original shape for very small values of σ by having a finite volume under the surface. This makes the Choi-Williams kernel and *CWWD* more stable than the hyperbolic kernel and *HyD* over extreme values of the control parameters β and σ.

At this point, it is appropriate to summarise all the trade-off(s) that have been stated in Chapters 3, 4 and in this chapter so that important remarks can be made. From Chapter 3, it can be concluded that for β in the range of $0.5 \leq \beta \leq 3.5$ the hyperbolic kernel is more effective than the CW kernel in terms of cross-term suppression and auto-term magnitude. For noise robustness, the range of β is $\beta \geq 3$. As stated in Chapter 3, for $20 \geq \beta \geq 0.5$, the hyperbolic auto-term magnitude is still acceptable even though it is lower than that of the CW auto terms. From Chapter 4, for $0.01 \leq \beta \leq 7.8$, the hyperbolic wavelet has a finer scale resolution ω_d than that of the CW wavelet. From the obtained results in this chapter, for effective signal detection, the useful range of β is $3 \leq \beta \leq 10$. However, the performance of the *HyD* is still compatible to that of the *CWWD* for $\beta \leq 50$. Thus, the useful range of β, taking into consideration the auto-term magnitude, auto-term resolution, cross-term suppression, noise robustness, scale resolution and effective signal detection, is $3 \leq \beta \leq 10$ and the practical or applicable range of β is $0.5 \leq \beta \leq 50$. However, the only disadvantage of the hyperbolic kernel is that its auto-term resolution is poorer than that of the CW kernel and most of the MTE kernels as was explained in Chapter 3.

This section has covered a number of important topics of the chapter. Firstly, Moyal's formula for the Wigner-Ville time-frequency distribution was stated. Then, the discrete Moyal formula for the *GNKD* was derived based on Moyal's formula for the WV distribution. After that, the *SNR* of the *GNKD* and performance comparison of the *HyD* and *CWWD* were discussed. The next section studies the *CORR*, *CWD* and *GNKD* by calculating their *SNR*'s.

5.4 Performance Comparison of Some Time-Frequency Signal Detectors

Section 5.3.3 derived an expression of the *SNR* of the *GNKD* in detail and analysed the physical meanings of the energy ratio X_1. Relative performance of the *HyD* and *CWWD* was successfully measured based on geometrical characteristics of the hyperbolic and Choi-Williams kernels respectively. In this section, performance of three signal detectors namely, *CORR*, *CWD* and *GNKD* will be estimated as a function of $X_1 = \frac{B_0}{A_0}$ and $X_2 = \frac{M_0}{N_0}$ under the general case and special cases. The *SNR* expression of the *GNKD* derived in Section 5.3.3 (Eq. (5.3.22)) is employed to determine its performance. The *SNR* expressions of the *CWD* and *CORR* have been given in the literature and will be used to compare their performance with that of the *GNKD*.

5.4.1 Performance of the Cross-Correlator Signal Detector (*CORR*)

The performance of the *cross-correlator* method, known as the matched-filter method, is considered as the best method in binary signal detection since it provides the best *SNR* [31, 74]. The statistical function η is given by

$$\eta_{CORR} = \int_t f(t) \cdot s^*(t)\, dt \text{ , where } -\infty \le t \le +\infty. \tag{5.4.1}$$

The *SNR* of the *cross-correlator* detector is given by [32]

$$SNR_{CORR} = \sqrt{A_0 / N_0} \tag{5.4.2}$$

where A_0 and N_0 are the signal energy and noise variance respectively.

The *SNR* of the *CORR* is not affected by the energy difference between the even and odd samples of the signal (X_1) as in the case for the *CWD* with a unity kernel function as will be discussed in the next section. For the case of a non-unity kernel signal detector, the effects of the absolute energy difference between the even and odd samples of the digital input signal and noise (X_2) are included as will be shown in Section 5.4.3.

5.4.2 Performance of the Wigner-Ville Detector

The performance of the Wigner-Ville detector (*CWD*) was studied by Kumar and Carroll for the continuous and discrete cases [18, 31, 32, 79]. The *SNR* of the Wigner-Ville signal detection is given by Eq. (5.4.3) as

$$SNR_{CWD} = \sqrt{\frac{A_0}{N_0}} \cdot \frac{1+\left(\frac{B_0}{A_0}\right)^2}{\sqrt{1+3\cdot\left(\frac{B_0}{A_0}\right)^2}} \tag{5.4.3}$$

The SNR_{CWD} of the Wigner-Ville time-frequency signal detector is clearly smaller than the SNR_{CORR} of the *cross-correlator* detector given by Eq. (5.4.2) due to the effects of the ratio $X_1 = \frac{B_0}{A_0}$.

5.4.3 Performance of the General Non-Unity Kernel Signal Detector (*GNKD*)

The performance of the *GNKD* was briefly estimated in Section 5.3.3. In this section, its performance under general and special cases, such as for small values of X_1 and X_2, is discussed. Relative performance of the *GNKD*, *CORR* and *CWD* is also estimated by taking ratios of their *SNR*'s to form the loss factor Q. The larger the value of the Q factor, the better the performance of the relevant signal detector.

From Eq. (5.3.22), it is evident that the *SNR* of the *GNKD* depends on X_1 and X_2, which clearly shows the effects of a noise source on the performance of the detector. It should be noted that in the case of the Wigner-Ville distribution employing a unity kernel, the effects of the noise ratio X_2 are not apparent [31]. In addition, the effects of X_1 range from the first power to the third-power terms as shown by Eq. (5.3.22). The noise ratio X_2 is of the first power only which significantly affects the *SNR* of the *GNKD*.

If $X_2 = \frac{M_0}{N_0}$ is very small, i.e. the noise energy difference is most evenly distributed among its even and odd samples or the noise bandwidth is small, then Eq. (5.3.22) becomes

$$SNR_{GNKD} = \frac{\sqrt{2}\cdot(A_{0F}^2 + B_{0F}^2)}{A_0\sqrt{N_{0FF}\cdot A_{0FF}\cdot\left\{1+\frac{B_0}{A_0}+2\left(\frac{B_0}{A_0}\right)^2\right\}}} \tag{5.4.4}$$

After separating the kernel weighting function, we obtain

$$SNR_{GNKD} = \sqrt{\frac{A_0}{N_0}}\cdot\frac{\sqrt{2}\cdot\left[1+\left(\frac{B_0}{A_0}\right)^2\right]}{\sqrt{1+\frac{B_0}{A_0}+2\cdot\left(\frac{B_0}{A_0}\right)^2}}, \text{ where } X_2 \text{ is very small.} \tag{5.4.5}$$

If $X_2 = \frac{M_0}{N_0}$ is not small, then SNR_{GNKD} will be further reduced and the performance of the *GNKD* is degraded. It is also important to note that the *SNR* of the *GNKD* has been calculated under the special conditions of $p_k = p_l$, $q_k = q_l$ and $q_l = q_{ll}$ stated in Section 5.3.3. This means that only the auto terms in the summations (see Appendix B) are included and interactions among them are ignored. The performance in this case can be considered as the lower limit performance of the detector. For the general case, the *SNR* of the *GNKD* is given by Eq. (5.4.6)

$$\begin{aligned} SNR_{GNKD} &= \sqrt{\frac{A_0}{N_0}}\cdot\frac{\sqrt{2}\cdot\left[1+\left(\frac{B_0}{A_0}\right)^2\right]}{\sqrt{1+\frac{B_0}{A_0}+2\left[\left(\frac{B_0}{A_0}\right)^2+\frac{M_0}{N_0}\cdot\left(\frac{B_0}{A_0}+\left(\frac{B_0}{A_0}\right)^3\right)\right]}} \\ &= \sqrt{\frac{A_0}{N_0}}\cdot\frac{\sqrt{2}\cdot(1+X_1^2)}{\sqrt{1+X_1+2\cdot[X_1^2+X_2\cdot(X_1+X_1^3)]}} \end{aligned} \tag{5.4.6}$$

The 3-D graphical presentation of the normalised *SNR* of the *GNKD*, SNR_{GNKD}, as a function of $X_1 = \frac{B_0}{A_0}$ and $X_2 = \frac{M_0}{N_0}$, is displayed in Figure 5.4.1. It should be noted again that for the case of the non-unity kernel time frequency signal detector, the effects of noise are taken into account which reduces the performance of the detector. From Figure 5.4.1, the minimum SNR_{GNKD} is -0.6602 dB at ($X_1 = 0.63$, $X_2 = 1$), i.e. when the energy difference of the even and odd samples of the noise is equal to its energy. In other words, when the noise is energy balanced.

Figure 5.4.2 and Figure 5.4.3 show the absolute and normalised loss factors of *GNKD/CORR*, *GNKD/CWD* and *CWD/CORR* respectively as a function of X_1. The absolute plot of the SNR_{GNKD} in Figure 5.4.2 has the same shape as that of Figure 5.4.3 except that its maximum value is $\sqrt{2} = 1.414$, i.e. the SNR_{GNKD} is improved by a factor of $\sqrt{2}$ or about 41.4% compared with the *CWD* and *CORR*. From these figures, detailed comparisons of the three signal detectors are shown clearly in Table 5.4.1.

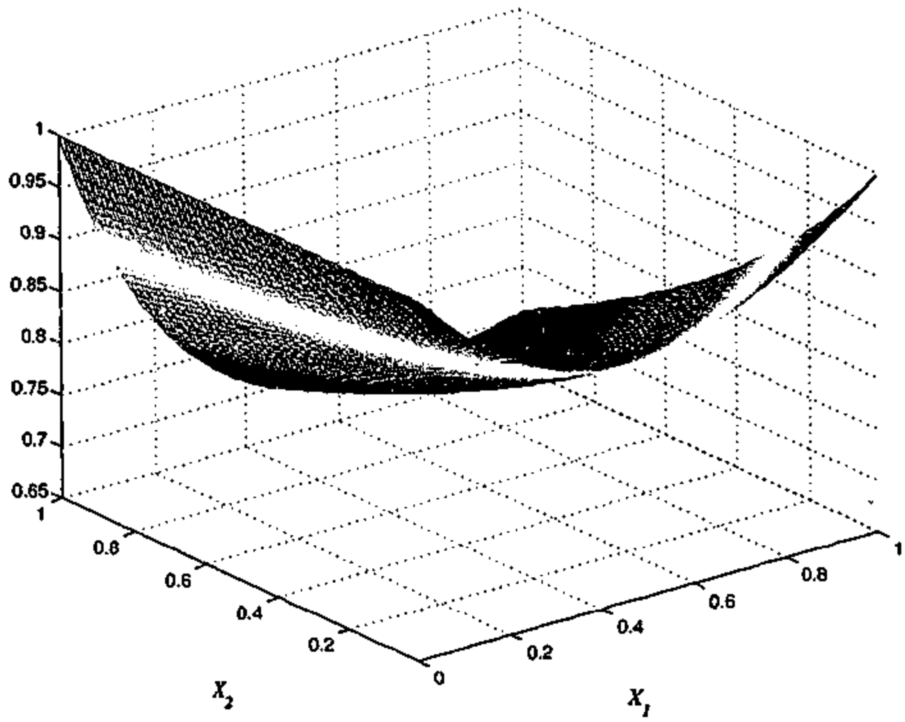

Figure 5.4.1: Normalised SNR_{GNKD} of the *GNKD* (Eq. (5.4.6)) as a function of X_1 and X_2. The optimum performance of the *GNKD* is obtained by having X_1 in the range of X_1 of $0.0 \leq X_1 \leq 0.2$ or $0.9 \leq X_1$ and $X_2 \leq 0.2$.

The Q factor of the *GNKD* and *CWD* is then given by the ratio of Eq. (5.4.5) to Eq. (5.4.3) (when X_2 is small)

$$Q(GNKD/CWD) = \frac{SNR_{GNKD}}{SNR_{CWD}} = \sqrt{2} \cdot \sqrt{\frac{1+3\cdot\left(\frac{B_0}{A_0}\right)^2}{1+\frac{B_0}{A_0}+2\cdot\left(\frac{B_0}{A_0}\right)^2}} = \sqrt{2} \cdot \sqrt{\frac{1+3\cdot X_1^2}{1+X_1+2\cdot X_1^2}} \tag{5.4.7}$$

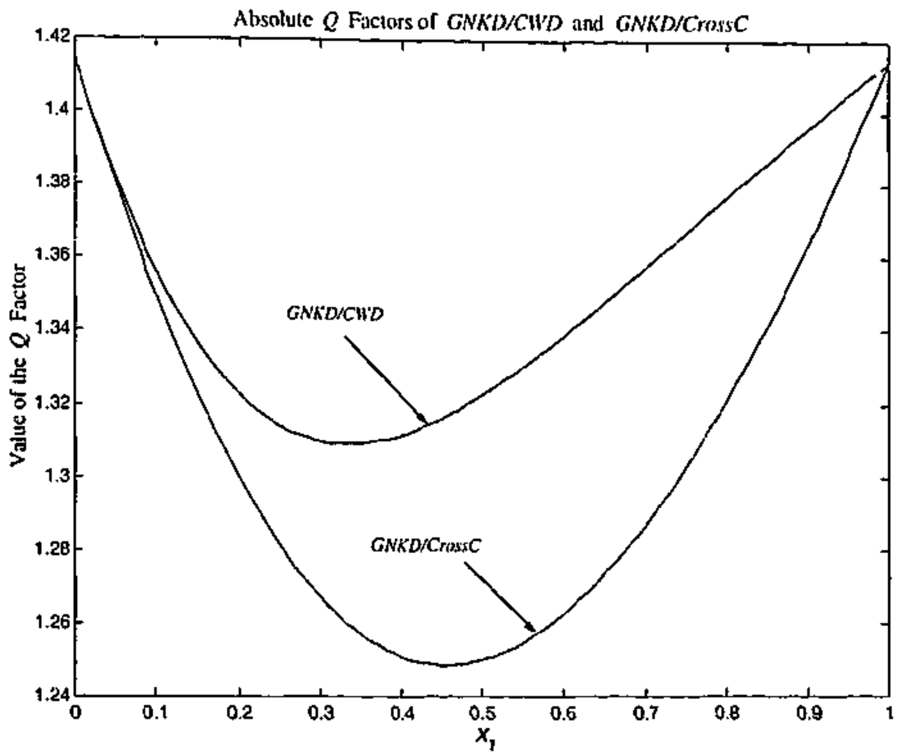

Figure 5.4.2: Loss factor Q (*GNKD/CWD*) (Eq. (5.4.7)) and Q (*GNKD/CrossC*) as a function of X_1. The typical range of X_1 for satisfactory performance is $0.0 \le X_1 \le 0.1$ or $0.9 \le X_1 \le 1.0$ when X_2 is small.

It should be noted that the *HyD* and *CWWD* belong to the *GNKD* class in which the kernel function is non-unity. The *CWD* is a special of the *GNKD* in which the unity kernel function is employed. From Eq. (5.4.7), *SNR* of the *GNKD* is about 41.4% (by a factor of $\sqrt{2}$) higher than that of the *CWD* which clearly reveals the advantage of using non-unity kernel time-frequency distributions.

When X_2 is small, the loss factor of the *GNKD* and the classical *CORR* is given by

$$Q(GNKD/CORR) = \frac{SNR_{GNKD}}{SNR_{CORR}} = \frac{\sqrt{2}\cdot\left[1+\left(\frac{B_0}{A_0}\right)^2\right]}{\sqrt{1+\frac{B_0}{A_0}+2\cdot\left(\frac{B_0}{A_0}\right)^2}} = \frac{\sqrt{2}\cdot(1+X_1^2)}{\sqrt{1+X_1+2\cdot X_1^2}} \tag{5.4.8}$$

It was found in Section 5.3.3 that the ratio of the bandwidth (Eq. (5.3.23)) to the total signal energy (Eq. (5.3.11)), X_1, determines the performance of a time-frequency signal detector. Decreasing the bandwidth and increasing the energy of the signal lowers this ratio and leads to better performance. For good performance, typical ranges of X_1 of $0.0 \leq X_1 \leq 0.2$ or $0.8 \leq X_1 \leq 1.0$ if $0.0 \leq X_2 \leq 0.2$ are required as can be seen in Figure 5.4.1. Thus, any value of X_1 in the range of $0.2 < X_1 < 0.8$ will considerably lower the *SNR* of the detector and should not be used. If there is energy balance in the odd and even samples of the input signal, i.e. X_1 is close to zero, then the signal is best detected.

When there is energy unbalance among the odd and even samples of the signal, i.e. there is clear dominance of even over odd samples or odd over even samples, X_1 will be close to unity. In this case, if the noise energy is evenly distributed then best detection can be achieved. Otherwise, the signal is not effectively detected.

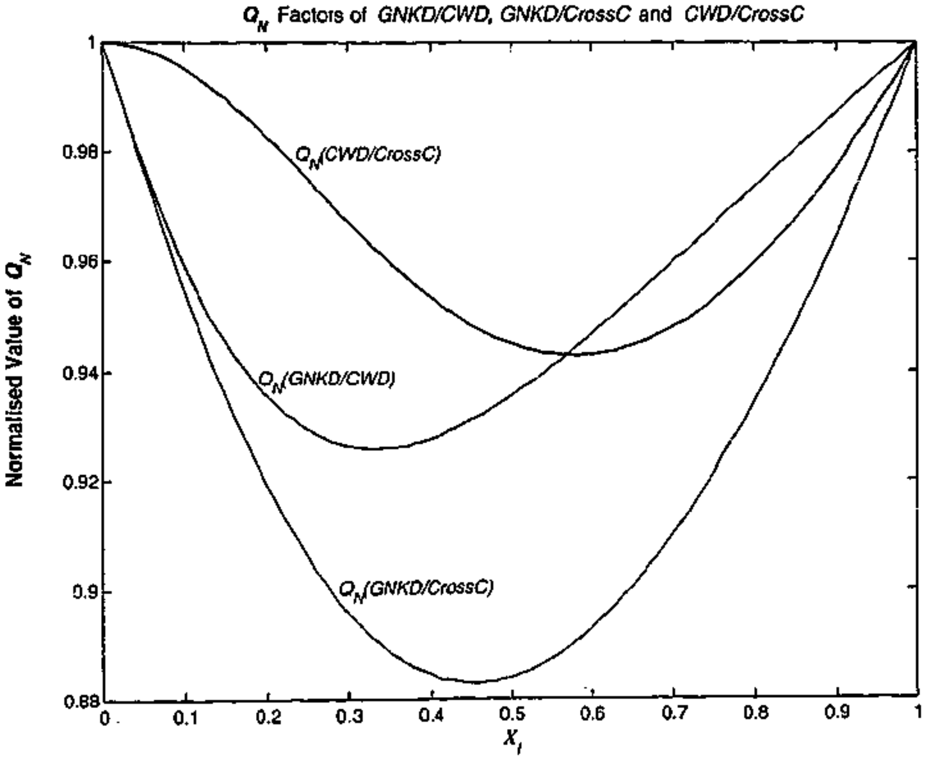

Figure 5.4.3: Normalised Q_N (*GNKD/CWD*) (Eq. (5.4.7)), Q_N (*GNKD/CORR*) (Eq. (5.4.8)) and Q_N(*CWD/CORR*) (ratio of Eq. (5.4.3) to Eq. (5.4.2)) as a function of X_1. For the *CWD*, the typical range of X_1 can be extended to $0.0 \leq X_1 \leq 0.3$. The maximum value of each Q factor was used as the normalisation factor.

Table 5.4.1: Worst performance ratio of the *GNKD* to *CWD* (Eq. (5.4.7)), the *CWD* to *CORR* (ratio of Eq. (5.4.3) to Eq. (5.4.2)) and the *GNKD* to *CORR* (Eq. (5.4.8)) as a function of X_1 as seen from Figure 5.4.3. The best performance is obtained at $Q_N = 1$ with the corresponding $SNR = 0\ dB$.

Signal Detector Comparison	Worst Performance		
	$X_1 = \frac{B_0}{A_0}$	Normalised Loss factor Q_N	*SNR* of Q_N (*dB*)
GNKD/CWD	0.35	0.9258	−0.67
CWD/Cross-Correlator	0.6	0.9428	−0.5
GNKD/Cross-Correlator	0.45	0.8829	−1.08

The performance of the *CORR*, *CWD* and *GNKD* has been compared and it is clear that the *GNKD* has the largest absolute *SNR*. The next section studies the effects of sampling on the performance of time-frequency signal detectors using typical signals.

5.4.4 Some Typical Examples

In Sections 5.4.1–5.4.3, the performance of the *CORR*, *CWD* and *GNKD* were theoretically estimated and compared by using the discrete Moyal formula derived in Section 5.3.2. In this section, the experimental aspects of detection performance and the effects of sampling on the input signal are examined using the *GNKD*. Moreover, particular attention is given to how the energy ratio X_1 varies with different values of the sampling interval Δt. Since X_1 is the energy ratio of the even and odd samples of the digital input signal, its value strongly depends on the type of signal and the sampling interval Δt. Some typical and popular signals in practice are examined such as a sinusoid at 50 *Hz* ($sin(2\pi \cdot 50t)$), decaying exponential $exp(-t)$, an exponentially decaying sinusoid $sin(t) \cdot exp(-t)$, chirped $cos(C \cdot t^2)$, ECG and speech.

As was mentioned in Section 5.3.3, for digital input signals, the sampling interval does affect the value of the energy ratio X_1 of the signal. A number of waveforms have been digitised at different sampling rates and the experimental results are summarised in Table 5.4.2. The sampling interval should be small enough to obtain small values of X_1. In this case, the sampling frequency is set to be about four times larger than the critical Nyquist frequency of the input signal. From Table 5.4.2, it appears that sinusoidal signals can be efficiently detected using time-frequency signal detectors.

The transient signal $exp(-t)$ has a large-valued X_1 which can yield poor SNR if $\Delta t \geq 500$ ms. Sinusoidal signals can be very efficiently detected with very small-valued X_1. It is important to emphasise that for periodic signals, the signal sampling interval should be chosen long enough so that X_1 can be correctly estimated. The exponentially decaying sinusoidal signal has the largest X_I of 0.14 at the worst sampling interval $\Delta t_w = 0.6$ s with the corresponding worst $SNR_w = 0.935$ as can be seen in Table 5.4.2, where the subscript "w" indicates the worst case and "b" the best case. The ECG and speech signals appear to have small-valued X_1 which might suggest that these signals can be successfully detected using the *GNKD*. The non-stationary chirp signals can be very efficiently detected [illegible]sing the *GNKD* with the worst and best *SNR*'s are relatively close to 0.99 and 1.0 [illegible]rom Figure 5.4.3, the Q_N factor in detecting signals that have $X_1 \leq 0.1$ is [illegible] range of $Q_N \geq 0.95$ which is satisfactory.

Table 5.4.2: The best and worst cases in detection of some typical [illegible]NKD in terms of normalised SNR_{GNKD} in Figure 5.4.1 when X_2 is small. T[illegible]w" and "b" indicate the worst and best cases respectively.

	Worst Case			Be[illegible]		
Signal	$\Delta t_w (ms)$	$X_{1(w)}$	SNR_w (dB)	Δt_b (ms)	$X_{1(b)}$	SNR_b (dB)
$sin(2\pi \times 50n)$	5.00	1.00	0.75	2.00	0.00	1.00
$exp(-n)$	500	0.462	0.65	10	0.01	0.98
$exp(-n)\cdot sin(n)$	600	0.14	0.935	100	0.04	0.975
$cos(2\pi n \cdot 0.125 \cdot (n/M))$	800	0.008	≈ 0.99	0.1	5.0×10^{-7}	≈ 1.00
ECG (averaged over 12 channels)	The sampling intervals for these cases are fixed. There are no worst or best cases for these signals.			1.00	4.75×10^{-6}	1.00
Speech [85] (English vowels "a", "e", "o", "u", "i" and the sound "sh")				—	3.7×10^{-4}	1.00

From Table 5.4.2, it should be noted that for the exponentially decaying sinusoid, the worst and best detection of the signal are quite similar which suggests that the signal can be well detected using the *GNKD*. Based on the performance of the *GNKD*, it is evident that stationary signals such as sinusoids can be effectively detected using time-frequency signal detectors. Detecting non-stationary signals such as the decaying exponential $exp(-t)$, chirp and exponentially decaying sinusoidal $sin(t)\cdot exp(-t)$ signals is dependent on the sampling interval used to sample the signal. If the sampling interval Δt is fast enough, then the detection process will be effective.

5.5 Conclusion

This chapter reports on some investigations in the field of time-frequency signal detection.

Firstly, the discrete Moyal formula has been derived for the general case in which the kernel function is not a unity kernel. The performance of the general non-unity kernel signal detector (*GNKD*) has been examined by using the discrete Moyal formula to obtain the *SNR* of the statistical function η. It has been shown that the *GNKD* perform better than the Wigner-Ville detector (*CWD*) by increasing its *SNR* by a minimum factor of $\sqrt{2}$. The performance of the correlator detector has also been examined and compared with that of the *CWD* and *GNKD*. It has been found that the hyperbolic detector *HyD* and Choi-Williams detector (*CWWD*) can improve the *SNR* over the *CWD* by a factor Q in the range of 22.5 $dB \leq Q_{Hy} \leq 24.8$ dB and 22.5 $dB \leq Q_{CW} \leq 23.99$ dB respectively over the typical range of $3 \leq \beta \leq 10$. From the results obtained in Chapters 3, 4 and in this chapter, the applicable range of β is $0.5 \leq \beta \leq 50$ for satisfactory performance on signal detection, scale resolution (discussed in Chapter 4), cross-term suppression, auto-term magnitude and noise robustness (discussed in Chapter 3) at the expense of having a poor auto-term resolution (discussed in Chapter 3). This is an important trade-off among crucial parameters that significantly affects the performance of a kernel.

Secondly, a new signal detector, the hyperbolic time-frequency signal detector, has been proposed and investigated. The new detector performs better than the famous *CWWD* by improving the *SNR* by 18% for $3 \leq \beta \leq 10$, independent of the input signal because of the nature of the weighting functions of the CW and hyperbolic detectors.

Thirdly, the performance of the *GNKD* using a number of typical signals has been examined. It has been shown that the sampling interval can affect the performance of the *GNKD* by varying the energy ratio $X_1 = B_0/A_0$. It has also been observed by simulation that sinusoidal and chirped signals can be efficiently detected with satisfactory *SNR*. Transient signals can be efficiently detected by using a suitable sampling interval. Physiological signals such as the ECG and speech can be successfully detected with the normalised *SNR* in the approximate range of 0.99 to 1.00.

Chapter 6: THE HYPERBOLIC WAVELET POWER SPECTRA OF TYPICAL SIGNALS

This chapter introduces the wavelet power spectrum as a useful technique to study signal characteristics in which the hyperbolic and *sym3* wavelets are employed. The hyperbolic wavelet power spectrum technique is employed for typical signals including ECG, sinusoidal and transient exponential. Chapter 7 develops this technique for practical non-stationary signals such as music and speech. Comparisons between the Fourier power spectrum technique and hyperbolic wavelet power spectrum technique are also made.

6.1 Theoretical Background of the Wavelet Power Spectrum Technique

Analyses of stationary signals have been carried out over many decades using the Fourier transform. The basis of the Fourier transform is that any function can be represented by a sum of a number of complex exponential functions, in other words, sinusoids and co-sinusoids. The discrete Fourier transform is defined as [41]

$$\hat{X}(k)=\sum_{n=0}^{M-1} x(n)\cdot e^{-jnk/M} \tag{6.1.1}$$

where $x(n)$ is a discrete input signal, $\hat{X}(k)$ is its Fourier transform, M is the length of $x(n)$ and k is the frequency variable in rad/s.

The merits of the Fourier transform method were discussed in detail in Chapter 1 at the beginning of the thesis. Based on this discussion, it is clear that there are a number of issues that need to be taken into account when using the Fourier transform. First, the Fourier frequency spectrum of a signal $f(t)$ is obtained by averaging its values over an infinite time interval since the Fourier kernel function is exponential. As a result, if the signal spectrum varies with time (as for non-stationary signals) then fine spectral details will be lost. Thus, for the Fourier transform to be accurate, the input signal should be stationary or wide-sense stationary which means that its statistical properties do not vary with time. Secondly, signals, in general (from linear and non-linear systems), consist of a sum of different harmonics which interact with each other to cause quadratic phase coupling [86, 87]. The significant limitation of the Fourier transform method is that it cannot provide information on interactions of different harmonics. Thus, estimating the Fourier transform of a signal only reveals information about the frequency spectral contents of each harmonic component over a specified frequency range averaged over the entire time horizon.

To understand a signal, its energy distribution or energy density must be examined. Higher-order statistical techniques using the Fourier transform such as the power spectrum and bispectrum have been successfully employed to study signal energy distributions [2, 4]. In using higher-order statistical techniques, correlation functions and the Wiener-Khinchin theorem play important roles. The general correlation function is defined by [86, 87]

$$R_m(\tau_1,\tau_2,\ldots,\tau_m)=\sum_{n=0}^{M-1} x(n)\cdot x(n+\tau_1)\cdot x(n+\tau_2)\cdot\ldots\cdot x(n+\tau_m) \tag{6.1.2}$$

where τ_i is the lag parameter in the range of $0-(M-1)$, m is the order of the correlation function, $m=1$ corresponds to the auto-correlation function (whose Fourier transform is the power spectrum) and n is the discrete index of the input signal $x(n)$.

The Wiener-Khinchin theorem [41] is given by

$$P(\omega)=\sum_{\tau=0}^{\tau=M-1} R_1(\tau)\cdot e^{-j\omega\tau} \tag{6.1.3}$$

where $R_1(\tau)=\sum_{n=0}^{M-1} x(n)\cdot x(n+\tau)$ is the auto-correlation function and $P(\omega)$ is the power spectrum of the discrete input signal $x(n)$.

From Eq. (6.1.3), the power spectrum can be rewritten in a simpler form

$$P(\omega)=\hat{X}(\omega)\cdot\hat{X}^*(\omega)=\left|\hat{X}(\omega)\right|^2 \tag{6.1.4}$$

where the symbol "*" indicates the complex conjugate operation.

It should be noted that Eq. (6.1.4) is directly derived from Eq. (6.1.3) and the former employs a product of two Fourier transforms of the input signal. The power spectrum is usually estimated by using Eq. (6.1.4) rather than Eq. (6.1.3) to reduce its computational burden. Other higher-order frequency functions such as the bispectrum and trispectrum can be formed based on Eq. (6.1.4), which is also used to form the analogous wavelet power spectrum (*WPS*) to improve calculation efficiency.

From Eq. (6.1.4), it is evident that the phase information is suppressed and only the magnitude information is given by the power spectrum. Therefore the power spectrum is not unique since there might be more than one signal which has identical magnitude but different phase information. Thus, the second-order statistical function, i.e. the bispectrum, is employed to study signals that have identical power spectra. The bispectrum of a discrete input signal $x(n)$ is the 2-D Fourier transform of its tri-correlation function $R_2(\tau_1, \tau_2)$ and is given by

$$B(f_1, f_2) = \sum_{\tau_2=0}^{M-1} \sum_{\tau_1=0}^{M-1} R_2(\tau_1, \tau_2) \cdot exp\left[-j2\pi(f_1\tau_1 + f_2\tau_2)\right] \tag{6.1.5}$$

where τ_1 and τ_2 are the lag parameters, $R_2(\tau) = \sum_{n=0}^{M-1} x(n) \cdot x(n+\tau_1) \cdot x(n+\tau_2)$ is the tricorrelation function and $B(f_1, f_2)$ is the bispectrum of the discrete input signal $x(n)$.

Eq. (6.1.5), like Eq. (6.1.3), can be rewritten in a simpler form as

$$B(f_1, f_2) = \hat{X}(f_1) \cdot \hat{X}(f_2) \cdot \hat{X}^*(f_1 + f_2) \tag{6.1.6}$$

The general expression for higher-order statistics spectrum, *HOSP*, is obtained using the following expression

$$HOSP(f_1, f_2, ..., f_m) = \hat{X}(f_1) \cdot \hat{X}(f_2) \cdot ... \cdot \hat{X}^*(f_1 + f_2 + ... + f_m) \tag{6.1.7}$$

The bispectrum decomposes the skewness or odd-order asymmetries of the input signal since three Fourier products are involved in Eq. (6.1.6) [88]. It also provides information on interactions of the frequency components f_1 and f_2 in the 2-D frequency plane, (f_1, f_2).

The power spectrum and bispectrum have been shown to be very effective in studying chaos, non-linear behaviour and turbulence of wide-sense stationary signals as reported in many studies [1-4, 89, 90]. They have also been used to examine non-stationary signals by dividing them into typically small segments of 1,024 samples for cases of the ECG and Duffing oscillator. In this case, the effectiveness of the power spectrum and bispectrum has been clearly shown by Chandaran [1] and Lipton [4].

However, for other non-stationary signals such as plasma phenomena, speech, underwater signals, whale sounds and music, the power spectrum and the bispectrum are not suitable tools since they suppress fine details of the energy distribution of the input signal. Thus there is a need for a time-frequency power spectrum analysis which provides information on how the energy of a non-stationary signal is distributed in time and frequency domains.

The time-frequency power spectrum technique employs both the time and frequency axes to display the spectrum of a non-stationary input signal. The additional time axis is required to track any changes of the signal spectrum over time and this axis enables detection of instantaneous behaviour of the signal. One typical time-frequency analysis tool is the wavelet power spectrum which uses the wavelet transform. The wavelet transform can be regarded as being analogous to the Fourier transform, but with a different kernel function. For the Fourier transform method, the kernel function is an exponential function and for the wavelet transform method there exist various arbitrary kernel functions which must satisfy admissibility constraint(s).

The wavelet transform method, like the Fourier transform method, assumes that any function can be represented by a sum of an arbitrary set of wavelet functions $\psi(t)$ (which satisfy the admissibility constraints) [51]. The Fourier transform might be considered a special case of the wavelet transform in which the mother wavelet function $\psi(t)$ is replaced by the complex exponential function, $\psi_{Fourier}(t) = exp(-j\omega t)$, if this function satisfies the admissibility constraints as will be discussed later. The mother wavelet function, $\psi(t)$, sometimes called the basic wavelet function, is the fundamental function in any expansion of the wavelet transform. The wavelet transform of the input signal $x(t)$ with the mother wavelet function $\psi(t)$, which was defined in Chapter 4, is given as [51, 52]

$$WT(a,b) = \int_{-\infty}^{+\infty} x(t) \cdot \psi\left(\frac{t-b}{a}\right) dt \tag{6.1.8}$$

where a and b are the scale and time indices respectively, and $WT(a, b)$ is the wavelet transform function of the input signal $x(t)$. For the Fourier transform, $\psi(\cdot)$ is the exponential function $exp(-j\omega t)$ which has infinite time support. To obtain finite time support for the wavelet transform, an appropriate mother wavelet function must be employed.

There are two mother wavelets that will be employed in this chapter. The orthogonal *sym3* wavelet, provided in MATLAB software, can be regarded as a useful symmetrical wavelet whose scale function exists and the proposed hyperbolic wavelet of the symmetrical "crude" wavelet group as was studied in Chapter 4.

To be a valid mother wavelet function, the function must satisfy the following admissibility constraint which is given by [51, 52]

$$\int_{-\infty}^{+\infty} \psi(t)dt = 0 \tag{6.1.9}$$

The above condition is satisfied by the complex exponential function since it can be written as a sum of two sinusoids. Thus, it can be said that the Fourier transform is a special case of the wavelet transform in which the mother wavelet is the complex exponential.

From Eq. (6.1.4), the wavelet power spectrum $WPS(t, \omega)$, analogous to the Fourier power spectrum $P(\omega)$ of an input signal $x(n)$, is given by [33, 34]

$$WPS(t,\omega) = WT(t,\omega) \cdot WT^*(t,\omega) = |WT(t,\omega)|^2 \tag{6.1.10}$$

where $WT(t, \omega)$ is the wavelet transform of the input signal $x(n)$.

The main difference between the wavelet transform and the Fourier transform is that the wavelet transform examines the frequency contents of the signal over a short time period since its mother wavelet function has finite-time support. By contrast, the Fourier transform averages the frequency contents of the signal over an infinite time interval by the effects of $sin(\cdot)$ and $cos(\cdot)$ functions. The time-support range of most wavelet functions (the hyperbolic, the Choi-Williams or Mexican-hat and Morlet wavelets for example) is approximately 10- to 20-unit time index (as was seen in Chapter 4). Thus, by employing the wavelet transform, it is possible to observe instantaneous behaviour of the signal, which is vital in studying the signal characteristics and predicting its future behaviour.

In addition, the *WPS* technique gives the energy density of the input signal in both time and frequency domains, whereas the Fourier power spectrum displays the energy contents of the signal in the frequency domain only. The combination of time and frequency domains yields local images of the input signal energy contents and thus it is possible to carry out in-depth study of the signal by examining its instantaneous behaviour. Since there exist many different wavelet functions, the corresponding wavelet transforms also exist and each wavelet function has different characteristics which means that they can be used for different specific applications. The hyperbolic wavelet function, which was proposed in Chapter 4, is employed in this chapter to further demonstrate the usefulness of the hyperbolic kernel family. The hyperbolic wavelet has been shown to have a fine scale resolution which is suitable for studying signals that do not have broad power spectra such as transients which was concluded in Chapter 4. The hyperbolic wavelet function also has a small number of computed scales which means it can be used for compression purposes. This reduces the computational burden of the hyperbolic wavelet power spectrum.

Milligen and others [33, 34] showed how the *WPS* and wavelet bispectrum techniques could be used to study chaos and turbulence which provide the foundation for the research reported in this chapter. In their studies, they showed that the wavelet bispectrum could be utilised to effectively study chaos. They showed that the main problem with the wavelet bispectrum was that there are four dimensions that need to be simultaneously expressed. Thus, the concept of slicing the wavelet bispectrum at separate frequencies was employed which was shown to be successful provided that the behaviour of the signal could be predicted. Farge and others [35] showed that the wavelet transform method could be used to study turbulence by detecting edgy behaviour in its time-frequency spectrum.

The *WPS* technique is much simpler than the wavelet bispectrum technique since there are only three quantities, i.e. time, scale and magnitude, that need to be simultaneously displayed. One major advantage of the *WPS* technique over the Fourier power spectrum technique is that the signal energy distribution is shown in 3D graphs which do not suppress the phase information as it is the case of the power spectrum, i.e. the phase information is included as a function of time.

Jubran and Hamdan [91] used the Gaussian wavelet transform to study the behaviour of flow induced vibration and cross flow in a cylinder. They compared the performance of a number of different mother wavelets including the Morlet wavelet, Daubechies wavelets and the Gaussian wavelet, and then concluded that the Gaussian wavelet was the most suitable wavelet for this particular application. In that paper, only the wavelet transform of the input data was estimated, while the *WPS* and bispectrum techniques were not considered.

This chapter aims to create a gallery of the hyperbolic and *sym3* wavelet power spectra of typical signals. Comparisons of the wavelet power spectra of these signals will be made to validate the hyperbolic wavelet power spectrum. However, quantitative comparisons between the *sym3* and hyperbolic wavelets are not the prime purpose of the chapter. Unlike the work done by Milligen and others, the work reported in this chapter explores further the effectiveness and usefulness of the *WPS* technique in studying instantaneous behaviour and energy distributions of non-stationary signals.

The chapter is organised as follows. Section 6.2 is the main section in which the wavelet power spectra of several signals are displayed. The *WPS* of a periodic sinusoidal signal is studied first in Section 6.2.1 as this is the most common and well-known signal in signal processing. Section 6.2.2 examines the popular exponential signal $exp(-t)$. Section 6.2.3 calculates the *WPS* of an exponentially decaying sinusoid $sin(t) \cdot exp(-t)$. The Duffing oscillator is studied in Section 6.2.4 including Periods 1, 2, 4 and chaotic state. The Fourier power spectra of these signals are also given to validate results drawn by using the wavelet power spectrum technique.

6.2 The Hyperbolic and sym3 Wavelet Power Spectra of Typical Signals

In this section, characteristics of a number of typical signals including sinusoids, exponentially decaying sinusoids, Duffing oscillator, ECG are examined by using the wavelet power spectrum (*WPS*) technique. Instantaneous energy distributions of these signals are continuously monitored so that their characteristics can be successfully revealed. The MATLAB software package has been extensively used to display various signal wavelet power spectra.

The main reasons that the two wavelets, *sym3* and hyperbolic, are used to calculate the wavelet power spectra of various signals are firstly, to validate the hyperbolic wavelet power spectrum technique as an effective tool for signal analysis. Secondly, to verify the correctness and effectiveness of the newly proposed hyperbolic wavelet, which is one of the main research topics of the thesis.

The main advantages of the *sym3* wavelet over the hyperbolic wavelet are first, the former has a larger number of possible calculated scales than those of the hyperbolic wavelet. This enables the *sym3* wavelet to cover all the necessary scales for different signals. However, one main advantage of having a small number of possible calculated scales is that the wavelet power spectral calculation process is less time consuming. Second, the scaling function of the *sym3* wavelet exitsts which is significantly different from other wavelets in the "crude" wavelet group. In this chapter, both the *sym3* and hyperbolic wavelet power spectra will be estimated and discussed in detail. For some signals, where the *sym3* and hyperbolic wavelet power spectra are almost identical, only one wavelet power spectrum will be displayed.

The purpose of creating a gallery of the wavelet power spectra of signals is to recognise wavelet power spectra of different signals so that analyses of unknown signals can be examined based on the known wavelet power spectra. Thus, qualitatively, the general patterns of the wavelet power spectra are mainly focused. Quantitative details of signals such as signal frequency, although can be estimated, is not strongly emphasised in this research.

To compare different wavelet power spectra, which are 2-D matrices of complex elements, normalised wavelet power spectra are employed. The normalisation process is carried out by dividing every member of the wavelet power spectrum matrix by the maximum magnitude of the matrix elements. For all graphs in this chapter and Chapter 7, the "samples" axis, which is used to indicate the sample number of the input signal, and the "Time Index" axis, which is used in the wavelet context for time expansion, are identical except that they are used in different context. Contour scales have also been added to all graphs where appropriate so that energy density levels can be identified.

By using the *WPS* technique, harmonics and sub-harmonics are displayed. A harmonic peak is recognised by an island of closed contour curves with the minimum scale of 0.7 on the normalised scale [113]. If the scale is less than 0.7, the peak is considered as a sub-harmonic. This convention is applied in this chapter and Chapter 7 to recognise harmonic and sub-harmonic peaks. For all graphs in this chapter and the next chapter, the "Scale Index" and "Time Index" correspond to a and b respectively as these notations were used in Chapter 4. Thus, interchangeably, it should be understood that a and b are defined as "Scale Index" and "Time Index" respectively.

6.2.1 The Wavelet Power Spectrum of a Sinusoid

In this section, contour plots of WPS_{sym3} and WPS_{hyp} of a periodic sinusoid $sin(2\pi t)$ are given in Figure 6.2.1 and Figure 6.2.2 respectively in which periodicity can be identified by the following points

1. There is a clear boundary between the peaks of the signal which indicates strong periodic behaviour. In addition, the energy is mainly concentrated at harmonic peaks and there is no broad energy distribution over a wide scale range. The harmonic peaks are located at the approximate scale of $a \approx 50$.
2. Contour curves are closely spaced and there are a large number of bounded small-radius contours towards the harmonic peaks.
3. The energy is discretely and uniformly distributed. The most important and recognisable feature of a periodic signal is that its energy distribution is repetitive. It can be seen that the discrete peaks of the input signal are clearly displayed by the *WPS* technique. Thus, it is evident that for periodic signals, their wavelet power spectra are not broad and smeared but discontinuous and exhibit distinctive peaks.

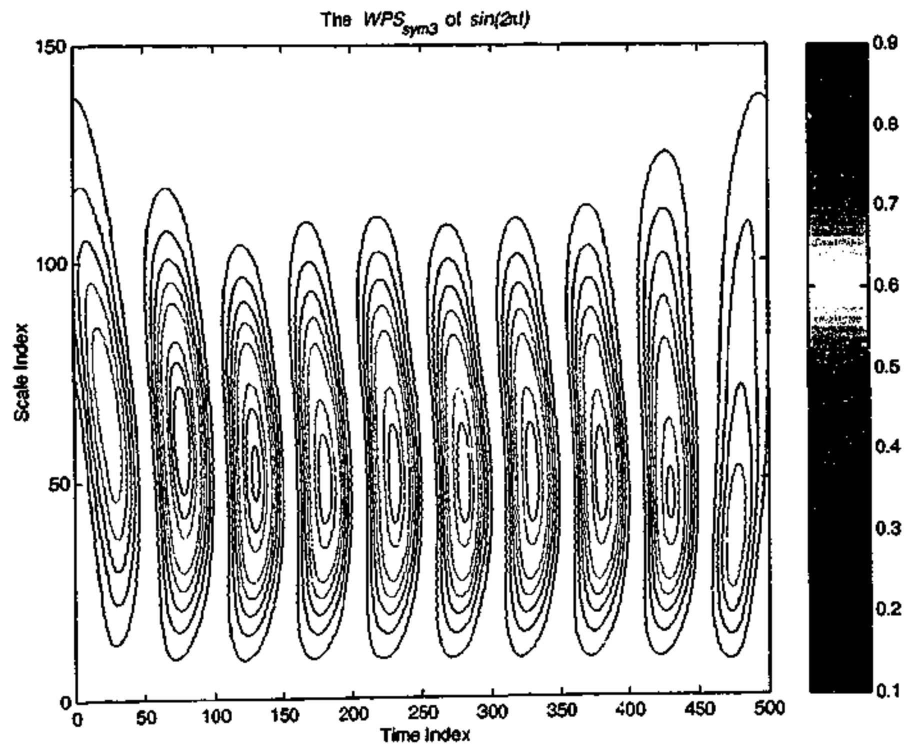

Figure 6.2.1: Contour plot of the WPS_{sym3} of $sin(2\pi t)$ signal

The WPS_{hyp} in Figure 6.2.2 shows distinctive peaks which indicate strong periodic behaviour of the signal as expected. Compared with the WPS_{sym3} given in Figure 6.2.1, the WPS_{hyp} requires a smaller number of scales which improves the computational efficiency of the hyperbolic time-frequency power spectrum. From Figure 6.2.1 and Figure 6.2.2, the scale ranges of a harmonic peak using the *sym3* and hyperbolic wavelets are approximately $20 \leq a_{sym3} \leq 130$ and $5 \leq a_{hyp} \leq 45$ respectively, where a_{sym3} and a_{hyp} are the scale indices of the *sym3* and hyperbolic wavelets respectively. As can be seen, the WPS_{sym3} and WPS_{hyp} are consistent which validates the hyperbolic wavelet and therefore the hyperbolic wavelet power spectrum technique.

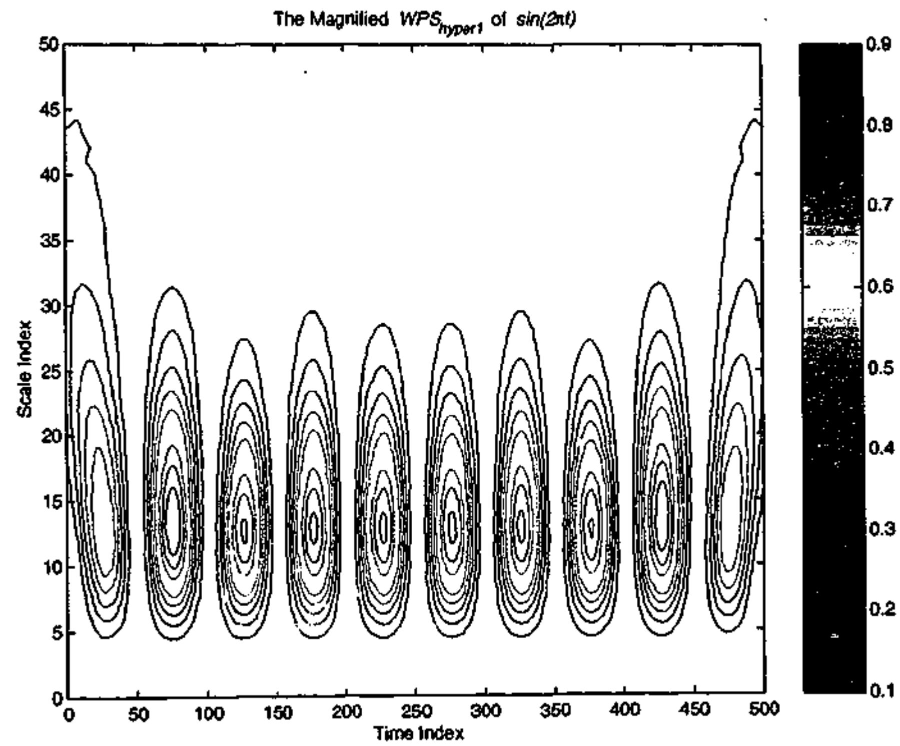

Figure 6.2.2: Contour plot of the WPS_{hyp} of $sin(2\pi t)$ signal

The relationship between the centre scale, a_{centre}, and the frequency of the signal, f_{signal}, is given by [99]

$$a_{centre} = \frac{f_{samp} \cdot \omega_0}{2\pi \cdot f_{wsamp}} \cdot \frac{1}{f_{signal}} \text{ or } f_{signal} = \frac{f_{samp} \cdot \omega_0}{2\pi \cdot f_{wsamp}} \cdot \frac{1}{a_{centre}} \tag{6.2.1}$$

where f_{wsamp}, f_{samp} and ω_0 are the wavelet sampling frequency, sampling frequency of the signal and centre frequency of the wavelet respectively.

From Eq. (6.2.1), the signal frequency f_{signal} can be estimated based on the centre scale a_{centre} specified by the wavelet power spectrum. The sampling frequency of the signal f_{samp} is usually about 100 times larger than the wavelet sampling frequency f_{wsamp} since the number of required sampling points for a wavelet function is much smaller than that for a signal as discussed in Chapter 4. For a sinusoidal input signal of $sin(2\pi t)$, using the *sym3* wavelet with the centre frequency $\omega_0 = 3$ rad/s, $a_{centre} \approx 50$, Figure 6.2.1 and Eq. (6.2.1), the signal frequency f_{signal} can be approximately estimated as

$$f_{signal} \approx \frac{100 \cdot 3}{2\pi} \cdot \frac{1}{50} \approx 0.96\ Hz \tag{6.2.2}$$

Similarly, the signal frequency f_{signal} of the signal $sin(2\pi t)$ can be approximately estimated using the hyperbolic wavelet with the centre frequency $\omega_0 = \beta = 1$ rad/s, $a_{centre} \approx 13$, Figure 6.2.2 and Eq. (6.2.1), we obtain

$$f_{signal} \approx \frac{100 \cdot 1}{2\pi} \cdot \frac{1}{13} \approx 1.2\ Hz \tag{6.2.3}$$

From the calculations performed in Eqs. (6.2.2) and (6.2.3), it is clear that the frequency of the signal can be estimated. It should be noted that the percent error could be as high as 20% (Eq. (6.2.3)) as in the case of the hyperbolic wavelet. Thus, quantitatively, the signal frequency can be more accurately estimated by using the Fourier power spectrum technique.

6.2.2 The Wavelet Power Spectrum of an Exponential Transient Signal

Exponential signals are common responses of first- and second-order linear circuits. This is the main reason why it is included in this research. The WPS_{sym3} of a transient signal $exp(-t)$ is given in Figure 6.2.3.

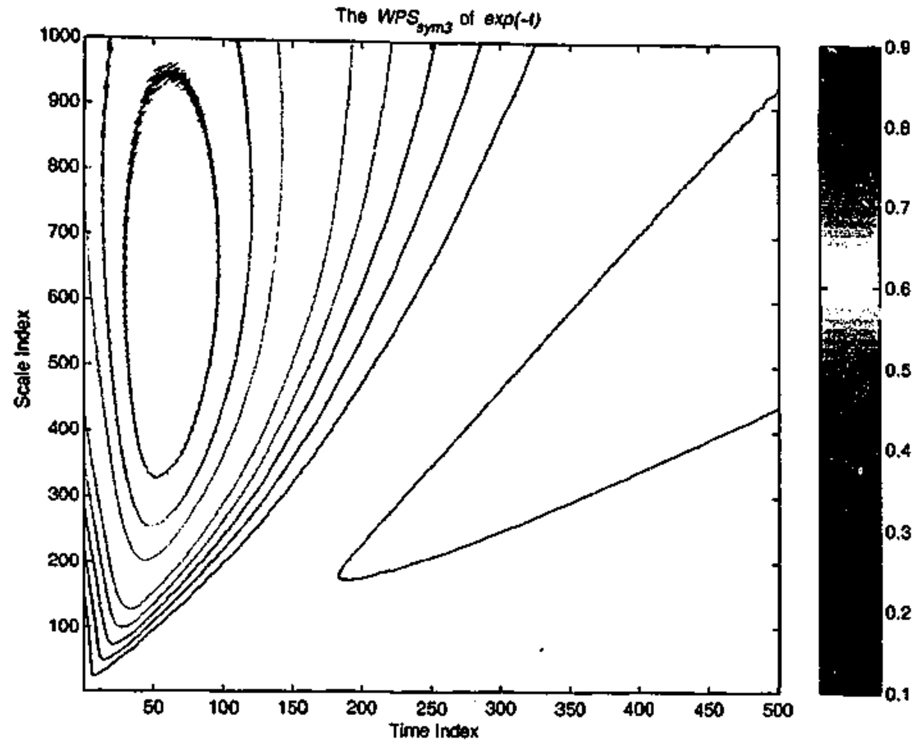

Figure 6.2.3: Contour plot of the WPS_{sym3} of the $exp(-t)$ signal

The WPS_{sym3} of an exponential signal $exp(-t)$ is broad and there seems to exist one distinctive peak represented by the smallest-radius contour curve with a large contour scale in Figure 6.2.3. The contour plot of the exponential transient signal is not closely spaced as it was in the case of a sinusoid. In fact, the number of closed contours is less than that of the sinusoid even though they both have dominant peak(s).

The decaying rate or time constant of the exponential signal is approximately the scale difference between the centres of two adjacent contours. From Figure 6.2.3, the scale difference between the inner-most contour curve and the second inner-most curve is about 60 which corresponds to the decaying rate of 0.8 by using Eq. (6.2.1). The accuracy of the estimation can be improved by taking the scale-difference average of all adjacent-contour pairs which yields the time constant of about 0.99. This is the expected time constant of the investigated exponential signal.

It should be noted that the energy of an exponential signal is not concentrated at one particular scale, but instead, spreading over a wide scale range as can be seen in Figure 6.2.3. The exponential energy tends to form closed contours, but this process appears to be very slow, i.e. contour curves have near-infinitely large radii which reflect the nature of exponential transient signals. The WPS_{hyp} is given in Figure 6.2.4.

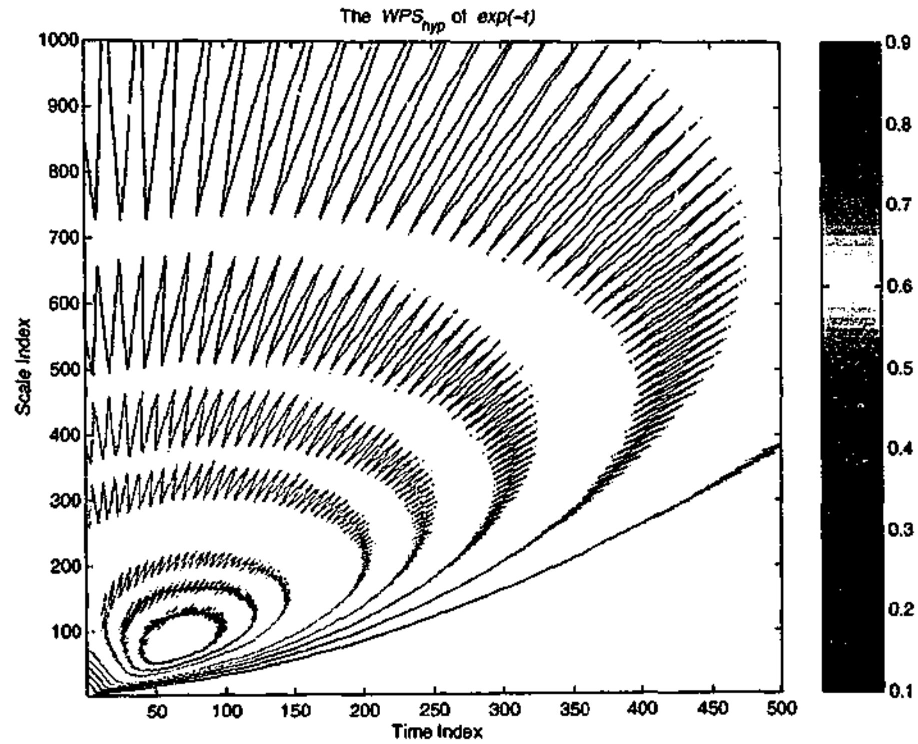

Figure 6.2.4: Contour plot of the WPS_{hyp} of an $exp(-t)$ signal

The time constant of the exponential signal can be similarly estimated using the same method applied to the hyperbolic wavelet. From Figure 6.2.4, the scale difference between two inner-most adjacent contour curves is approximately 20, which corresponds to the time constant of 0.78. By taking the scale-difference average, we obtain the estimated time constant of the exponential signal of about 0.95 which is close to the expected result.

Even though the time constant of a transient signal can be estimated with reasonable precision, it is still hard to estimate since the centre of each contour curve is sometimes hard to determine. For example, for the case of the *sym3* wavelet, since the centres of all contour curves cannot be clearly displayed, the distance between two adjacent curves along the vertical line is taken instead. For the hyperbolic wavelet, since most of the curves are clearly displayed, the scale difference between two adjacent centres can be effectively estimated. Secondly, this process is lengthy and tedious since the average value of the scale differences of all contour curves has to be calculated.

The WPS_{hyp} in Figure 6.2.4 is consistent with the WPS_{sym3} given in Figure 6.2.3 in which large-radius contours are detected. The peak is detected by the smallest contour curve. Even though the peak contour curve has a high scale, it is not filled which illustrates the main difference between sinusoidal signals and exponential signals. For the former, all peak contour curves are filled, whereas that is not the case for the latter. If the contour curves are filled, then the signal energy tends to be more concentrated around the peak which implies periodic characteristics. It should also be noted that the WPS_{hyp} is calculated over a smaller scale range than that of the WPS_{sym3}.

The compression ratio is estimated by counting the number of contour curves up to a certain scale. From Figure 6.2.3 and Figure 6.2.4, the scale ranges of the inner-most contour curves of the WPS_{sym3} and WPS_{hyp} are approximately $320 \le a \le 950$ and $50 \le a \le 120$ respectively. This yields the compression ratio in the range of about 6.4 and 7.9. The wavelet power spectra of periodic sinusoidal and exponential signals are used to establish a basis for further studies on other signals as shall be seen later.

6.2.3 The Wavelet Power Spectrum of an Exponentially Decaying Sinusoidal Signal

This section examines the *WPS* of an exponentially decaying sinusoidal signal $exp(-t)\cdot sin(2\pi t)$. The periodic and transient components of this signal were separately studied in Sections 6.2.1 and 6.2.2 respectively. The contour plot of its WPS_{sym3} is given in Figure 6.2.5. The WPS_{hyp} is similar to the WPS_{sym3} and is not given in this case.

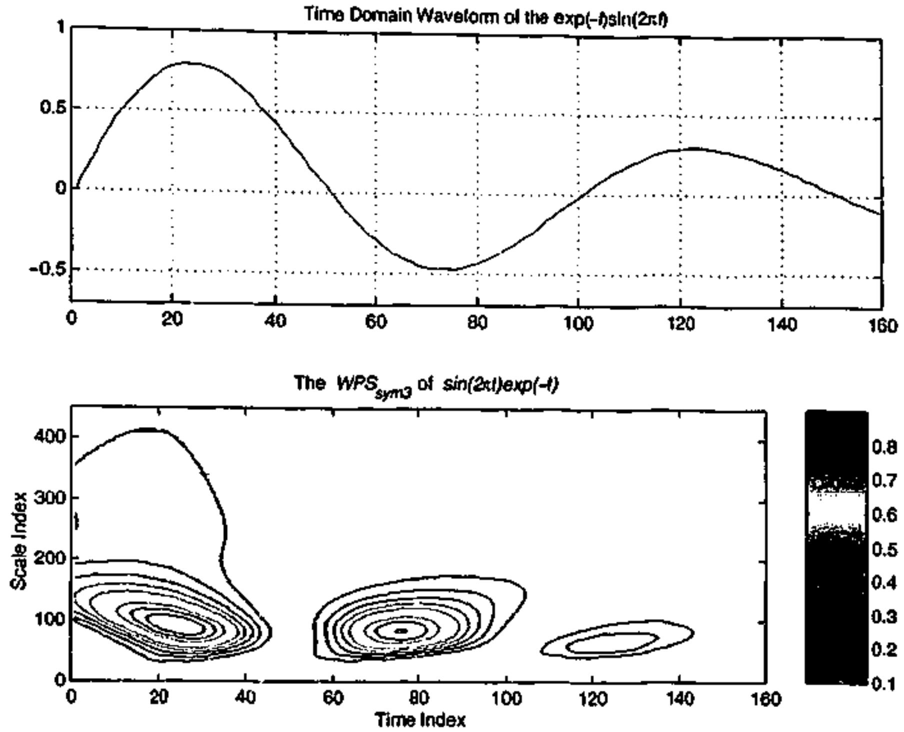

Figure 6.2.5: Time-domain waveform and contour plot of the WPS_{sym3} of an exponentially decaying sinusoid $exp(-t)\cdot sin(2\pi t)$. The WPS_{hyp} is very similar to the WPS_{sym3}.

It is expected that this signal have a combination of transient and periodic characteristics which hopefully can be detected by the *WPS* technique. From Figure 6.2.5, it is evident that the energy density is densely concentrated at the scale of about 100 for three harmonic peaks and decays to zero as the signal reaches steady state. There are three dominant and distinctive peaks in the signal whose positions correspond to those shown by the time-domain waveform. These peaks are clearly detected by using the *WPS* technique and indicated by three closed contours which represent periodic characteristics of the signal. The diminishing of energy as the signal reaches steady state indicates that the final value of the signal is zero. It should also be noted that the number of contour curves surrounding the peaks decrease as the time index increases which reflects transient characteristics in the signal. In particular, the number of contour curves in the third peak is only two at low scales compared with nine curves at high scales for the first and second peaks. Thus, it might be suggested that exponentially decaying sinusoidal signals with a zero final value can be well recognised by using the *WPS* technique.

If the final value of an exponentially decaying sinusoidal signal is non-zero, its *WPS* is expected to be broad since the peaks are now smoothened by a broad energy distribution of the DC component [41]. As a result, the *WPS* of an exponentially decaying sinusoid with a non-zero final value has only one major peak whose contours are not closely spaced as can be seen in Figure 6.2.6. It should be noted that the transient characteristics of the signal are indicated by large-radius contour curves as was seen in Section 6.2.2 for the case of an exponential signal.

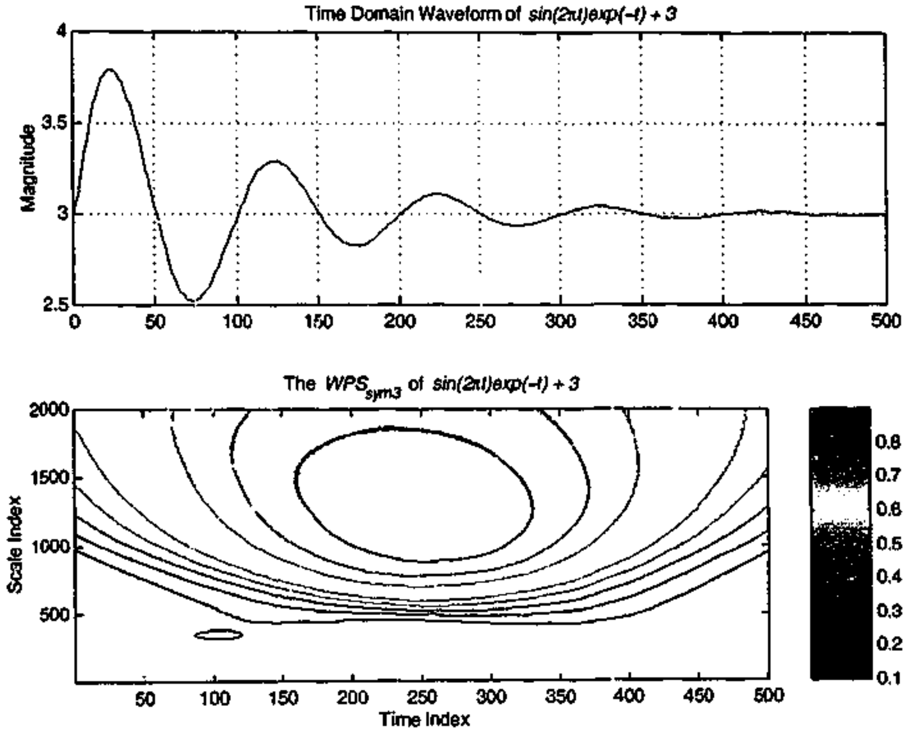

Figure 6.2.6: Contour plot of the WPS_{sym3} of the $sin(2\pi t)\cdot exp(-t) + 3$ signal

The main difference between an exponentially decaying sinusoid with a non-zero final value and an exponential signal is that for the former the radii of energy contours are not infinitely large but finite. This reflects periodic characteristics in the former signal whose *WPS* tends to form islands of closed contour curves. For comparison purposes, the WPS_{hyp} of the signal $sin(2\pi t)\cdot exp(-t) + 3$ is shown in Figure 6.2.7.

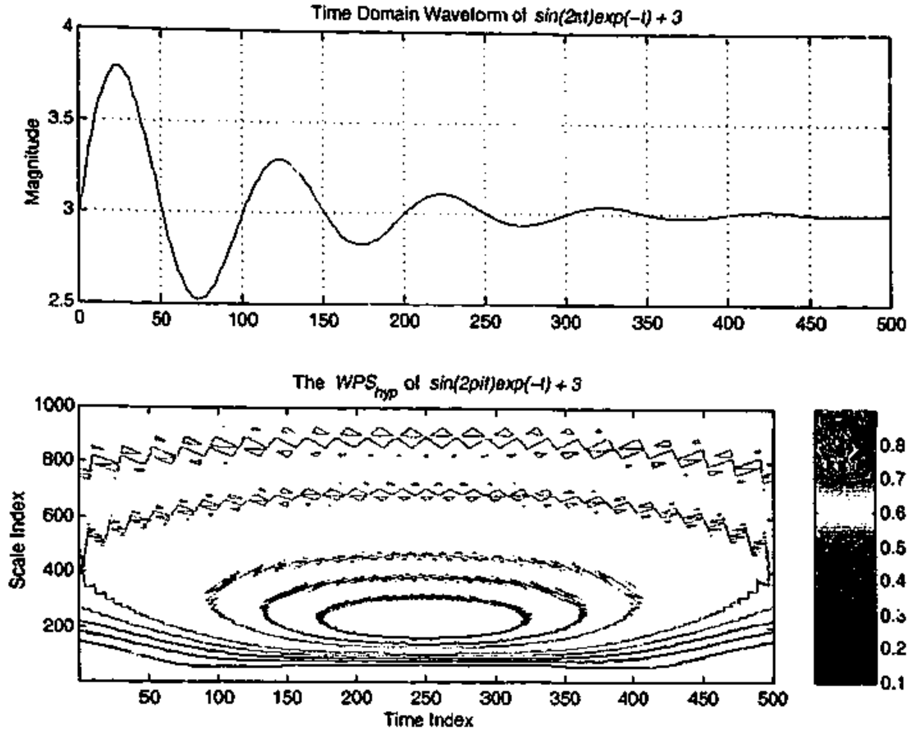

Figure 6.2.7: Time-domain waveform and contour plot of the WPS_{hyp} of the $sin(2\pi t)\cdot exp(-t)$ + 3 signal

From Figure 6.2.7, the upper parts of the contour curves have thick edges which indicate an energy-smearing phenomenon. It should not be incorrectly concluded that the signal is chaotic since the contour curves are densely located in the time-frequency plane which indicates discrete energy distribution. These features allow distinctive differentiation between exponentially decaying sinusoids with a non-zero final value and chaotic signals such as the ECG which will be examined later. The scale ranges in Figure 6.2.6 and Figure 6.2.7 are different since the WPS_{hyp} in Figure 6.2.7 is magnified so that its contour curves can be clearly displayed. It is clear that the hyperbolic wavelet is more efficient than the *sym3* wavelet in which more contour curves are displayed over the same scale range.

From Chapter 4, it was reported that the hyperbolic wavelet is most suitable for transient signals. By comparing Figure 6.2.6 and Figure 6.2.7, it is clear that the WPS_{hyp} can display more contour curves than the WPS_{sym3} due to the former has a finer scale resolution and a smaller total number of calculated scales. This fact was also shown in Section 4.3.6 when calculating the hyperbolic, Morlet and Choi-Williams wavelet power spectra of the English vowel "e".

In addition, by comparing Figure 6.2.3, Figure 6.2.4, Figure 6.2.6 and Figure 6.2.7, it might be suggested that the hyperbolic wavelet is a symmetrical function whose wavelet power spectra are perfectly symmetrical about the vertical line. Thus, graphical representations of the hyperbolic wavelet power spectra are better displayed than using the *sym3* wavelet. As a result, it is easier to differentiate the exponential signal from the exponentially decaying sinusoidal signals with zero and non-zero values by using the hyperbolic wavelet power spectrum. Further, one advantage of the hyperbolic wavelet over the *sym3* wavelet is that the former can reveal more information over an identical scale range than the latter due to the compression ability of the hyperbolic wavelet of having a smaller total number of calculated scales. This also increases the efficiency of the hyperbolic wavelet power spectrum calculation process.

6.2.4 The Wavelet Power Spectrum of Duffing Oscillator

Duffing oscillator has been popular in signal processing because of its simplicity [1]. In this section, the Duffing oscillator is studied by calculating its wavelet power spectra of Periods 1, 2, 4 and chaotic state. From this, it is possible to determine how its energy is distributed and therefore deducing the system characteristics and detecting possible transition(s) into the chaotic region.

The equation governing Duffing oscillator is given as [1]

$$\ddot{u} + \gamma\dot{u} - 0.5(u - u^3) = Fcos(\omega t) \tag{6.2.4}$$

where $\gamma = 0.168$, $\omega = 1$, $u(t)$ is the displacement function of the time t and F is the driving function. For Period 1, $F_1 = 0.05$, Period 2, $F_2 = 0.178$, Period 4, $F_4 = 0.197$ and for chaotic state, $F_{chaos} = 0.21$. The initial conditions used for the system were $[u \;\; \dot{u}] = [0 \;\; 1]$.

6.2.4.1 Duffing Period 1

Duffing Period 1 waveform can be regarded as a genuine periodic signal whose energy is concentrated over the high-frequency range (low scale range). Duffing Period 1 time-domain waveform and its *WPS* using the hyperbolic and *sym3* wavelets are given in Figure

6.2.8. It should be noted that the energy is uniformly and repetitively distributed over the whole range of the time index which means each segment of the input data points is almost identical. This means that there is no degree of disorder or chaos in Duffing Period 1 waveform.

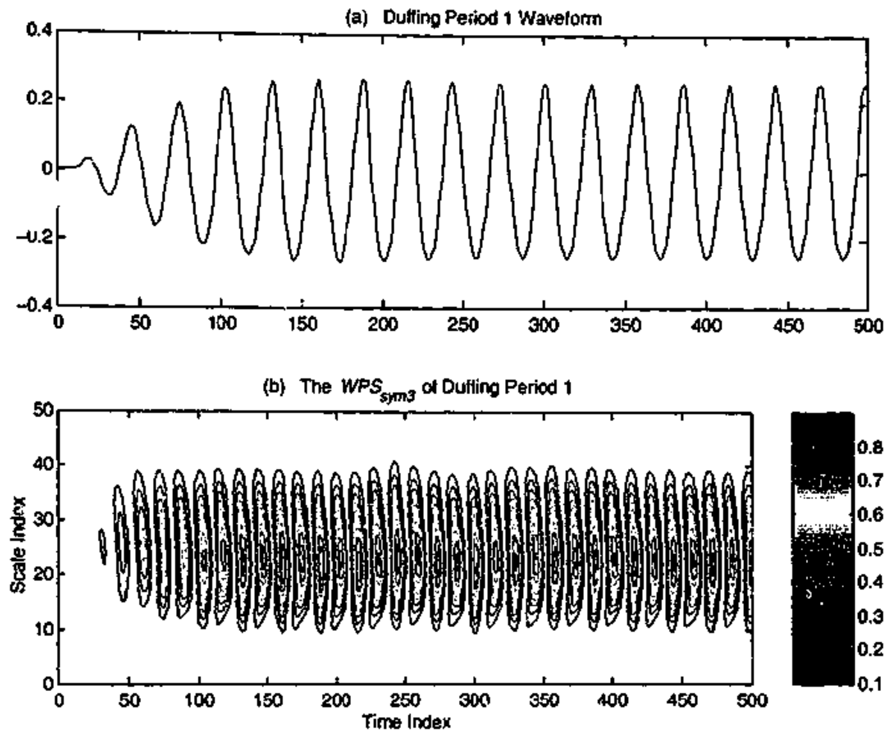

Figure 6.2.8: Time-domain waveform and contour plot of the WPS_{sym3} of Duffing Period 1. The WPS_{hyp} of Duffing Period 1 waveform is very similar to the WPS_{sym3} except that it is displayed on a different scale range.

From Figure 6.2.8, periodic (harmonic) peaks of Duffing Period 1 waveform are clearly separated which is similar to the case of the sinusoidal signal $sin(t)$ studied in Section 6.2.1. Apart from a minor drift of the signal at the beginning where the time index is roughly $b \leq 300$, there is no difference in the amount of energy density over time, which suggests that Duffing Period 1 waveform is near periodic. The duration during which the periodic peak occupies is short which suggests that the signal energy is highly concentrated. This fact has been well known and extensively reported in the literature [1]. Thus the correctness and consistency of the proposed *WPS* technique are validated. It should be noted that the symbol b is used as the time index as explained earlier in Section 6.1. The symbol a, which was used in Section 6.2.1, will be used as the scale index. These notations will be conveniently used in this chapter and the next chapter.

6.2.4.2 Duffing Period 2

The driving force used for Duffing Period 2 waveform in Eq. (6.2.4) is $F_2 = 0.178$. The time-domain waveform and its WPS_{sym3} are displayed in Figure 6.2.9. The WPS_{hyp} of Duffing Period 2 waveform is given in Figure 6.2.10.

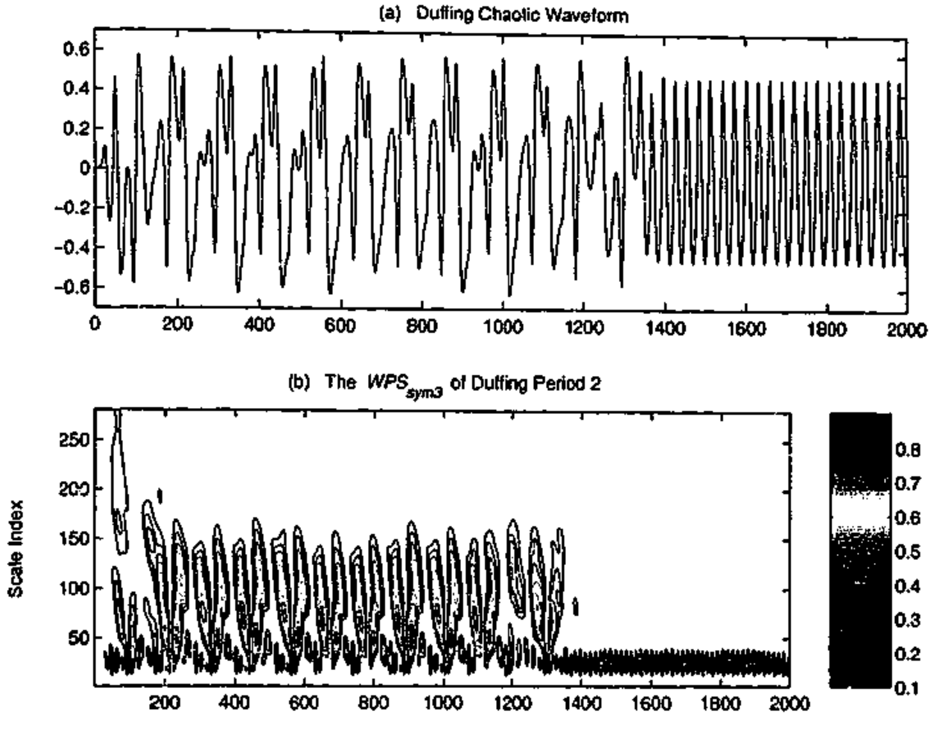

Figure 6.2.9: Time-domain waveform of Duffing Period 2 and contour plot of its WPS_{sym3}

The time-domain waveform of Duffing Period 2 shows an early sign of deviation from periodicity in which the number of detected sub-harmonics is large. The signal regains its near-periodic characteristics at the time index $b \approx 1{,}400$. For $b \geq 1{,}400$, the signal exhibits similar characteristics to those of Duffing Period 1 waveform which suggests that periodicity is dominant. However, since its contour scale is lower than that of Duffing Period 1 waveform, it can be suggested that Duffing Period 2 waveform does not completely regain its periodicity as can be seen in Figure 6.2.9. Figure 6.2.11 and Figure 6.2.12 show the magnified contour plot of the WPS_{sym3} and WPS_{hyp} for $2{,}000 \geq b \geq 1{,}000$ and $0 \leq b \leq 1{,}100$ respectively.

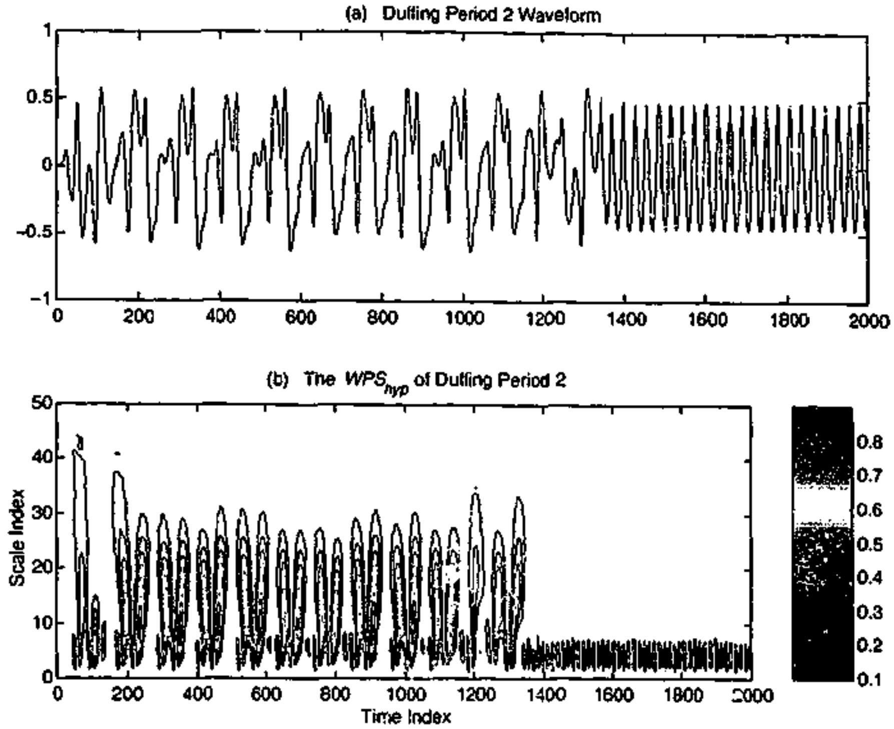

Figure 6.2.10: Time-domain waveform and contour plot of the WPS_{hyp} of Duffing Period 2

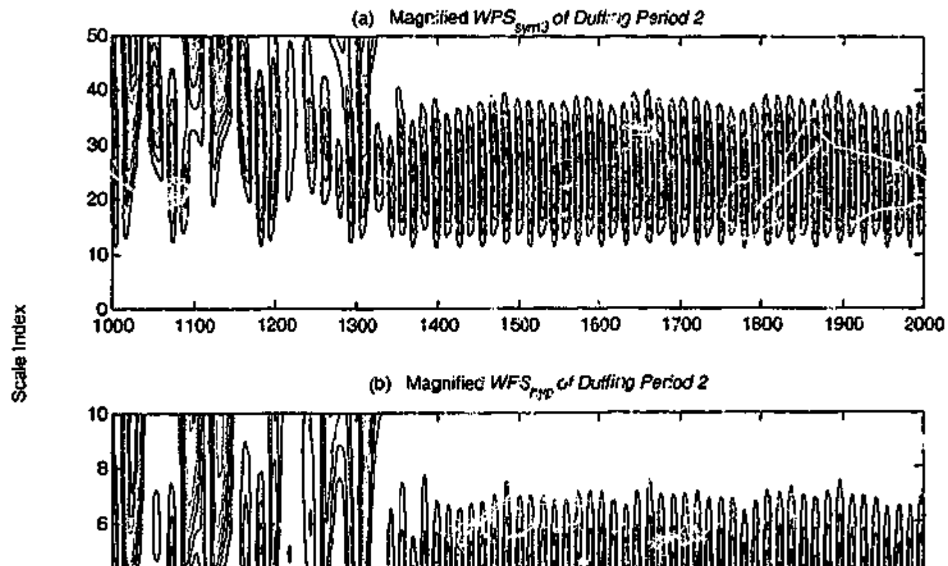

Figure 6.2.11: Magnified contour plot of Figure 6.2.9 and Figure 6.2.10 for $2{,}000 \geq b \geq 1{,}000$

The near periodicity of Duffing Period 2 waveform is detected by a series of closed and filled contours (at high scales) with repetitive patterns over time. Although there are discrete and filled contours, more sub-harmonics are detected by the *WPS* technique in Duffing Period 2 waveform than in Duffing Period 1 waveform. In addition, the contour scale of the wavelet power spectrum of Duffing Period 2 waveform is lower than that of Duffing Period 1 waveform which suggests that the former has a broader energy distribution than that of the latter.

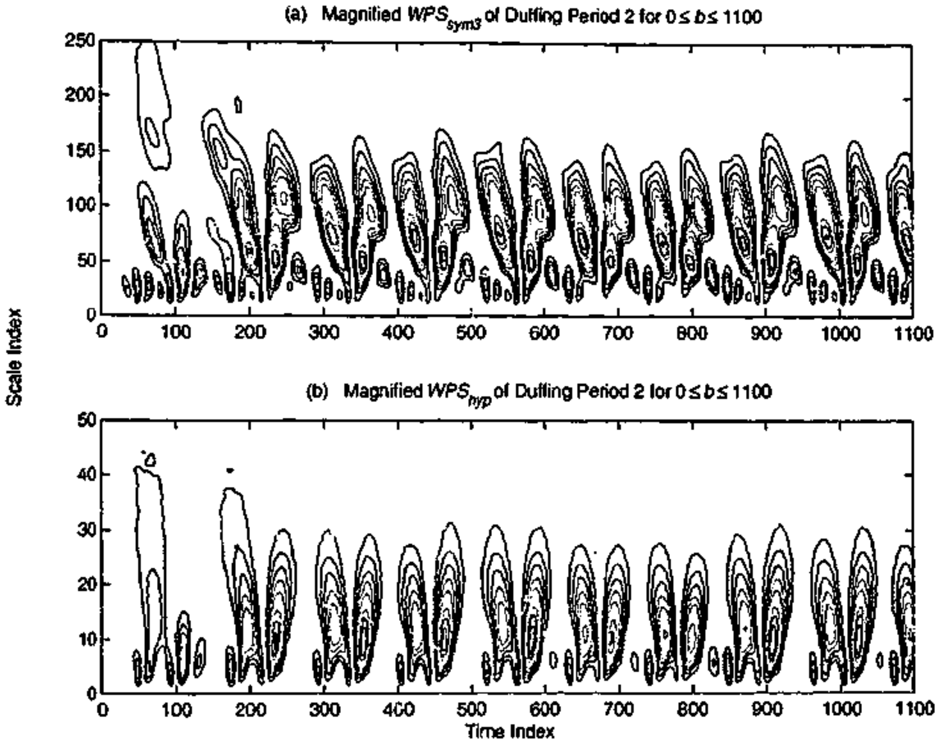

Figure 6.2.12: Magnified contour plots of the WPS_{sym3} and WPS_{hyp} of Duffing Period 2 waveform for $0 \le b \le 1{,}100$. The colour contour scale is similar to those given in Figure 6.2.9 and Figure 6.2.10.

From Figure 6.2.12(b), it is clear that for the hyperbolic wavelet, main details of the WPS_{hyp} are successfully displayed. However, minor fine details are missing, i.e. sub-harmonics are not clearly shown as in Figure 6.2.12(a), which are mainly due to the effects of having a small total number of calculated scales or the compression ability of the hyperbolic wavelet. However, more importantly, the WPS_{sym3} and WPS_{hyp} are consistent and can be employed to successfully detect periodicity and deviation from periodicity of Duffing Period 2 waveform.

6.2.4.3 Duffing Period 4

Duffing Period 4 waveform can be regarded as a transition state from periodicity to chaos [1] which hopefully will be detected by using the wavelet power spectrum technique. The time-domain waveform and its WPS_{sym3} of Duffing Period 4 are given in Figure 6.2.13. The WPS_{hyp} is given in Figure 6.2.14.

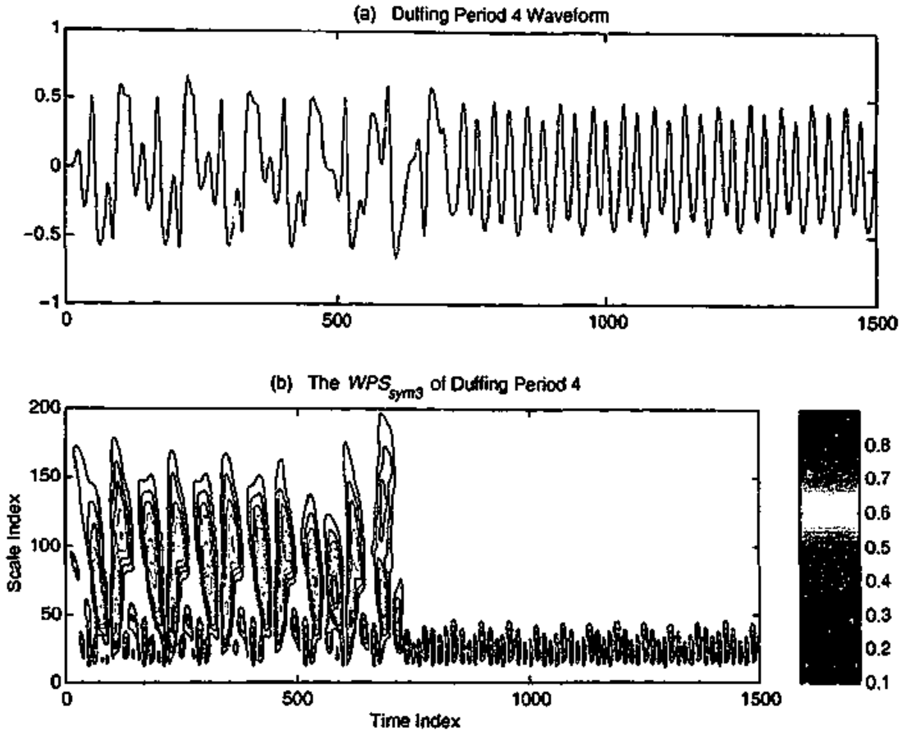

Figure 6.2.13: Duffing Period 4 time-domain waveform and contour plot of its WPS_{sym3}

Similar to Duffing Period 2 waveform, Duffing Period 4 waveform exhibits early deviation from periodicity with 2 major harmonics and a number of sub-harmonics for $b \leq 700$. Transitions into the chaotic region is signalled by continuous closed contours at $b \approx 700$ with a wider scale range as can be seen in Figure 6.2.13. The continuity of energy indicates that the waveform has entered into the chaotic region. However, chaotic components of the waveform are not strong enough since periodic components are still present. For $b \geq 700$, the signal partially regains its periodicity with low contour-scale curves. All of the above features can be clearly seen in the magnified contour plots of the WPS_{sym3} and WPS_{hyp} shown in Figure 6.2.15.

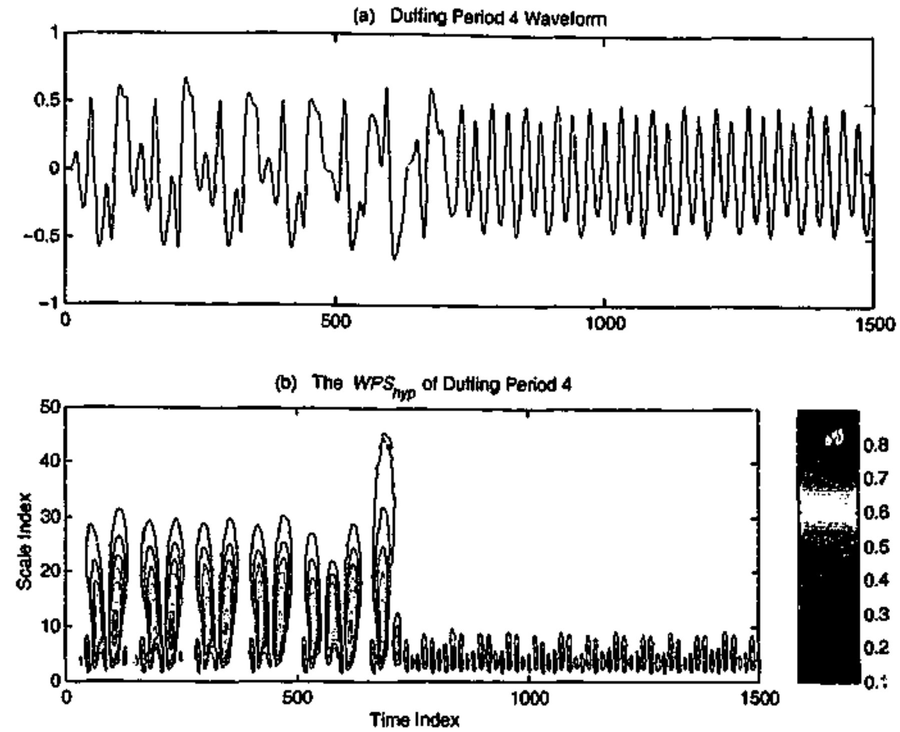

Figure 6.2.14: Time-domain waveform and contour plot of the WPS_{hyp} of Duffing Period 4

The main difference of Duffing Period 4 waveform from Duffing Period 2 waveform is that after the signal regains its periodicity, a number of sub-harmonics at low scales are still detected in the former. Whereas, for the latter, repetitive energy patterns at high contour scales are detected which are similar to Duffing Period 1 waveform. This clearly shows that Duffing Period 2 waveform is more periodic and stable than Duffing Period 4 waveform.

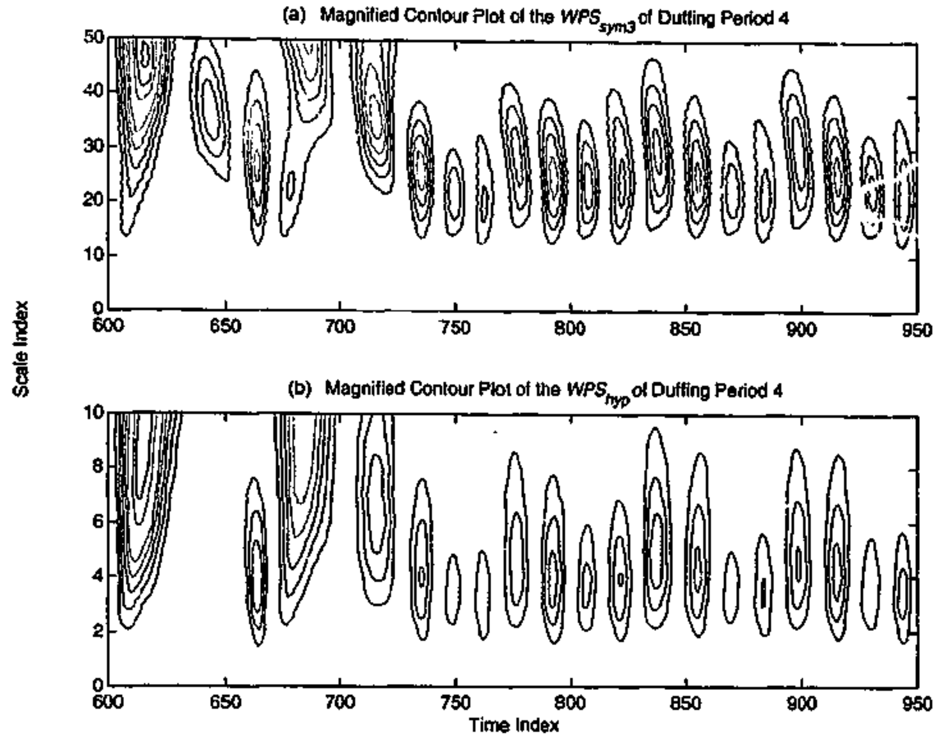

Figure 6.2.15: Magnified contour plots of the WPS_{sym3} and WPS_{hyp} of Duffing Period 4 waveform which shows the transition into the chaotic region at $b \approx 700$. For $b \geq 700$, Duffing Period 4 waveform partially regains its periodicity by having repetitive closed contour curves. However, these curves have low contour scales which suggests that their energy is not densely distributed. Thus, the waveform is vulnerable to chaotic behaviour.

From Figure 6.2.13 and Figure 6.2.14, it can be seen that the hyperbolic and *sym3* wavelet power spectra of Duffing Period 4 waveform are consistent in which the transition into the chaotic region is detected at $b \approx 700$. For $b < 700$, near-periodic behaviour is detected in the waveform and for $b > 700$ a mixture of chaotic and periodic components are detected. However, due to the compression effects of the hyperbolic wavelet, some sub-harmonics are suppressed. It is important to note that by using the wavelet power spectrum time-frequency technique, it is possible to determine when the waveform enters into the chaotic region hence the transition region of the waveform can be clearly identified.

6.2.4.4 Duffing Chaotic

For Duffing chaotic, the driving force has the value of $F_{chaos} = 0.21$. The time-domain waveform and its WPS_{sym3} are shown in Figure 6.2.16. The WPS_{hyp} is given in Figure 6.2.17. The magnified versions of these figures are given in Figure 6.2.18 and Figure 6.2.19 respectively.

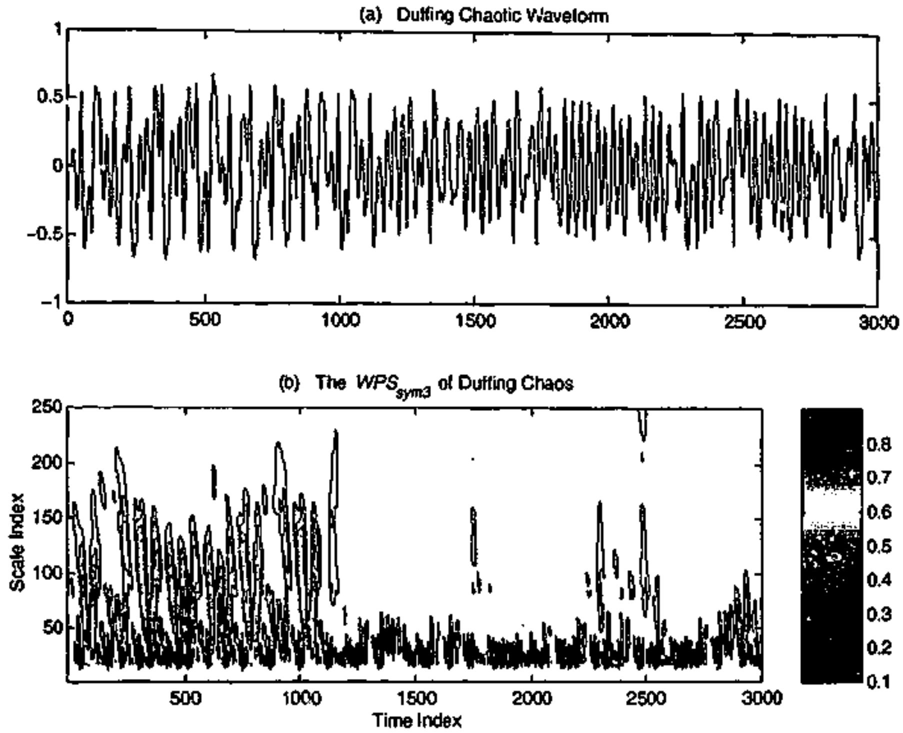

Figure 6.2.16: Time-domain waveform of Duffing chaotic and contour plot of its WPS_{sym3}

From Figure 6.2.16 and Figure 6.2.17, Duffing chaotic waveform exhibits early deviation from periodicity as was detected in Duffing Period 2 and 4 waveforms. However, instead of regaining its periodicity, Duffing chaotic waveform remains chaotic after the transition into its chaotic region. Duffing chaotic state is signalled by a non-repetitive and broad energy distribution in which distinctive peaks are unevenly distributed for time index less than 1,100, which is the transition region of the waveform. All of these features can be seen in Figure 6.2.18(a) and Figure 6.2.19(a).

From Figure 6.2.16 and Figure 6.2.18, for time index greater than 1,100, it is evident that the waveform has entirely entered into the chaotic state in which its fundamental periodic components have disappeared as compared to Duffing Period 1, 2 and 4 waveforms. The energy is unevenly distributed over the scale range of 20 to 250. There is no particular energy concentration in any region of the spectrum (due to low contour scales), which is broadly distributed (approximately at the time indices of 1,100, 1,700 and 2,500 in Figure 6.2.17(b)), which strongly suggests that the waveform is chaotic. It should be noted that in this case, even though the curves have low contour scales, they still represent repetitive energy patterns which means that for time index greater than 1,100 Duffing chaotic waveform still possesses weak periodicity. Thus, it can be suggested that Duffing chaotic waveform is superposition of periodic and chaotic components. The main difference between Duffing Period 4 waveform and Duffing chaotic waveform is that there are no disordered energy patterns in the former, whereas the contour scales of the latter vary with time which suggests disordered characteristics. Figure 6.2.18(b) and Figure 6.2.19(b) clearly show the magnified energy distribution of the waveform for time index greater than 1,100.

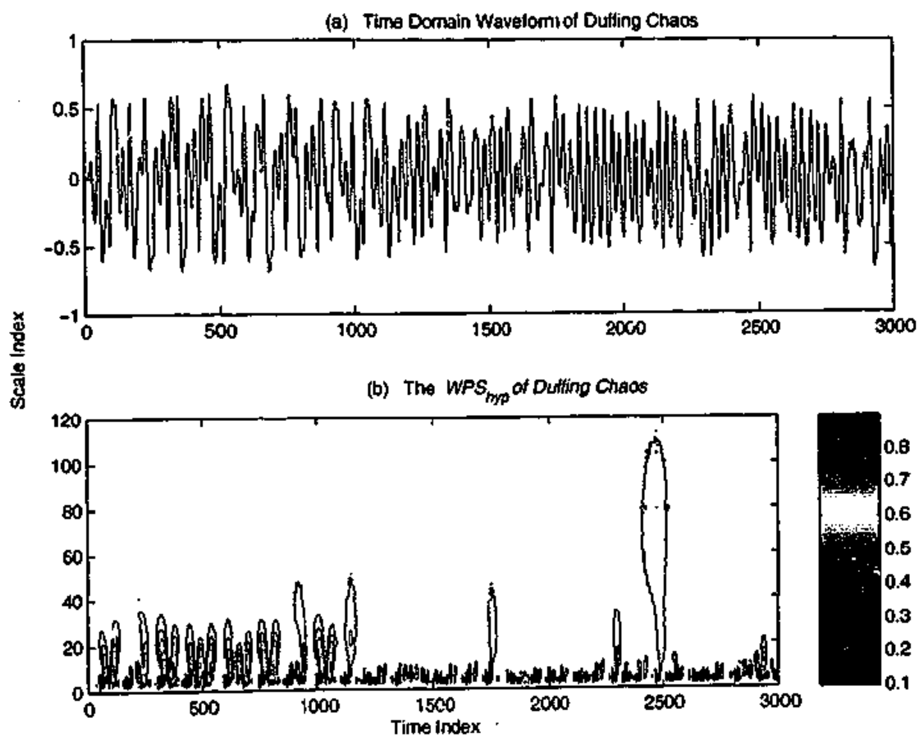

Figure 6.2.17: Time-domain waveform and contour plot of the WPS_{hyp} of Duffing chaotic

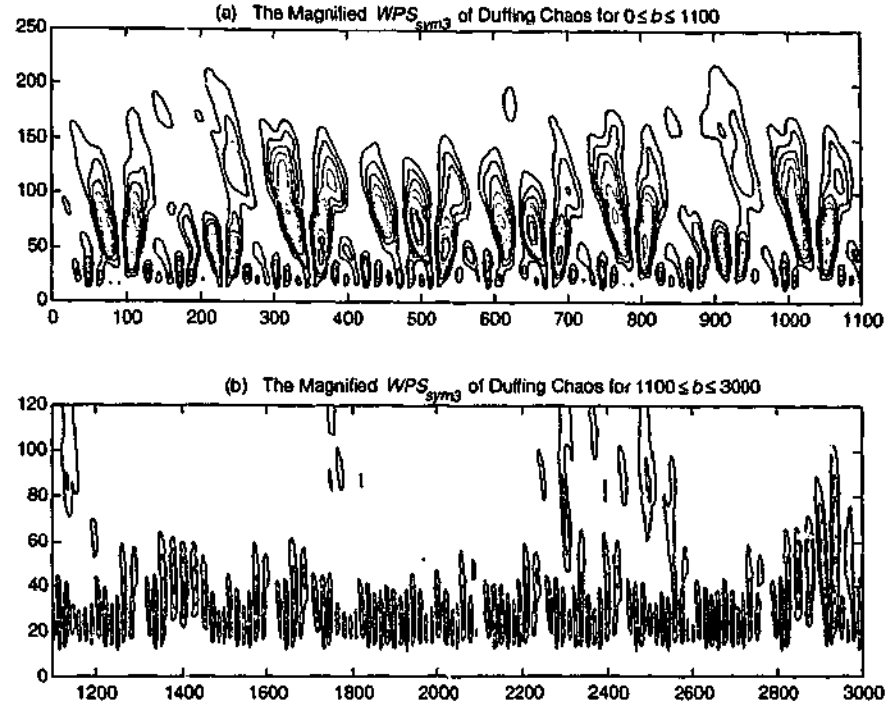

Figure 6.2.18: Magnified version of Figure 6.2.16(b), the horizontal and vertical axes are "Time Index" and "Scale Index" respectively as was in Figure 6.2.17

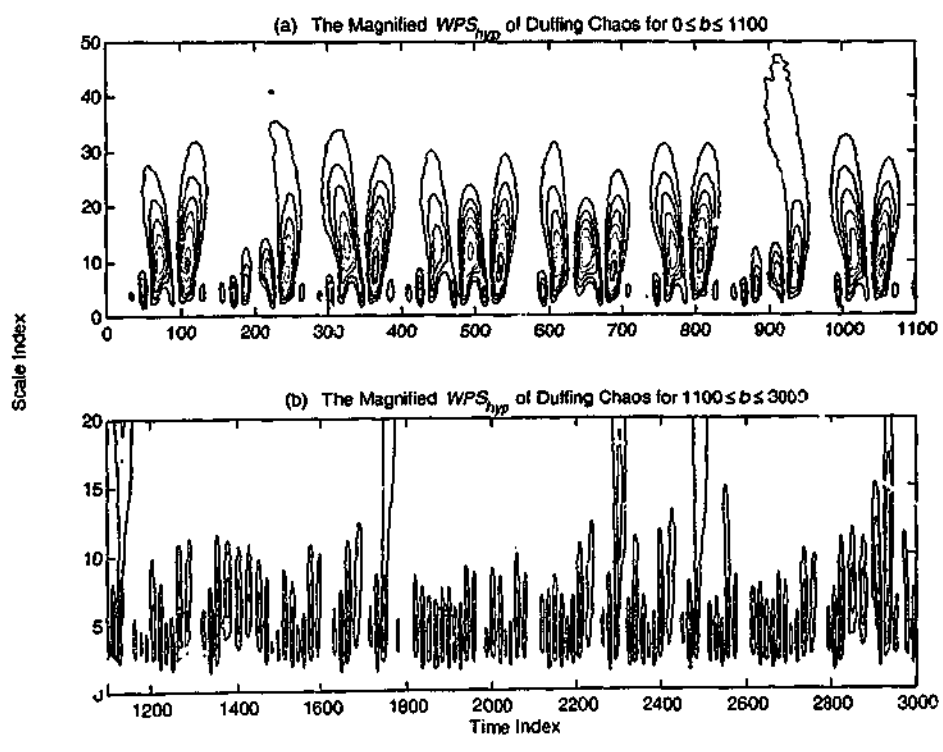

Figure 6.2.19: Magnified version of Figure 6.2.17(b)

Sections 6.2.1–6.2.4 examined signals whose characteristics have been well known. The next section calculates the WPS_{sym3} and WPS_{hyp} of the ECG signal. The Fourier power spectrum technique will also be used so that the effectiveness of the *WPS* technique can be assessed.

6.2.5 The Wavelet Power Spectrum of ECG Signal

The practical ECG (Human Electrocardiogram) is examined in this section due to its importance in medical diagnosis. The time-domain waveform of the ECG signal is displayed in Figure 6.2.20.

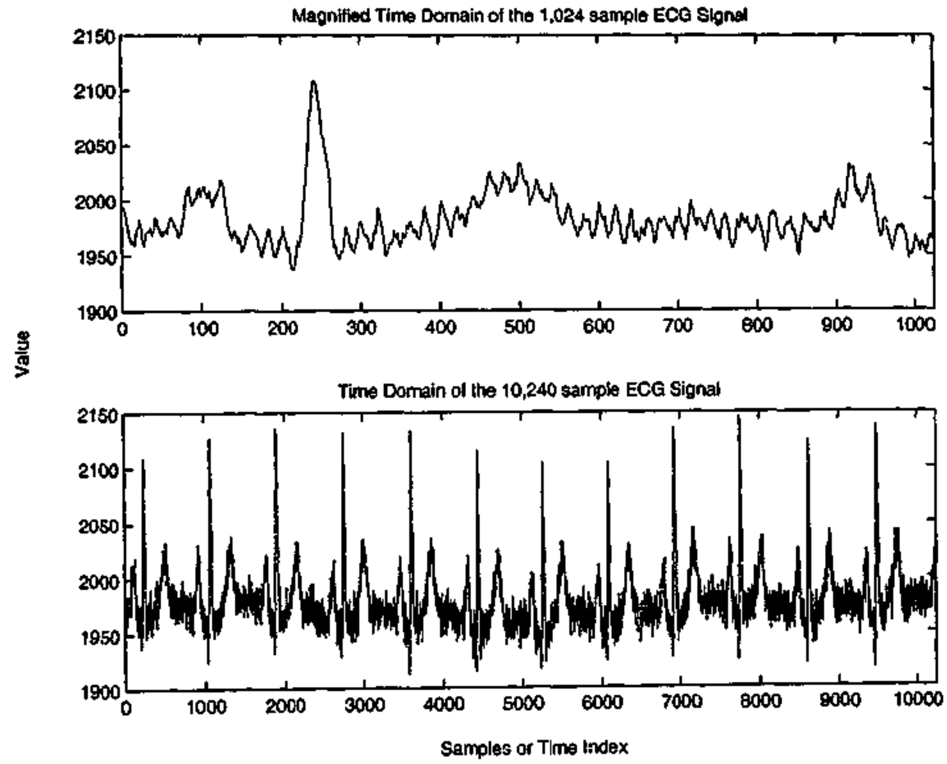

Figure 6.2.20: Time-domain waveform of the ECG

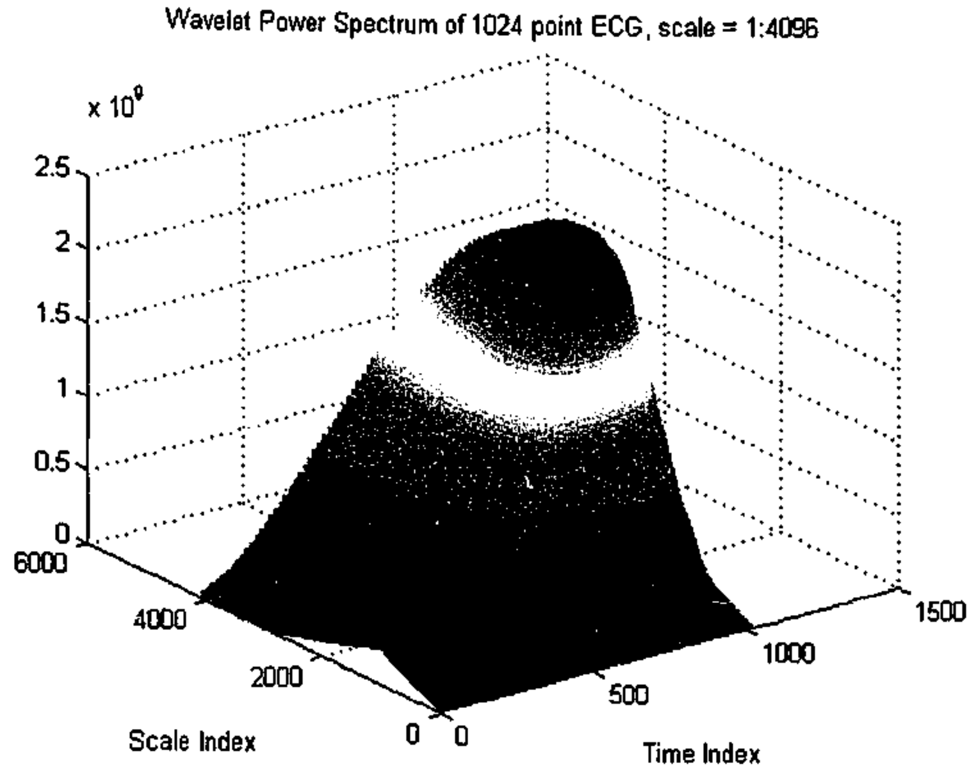

Figure 6.2.21: 3-D mesh plot of the WPS_{sym3} of 1024-point ECG

The WPS_{sym3} of ECG signal, which is shown in Figure 6.2.21 in 3-D plot and Figure 6.2.22 in contour plot, is similar to that of the exponential signal shown in Section 6.2.2 and quite similar to the WPS_{sym3} of the exponentially decaying sinusoid with a non-zero DC component shown in Section 6.2.3. For periodic and exponentially decaying sinusoidal signals, their wavelet power spectra all have dominant peaks and large-radius contours. The WPS_{sym3} of the ECG does not have identifiable peaks and its contours have very large radii which suggests that the ECG might be a transient type. The energy distribution of the ECG is spread over a wide scale range as can be seen in Figure 6.2.22. In addition, ECG contours are not sharp but thick in width.

The scale range of contour curves is worked out by estimating the corresponding scales of the lowest and the highest contours. For example, in Figure 6.2.22, the scale range of the *sym3* wavelet power spectrum will be $900 \leq a$, *not* $2{,}000 \leq a$ since the lower contour stretches down to the scale of 900 and 2,000 is its starting point. This method has been used in this chapter and Chapter 7 to work out the scale range of various wavelet power spectra.

From Figure 6.2.21, the ECG energy distribution is smooth and there are no abrupt changes in its energy density over the high- and low-valued region of the scale and time indices. By comparing the WPS_{sym3} in Figure 6.2.6 (the exponentially decaying sinusoid with a non-zero final value) and Figure 6.2.22 (ECG signal), it is evident that the energy density of the ECG is broader and distributed over a larger scale range of $900 \leq a \leq 3{,}500$ than that of an exponentially decaying sinusoid with a non-zero DC component ($500 \leq a \leq 1{,}800$). In addition, the inner-most contour curve has the scale range of $1{,}800 \leq a \leq 3{,}700$ for the ECG and $900 \leq a \leq 1{,}800$ (smaller scale range than the ECG) for the exponential signal, which suggests that the former might exhibit chaotic characteristics. However, since the energy patterns provided by the wavelet power spectral technique are almost similar, it is difficult to distinguish between these signals.

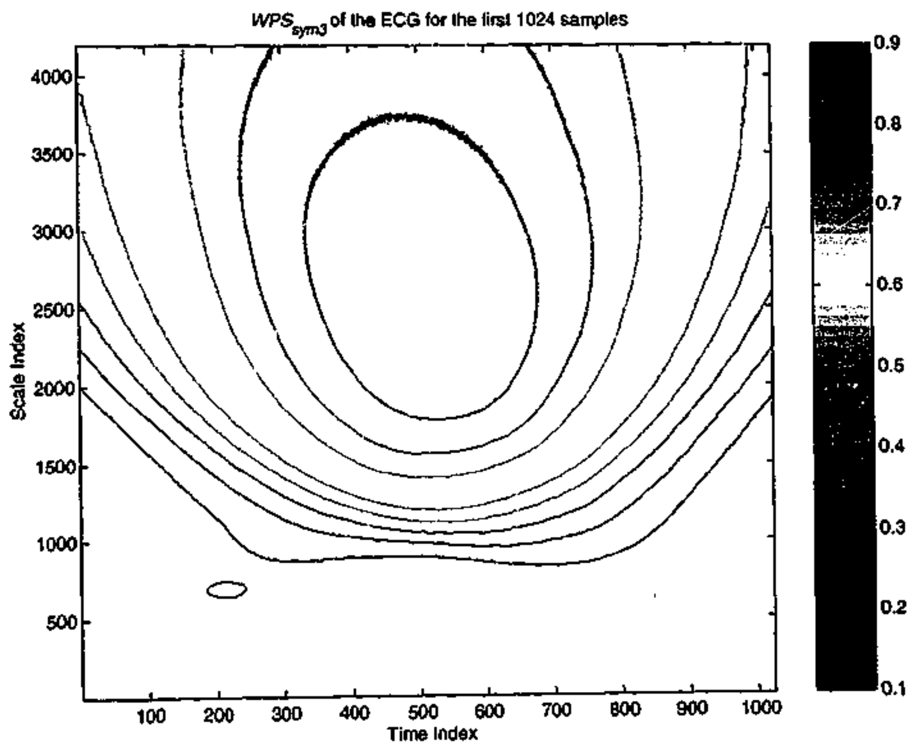

Figure 6.2.22: Contour plot of the WPS_{sym3} of 1,024-point ECG signal

For comparison purposes, the WPS_{hyp} is shown in Figure 6.2.23 in which it is symmetrically displayed. From Section 6.2.3, the WPS_{hyp} of an exponentially decaying sinusoid with a non-zero final value is very similar to that of the ECG. The only difference between the two wavelet power spectra is that the occupied scale range of the ECG ($120 \leq a \leq 1{,}800$ in Figure 6.2.23) is wider than that of an exponentially decaying sinusoidal signal with a non-zero final value ($70 \leq a \leq 850$ in Figure 6.2.7). This suggests that the ECG energy distribution is broader than that of the exponentially decaying sinusoid signal as explained earlier. In addition, the wavelet power spectrum of the exponential signal has more contour curves over the same scale range than those of the ECG wavelet power spectrum which further supports the above statement.

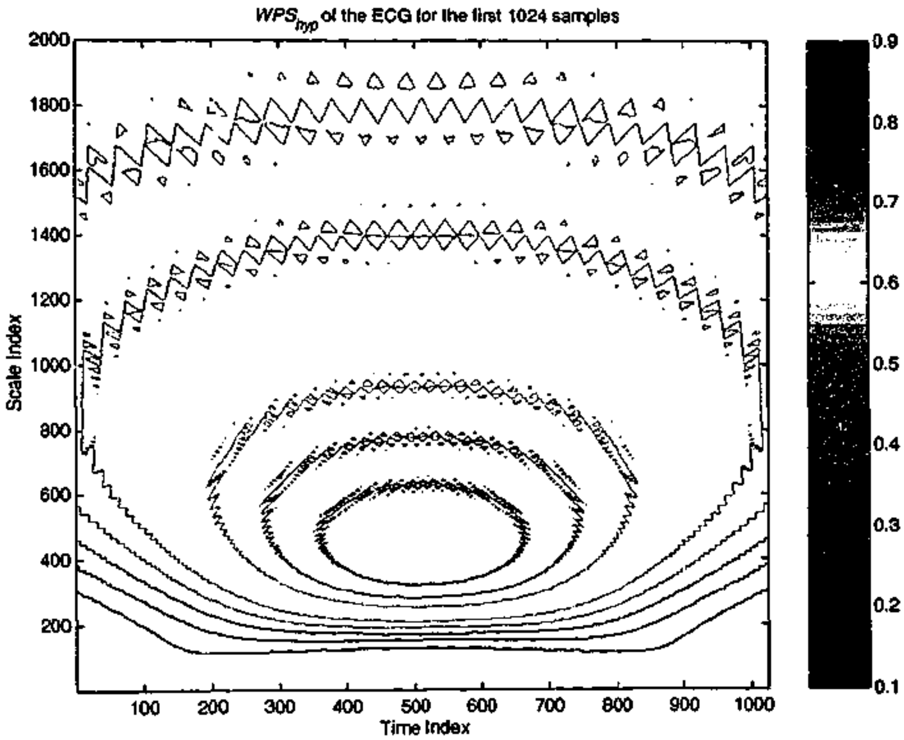

Figure 6.2.23: Contour plot of the WPS_{hyp} of the ECG

The scale ranges of the ECG and an exponentially sinusoidal signals, using the hyperbolic wavelet, are clearly different compared with those observed using the near-symmetric *sym3* wavelet which makes the differentiation of these signals easier. This indicates the usefulness of a perfectly-symmetrical wavelet such as the hyperbolic wavelet. Thus, the hyperbolic wavelet power spectrum technique can be used to study the ECG signal by examining the occupied scale range of its wavelet power spectrum. In general, however, the *WPS* technique is not very effective in studying the ECG signal since it cannot clearly differentiate between an exponentially decaying signal and the ECG signal due to their identical wavelet power spectra. This is a disadvantage of the *WPS* technique compared with the Fourier power spectrum technique as will be further explored in Chapter 7.

Although the wavelet power spectra of the ECG and $exp(-t)\cdot sin(t)+3$ signals are very similar, their Fourier power spectra, given in Figure 6.2.24 and Figure 6.2.25 respectively, are quite different which shows that the Fourier technique is more effective than the *WPS* technique in this case. Thus, depending on the particular application, the wavelet power spectrum technique or the Fourier power spectrum technique should be used to reveal the signal characteristics. In this case, it might be suggested that the ECG signal can be effectively studied using the Fourier power spectrum technique. This conclusion is consistent with the findings in [4]. However, transient signals such as $exp(-t)$ and $exp(-t)\cdot sin(t)+3$ can be successfully studied by employing the wavelet power spectrum technique since their Fourier power spectra are quite similar as can be seen in Figure 6.2.25.

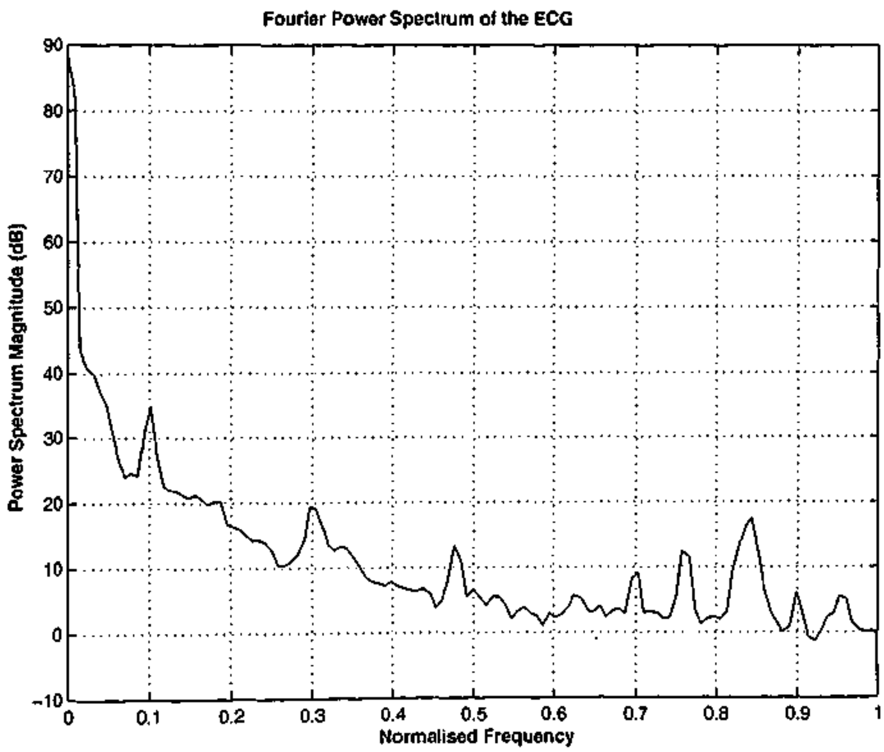

Figure 6.2.24: Fourier power spectrum of ECG signal. The frequency is normalised by dividing every frequency bin by the largest frequency bin in the signal spectrum. This method of frequency normalisation is used in this chapter and also in Chapter 7.

From Figure 6.2.24, it is evident that the ECG signal has a broad power spectrum which suggests that it is not periodic. This is consistent with the conclusions drawn by using the *WPS* technique from Figure 6.2.21 to Figure 6.2.23. Transient characteristics of the ECG signal, however, are not effectively detected by using the Fourier power spectrum technique, but can be detected by using the *WPS* technique.

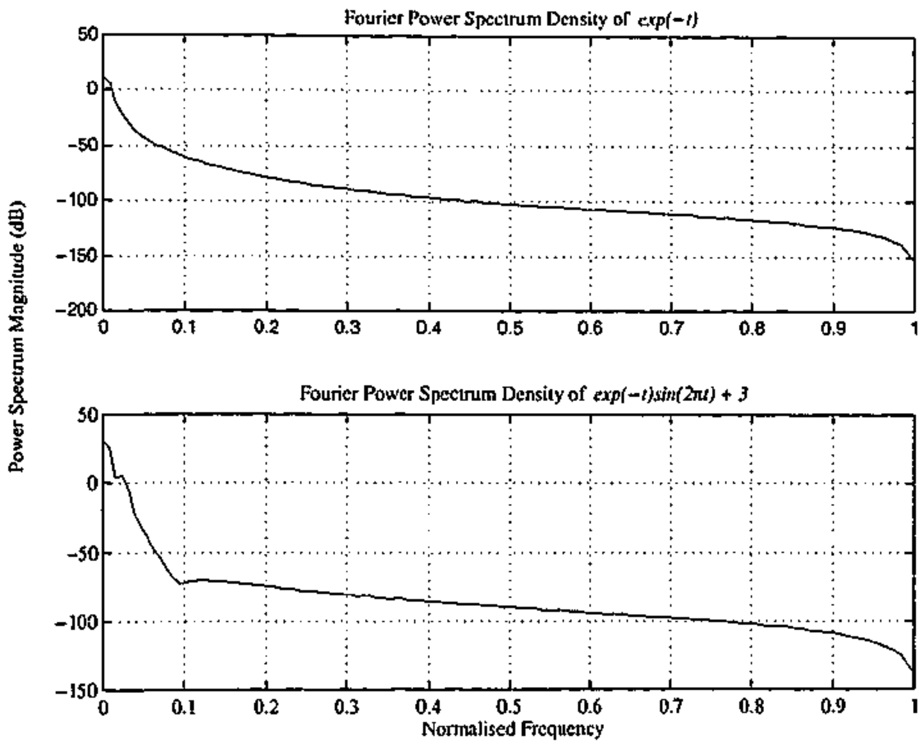

Figure 6.2.25: Fourier power spectra of $exp(-t)$ and $exp(-t)\cdot sin(t) + 3$ signals

From Figure 6.2.25, there is a minor discrepancy with the magnitude of the power spectrum of the $f(t) = exp(-t)\cdot sin(t) + 3$ signal at DC condition. The power of the signal $[f(t)]^2$ at DC condition is 9 which corresponds to about $10\cdot\log_{10}(9.0) \approx 9.5$ *dB*. The signal power calculated by MATLAB gives a value of about 30 *dB*. To work out the power at DC condition, we take the magnitude of the power spectrum at DC, which is 30 *dB*, and divide that by $\sqrt{2\pi}$ which yields about 12 *dB*. This corresponds to a percent error of about 26% with respect to 9.5 *dB*. It should be noted that this method of power spectral estimation by MATLAB is only approximate. In addition, it is the qualitative information of the graph that is of importance, not the quantitative details. However, to provide a satisfactory result, the Welch method of power spectral estimation can be used to obtain the power spectrum of the signal. This method yields the DC power of about 9.6 *dB* which corresponds to about 0.1%

percent error. However, calculation of the Fourier power spectra of signals is not the main emphasis of this chapter and we stop the discussion about power spectral estimation at this point.

The following table summarises notable characteristics of an exponential signal $exp(-t)$, an exponentially decaying sinusoidal signal with a zero final value $exp(-t)\cdot sin(t)$, an exponentially decaying sinusoidal signal with a non-zero final value $exp(-t)\cdot sin(t) + 3$ and the ECG signal.

Table 6.2.1: Characteristics of $exp(-t)$, $exp(-t)\cdot sin(t)$, $exp(-t)\cdot sin(t) + 3$ and ECG signals

Signal	Characteristics
$exp(-t)$	Exhibits transient behaviour, periodicity is not present, bounded energy
$sin(t)\cdot exp(-t)$	Exhibits transient and periodic behaviour, strong sinusoidal decaying characteristics, bounded energy, *WPS* contour curves have small radii; energy smearing is not present
$sin(t)\cdot exp(-t) + 3$	Exhibits transient behaviour, bounded energy due to closed contour curves. The density of contour curves is high which suggests that the signal is not chaotic. *WPS* contour curves have large radii; energy smearing is not present.
ECG	Exhibits transient and chaotic behaviour, bounded energy, *WPS* contour curves have large radii, energy smearing is present. The occupied *WPS* scale range of the ECG signal is wider than that of the $sin(t)\cdot exp(-t) + 3$ signal.

Throughout this chapter, comparisons between the hyperbolic and *sym3* wavelet power spectra have been made. It can be seen that the WPS_{hyp} of most signals have a smaller scale range than that of the WPS_{sym3} which significantly reduces the wavelet power spectral computational burden. In addition, the WPS_{hyp} converges faster than the WPS_{sym3} in the case of an exponential signal in which the signal energy distribution is more clearly displayed by using the hyperbolic wavelet than by using the *sym3* wavelet. One advantage of the WPS_{sym3} over the WPS_{hyp} is that the former can clearly display all sub-harmonics of signals, whereas, some sub-harmonics are missing if the latter is used due to the compression effects (was studied in Chapter 4) as will also be demonstrated in Chapter 7. This can be seen as a trade-off between efficiency and fine-detail display of the hyperbolic and *sym3* wavelets.

6.3 Remarks

Five different signals have been studied in detail in this chapter including a sinusoid, an exponential transient, an exponentially decaying sinusoid with a zero and non-zero DC component, Duffing oscillator and the ECG. We have employed contour plots of their wavelet power spectra to recognise the presence of periodic, transient and chaotic characteristics. The following remarks are drawn from the numerical simulation and analysis in this chapter

1. If the contour plot of the *WPS* consists of a repetitive and discrete sequence of islands of closed and filled contour curves at high scales, then it is effectively periodic. Typical examples are a sinusoidal signal in Section 6.2.1, Duffing Period 1 and 2 waveforms in Section 6.2.4;

2. If the contour plot of the *WPS* consists of one peak and the contour curves have very large radii which indicate a slow converging rate, then it is of a transient type. A typical example is an exponential signal in Section 6.2.2. If a small number of islands of distinctive closed contour curves are present, then the signal is transient-periodic, e.g. an exponentially decaying sinusoid with a zero final value in Section 6.2.3;

3. If the contour plot of the *WPS* consists of bounded contour curves with large radii * then the signal is either transient-periodic (quasi periodic), e.g. an exponentially decaying sinusoid with a non-zero DC component in Section 6.2.3, or chaotic, e.g. the ECG. To differentiate between these two cases, the occupied scale ranges of their wavelet power spectra are used. If the *WPS* of the signal covers a large scale range (approximately larger than 2,000 scales), then the signal is chaotic, e.g. the ECG. Otherwise, it is transient-periodic;

4. If the *WPS* is continuous (ECG signal) and there are changes in its energy patterns, i.e. the energy distribution varies for different data segments, that means the signal is possibly in the transition into the chaotic region, e.g. Duffing Period 4 waveform;

* These radii are smaller than those discussed in Case 2.

5. If the contour plot of the *WPS* is non-repetitive and contour curves have low scales, then the signal has entered into the chaotic region, e.g. Duffing chaotic waveform. It should be noted that for cases 4 and 5, the occupied wavelet power spectral scale range of the signal is wide and its energy density is broad and continuous.

This chapter has established a gallery of the WPS_{hyp} and WPS_{sym3} which are necessary for the next chapter in investigating speech and music signals using the hyperbolic and *sym3* wavelets.

6.4 Conclusion

The contribution of this chapter is to establish a gallery of the proposed WPS_{hyp}. Five typical signals have been examined including a sinusoid, an exponential transient, an exponentially decaying sinusoid with a zero and non-zero DC component, Duffing oscillator and the ECG. The Fourier power spectra of these signals have also been displayed and compared with their wavelet power spectra. The Duffing oscillator has been examined in detail and the observed results are consistent with previous results reported in the literature. The ECG signal appears to exhibit chaotic behaviour in which smooth, broad and no major harmonic peaks were detected in its wavelet power spectrum.

Although the *WPS* technique has an advantage (over the Fourier method) of showing signal energy distributions in time and frequency domains, it is not effective when examining the ECG signal. In fact, the WPS_{hyp} and WPS_{sym3} of the ECG and the exponentially decaying sinusoid with a non-zero DC component signals were almost identical. In contrast, the Fourier power spectra of these signals can be differentiated which is a disadvantage of the *WPS* technique. Thus, the most appropriate technique is identified based on the application and the nature of the signal. The next chapter carries out investigations on speech and music signals by using the hyperbolic and *sym3* wavelet power spectrum techniques.

Chapter 7: THE HYPERBOLIC WAVELET POWER SPECTRA OF MUSIC AND SPEECH SIGNALS

This chapter studies two different non-stationary signals – music and speech – by examining their hyperbolic and *sym3* wavelet power spectra. These wavelet power spectra will be compared and discussed. Background on the wavelet and Fourier transform methods was discussed in detail in Chapter 6. It will be shown that music and speech signals can be effectively studied by using the wavelet power spectrum technique.

7.1 The Wavelet Power Spectrum of Musical Signals

Music signals are studied in detail in this section using the hyperbolic wavelet power spectrum technique. Musical sounds have been studied by a number of researchers on music multi-dimensional scale analysis [70-72], music classification [103, 104, 111, 112] music identification [105] and music recognition by using the continuous wavelet transform [112].

Classifying different piano sounds was studied by Delf and Jondral by using a number of time-frequency techniques such as the Short-Time Fourier Transform (STFT) and the Wigner-Ville (WV) time-frequency distribution. As was reported in Chapter 3, the WV kernel was not effective since it generates unnecessary cross terms in the time-frequency plane which provides misleading information about the signal. Thus, the Wigner-Ville time-frequency power spectrum technique is not effective. Even though the WV time-frequency technique was used in [103], instantaneous behaviour of musical signals was not emphasised in the paper. The STFT is also not an effective time-frequency technique since its time and frequency resolutions are coarse [8]. The technique used in this chapter is the new hyperbolic time-frequency power spectrum technique, which was shown in Chapter 3 to be more effective than the WV time-frequency technique in terms of cross-term suppression, auto-term resolution and noise robustness. The *sym3* wavelet power spectrum technique is also employed as was done in the previous chapter.

Hamdy, Tewfik, Chen and Takagi [104, 111] reported a music classification method using a statistical technique of calculating skewness, entropy, the first- and second-order statistics of different musical sounds such as jazz, rock, pop and then estimated the appropriate threshold so that these sounds can be clearly distinguished. The main limitation on this method is that it does not make use of the instantaneous information of the signal which yields its characteristics and hence allows effective signal classification.

The most relevant work to this chapter is the work by Olmo, Dovis, Benotto, Calosso and Passaro [112] on using the continuous wavelet transform to detect different tones in music. The authors designed the new wavelet, called the Log-Morlet wavelet, and then showed that the new wavelet was capable of recognising different harmonics and tones in music waveforms. This method used the same principle in detecting edges and abrupt changes in an input signal as was reported in [52]. The main difference of this work to our work is that we employ the hyperbolic wavelet power spectrum, not the continuous wavelet transform, for music recognition and to study music characteristics. Although time and frequency localisation was employed in [112], it was not as strongly emphasised as in our work reported in this chapter. In particular, all harmonic and sub-harmonics will be identified in both time and frequency domains which allows effective music recognition. As was shown in Chapter 4, the hyperbolic wavelet has a finer scale resolution than both the CW (Mexican-hat) and Morlet wavelets which subsequently yields finer time and frequency resolutions in the hyperbolic wavelet power spectrum. This is the main reason why the hyperbolic wavelet is employed in this section for music investigation.

The main objective of this section is to study musical signals by observing their instantaneous behaviour which is the main difference from other methods reported in the literature. From that, their characteristics can be thoroughly studied, hence, it is possible to conclude whether the signal is chaotic, transient, periodic or transient periodic. The hyperbolic wavelet power spectrum technique, as was done in Chapter 6, is used to study signal characteristics, whereas other researchers concentrated on time-frequency distribution techniques such as using the Wigner-Ville distribution [33, 34] and Choi-Williams distribution [28].

Two musical files *accb32.dat* (refers to accordion sound) and *clab42.dat* (refers to clarinet sound) are used as input signals. The files were sampled at 44.1 *kHz* for a time period of 2 seconds. Since music files have a large number of data samples of 44,100, they are divided into small sets of 2,048 samples and each set consists of two segments of 1,024 samples. The wavelet power spectrum of each set is calculated. The characteristics of various music segments are examined in detail. To accurately examine music signals, however, the entire segment of 44,100 samples is used to calculate their wavelet power spectra so that chaotic behaviour or non-uniform energy distributions can be successfully detected.

The input data are carefully chosen so that noise and other interference sources are eliminated, i.e. the first few segments of a music source file are ignored and the subsequent segments are used as an input signal. Usually, after the first 5,000 data samples, music samples are suitable for signal analysis. To further strengthen the correctness of the obtained results, the Fourier power spectrum technique is employed so that consistency between the wavelet power spectrum and Fourier power spectrum techniques is validated.

It should be noted that as stated in Chapter 6, a harmonic peak can be identified by having a colour contour scale of larger than 0.7 and a sub-harmonic peak has a contour scale of smaller than 0.7, but usually, it is less than 0.5. The lower the order of a sub-harmonic, the lower the contour scale. This method has also been employed by Olmo *et al* [112] and Sussaman and Karsh [113] by means of energy separation.

7.1.1 The Wavelet Power Spectrum of an accordion music signal

Various segments of the accordion music signal are examined in this section so that its instantaneous behaviour can be detected. For data samples from 1,024×5 to 1,024×7 and from 1,024×30 to 1,024×32, the wavelet power spectra of each set are given in Figure 7.1.1 and Figure 7.1.2 respectively.

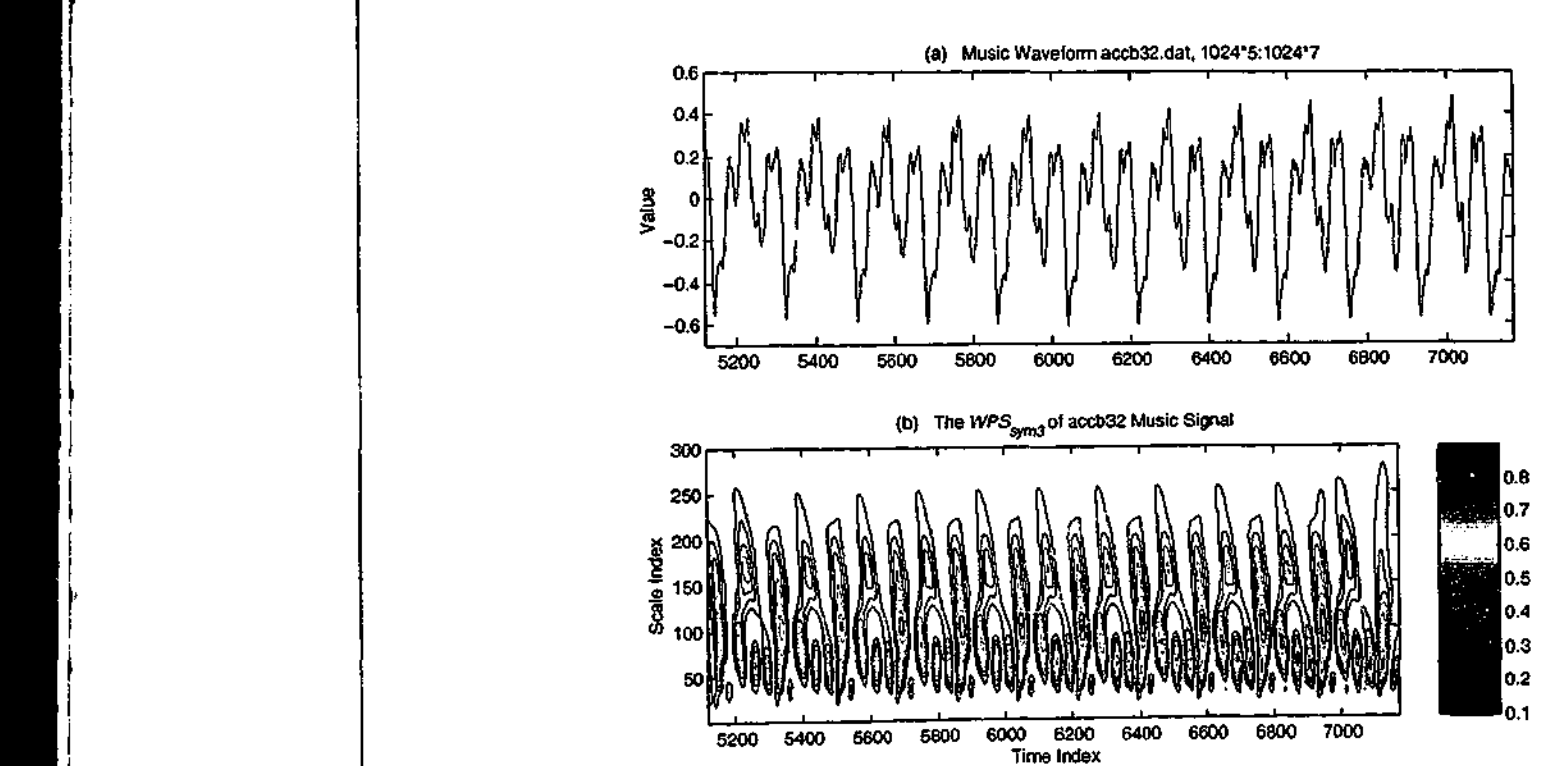

Figure 7.1.1: Time-domain waveform and contour plot of the WPS_{sym3} of the accordion music, data samples are in the range of 1,024×5:1,024×7

As can be seen from Figures 7.1.1 and 7.1.2, the 2,048-sample music sets seem to exhibit periodic behaviour in which their wavelet power spectra show repetitive energy patterns over time. The energy is mainly concentrated at the harmonic peaks and there are three sub-harmonics that can be clearly identified. This information can be used to classify the specific characteristics of the accordion signal. Thus, sounds from different musical instrument can be distinguished by examining their wavelet power spectra. The WPS_{hyp} of the accordion music signal of the previous data sets are given in Figure 7.1.3.

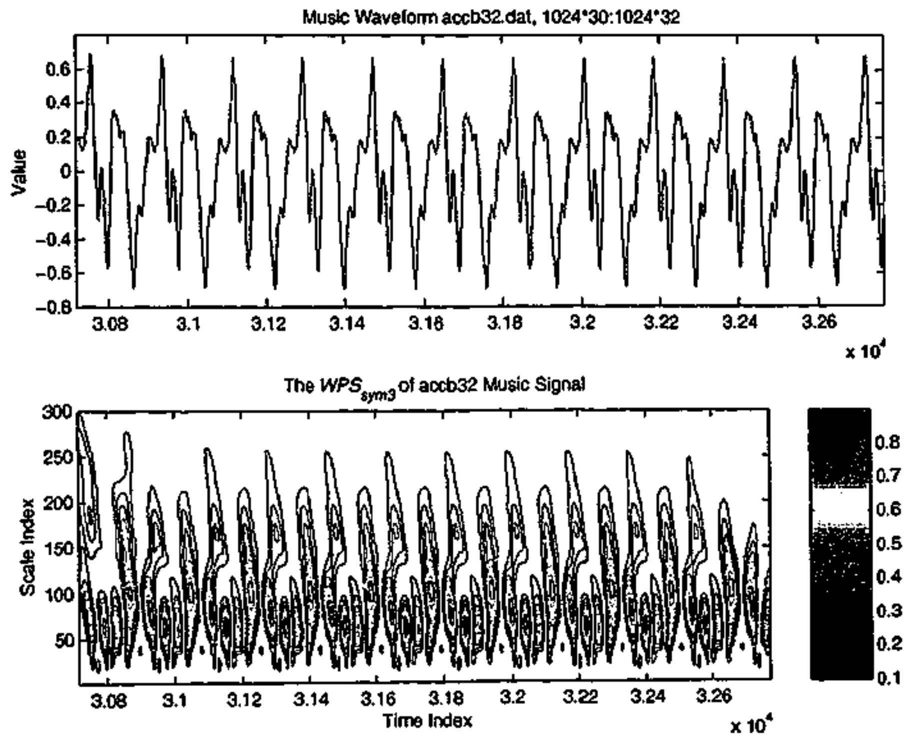

Figure 7.1.2: Time-domain waveform and contour plot of the WPS_{sym3} of the accordion music signal for data samples from 30×1,024:32×1,024

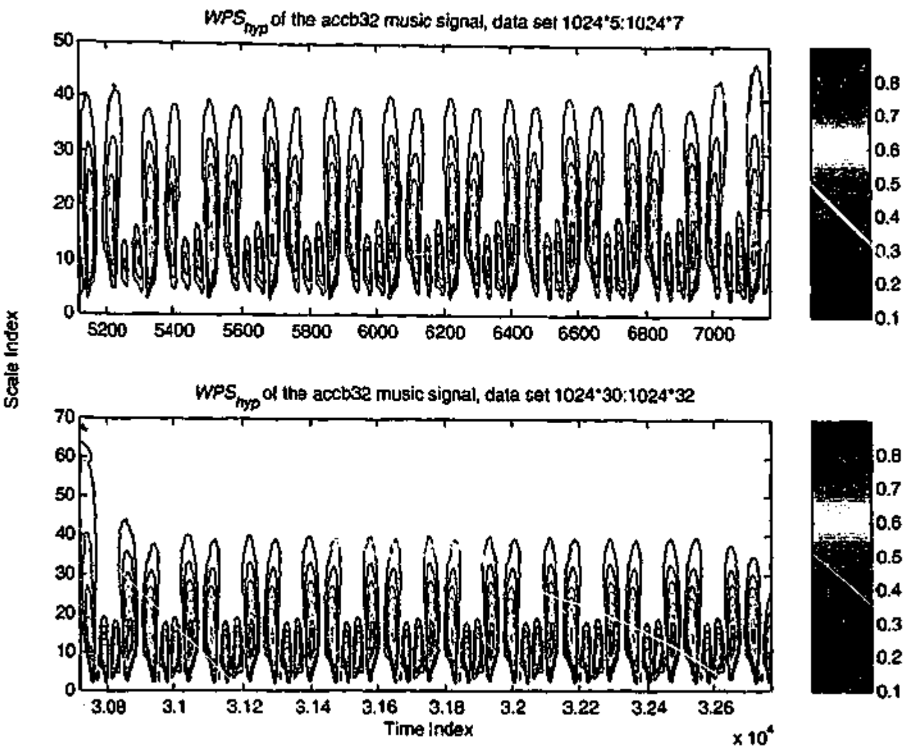

Figure 7.1.3: Contour plots of the WPS_{hyp} of the accordion music signal of the two data sets

As can be seen from Figure 7.1.3, the hyperbolic wavelet power spectra of the accordion music signal are consistent with the WPS_{sym3} given in Figures 7.1.1 and 7.1.2 in which instantaneous periodic characteristics of the signal are successfully detected with bounded contours and repetitive patterns. The WPS_{hyp} and WPS_{sym3} are both very similar which suggests that this music signal can be effectively studied using the hyperbolic and *sym3* wavelet power spectrum techniques.

From simulations, since the wavelet power spectra of various sets of the signal are very similar, it might be suggested that the musical accordion signal is periodic. To validate this conclusion, the WPS_{sym3} of the entire music signal is calculated and given in Figure 7.1.4.

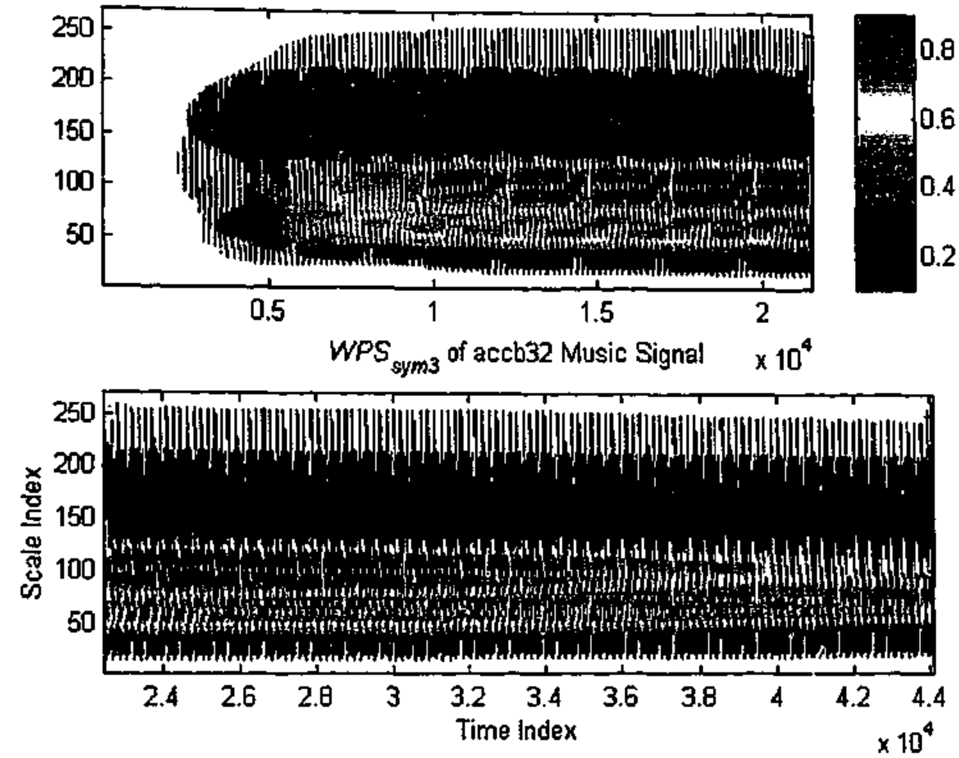

Figure 7.1.4: Contour plot of the WPS_{sym3} of the accordion music signal for the time index from 1 to 44,100

As can be seen from Figure 7.1.4, there are three main peaks present in the wavelet power spectrum of the signal. For the first 30,000 samples, periodic behaviour is dominant as the peaks are clearly shown and the signal energy is evenly distributed as seen in the time-index magnified versions of Figure 7.1.4, which are Figure 7.1.1 and Figure 7.1.2. The third peak (located at the scale $a \approx 80$) disappears after the first 30,000 samples and there is uneven matching in energy distribution between the two strongest peaks (approximately located at scales of 60 and 100) which reflects the instability of the signal. From Figure 7.1.4 and Figure 7.1.11 (Fourier power spectrum of the accordion signal) it might be suggested that the music signal has entered into the chaotic region due to its broad and continuous Fourier power spectrum and uneven energy density as the time index increases.

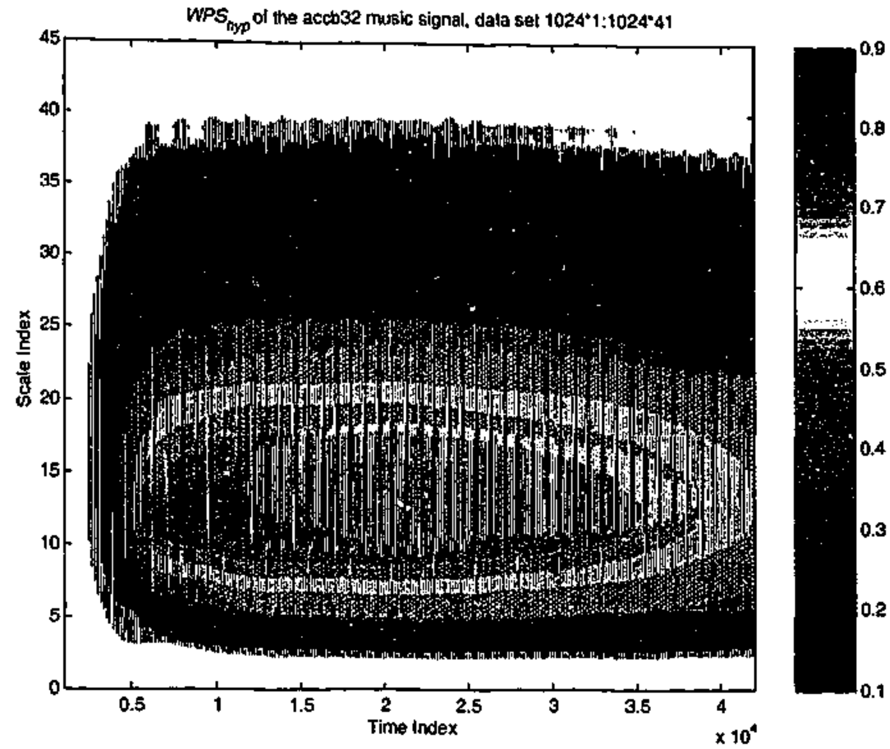

Figure 7.1.5: Contour plot of the WPS_{hyp} of the accordion music signal

The WPS_{hyp} of the accordion music signal, given in Figure 7.1.5, agrees with the WPS_{sym3} given in Figure 7.1.4 in which both wavelets can detect instantaneous periodic characteristics of the signal by having high contour scales. It should be noted that the compression ratio of about 250/38 ≈ 6.5 is observed in Figure 7.1.5 which results in a smaller scale range in this figure. From Figure 7.1.5, the WPS_{hyp} consists of energy-density layers (approximately having the same center) at different scales. Especially, the harmonic peak (represented by the inner-most layer) approximately terminates at the time index $b \approx$ 30,000 which reflects the discontinuity of this musical signal. Other harmonic peaks (at lower scales due to a brighter colour on the colour scale) of the signal subsequently terminate at the time indices of 35,000 and 38,000. However, the sub-harmonic peaks (at very low scales) are present in the WPS_{hyp} for all time indices.

As seen from Figure 7.1.4 and Figure 7.1.5, the wavelet power spectrum of the accordion signal varies for different values of time index, i.e. harmonic peaks decay over time which indicates uneven energy density and discontinuity in the signal. In addition, different segments of the signal, which were recorded from one musical instrument, have different wavelet power spectra. Thus, although the signal appears to be periodic in its time-domain waveforms, it is disordered or chaotic as it has been shown in this section and in other findings [106-110].

A clarinet music signal is studied in the next section.

7.1.2 The Wavelet Power Spectrum of a clarinet music signal

The wavelet power spectra of various segments of the clarinet signal are given in Figure 7.1.6 and Figure 7.1.7.

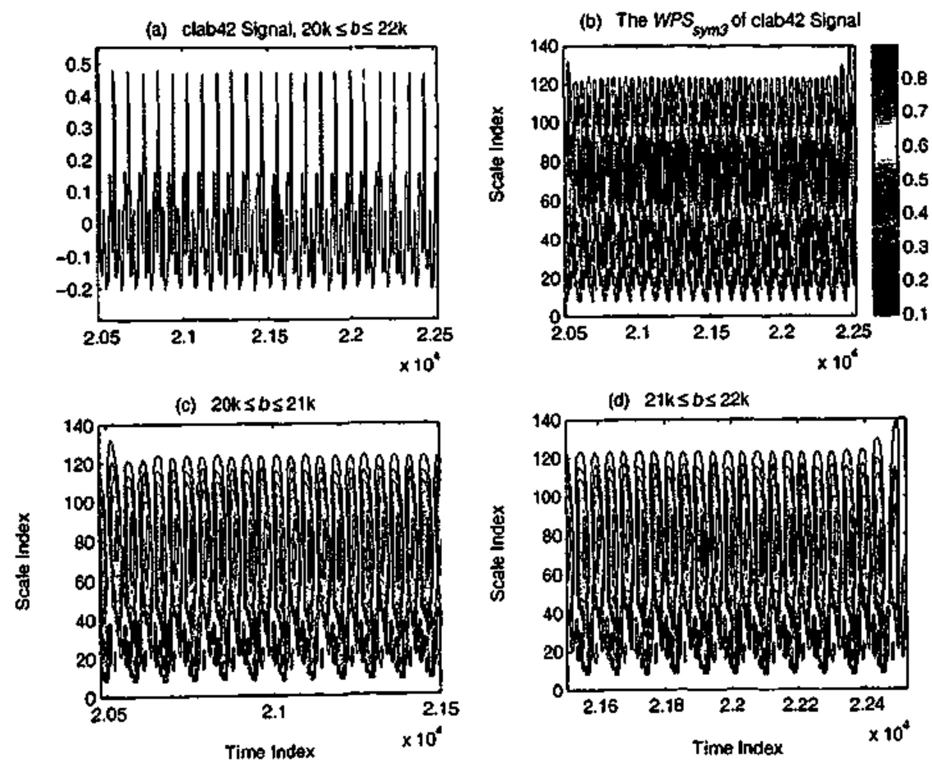

Figure 7.1.6: Time-domain waveform (a) of the clarinet music data samples from 1,024×20 to 1,024×22 samples and contour plot of its WPS_{sym3} (b). The magnified contour plots of WPS_{sym3} are given in (c) and (d).

The first data set is chosen from the 1,024×20 sample to 1,024×22 sample. As can be seen in Figure 7.1.6, the wavelet power spectrum consists of two distinctive peaks which suggest that this set of the signal is periodic. There are two harmonic peaks and three sub-harmonics in "one period" of the waveform of about 80 samples or 1.8 *ms*. In Figures 7.1.6(c) and (d), the peaks (which can be identified by high contour scales in the graph) are located at the scale of $a \approx 70$ and the sub-harmonics at the scale of $a \approx 20$. There is also continuity of energy because one peak and one sub-harmonic are located at the same time index.

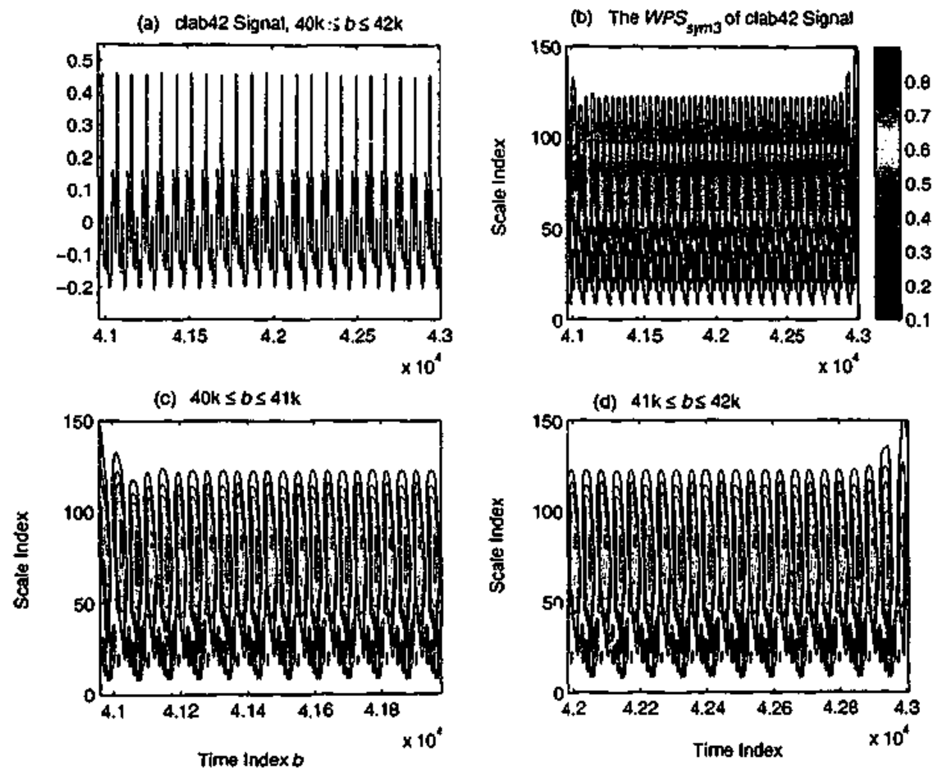

Figure 7.1.7: Time-domain waveform of the 1,024×40 to 1,024×42 samples (a) and contour plot of WPS_{sym3} (b) of the clarinet musical signal. The magnified graphs of (b) are given in (c) and (d).

The second data set is chosen from the 1,024×40 sample to 1,024×42 sample. This set, as seen in Figure 7.1.7, exhibits similar characteristics to those of the first set, except that the energy concentration at the main harmonic peak is reduced due to a lower contour scale. The energy patterns of the harmonics and sub-harmonics are unchanged. The WPS_{hyp} of two separate data sets 1,024×20:1,024×22 and 1,024×40:1,024×42 are given in Figures 7.1.8(a), (b) and (c), (d) respectively.

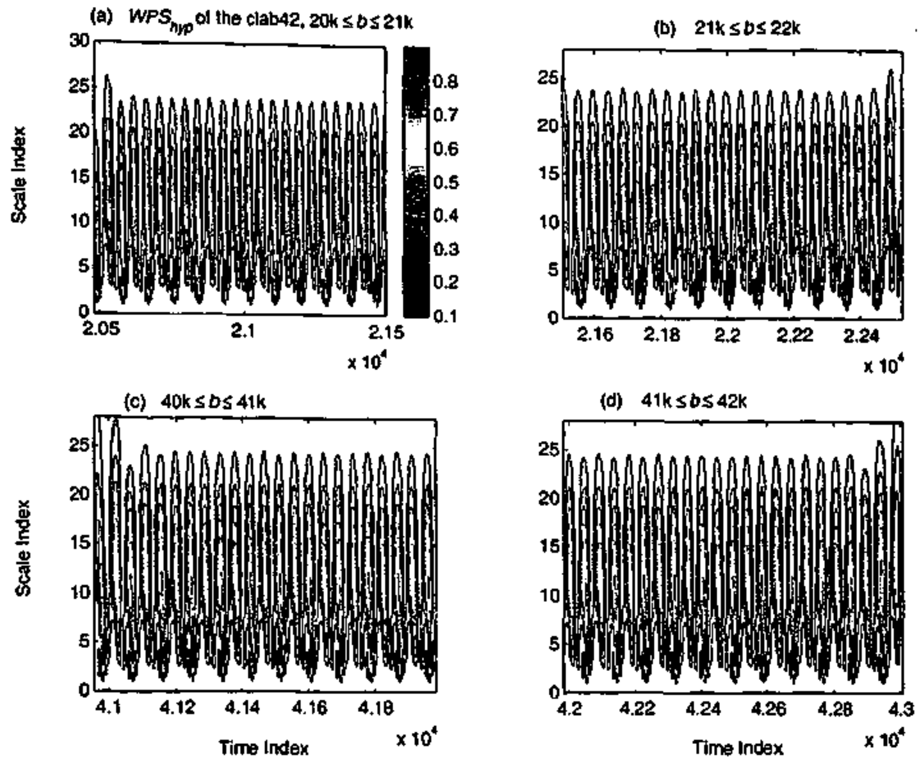

Figure 7.1.8: Contour plots of the WPS_{hyp} of two separate sets of the clarinet music signal. The time index is expanded so that the WPS_{hyp} can be clearly displayed.

The WPS_{hyp} of this data set is consistent with its WPS_{sym3} in which major harmonic and sub-harmonic peaks are successfully detected. However, as seen from Figure 7.1.8, the WPS_{hyp} cannot display some sub-harmonics of the clarinet music signal as it can in the case of the accordion signal. The WPS_{sym3} successfully displays all fine details of the signal as can be seen in Figure 7.1.6. This is one disadvantage of the hyperbolic wavelet compared with the *sym3* wavelet. However, the WPS_{hyp} is compressed to the highest scale of 25 which is about six times smaller than that of the WPS_{sym3}. Thus, calculation time of the WPS_{hyp} can be significantly reduced which makes it more efficient than the WPS_{sym3}.

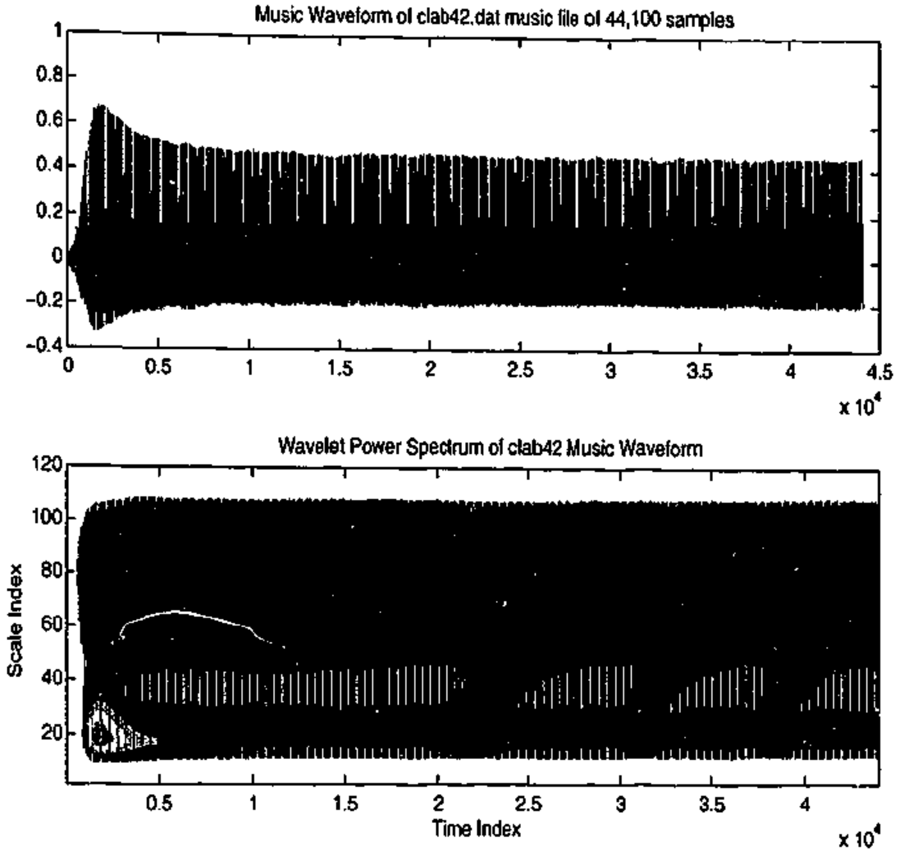

Figure 7.1.9: Time-domain waveform and contour plot of the WPS_{sym3} of the clarinet signal. The WPS_{sym3} shows early "periodic" behaviour due to high contour colour scales of the energy density which represents one rapid-decaying peak. This can be understood as a burst of energy or an abrupt change in the signal energy. After that, there is no major harmonic peak in the signal and sub-harmonic peaks at low contour scales, indicated by the blue colour which corresponds to the maximum scale of 0.3, are dominant. At the 15,000th sample, a change in energy is detected. From the time index of 25,000 onwards, sub-harmonics are discontinued and they completely terminate at the time index $b \approx 37{,}000$. This might suggest that the signal is disordered since its components disorderedly vary with time. The contour scale of this figure is similar to that given in Figure 7.1.7.

Figure 7.1.9 displays the WPS_{sym3} of the entire clarinet music signal which shows one rapid-decaying "periodic" peak at the approximate scale of 20. This represents a burst of energy in the signal. The periodic peak disappears from approximately the 5,000th sample onwards and the signal appears to behave non-periodically. From the 15,000th sample onwards, chaotic behaviour appears to be dominant due to a broad energy distribution and uneven energy density of the signal.

Figure 7.1.10 displays the WPS_{hyp} of the clarinet music signal.

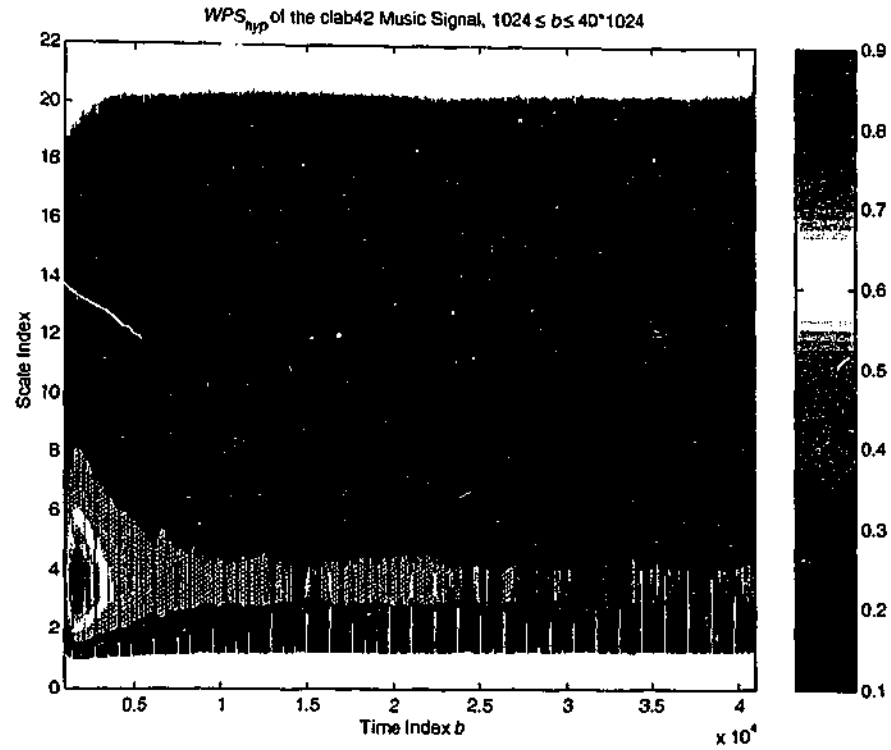

Figure 7.1.10: Contour plot of the WPS_{hyp} of the clarinet music signal. Periodic components are clearly displayed in the scale range of 1 to 8 compared with 10 – 30 in Figure 7.1.9.

From Figure 7.1.9 and Figure 7.1.10, it can be suggested that the hyperbolic and *sym3* wavelets can successfully detect early periodic characteristics of the clarinet music signal. Then, periodicity decays away which makes the wavelet power spectra of subsequent segments of the waveform to be unsymmetrical. Similar to the accordion musical signal, different segments of the clarinet music signal have different wavelet power spectra which reflects its disordered characteristics.

Section 7.1.1 and this section investigated the characteristics of two different musical signals, accordion and clarinet, using the hyperbolic and *sym3* wavelet power spectrum techniques in which the disordered characteristics of both signals have been successfully revealed. Their wavelet power spectra have been instantaneously displayed so that their behaviour could be effectively monitored. The results found in this chapter also agree with results found by other researchers [106-111]. The Fourier power spectra of music signals are given next to validate the results drawn by using the wavelet power spectrum technique.

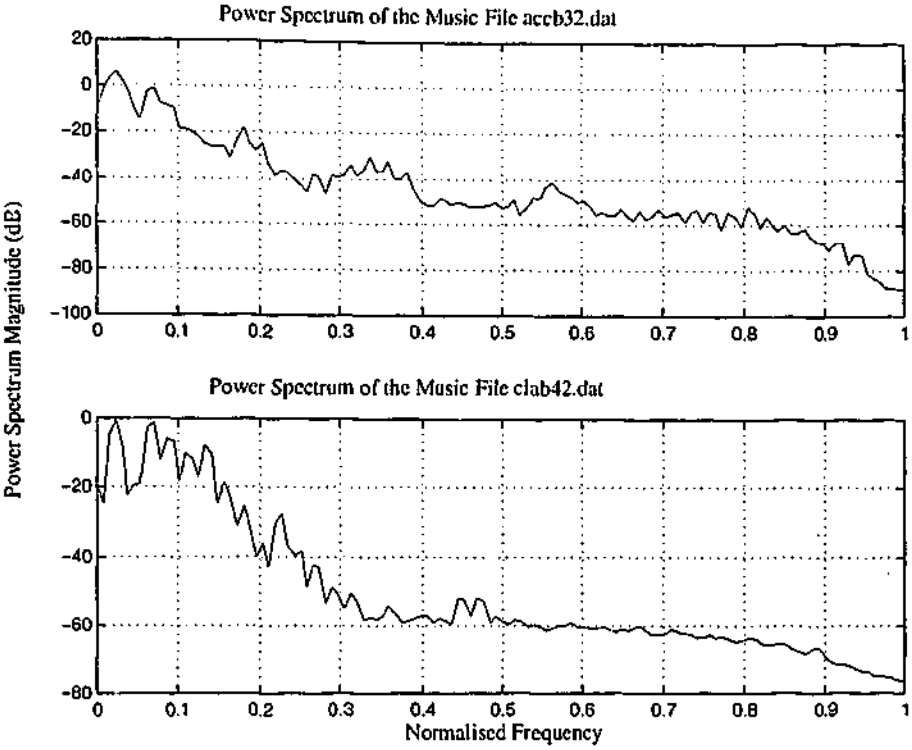

Figure 7.1.11: Fourier power spectra of the accordion and clarinet music signals. The frequency is normalised by dividing every frequency bin by the largest frequency bin in the signal spectrum. This rule is also applied for all subsequent graphs in this chapter which will not affect the general shape of the spectrum. The true maximum frequency of both the accordion and clarinet signals is 22.5 *kHz*.

From Figure 7.1.11, it can be seen that the Fourier power spectrum of the accordion signal is broad with harmonic peaks over the low-frequency range. These peaks, as explained before, are fundamental components of the signal. The true behaviour of the signal is based on the high-frequency components. In this case, the high-frequency spectrum is broad which suggests that the signal exhibit chaotic behaviour.

The power spectrum of the clarinet music signal has more distinctive periodic peaks than that of the accordion signal which means that the former is more stable than the latter. Over the high-frequency range, the power spectrum is broad (as it was the case for the accordion signal) which suggests that the clarinet music signal might exhibit chaotic behaviour. However, due to a large amount of periodic components over the low- and mid-frequency ranges, chaotic behaviour might not be dominant and the signal in this case can be said to be in a transition to chaos.

In this section, the Fourier power spectrum technique has been successfully used to study characteristics of musical signals. It is clear that the wavelet power spectrum technique and Fourier power spectrum technique are consistent. However, compared with the wavelet power spectrum, the Fourier technique does not show instantaneous behaviour of the signals over time. This is a disadvantage of the Fourier power spectrum technique since different segments of music signals have different power spectra. Thus, even though the Fourier technique can been used to study music, it is not effective compared with the wavelet power spectrum technique in this aspect. It should be recalled that in Chapter 6, the wavelet power spectrum was not an effective tool to study an exponentially decaying sinusoidal signal and the ECG but the Fourier technique was. Therefore, depending on the input signal, the appropriate technique is chosen. For unknown signals, which are often encountered in practice, both techniques should be employed so that the most suitable technique can be identified.

By studying the instantaneous characteristics of musical signals using the hyperbolic wavelet power spectrum technique, it is possible to classify different musical sounds. The hyperbolic wavelet power spectrum technique gives locations in time (time index b) and frequency (inverse of the scale a) of harmonic and sub-harmonic peaks which are unique for every signal. This is a major advantage of the wavelet power spectrum technique over the statistical technique [104, 111], the WV time-frequency power spectrum and the time-frequency STFT techniques [103]. The hyperbolic wavelet power spectrum technique has been shown to be an effective tool which promises useful applications in non-stationary signal classification. The main limitation on the time-frequency hyperbolic wavelet power spectrum technique is that for some signals, i.e. the ECG and the exponentially decaying sinusoid with a non-zero final value studied in Chapter 6, their wavelet power spectra are similar even though their harmonic peaks are located at different scales. This makes the classification process of the two signals difficult. Another limitation is that the hyperbolic wavelet power spectrum is intensive to compute and thus powerful computing tools are required to improve the computation speed. This issue will be dealt with in Chapter 8 along with the parallel computation of the second-order statistical bispectrum.

In the next section, speech waveforms of the English vowels and the sound "sh" are studied using the hyperbolic wavelet power spectrum, *sym3* wavelet power spectrum and Fourier power spectrum techniques.

7.2 The *WPS* of Speech Signals

Speech signals are examined in this section using both the wavelet power spectrum and the Fourier power spectrum techniques. Speech signals have been studied using time-frequency power spectrum analyses to detect formants over time [9, 39, 92]. Some popular kernels that have been used to study speech signals are the cone-shaped kernel [39], Choi-Williams [28] and signal-dependent Gaussian-shaped kernels [8, 9]. This section attempts to study characteristics of speech signals using the new hyperbolic wavelet and the *sym3* wavelet to detect periodic and chaotic behaviour. The speech signals in this section are of 4,096-sample long and they are the English vowels "a", "e", "i", "o", "u" and the sound "sh". Each of these signals will be individually examined. All graphs in this section have the time index of length 2,000, instead of 4,096, for magnification purposes so that harmonic and sub-harmonic peaks can be clearly seen. Figure 7.2.1 shows the time-domain signal and the WPS_{sym3} of the vowel "a".

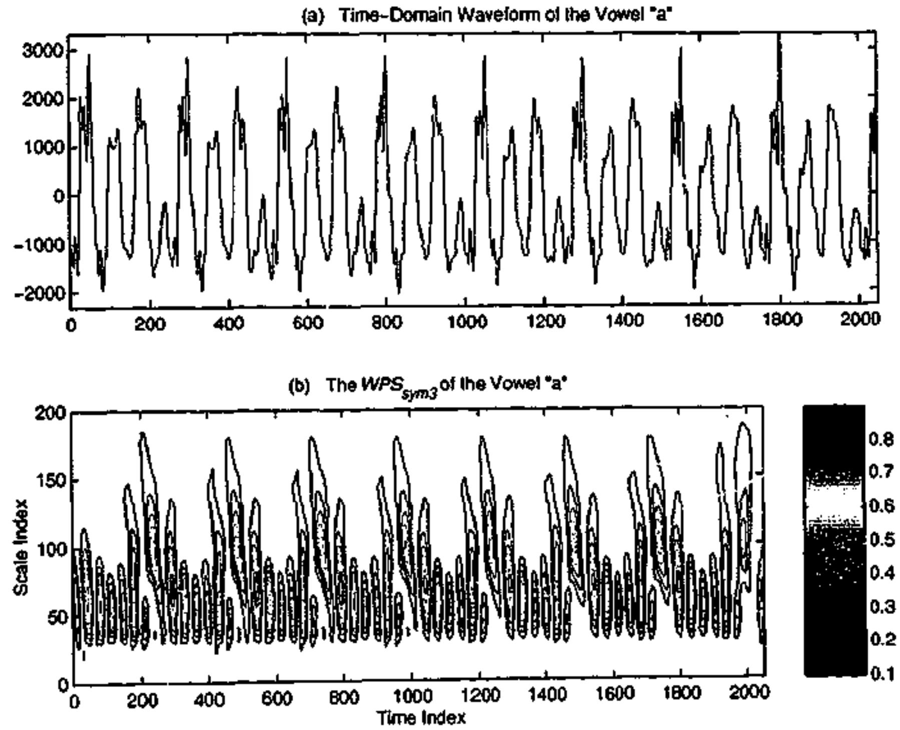

Figure 7.2.1: The speech time-domain waveform of the vowel "a" and contour plot of its WPS_{sym3}

This speech signal exhibits strong periodic behaviour by having a concentrated energy density at the peaks. Periodicity is strongly indicated by repetitive islands of closed contour curves in the wavelet power spectrum of the signal. Each harmonic peak is surrounded by a large number of sub-harmonics. The waveform seems to indicate periodic behaviour although the energy is not completely discrete as compared with the sinusoidal signal in Section 6.2.1 in Chapter 6. It should also be noted that there are no sub-harmonics associated with a harmonic peak simultaneously in time as it was the case for the clarinet music signal studied in Section 7.1.2. Thus, chaotic behaviour does not exist in this speech waveform.

The WPS_{sym3} of the vowel "e" is given in Figure 7.2.2. The speech signal of the vowel "e" is genuinely considered periodic because of its uniform energy distribution and repetitive energy patterns over time. The peaks are clearly displayed and their scale index location is almost unchanged. There are two main peaks in "one period" of the WPS_{sym3} of the signal which makes it significantly different from the vowel "a". This might suggest that the vowel "e" is more difficult for speech recognition due to its non-sub-harmonics characteristics.

The time-index magnified WPS_{hyp} of the vowels "a" and "e" are given in Figures 7.2.3(a) and (b) respectively which clearly show periodic peaks and sub-harmonics in the signals. These were also successfully detected by the *sym3* wavelet. However, due to its compression effects compared with the WPS_{sym3}, the WPS_{hyp} could not clearly display the sub-harmonics as can be seen in Figure 7.2.3(b) and Figure 7.2.2(b) for the vowel "e". This disadvantage has been reported earlier in Chapter 6. However, the overall shape of the WPS_{hyp} and WPS_{sym3} are consistent which validates the hyperbolic wavelet power spectrum technique and the effectiveness of the hyperbolic wavelet.

Figure 7.2.4 shows the WPS_{sym3} of the vowel "i".

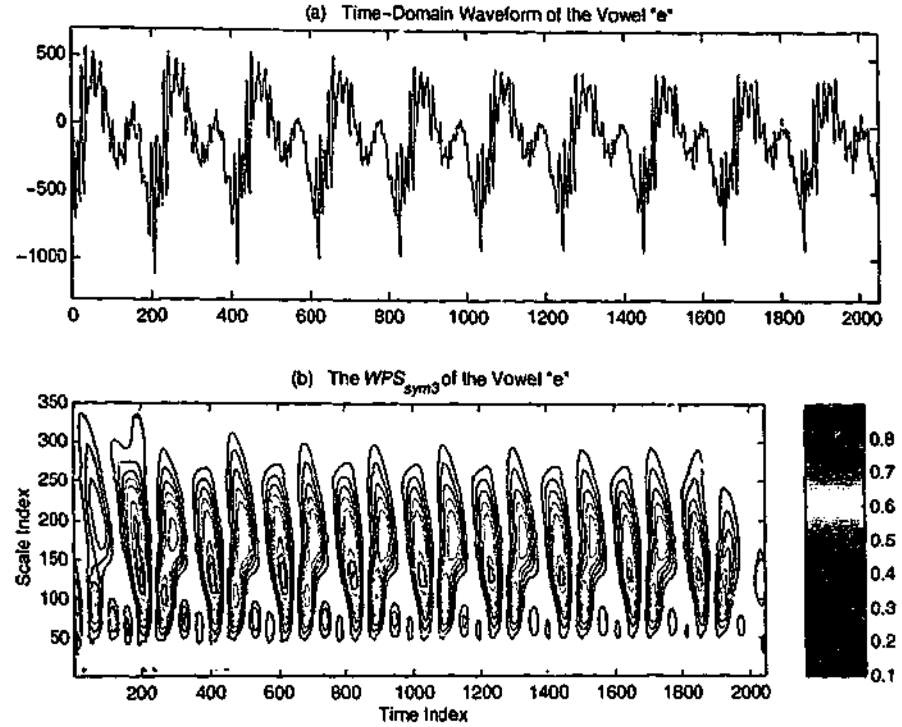

Figure 7.2.2: Speech waveform of the vowel "e" and contour plot of its WPS_{sym3}

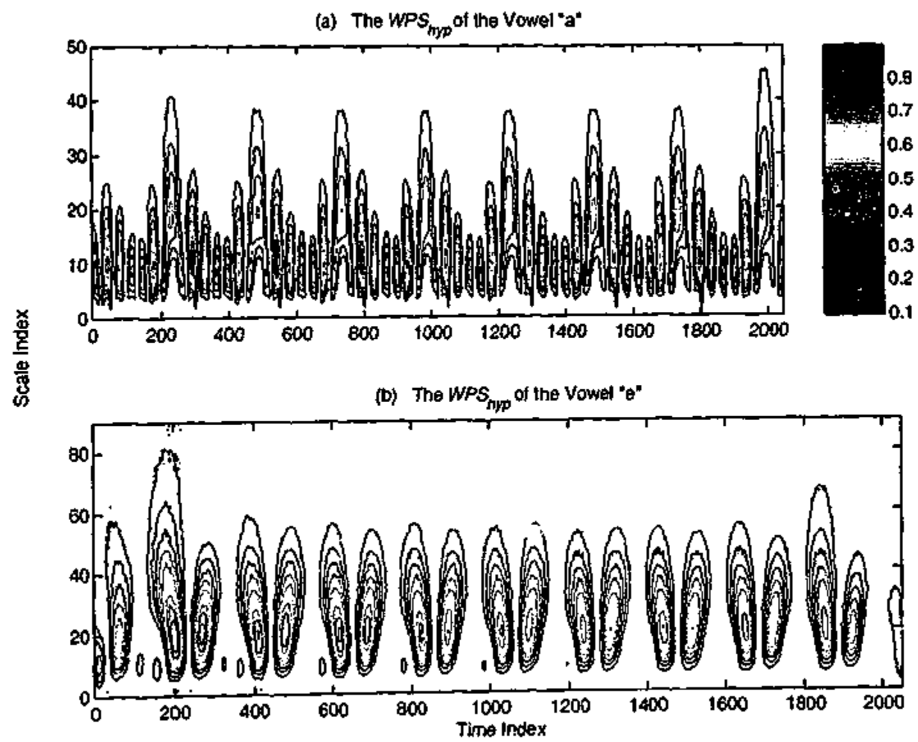

Figure 7.2.3: Contour plots of the WPS_{hyp} of the speech vowels "a" and "e"

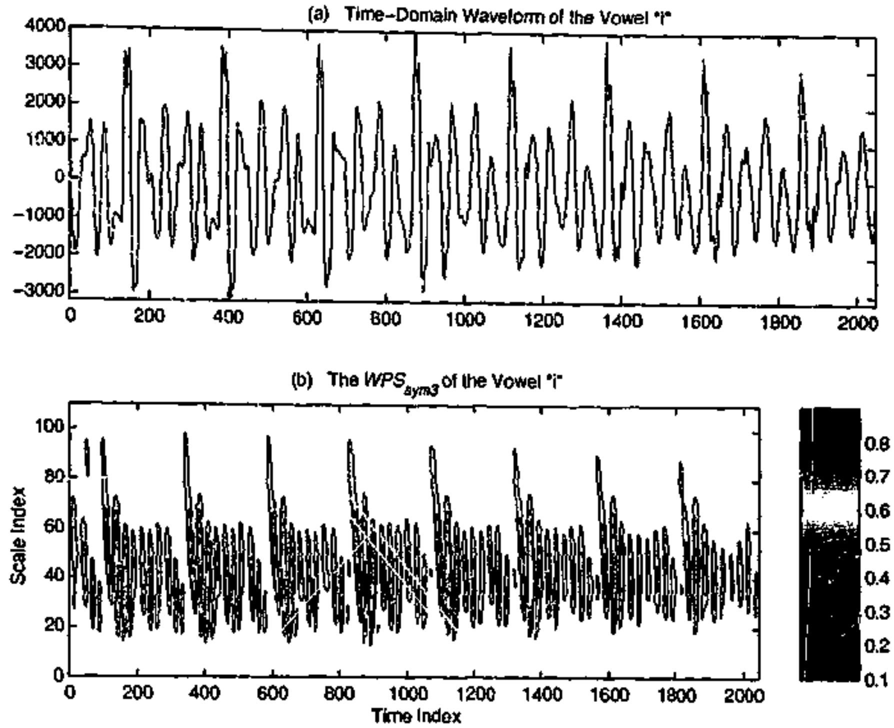

Figure 7.2.4: Speech time-domain waveform of the vowel "i" and contour plot of its WPS_{sym3}

As can be seen from Figure 7.2.4, the first part of the speech signal, which corresponds to the time index of less than 1,600, exhibits periodic behaviour, although "minor" chaotic behaviour is indicated by its broad energy distribution at a low contour scale. In the second part of the waveform which corresponds to the time index of greater than 1,600, fundamental harmonics are not fully displayed but only the sub-harmonics at low contour scales.

Although the majority of harmonics (contours at high scale) disappear from the 1,600th sample onwards, the sub-harmonics (contours are at low scales) of the signal are still repetitive at regular time intervals. In addition, the occupied scale range of 20 – 100 of the energy distribution is short which indicates strong periodicity in the signal. These two points suggest that the signal is periodic. This phenomenon, however, does not indicate chaos but indicates a change in the speech components of the signal which can be seen by changes in the time-domain waveform.

The contour plot of the WPS_{sym3} of the vowel "o" is displayed in Figure 7.2.5.

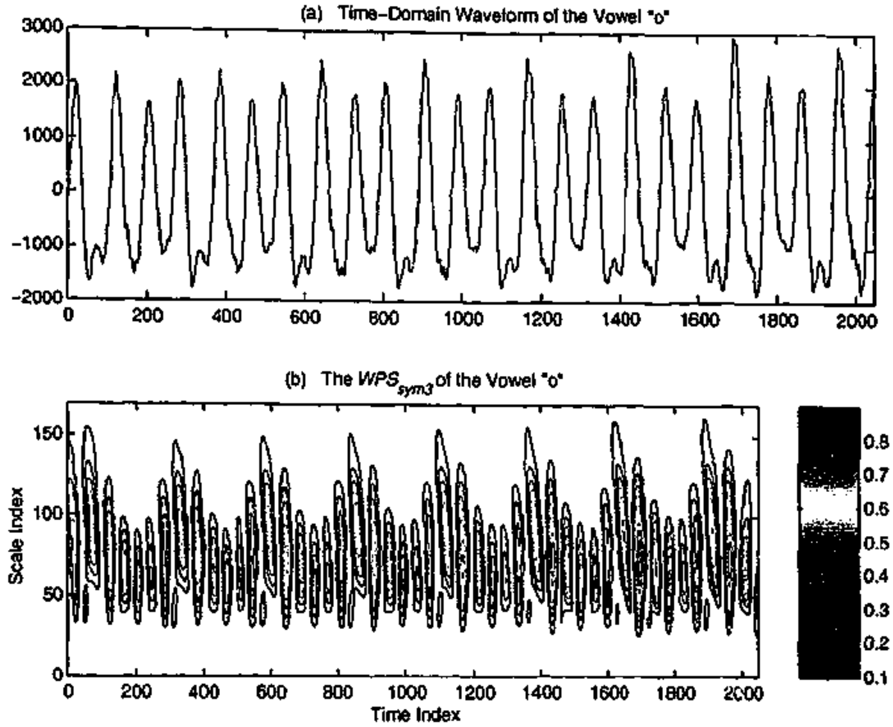

Figure 7.2.5: Speech time-domain waveform of the vowel "o" and contour plot of its WPS_{sym3}

For the speech signal of the vowel "o", the flow of energy from the 1,500th sample onwards indicates a change in its components in which the scale range slightly changes and the number of sub-harmonics decreases as the time index increases. The energy distribution of the signal is repetitive over regular time intervals which might suggest that the signal is largely dominated by periodic components.

It should also be noted that the occupied scale range of the energy distribution is short, which indicates strong periodicity, as was the case of the vowels "a", "e" and "i" investigated earlier. The energy concentration at the peaks remains almost unchanged which further validates the above statement.

The WPS_{hyp} of the vowels "i" and "o" are shown in Figures 7.2.6(a) and (b) respectively.

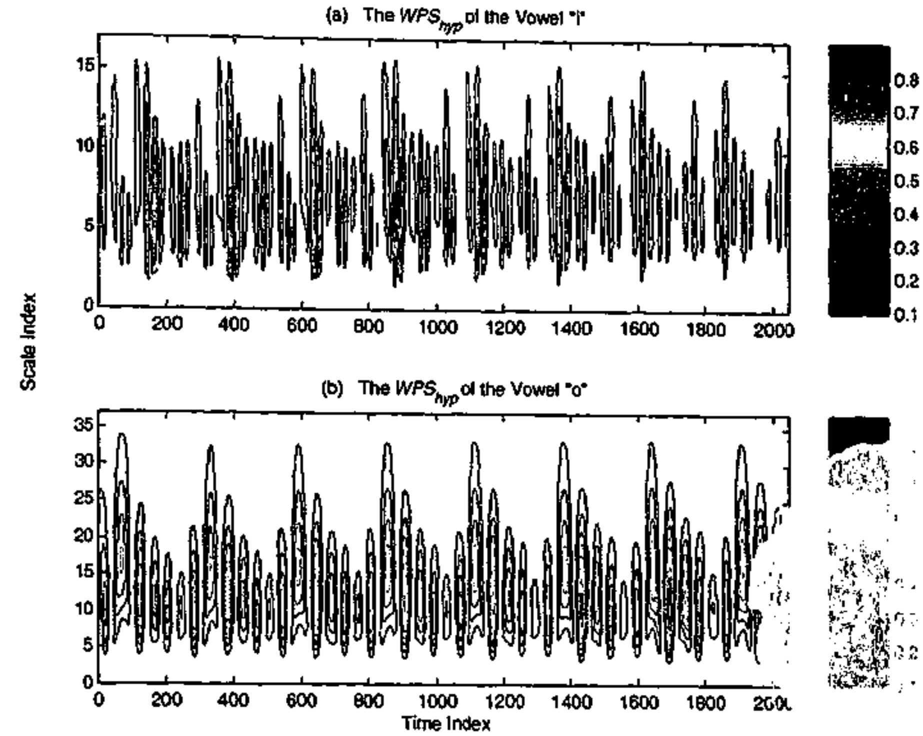

Figure 7.2.6: Contour plots of the WPS_{hyp} of the speech vowels "i" and "o"

As expected, the WPS_{hyp} of the vowels "i" and "o" are consistent with the WPS_{sym3} given in Figure 7.2.4 and Figure 7.2.5 in which periodicity is successfully detected. The energy patterns of the vowels "i" and "o", which are shown by using the hyperbolic wavelet, are consistent with those obtained by using the *sym3* wavelet. The occupied scale ranges of the WPS_{hyp} of the vowels "i" and "o" are short, which once again indicates strong periodic characteristics. It should also be noted that there are no changes in the scale range of the spectral components of the WPS_{hyp} of the vowels "i" and "o" as it is the case for the WPS_{sym3}. This is an advantage of the hyperbolic wavelet to the *sym3* wavelet which is mainly due to its perfect symmetry. It is also important to stress that if the energy density of a signal is distributed over a narrow scale range and even though there is component-variation in the signal, the signal is likely not chaotic since its energy is not broadly distributed. This fact should be clearly understood since there are a number of signals that have component-variation characteristics, however, are not chaotic.

The WPS_{sym3} of the vowel "u" is given in Figure 7.2.7.

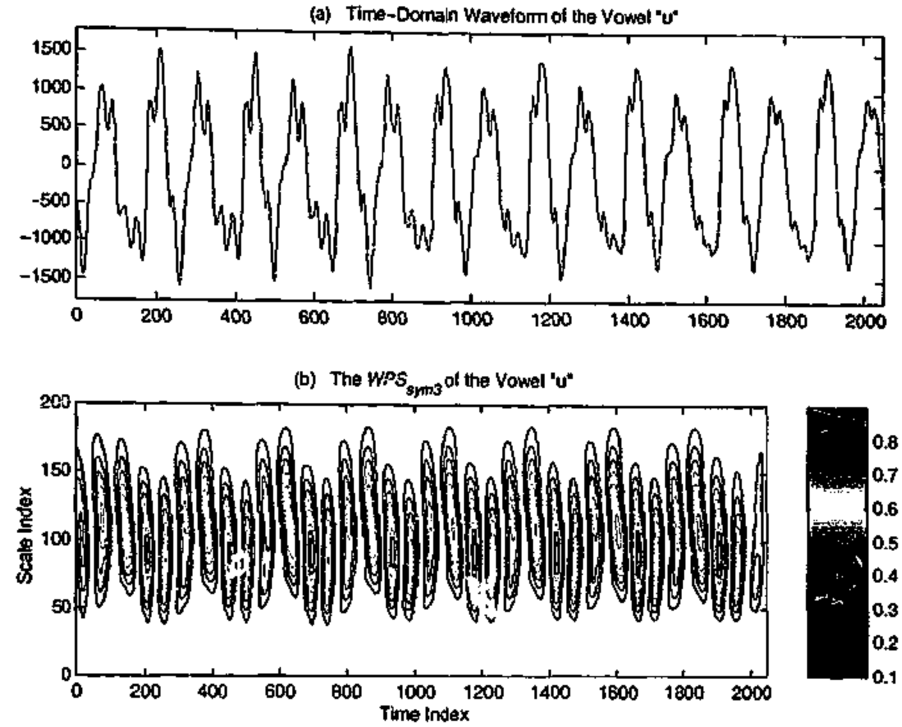

Figure 7.2.7: Speech time-domain of the vowel "u" and contour plot of its WPS_{sym3}

The WPS_{sym3} of the vowel "u" is discrete in which fundamental peaks at high contour scales are clearly shown, which shows strong periodicity in the signal. This signal can be regarded as similar to Duffing Period 1 waveform except that there are three sub-harmonics associated with the main harmonic located at the approximate scale of 100. The energy distribution is repetitive over regular time intervals. The WPS_{sym3} of the vowels "u" and "e" (given in Figure 7.2.2) are similar in which both wavelet power spectra do not consist of sub-harmonics which might suggest that they are difficult for speech recognition.

The contour plot of the WPS_{sym3} of the "sh" sound is given in Figure 7.2.8 whose magnified plot is given in Figure 7.2.9. Figure 7.2.8 and Figure 7.2.9 have identical contour scales.

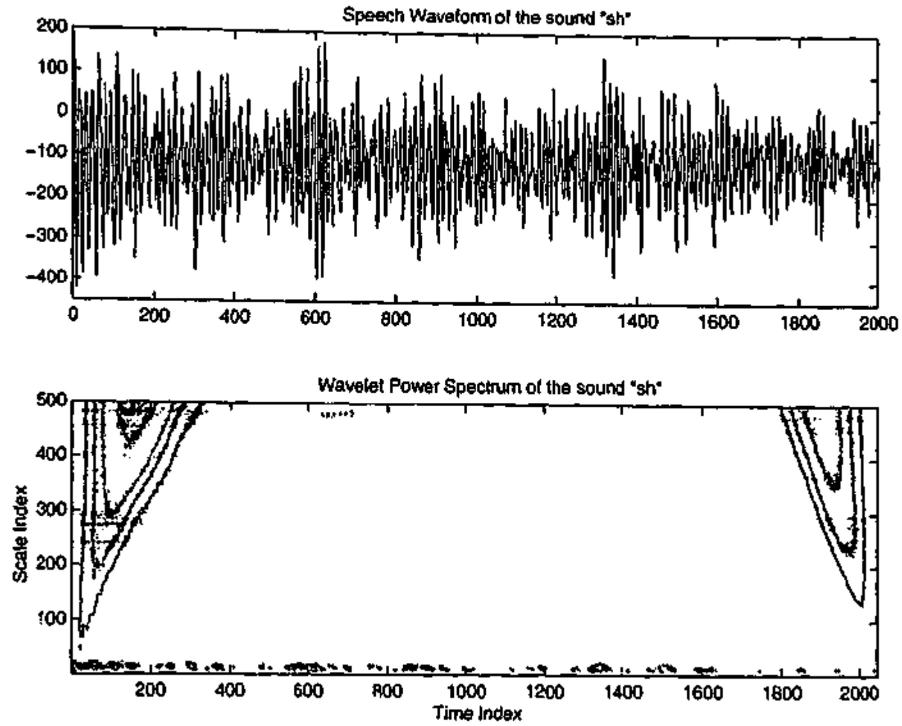

Figure 7.2.8: Speech time-domain waveform of the sound "sh" and contour plot of its WPS_{sym3}. The colour contour scale of the graph is given in Figure 7.2.9.

The sound "sh", due to its nature, is very fast and noisy. As can be seen from Figure 7.2.8, the usual discrete spectral components, as seen from the previous cases, disappear. Instead, there appears chaotic behaviour similar to Duffing Period 4 and Duffing chaotic waveforms studied in Section 6.2.4 in Chapter 6. However, harmonic peaks at very high scales (about larger than 500) are detected as can be seen in Figure 7.2.8 which suggests that the signal is not chaotic. The main characteristics of the signal are determined by the low-scale energy patterns whose magnified plot is given in Figure 7.2.9 in which the energy is unevenly distributed at the beginning of the signal. The samples in the middle of the signal are partly periodic and chaotic and there are no repetitive patterns in the wavelet power spectrum. This might suggest that the waveform is not periodic as it is usually the case for speech signals. The true characteristics of the waveform is determined when the magnified contour plot over the low-scale range is examined next.

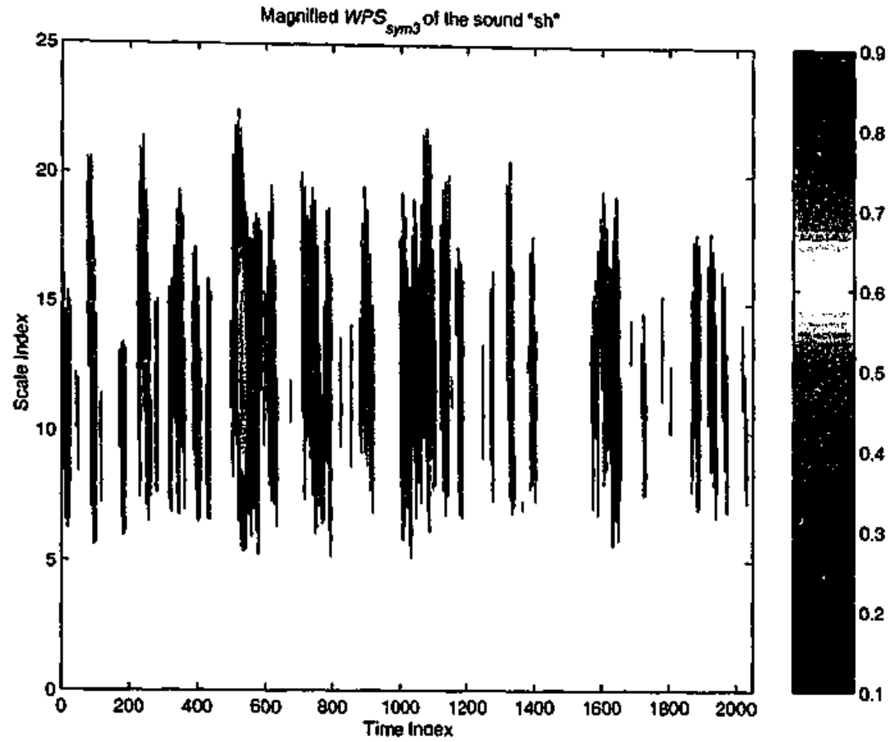

Figure 7.2.9: Magnified contour plot of the WPS_{sym3} of the speech sound "sh"

A comparison of Figure 7.2.9 to Figures 6.2.16 and 6.2.17 (Duffing chaotic waveform) shows that the "sh" sound speech signal is not chaotic since its energy patterns are sharp and discontinuous, whereas Duffing chaotic energy patterns are smooth. The sharp and fast-rising energy patterns of the "sh" sound signal do not indicate chaotic behaviour, but instead, indicate considerable sub-harmonics variation in the waveform. The fundamental peak of the sound at a high contour scale (Figure 7.2.8) is still present at $b \approx 550$ which validates the above suggestion that this signal is periodic. In addition, the scale range of this speech is short which strongly suggests that it is not chaotic. The wavelet power spectrum of the sound "sh" is similar to that of the vowel "i" (but having fewer harmonic peaks) in which spectral component-variation is present.

Contour plots of the WPS_{hyp} of the vowel "u" and the sound "sh" are shown in Figures 7.2.10(a) and (b) respectively.

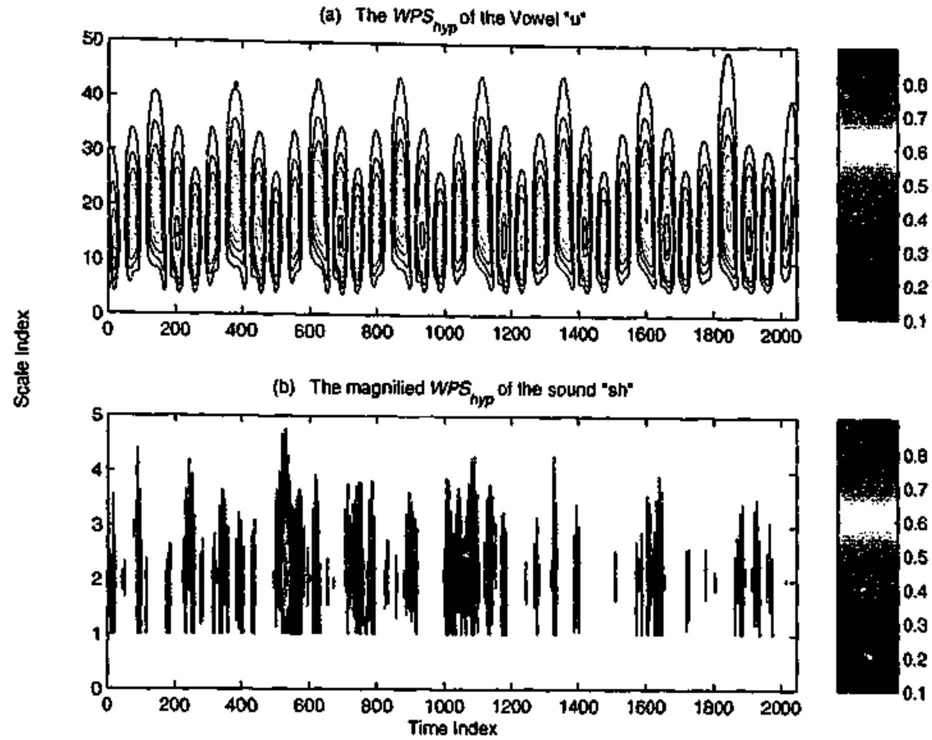

Figure 7.2.10: Contour plots of the WPS_{hyp} of the speech vowel "u" and the speech sound "sh"

The WPS_{hyp} of the vowel "u" is similar to the WPS_{sym3} in which three associated sub-harmonics are detected. Although periodic peaks are not clearly presented in the WPS_{hyp}, the speech "sh" waveform is not chaotic since its energy distribution is not smooth and distributed over a narrow scale range as can be seen in Figure 7.2.10(b). The energy of the "sh" sound signal appears to flow in bursts at uneven time intervals which means there is component-variation in the signal.

The new wavelet power spectrum technique, in particular using the hyperbolic and *sym3* wavelets, has been effectively used to study speech signals. It has been shown that true characteristics of various musical and speech signals can be effectively studied by examining their hyperbolic and *sym3* wavelet power spectra. The Fourier power spectrum technique is now employed to examine these signals so that the effectiveness of each technique can be clearly identified. The Fourier power spectra of all the speech vowels and the sound "sh" are shown in Figure 7.2.11 and Figure 7.2.12 respectively.

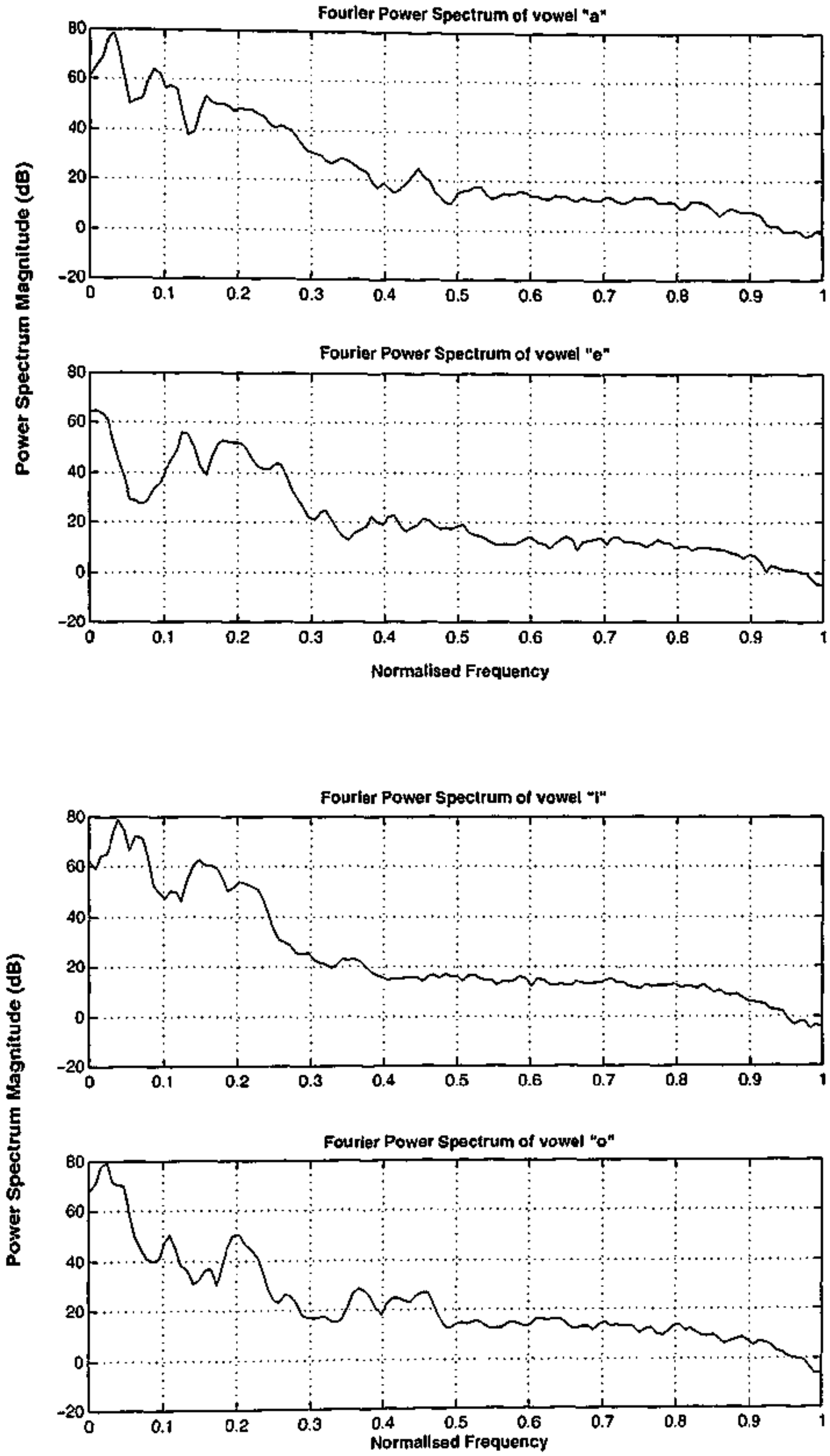

Figure 7.2.11: The Fourier power spectra of the vowels "a", "e", "i" and "o"

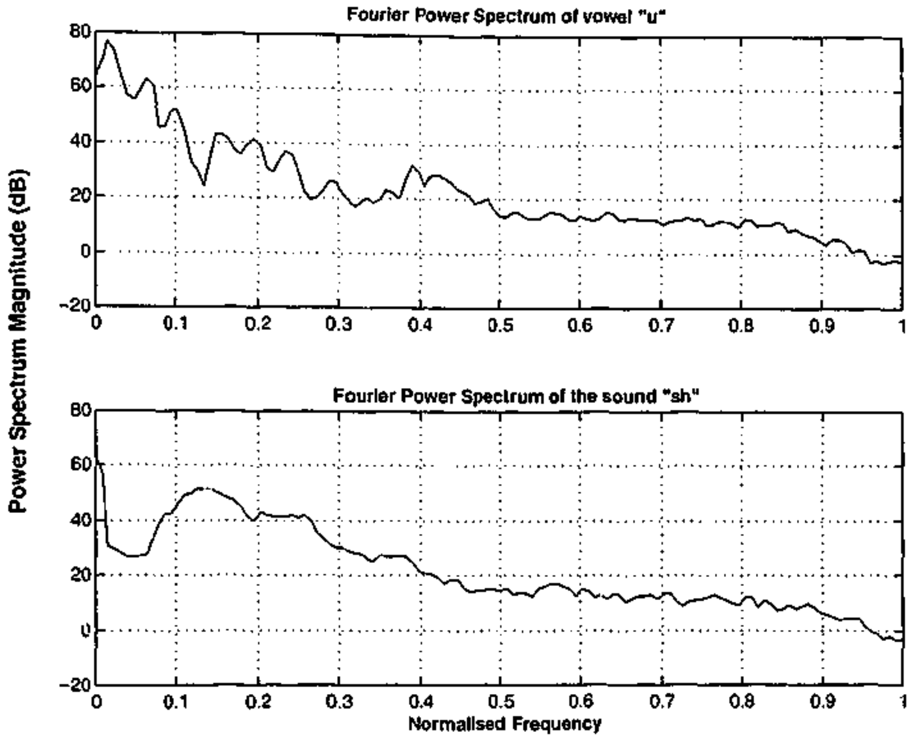

Figure 7.2.12: Fourier power spectra of the vowel "u" and the sound "sh". The frequency of each signal is normalised by dividing every frequency bin by the largest frequency bin of the signal spectrum. Thus, the shape of the power spectrum will not be changed by the normalisation process.

As can be seen from Figure 7.2.11 and Figure 7.2.12, the Fourier power spectra of speech signals exhibit distinctive peaks which suggests that they are periodic. This conclusion is consistent with the results obtained using the wavelet power spectrum technique. The "sh" sound, although has a broad Fourier power spectrum, still does have one distinctive peak at a high contour scale. Evidently, by using the Fourier power spectrum technique, it is not possible to detect instantaneous behaviour of the signal as it can be done using the wavelet power spectrum technique. This makes the wavelet transform method more applicable and effective than the Fourier transform method.

7.3 Remarks

Music and speech signals have been studied in this section by using the hyperbolic wavelet power spectrum technique, *sym3* wavelet power spectrum technique and the Fourier power spectrum technique.

For music signals, their wavelet power spectra should be calculated by using the entire signal of 44,100 samples to determine the true characteristics. The wavelet power spectra of data sets of 2,048 samples only reveal the instantaneous characteristics of the signals. Thus, music signals are quite different to other signals in which long data records are required so that misleading information about the signals can be avoided.

Speech signals of the English vowels and the sound "sh" are periodic signals with discrete and repetitive harmonic and sub-harmonic peaks. There is component-variation within speech signals. However, they are not chaotic since their energy is not broadly distributed over a wide scale range even though there is component-variation.

One disadvantage of the wavelet power spectrum technique is that for some cases such as music signals, the input signal must be taken long enough to effectively detect its characteristics. Usually, the number of data samples for input signals is larger than 10,240.

Another disadvantage of the wavelet power spectrum technique is that for some cases, the Fourier power spectrum is required to differentiate signals with similar characteristics, which could not be clearly revealed by the wavelet power spectrum technique. For example, the wavelet power spectra of the ECG and exponentially decaying sinusoidal with a non-zero final value signals are identical but their Fourier power spectra are quite different. However, for other cases, the Fourier power spectrum technique is not effective compared with the wavelet power spectrum technique, e.g. music and speech. Therefore, depending upon specific applications, the appropriate technique is used. In general, both techniques should be employed to effectively reveal the signal characteristics.

7.4 Conclusion

Music and speech signals have been examined using the wavelet power spectrum and the Fourier power spectrum techniques. Music signals behave chaotically even though the wavelet power spectra of different sets of the signals appear to be instantaneously periodic. This can be explained by the fact that harmonic peaks of musical signals abruptly disappear over time which reflects the disordered characteristics of these signals. The findings in this chapter agree with results found by other researchers [103-111].

Speech signals have been found to be periodic with strong harmonic peaks and several sub-harmonics. It has been shown that remarks drawn by using the wavelet power spectrum and Fourier power spectrum techniques are consistent. Some speech signals exhibit component-variation but they are not chaotic since their wavelet power spectra are not broad and occupy narrow scale ranges. The wavelet power spectrum technique seems to be more effective than the Fourier power spectrum technique in studying music and speech signals. On the other hand, it is not effective in studying the ECG signal as was discussed in Chapter 6.

The contribution of this chapter is that the hyperbolic wavelet power spectrum is employed as a new technique to effectively study music and speech signals. This might lead to a new way of using more sophisticated time-frequency techniques for signal analysis and other aspects of signal processing such as signal detection (Chapter 5), time-frequency higher-order statistics such as time-frequency bispectrum and speech recognition.

The hyperbolic wavelet is more efficient than the *sym3* wavelet because its total number of scales is smaller than that of the *sym3* wavelet due to the compression effect. In addition, for some speech signals, the hyperbolic wavelet is more effective than the *sym3* wavelet by clearly displaying the signal harmonics and sub-harmonics. The hyperbolic wavelet power spectrum technique is also able to focus on main harmonics of an input signal which leads to a disadvantage of missing fine sub-harmonics. Overall, the hyperbolic wavelet is an efficient and suitable wavelet for the wavelet power spectrum technique.

The next chapter applies parallel computing to calculate the bispectrum and time-frequency power spectrum by using a parallel computer.

Chapter 8: PARALLEL COMPUTATION OF THE BISPECTRUM AND HYPERBOLIC TIME-FREQUENCY POWER SPECTRUM

Chapters 6 and 7 examined the wavelet power spectrum technique and mentioned the bispectrum technique as effective tools for signal analysis, especially, in detecting chaos and non-linearity. Although the bispectrum and hyperbolic time-frequency power spectrum techniques are effective, these techniques require a large amount of computation time if using a serial-processing computer which limits their practical applications. Thus powerful computing tools are required to improve their efficiency so that they can be widely used. Parallel computing is one such tool. In parallel computing, a large task is split into smaller tasks so that they can be concurrently executed by multiple independent processors.

This chapter shows that parallel computing is an appropriate and effective tool to solve computationally intensive tasks in signal processing. The bispectrum and hyperbolic time-frequency power spectrum are two typical applications that are investigated by using parallel programming. In this chapter, background of the bispectrum is given in Section 8.1. The hyperbolic time-frequency power spectrum was studied in detail in Chapter 3 and its basic relations are briefly repeated in this chapter.

The bispectrum and hyperbolic time-frequency power spectrum are closely related by the auto-correlation function which was used to calculate the power spectrum in Chapter 3, Section 3.2, which is the main reason why the bispectrum is included in this chapter. The theoretical background of the bispectrum is given first to form a fundamental background for the hyperbolic time-frequency power spectrum which will be discussed in the next section.

8.1 Theoretical Background of the Bispectrum

Parallel programming and parallel machines have been extensively studied and used over many decades mainly for predicting weather patterns [93] and in image processing, [94, 95] where tasks, which are computationally expensive, are executed. However, to the best of our knowledge, parallel programming techniques have not been widely used in the field of higher-order statistics and higher-order spectra such as the bispectrum. Recently, two papers were published applying parallel computing in estimating the bispectrum. The first paper [96], which was published in 1991, reported the performance of an 8-CPU shared-memory CRAY Y-MP machine and 1024-CPU distributed-memory n-CUBE machine in calculation of the bispectrum. In particular, the speedup factor of the bispectrum calculation process was measured and compared for different machine configurations in which a near super-linear speedup factor was obtained. The second paper [97] proposed an algorithm to estimate higher-order moments using the MASPAR-1 machine, which is a SIMD (Single Instruction Multiple Data) machine.

In this chapter, the bispectrum is estimated by using two different methods namely *direct* and *indirect*. The *direct* method employs the 1-D FFT algorithms and the *indirect* method employs the 2-D FFT algorithm to estimate the bispectrum. Both methods are implemented by using 2 different parallel programming techniques: semi- and full-automatic or the Power C Analyser (PCA). The Silicon Graphics Power Challenge Multiprocessor System (with 12 CPUs) is used to run the parallel codes.

The *direct* method is used to estimate the bispectrum [86, 87], which was given by Eq. (6.1.6) in Chapter 6, and is repeated here by Eq. (8.1.1)

$$B(f_1, f_2) = \hat{X}(f_1) \cdot \hat{X}(f_2) \cdot \hat{X}^*(f_1 + f_2) \tag{8.1.1}$$

where $\hat{X}(\cdot)$ is the 1-D Fourier transform of a given discrete series $x(n)$ of M samples. For more information on the bispectrum, the interested reader should consult references [86, 87].

The *indirect* method uses the 2-D Fourier transform (calculated using the 2-D FFT technique) of the tricorrelation function $R_{xxx}(\tau_1,\tau_2)$

$B(f_1,f_2) = F\{R_{xxx}(\tau_1,\tau_2)\}$, where

$$R_{xxx}(\tau_1,\tau_2) = \sum_{n=0}^{M-1} x(n)\cdot x(n+\tau_1)\cdot x(n+\tau_2);\ \tau_1,\tau_2 = 0, 1, 2, \ldots, M-1 \qquad (8.1.2)$$

The bispectrum has been shown to be a useful tool in the study of chaos and behaviour of non-linear systems [1, 98, 99]. However, since the Fourier transform is most suitable for wide sense stationary signals only (as stated in Chapter 5), the bispectrum therefore has its limitations when the input signal is non-stationary. If the input signal is not wide sense stationary, then the frequency contents of the signal (or energy density) will change with time. By using the Fourier transform, fine-detailed information of the signal energy density will be lost since the Fourier transform method averages the energy density over an infinite time interval. It has been observed that the fine-detailed information of the energy density is useful to detect the transitions to chaos and turbulence [33, 34] as was shown in Chapters 6 and 7. To be able to observe time and frequency variation of the energy density of a non-stationary signal, a time-frequency power spectrum technique is employed. The time-frequency power spectrum technique displays the energy density of a signal as a function of time (t) and frequency (f) which means that the signal characteristics can be instantaneously monitored. The general formula of the time-frequency power spectrum is given by Eq. (8.1.3)

$$P(t,\omega) = \frac{1}{4\pi^2}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \underbrace{\left[e^{-j\theta(t-u)}\cdot\Phi(\theta,\tau)\right]}_{W(t-u,\tau)}\cdot e^{-j\tau\omega}\cdot R_{t,1}(t,\tau)\,du\,d\tau\,d\theta \qquad (8.1.3)$$

where $x(n)$ is the input signal, $\Phi(\theta,\ \tau)$ is the kernel function and $R_{t,1}(t,\tau) = x\left(u+\frac{\tau}{2}\right)\cdot x^*\left(u-\frac{\tau}{2}\right)$ is the local auto-correlation function. The formulas of the first-order hyperbolic and other kernels were given in Chapter 3. More details on the general time-frequency power spectrum may be found in [9, 28].

The discrete version of Eq. (8.1.3) is given by Eq. (8.1.4) [28]

$$DTF(n,k)=2\sum_{\tau=-\infty}^{\infty} e^{-j2\pi k\tau/M}\sum_{\mu=-\infty}^{\infty} W(\mu,\tau)\cdot x(n+\mu+\tau)\cdot x^{*}(n+\mu-\tau) \qquad (8.1.4)$$

As was evident from Eqs. (8.1.3) and (8.1.4), the amount of required computation of the time-frequency power spectrum is very large due to the triple integral and the two running variables t and ω. The resultant serial program consists of four nested `for` loops. If the input signal has 256 samples, then the number of required iterations will be $(256)^4$ = 4,294,967,296 which is a large number which makes the calculation process time consuming if it is run serially. For the bispectrum, to calculate the bispectrum of the ECG using 10,240 samples would take more than a day running on a normal PC. Thus, parallel computing is employed with the hope that it can reduce the amount of computation time of these tasks.

Sequential C programs were written first based on Eqs. (8.1.1) and (8.1.2) to estimate the bispectrum using the *direct* and *indirect* methods respectively. Then semi-automatic and full-automatic parallel (PCA) programs were constructed based on the sequential programs. Semi-automatic programs are obtained by inserting `#pragma` directives into the sequential program at appropriate points. This technique is based on a coarse-grained technique whereas the PCA method is based on a fine-grained technique. Also, arrays and loop parameters of the sequential program are controlled so that they can be independently run by different CPUs to avoid data dependency. The PCA parallel technique is activated by running the `-pca` flag of the Power C compiler.

For the hyperbolic time-frequency power spectrum, it should be noted that due to the immaturity of parallel compilers, the PCA or full-parallel method might not provide satisfactory results. The PCA method does not efficiently parallel the serial program due to data dependency even though the program is free from data dependency. The compiler only parallels a loop or breaks it into smaller tasks if and only if it knows that the loop is free from data dependency. This is usually the case for single loops, but not for nested loops which contains most of the computational burden. Thus, semi-automatic parallel method seems to be the most appropriate way to apply parallel programming although it requires in-depth understanding of the programming language and the structure of the parallel system.

This chapter focuses on the effectiveness of the Silicon Graphics Power Challenge multiprocessor shared-memory MIMD (Multiple Instruction Multiple Data) machine (HOTBLACK)* in calculation of the bispectrum and the hyperbolic time-frequency power spectrum. Each CPU can be considered as an independent processor with a separate local memory and cache. To effectively program the machine, it is important to arrange the loop parameters and data structure inside the programs so that they are suitable for the specific configurations of the computer. This is the most difficult part of parallel programming in which the programmer must understand the configuration of the particular machine. The semi- and full-automatic parallel methods are employed to run the programs so that their speedup factors can be accurately measured and compared. The semi-automatic method employs the coarse-grained parallel technique which usually gives a better measured-speedup factor. The full-automatic method employs the fine-grained parallel technique which parallels a large number of small loops and usually results in higher parallel overhead compared with the semi-automatic method. For the hyperbolic time-frequency power spectrum calculation process, the semi-automatic method is only used. The full-automatic PCA method is not employed since it results in very large parallel overhead which causes a poor measured-speedup factor. The effectiveness of the semi- and full-automatic parallel methods in calculating the bispectrum is examined in the following section.

8.2 Parallel Computation of the Bispectrum

In the HOTBLACK parallel computer, there are 12 parallel programs with different `numthread(N)` to run on 12 different processors on the system. `numthread(N)` is a parallel directive from Silicon Graphics that allows the program to be executed in parallel using N independent CPUs. For example, if $N = 3$ then the program will be executed in parallel using only 3 CPUs. To ensure efficient compilation, the programs are submitted into a batch queue to obtain more `CPU_time`, `memory_use` and `stack_data_size` quota. Four script files have been written to run the programs under the UNIX operating system.

* HOTBLACK is a local name of the machine at Monash University.

The measured-speedup factor is estimated as

$$\text{Measured - Speedup Factor} = \frac{\text{Sequential\_Time}}{\text{Parallel\_Time}} \tag{8.2.1}$$

where the Sequential_Time is the real CPU time used to run the sequential source code and the Parallel_Time is the real CPU time of the slowest CPU employed to run the parallel code.

The efficiency of a parallel program is estimated as

$$\text{Parallel\_Efficiency} = \frac{\text{Measured\_Speedup}}{\text{Ideal\_Speedup}} \tag{8.2.2}$$

Theoretically, the ideal or super-linear measured-speedup factor of a program is defined as N if the parallel program is run using N CPUs [101]. Practically, the measured-speedup factor is always less than the super-linear speedup factor due to parallel overhead implying that parallel efficiency is less than 100%.

To ensure consistency between parallel and sequential programming techniques, their resultant output files are compared and it has been observed that they are identical. If the files are not identical, data dependency must have occurred in the sequential source code. The difference between the fine-grained and coarse-grained parallel techniques is that, for the former, which is used by the PCA method, only small repeated loops are paralleled which yields more than one parallel loop in a parallel program and thus results in more parallel overhead. For the coarse-grained parallel technique, the largest parallel loop that has the largest workload in the program is paralleled. Usually, there is only one largest repetitive loop in the program. Since the number of paralleled loops in the coarse-grained technique is much less than that in the fine-grained technique, the amount of parallel overhead in the former is much less than that in the latter.

For comparison, the size of each segment of the *direct* method is 2,048 data points which is twice that of the *indirect* method of 1,024 data points (since 1,024 data-point segments are still not large enough for the *direct* method, 2,048-data-point segments are used instead, also longer segment size of up to 10,240 points can be used). From simulation results, it has been observed that the serial program of the *direct* method took approximately

23 seconds to run compared to 587 seconds running time of the *indirect* method although the segments in the latter method are half as long. Thus the *direct* method (Eq. (8.1.1) or equivalent) is more computational efficient than the *indirect* method. The measured-speedup factors of the semi- and full-automatic parallel methods are plotted as a function of the number of processors N in Figure 8.2.1; the parallel efficiency of each technique is plotted in Figure 8.2.2. In these figures, Direct and Indirect are the measured-speedup factors using the semi-automatic parallel method for the *direct* and *indirect* methods respectively. Direct PCA and Indirect PCA are measured-speedup factors obtained using the PCA full-automatic parallel method for the *direct* and *indirect* methods respectively.

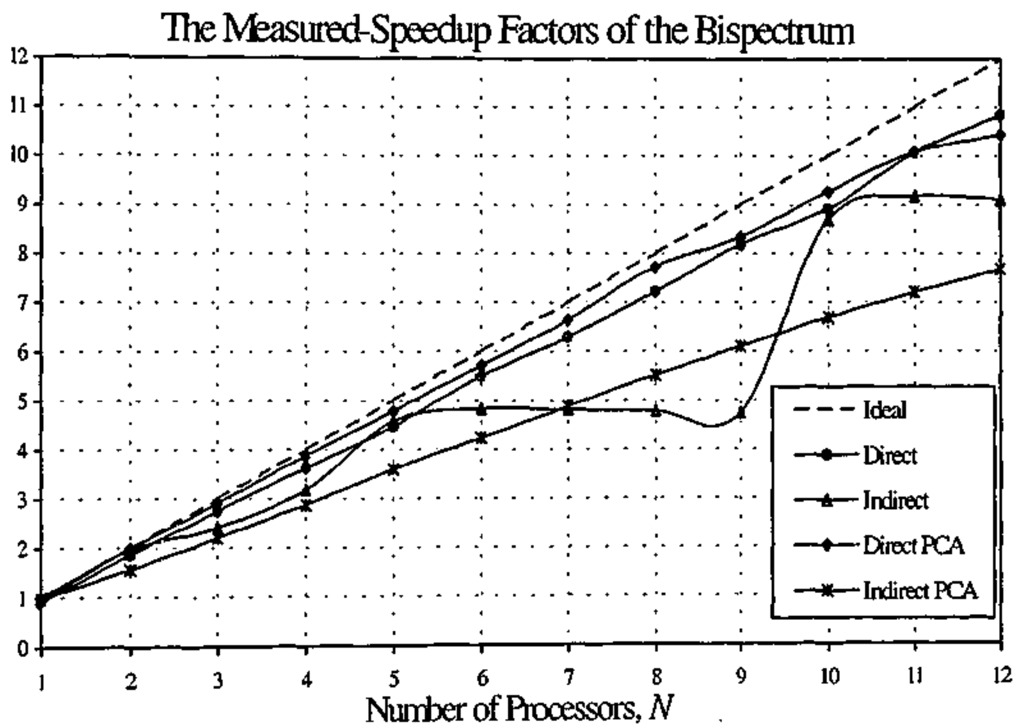

Figure 8.2.1: The measured-speedup factors (based on Eq. (8.2.1)) of the semi- and full-automatic methods using the *direct* method (Eq. (8.1.1)) and *indirect* method (Eq. (8.1.2)) as a function of N

As N increases, the amount of parallel overhead increases due to synchronisation and waiting time of the processors. However, near super-linear measured-speedup factors are obtained by using the semi- and full-automatic parallel methods for the *direct* method as seen in Figure 8.2.1. For the *direct* method using the PCA parallel method and for large values of N (for instance, at N = 12), the measured-speedup factor starts to decrease which illustrates the limitation of the fine-grained parallel technique: large amount of parallel overhead are generated for large values of N which lowers the measured-speedup factors (the Direct_PCA and Indirect_PCA curves in Figure 8.2.1). For $N \leq 12$, the PCA method provides a near-linear measured-speedup factor which indicates that the fine-grained technique is more suitable for the *direct* method (using the PCA method) than the coarse-grained technique. If the segment size increases further, near-linear measured-speedup factors might not be obtained for the *indirect* method due to a long waiting time as explained in the following section.

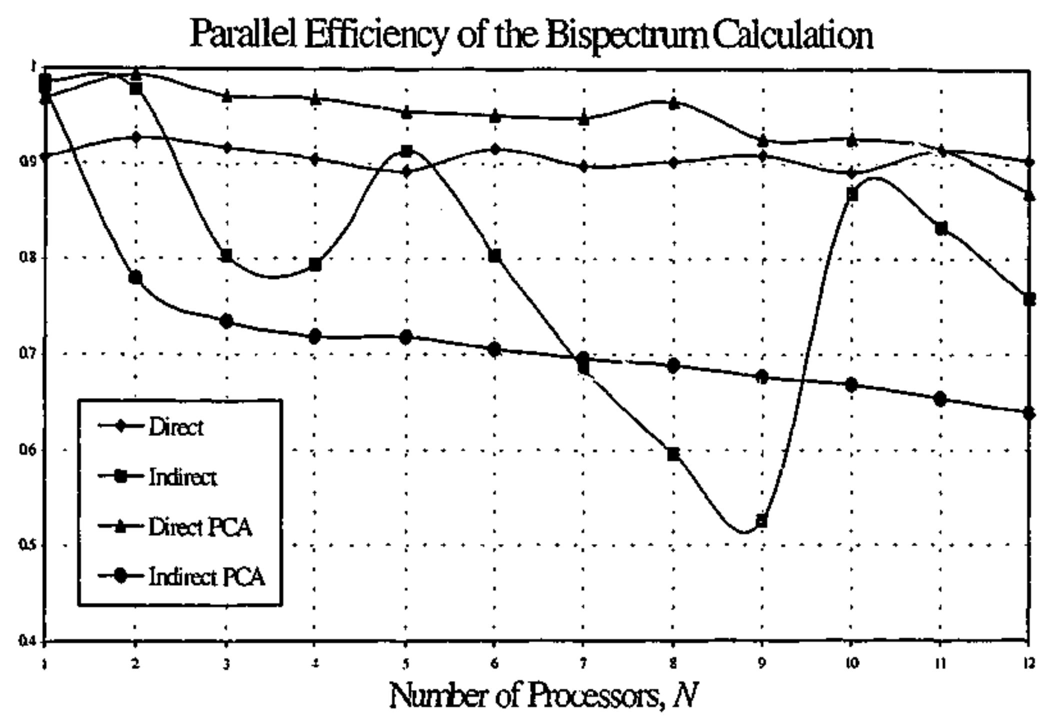

Figure 8.2.2: Parallel efficiency (based on Eq. (8.2.2)) of the semi- and full-automatic parallel methods versus the number of processors

For the *indirect* method using the semi-automatic parallel method, near super-linear measured-speedup factors are obtained only with some specific numbers of CPUs which are multiple of the loop size of 10. That means when N = 1, 2, 5 or 10, near super-linear measured-speedup factors will be obtained. For other values of N, since the associated work with each iteration of the loop is large, there will be "unemployed" processors waiting for other processors to complete their tasks. For example, if N = 6 all six CPUs will be assigned to the first six iterations of the loop. After finishing the 6 iterations, four of the six CPUs will be used to complete the remaining 4 iterations and two CPUs have to wait ("spin") until the iterations are finished. Since the associated work of each iteration is large, this results in a long waiting time and thus the measured-speedup factors and parallel efficiency will be lowered as illustrated for cases of $N = 3, 4, 6, 7, 8, 9, 11$ and 12 in Figure 8.2.1 and Figure 8.2.2. Hence, if N is not a factor of the loop size (which is 10) and if the work of each iteration is large, increasing N will increase parallel overhead and constrain the measured-speedup factor.

For the *indirect* method applying the PCA parallel method, the measured-speedup factor linearly increases as N increases although with lower values compared to the case of semi-automatic parallel method due to high parallel overhead in several small parallel loops. However, the performance of the PCA method is predictable. From Figure 8.2.1, in contrast with the *direct* method, the coarse-grained technique is more suitable for the *indirect* method than for the *direct* method since a better measured-speedup factor is achieved. However, to obtain satisfactory performance, N must be chosen to be a factor of the loop size. Table 8.2.1 compares the maximum measured-speedup factors of the semi- and full-automatic parallel methods applied to the *direct* and *indirect* methods.

Table 8.2.1: Maximum measured-speedup factor comparison of the semi-automatic and PCA parallel methods (using the *direct* and *indirect* methods) for the bispectral parallel computation process

Method	Semi-automatic measured-speedup factor	Full-automatic PCA measured-speedup factor
Direct calculation of the bispectrum	10.84 at $N = 12$	10.44 at $N = 11$
Indirect calculation of the bispectrum	9.17 at $N = 11$	7.67 at $N = 12$

From Figure 8.2.1 and Figure 8.2.2, the bispectrum measured-speedup factor increases but its parallel efficiency decreases as N increases. Thus, there ought to be a realised speedup factor or effective-speedup factor that considers both the measured-speedup factor and parallel efficiency by forming their product as a function of N. The effective-speedup factor reflects the effectiveness of a parallel calculation procedure and should be considered important. Thus, when choosing a particular value of N for designing purposes, the measured-speedup factor, parallel efficiency and effective-speedup factor should all be considered. Depending on a particular application (whether the measured-speedup factor or parallel efficiency is the first priority), the appropriate parameter is used. If both of them are required, the effective-speedup factor is employed. The effective-speedup factor of the bispectrum is displayed in Figure 8.2.3.

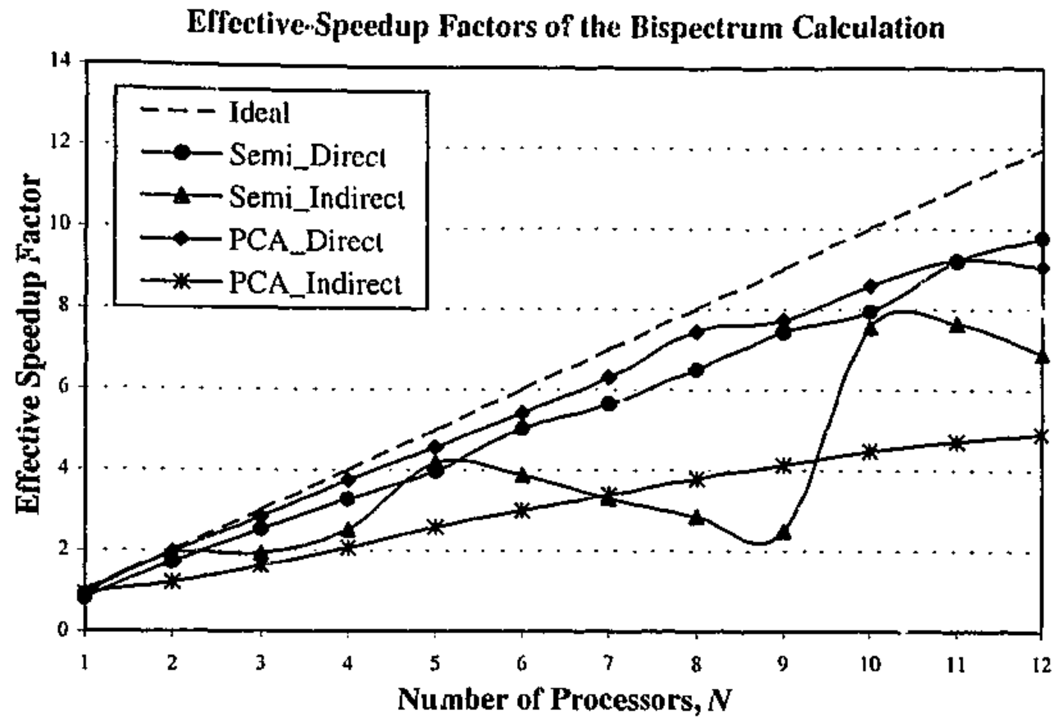

Figure 8.2.3: The effective-speedup factors of the bispectrum calculation. The measured-speedup factors are not included here for simplicity.

This section has presented the results of parallel computation of the bispectrum, the next section shows that the hyperbolic time-frequency power spectrum can also be efficiently calculated by using parallel computing.

In this thesis, two signal-processing techniques have been proposed. Firstly, the hyperbolic kernel and secondly, the hyperbolic wavelet which can be generated from the hyperbolic kernel. The bispectrum has not been thoroughly studied in this thesis and yet it has been calculated using a parallel computer to reduce its computational burden. The hyperbolic time-frequency power spectrum will be estimated using a parallel computer in the next section. However, the hyperbolic wavelet transform and wavelet power spectrum have not been mentioned or calculated using a parallel computer up to this point. The main reason for this is that the wavelet transform and wavelet power spectrum have been extensively estimated using a parallel computer in the literature [114-126] and thus this chapter will not pursuit further the already-established results. It has been found that the wavelet transform can be efficiently estimated with a near-linear measured-speedup factor [114-117, 119, 120, 121, 123, 124] by using a parallel computer.

The wavelet transform and wavelet power spectrum are time-frequency techniques and thus they are closely related to the time-frequency power spectrum technique via the local auto-correlation function (Eq. (8.1.3)). The wavelet power spectrum technique has been shown to be useful in studying signal characteristics in Chapters 6 and 7. Some of the parallel techniques used to estimate the wavelet such as the coarse-grained parallel technique [117, 121] will be employed to estimate the hyperbolic time-frequency power spectrum.

8.3 Parallel Computation of the Hyperbolic Time-Frequency Power Spectrum

In Section 8.2, the bispectrum was estimated using a super parallel computer and that a near-ideal measured-speedup factor was achieved in estimating the bispectrum. In this section the computer is utilised to estimate the hyperbolic time-frequency power spectrum. From that, suggestions can be made whether the hyperbolic time-frequency power spectrum is suitable for parallel computing analysis. Fundamental background of the hyperbolic time-frequency power spectrum was given in Section 8.1 from which the serial program to calculate the hyperbolic time-frequency power spectrum is constructed.

The coarse-grained parallel technique is used for the semi-automatic method and the fine-grained parallel technique is employed for the full-automatic PCA method as was applied in the case of the bispectrum calculation process. However, only the semi-automatic parallel method is investigated due to inefficient coding of the C annotator (full-automatic PCA method). The compiler mistook the inner loop as the most efficient loop for parallel programming, in other words, it inefficiently employed the fine-grained parallel technique which resulted in long parallel overhead and lowered the measured-speedup factor. For the semi-automatic parallel method, the structure of the parallel program is manually constructed thus the coarse-grained parallel technique can be efficiently employed which gives a better measured-speedup factor and parallel efficiency. This agrees with previous work by Suzuki *et al* [117] and Lihua and Misra [121].

Key factors that affect the performance of parallel programs are parallel load (the load among the processors should be evenly balanced), parallel overhead (the amount of communication among the processors should be minimised) and data dependency. There are four nested `for` loops in the program which require a large amount of computation. The measured-speedup factor is obtained by using 256-sample `for` loops. Since there are 4 nested `for` loops, the number of iterations would be $(256)^4$ = 4,294,967,296 iterations or approximately $4.3 \cdot 10^9$. Because the `stack_size` and `memory_size` quotas on the parallel computer on which all experiments were run are limited, a larger loop size (for example 1024 or larger) could not be performed.

The coarse-grained parallel technique is utilised in the outer-most loop of the program by dividing it into smaller tasks. Each independent task has 3 nested `for` loops, which can be concurrently executed by independent processors. The outer-most loop is run in parallel since it is the largest loop in the program, which requires the largest amount of computing power. Each iteration of the loop is independent of each other which considerably reduces parallel overhead among the processors.

It should be noted that time-frequency power spectra of various kernels are different in the kernel functions only and their general architectural structures are unchanged. In addition, the associated computation of the general structure (as stated earlier) is too large compared with the calculation of the kernel which involves simple multiplication and division operations. Thus, if the hyperbolic time-frequency power spectrum can be efficiently estimated, other time-frequency power spectra using different kernels such as Choi-Williams or Wigner-Ville can also be efficiently estimated. The parallel measured-speedup factor and efficiency are estimated by using Eqs. (8.2.1) and (8.2.2) and given in Figure 8.3.1 and Figure 8.3.2 respectively. Comparisons of the measured-speedup and effective-speedup factors are given in Figure 8.3.3.

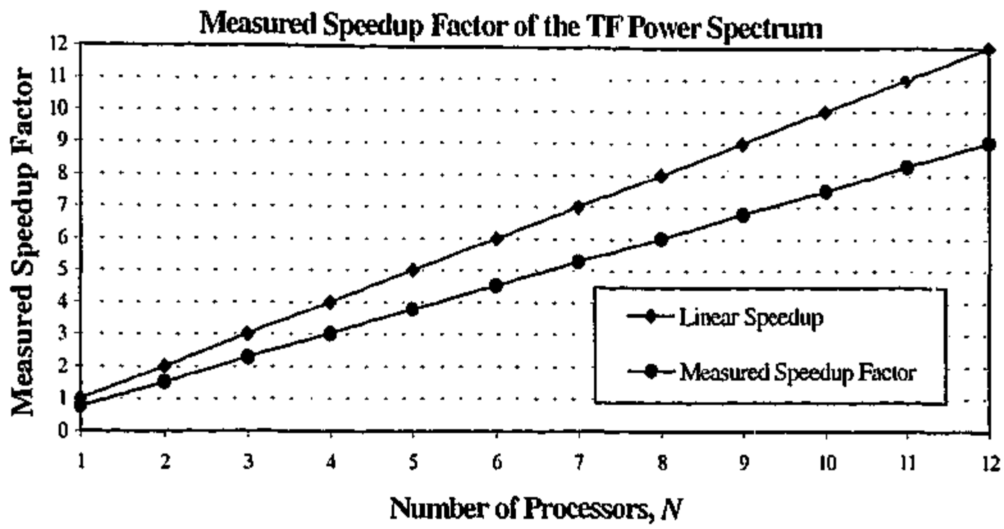

Figure 8.3.1: The measured-speedup factor of the hyperbolic time-frequency power spectral calculation process, the loop size is $M = 256$

The measured-speedup factor of the hyperbolic time-frequency power spectral parallel calculation process is evidently near super-linear. The measured-speedup factor linearly increases as there are no "humps" or sudden "jumps" in the curve as seen in Figure 8.3.1. This shows that the hyperbolic time-frequency power spectrum can be efficiently calculated using parallel computing. By comparing the measured-speedup factors of the hyperbolic time-frequency power spectrum and bispectrum, it is clear that the bispectrum can be more efficiently calculated than the hyperbolic time-frequency power spectrum. This is mainly due to the structure of the inner loops of the hyperbolic time-frequency power spectrum program because the number of `if` statements in the bispectrum program is less than that in the hyperbolic time-frequency power spectrum program which significantly slow the processors down.

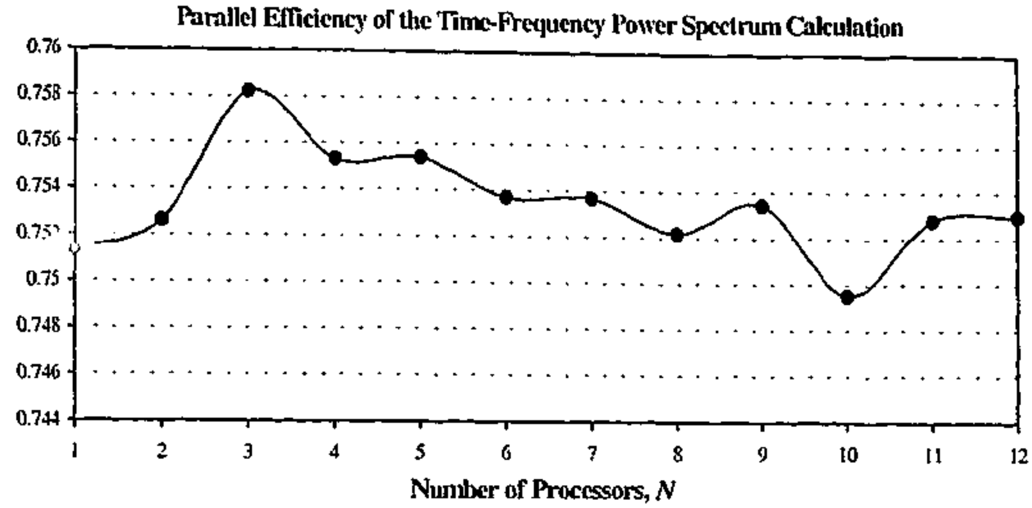

Figure 8.3.2: Parallel efficiency of the hyperbolic time-frequency power spectral calculation process, the loop size is 256

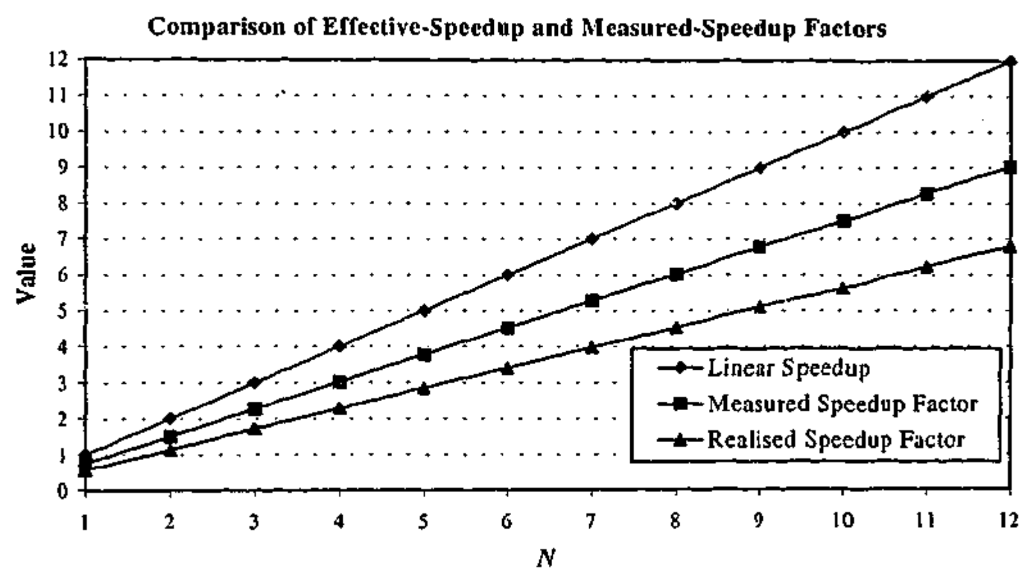

Figure 8.3.3: Comparison of the measured- and effective-speedup factors as a function of N

The hyperbolic time-frequency power spectral calculation process has lower parallel efficiency than the bispectral calculation (for the case of the *direct* semi-automatic and *direct* PCA full-automatic method only) process because the number of nested `for` loops inside the main loop of the former program (4 `for` loops) is larger than that in the latter program (3 `for` loops). In addition, for each nested `for` loop, there is a summing mechanism, which must be performed so that the results are ready for the next outer loop. As a result, there is idle time in an outer loop since its inner loops are not always ready for calculation if all of their iterations are not completed. This creates a large amount of unavoidable parallel overhead (on average about 10 seconds compared with 1 second in the case of the bispectral calculation process).

The parallel efficiency of the hyperbolic time-frequency power spectral calculation process attains its minimum value of 74.9% at $N = 10$ and its maximum value of 75.8% at $N = 3$. The parallel efficiency remains unchanged with the average value of 75.2% for other values of N. From Figure 8.2.2 (bispectrum parallel efficiency) and Figure 8.3.2 (hyperbolic time-frequency power spectral parallel efficiency), the bispectral parallel efficiency is higher than the hyperbolic time-frequency power spectral parallel efficiency (except for the case of the *indirect* semi-automatic and PCA full-automatic methods) as previously explained.

It should be noted that the efficiency of the hyperbolic time-frequency power spectral parallel calculation process, though lower than that of the bispectrum, is quite stable, unlike the case of the bispectral parallel calculation process. This is mainly due to the size of the input data sets. For the case of the bispectrum, there are 10 segments, each consisting of 1,024 samples and the total number of samples is 10,240. If the number of samples per segment is less than 1,024, then there are no significant conclusions on the signal characteristics (for example, the ECG) because the data set is not large enough. Therefore, 10 iterations are chosen. Since the number of iterations of the parallel loop is small, the number of processors used to run the bispectrum parallel program plays an important role in speeding up the calculation process. Although there are only 10 iterations, the associated work with each iteration is large which results in long waiting time for the processors as explained in Section 8.2 for the *indirect* method.

For the hyperbolic time-frequency power spectrum, the number of iterations for the main parallel loop is 256, which is a large number compared with the number of processors of the system (12 independent processors). Thus, the waiting time of one processor is considerably small compared with the calculation time of other processors. In other words, if one or more processors wait for others to finish their tasks, then the waiting time is always considerably less than the useful time. For example, if the number of processors is $N = 5$, then the number of iterations that each processor has to complete is 51 and there is 1 iteration left. After the processors finish their 51-iteration work then one of them has to finish the remaining iteration and other processors (four in this case) have to wait or spin until that remaining iteration is completed. The spinning time associated with one iteration is clearly much less than the required time to complete 51 iterations. This explains why the parallel efficiency of the hyperbolic time-frequency power spectral calculation process is more stable than that of the bispectral calculation process.

Overall, although the parallel efficiency of the hyperbolic time-frequency power spectral process (75.2%) is not as high compared with that of the parallel bispectral calculation process (about 90% or higher for the *direct* method), it has been shown that the hyperbolic time-frequency power spectrum can be efficiently calculated by using parallel computing which is an encouraging result. The parallel efficiency of the hyperbolic time-frequency power spectral calculation process can be improved by increasing the loop size.

8.4 Conclusion

The contribution of this chapter is to improve the speed of computation of the bispectrum and hyperbolic time-frequency power spectrum processes by using parallel computing. Near-linear measured-speedup factors of the bispectral parallel calculation process have been achieved by using the semi- and full-automatic (PCA) methods for the *direct* method. For the *indirect* method, parallel overhead gradually increases when $N \geq 6$ due to the specific loop structure of the serial program. However, for $N \leq 5$ or $N = 10$, near-linear measured-speedup factors were obtained for the *indirect* method by using the semi-automatic method. Thus it can be concluded that the *direct* method of bispectrum computation is more suitable for parallel programming than the *indirect* method. The PCA method can be used to achieve the measured-speedup factor of 7.67 at $N = 12$ (for the *indirect* method). However, the PCA method suffers from high parallel overhead for large values of N ($N \geq 12$) since a PCA parallel program (employing the fine-grained parallel technique) contains several small parallel loops.

For the hyperbolic time-frequency power spectral parallel calculation process, a near-linear measured-speedup factor has been obtained with the minimum efficiency of 75.13% at $N = 1$. The maximum efficiency of 75.58% was achieved when $N = 3$ at the measured-speedup factor of 2.3. At $N = 12$, the efficiency was 75.3% which corresponds to the measured-speedup factor of 9.03. The average efficiency is approximately 75.34%. It has been observed that the PCA method could not provide a suitable parallel solution to the serial program thus only the semi-automatic method was employed to parallel the hyperbolic time-frequency power spectrum serial program. From the obtained results, it appears that the hyperbolic time-frequency power spectrum can be efficiently calculated by using parallel computing. However, the efficiency of the process is not high compared with that of the bispectral calculation process. As stated earlier, other time-frequency power spectra using different kernel functions such as the Choi-Williams and Wigner-Ville can be efficiently estimated using parallel computing due to a small difference in the computational burden of these kernels.

The effective-speedup factors of the bispectrum and hyperbolic time-frequency power spectrum have also been obtained for completeness and for practical purposes. Since the efficiency is always less than unity, effective-speedup factors are always less than measured-speedup factors. As its name implies, the effective-speedup factor represents the true speedup factor gained by using a particular parallel system. This speedup factor can be used for applications in which both the measured-speedup factor and parallel efficiency are required so that the most suitable number of processors can be identified. It is important to realise that the more processors, the faster the computation can be carried out. On the other hand, simultaneously, parallel overhead will be increased. Thus, there must be a balance between the number of processors and parallel overhead in a parallel system. For the HOTBLACK system, it is clear that the calculation efficiency of the bispectrum and hyperbolic time-frequency power spectrum is always higher than 75% (except for the case of the *indirect* semi-automatic and PCA full-automatic methods) which shows that appropriate parallel methods have been employed and parallel overhead has been sufficiently small in our work. Future work can be carried out on different parallel systems to examine the effects of parallel overhead and the number of processors.

Chapter 9: CONCLUSIONS AND FUTURE RESEARCH

9.1 Summary and Conclusions

This research has explored the theoretical characteristics and some practical applications of the hyperbolic kernel family. The research can be divided into two parts: theory and application. The theoretical part consists of the first four chapters in which the hyperbolic kernel family and hyperbolic wavelet were investigated in detail in Chapter 3 and Chapter 4 respectively. In Chapter 4, the strong link between time-frequency kernels and wavelet functions was established which forms a foundation to expand the two areas of time-frequency and wavelet signal processing further. This is one of the major contributions of the thesis.

The application part of the research examines possible applications using the hyperbolic kernel and hyperbolic wavelet including signal detection in Chapter 5 and signal analysis using the hyperbolic wavelet power spectrum technique in Chapters 6 and 7. Chapter 8 presented parallel computing as a useful tool to improve the computational speed of the bispectrum and hyperbolic time-frequency power spectrum.

The first-order hyperbolic kernel of the hyperbolic kernel family has been found to be a simpler and more effective kernel than the popular Choi-Williams and multiform tiltable exponential kernels in some applications. The hyperbolic kernel has been shown to be cross-term effective and noise robust. However, the kernel has a coarse auto-term resolution which could result in weak support for the auto terms in the time-frequency plane. It has been shown that the new hyperbolic wavelet has a finer scale resolution compared with the Morlet and Choi-Williams wavelets. The hyperbolic wavelet has been shown to have the smallest total number of calculated scales among the three wavelets considered which enables compression ability. This significantly improves the efficiency of the hyperbolic wavelet power spectrum calculation process.

In signal detection, it has also been shown that the hyperbolic signal detector is more ive than the Choi-Williams, Wigner-Ville and matched-filter detectors by improving al-to-noise ratio by up to 40%. The discrete Moyal formula for non-unity kernels has en derived based on Moyal's formula for the unity kernel, i.e. the Wigner-Ville kernel. By deriving Moyal's formula for non-unity kernels, the effects of noise were considered and explored in detail using non-unity kernel signal detectors.

The hyperbolic wavelet power spectrum technique has been shown to be effective for stationary and non-stationary signal analysis. A gallery of the hyperbolic wavelet power spectra of various signals including ECG, speech, music, periodic sinusoidal and exponential has been established. Transitions into the chaotic region of the ECG and Duffing oscillatory signals have been successfully identified by using the hyperbolic wavelet power spectrum.

The trade-off among cross-term suppression, noise robustness, scale resolution, signal detection ability and auto-term resolution has been established. All of these features can be successfully achieved at the expense of having a poor auto-term resolution. In addition, as was shown in Chapters 4, 6 and 7, the hyperbolic wavelet power spectrum has been very effective in studying non-stationary signals with a finer scale resolution and a smaller scale range, compared with the Choi-Williams (Mexican-hat) wavelet and Morlet wavelet.

It has been shown that parallel computing can improve the efficiency of heavy computational tasks in signal processing such as the bispectrum and hyperbolic time-frequency power spectrum. Near-ideal speedup factors have been achieved by using the semi- and full-automatic parallel methods. The minimum parallel efficiency of calculating these tasks is 75% which shows that they can be efficiently calculated using a parallel computer.

In conclusion, the contributions of this thesis are:

- Proposal of the new hyperbolic kernel family,
- Discovery of the new hyperbolic wavelet which is generated from the first-order hyperbolic kernel,
- Identification of the link between the wavelet and time-frequency signal processing areas which enables expansion of time-frequency and wavelet analyses,
- Proposal of the new non-unity time-frequency detector, in particular the hyperbolic detector with improved *SNR*,
- Establishment of the trade-off among cross-term suppression, noise robustness, scale resolution, wavelet compression and signal detection against auto-term resolution
- Proposal of the hyperbolic wavelet power spectrum technique for signal analysis,
- Demonstration of the effectiveness of parallel computing in signal processing.

9.2 Future Research

The discovery of the hyperbolic kernel family has opened new research directions such as kernel design, wavelet theory, signal detection and signal analysis using the time-frequency power spectrum technique.

In this research, the first-order hyperbolic kernel $\Phi_{hyperbolic} = sech(\beta\theta\tau)$ has been extensively studied. However, other members of the family should be investigated. In particular, the third-order hyperbolic kernel $\Phi_{cubic\ hyperbolic} = [sech(\beta\theta\tau)]^3$, is an exciting kernel with useful features that are worth investigating. This kernel was mentioned a few times in the thesis and some of its properties such as volume under the surface and time-frequency power spectrum expression were given. However, its ability in suppressing cross terms, noise robustness, auto-term resolution and signal detection have not been examined. More work needs to be done so that the entire hyperbolic family kernel can be classified.

Chapter 4 showed that when new kernels were found, new wavelets could be generated. It would be interesting to investigate possible new wavelet functions that can be generated from the hyperbolic kernel family as it is not known whether there are useful wavelets that could be generated from the family.

Although there are many new research directions in the theoretical study of the hyperbolic kernel family, the applications of the family are also worth exploring. Detailed studies need to be done to establish a kernel that yields the corresponding time-frequency signal detector with a near perfect performance. Since the first-order hyperbolic kernel has been shown to be superior to the Choi-Williams kernel, this kernel could be one of the members of the hyperbolic family kernel or a kernel from another family.

Further work on parallel computing can be carried out by using the hyperbolic time-frequency bispectrum as a typical application. The associated computation of the time-frequency bispectrum is extensive since the spectrum must be able to "slide" to be successfully displayed. This area of research has not been done and promises interesting research work.

APPENDIX A : EXPLANATION OF THE SEVEN CONSTRAINTS

This section of the thesis attempts to explain meaning of the seven constraints, which were given in Section 3.1, so that deeper understanding on them can be gained. Detailed meaning of these constraints imposed on a time-frequency kernel will be given.

The general energy distribution, $P(t, \omega; \Phi)$, defined by Cohen [6, 9] is given as

$$P(t,\omega;\Phi) = \frac{1}{4\pi^2} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-j\theta t - j\tau\omega + j\theta u} \cdot \Phi(\theta,\tau) \cdot x^*\left(u - \tfrac{\tau}{2}\right) \cdot x\left(u + \tfrac{\tau}{2}\right) du\, d\tau\, d\theta \qquad \text{(A.1)}$$

where $x(t)$ is the input signal and τ is the lag parameter. For discrete input signal $x(n)$, τ is in the range of $0 - (M - 1)$ where M is the number of samples of the input signal $x(n)$.

Eq. (A.1) was derived by Cohen by using the operator characteristic method. This method uses two special operators, namely the time operator $\Im$ and the frequency operator Ω. The correspondence rules of the operators are [6, 8]

$$\begin{aligned} &\Im \rightarrow t && \Omega \rightarrow -j\frac{d}{dt} \\ &\Im \rightarrow j\frac{d}{d\omega} && \Omega \rightarrow \omega \end{aligned} \qquad \text{(A.2)}$$

The characteristic function, $M(\theta, \tau)$, is an average function of the complex signal $e^{j\theta t + j\tau\omega}$ in which θ and τ are associated with time t and frequency ω respectively. The time-frequency power spectrum can be obtained from the characteristic function using the following relation

$$P(t,\omega;\Phi) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} M(\theta,\tau)\cdot e^{-j\theta\Im - j\tau\Omega}\cdot\Phi(\theta,\tau)\ d\theta\ d\tau \tag{A.3}$$

The general formula of the characteristics function is given by

$$M(\theta,\tau) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{j\theta t + j\tau\omega}\cdot P(t,\omega)\ dt\ d\omega \tag{A.4}$$

where $P(t,\omega) = \frac{1}{\sqrt{2\pi}}\cdot \mathrm{Re}\left\{x^*(t)\cdot e^{j\omega t}\cdot \hat{X}^*(\omega)\right\}$ in which $\hat{X}^*(\omega)$ is the Fourier transform of $x(t)$ and "*" denotes complex conjugate operation. If $x(t)$ is real then $x^*(t) = x(t)$.

From Eqs. (A.3) and (A.4), the characteristic function is given as

$$M(\theta,\tau) = \left\langle e^{j\theta t + j\tau\omega}\right\rangle = \int_{-\infty}^{+\infty} x^*(t)\cdot e^{j\theta\Im + j\tau\Omega}\cdot x(t)\,dt \tag{A.5}$$

The exponential function involved in the operators can be split further into separate components using the Baker-Hausdorf theorem [6, 8, 100]

$$e^{j\theta\Im + j\tau\Omega} = e^{-j\theta\tau/2}\cdot e^{j\tau\Omega}\cdot e^{j\theta\Im} = e^{j\theta\tau/2}\cdot e^{j\theta\Im}\cdot e^{j\tau\Omega} \tag{A.6}$$

in which the change in the multiplication order of the operators results in an extra exponential term with a changing sign. The extra exponential terms are present since the operators $\Im$ and Ω do not commute and $e^{j\theta\Im + j\tau\Omega}$ is *not* equal to $e^{j\theta\Im}\cdot e^{j\tau\Omega}$ like ordinary variables [6]. In fact, the relation $\Im\Omega - \Omega\Im = j$ [6] exists. More details on this derivation may be found in [8].

In addition, by using the first translation of Eq. (A.2), we obtain

$$e^{j\tau\Omega}x(t) = e^{\tau\left(\frac{d}{dt}\right)}x(t) = x(t+\tau)\text{, where } e^{j\tau\Omega} \text{ is the translation factor [8].} \tag{A.7}$$

By combining Eqs. (A.5)–(A.7), the characteristic function can be rewritten as

$$M(\theta,\tau) = \int_{-\infty}^{+\infty} x^*(t)\cdot e^{j\theta\tau/2}\cdot e^{j\theta t}\cdot x(t+\tau)\,dt \tag{A.8}$$

Putting $u = t + \frac{\tau}{2}$ and $du = dt$ with τ as a constant parameter, Eq. (A.8) becomes

$$M(\theta,\tau) = \int x^*(u-\tfrac{\tau}{2})\cdot e^{j\theta u}\cdot x(u+\tfrac{\tau}{2})\,du \tag{A.9}$$

where the integral is taken from $-\infty$ to $+\infty$.

From Eqs. (A.3) and (A.9), the general fori · 'ı for the time-frequency power spectrum is obtained as given in Eq. (A.1).

There are a number of properties and corresponding constraints to ensure that time-frequency kernels are valid. The first property that we need to consider is that the energy distribution of a particular signal should not be complex, i.e. the energy distribution $P(t, \omega)$ should be real. Thus, the first constraint is that the kernel must be real and even, that means

$$\Phi(\theta, \tau) = \Phi^*(-\theta, -\tau) \qquad \text{(Constraint 1)}$$

The energy distribution is also required to satisfy the shift properties with respect to the time and frequency domains respectively, that means the following must be satisfied

$P(t, \omega; \Phi) = P(t - t_{01}, \omega; \Phi)$

and

$P(t, \omega; \Phi) = P(t, \omega - \omega_0; \Phi)$

where t_{01} and ω_0 are arbitrary constants in time and frequency domains respectively.

Therefore, the constraint imposed on the kernel is that the kernel must be independent of both time and frequency. That means the following constraints must be met,

$\Phi(\theta, \tau)$ does not depend on the time t, (Constraint 2)

$\Phi(\theta, \tau)$ does not depend on the frequency ω. (Constraint 3)

Since the time-frequency power spectrum is an instantaneous energy distribution, the sum of $P(t, \omega; \Phi)$ over the time and frequency domains must be the total energy distribution of the input signal $x(t)$, that means

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty} P(t,\omega;\Phi)\, d\omega = \left|x(t)\right|^2 \tag{A.10}$$

and

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty} P(t,\omega;\Phi)\, dt = \left|\hat{X}(\omega)\right|^2 \tag{A.11}$$

The total power spectrum of the input signal $x(t)$ is integrated with respect to the frequency variable ω in Eq. (A.10). To equate the left-hand side to the right-hand side of Eq. (A.10), two conditions must be satisfied: first, the kernel should not have any effect on the energy distribution since it is the power spectrum of the input signal itself and second, the "frequency variable ω" of the kernel must be eliminated. It should be noted that θ is the time variable and τ is the frequency variable used in the characteristic operator method.

To fulfill the above two requirements, we must have

$\Phi(\theta, 0) = 1$ for all θ. (Constraint 4)

Similarly, we obtain another kernel constraint based on the property given in Eq. (A.11)

$\Phi(0, \tau) = 1$ for all τ. (Constraint 5)

Next, firstly, it is required that the average frequency of the energy distribution $P(t, \omega; \Phi)$ and the instantaneous frequency of the signal should be equal at a certain time. Secondly, the group delay and the average time of the distribution $P(t, \omega; \Phi)$ should be identical in value at a certain frequency. Mathematically, the above two conditions can be written as [13, 14]

$$\frac{\int_{-\infty}^{+\infty} \omega \cdot P(t,\omega;\phi)\, d\omega}{\int_{-\infty}^{+\infty} P(t,\omega;\phi)\, d\omega} = \mathrm{Im}\left\{\frac{d}{dt}\ln x(t)\right\} \tag{A.12}$$

and

$$\frac{\int_{-\infty}^{+\infty} t \cdot P(t,\omega;\phi)\, dt}{\int_{-\infty}^{+\infty} P(t,\omega;\phi)\, dt} = -\mathrm{Im}\left\{\frac{d}{d\omega}\ln \hat{X}(\omega)\right\} \tag{A.13}$$

where $\hat{X}(\omega)$ is the 1-D Fourier transform of the input signal $x(t)$ and $\ln(\cdot)$ denotes the natural logarithmic of the function $(\cdot)$ to the base e.

To meet the requirement given in Eq. (A.12), constraint 4 is required so that the time-frequency power spectrum $P(t, \omega; \Phi)$ is independent of the "frequency variable" of the kernel which is τ. Recall that τ and θ are the frequency and time variables respectively of the kernel functions $\Phi(\theta, \tau)$ as have been used in the characteristic operator method. Thus, this property covers the property previously given in Eq. (A.10).

The right-hand side of Eq. (A.12) can be rewritten as

$$\mathrm{Im}\left\{\frac{d}{dt}\ln x(t)\right\} = \mathrm{Im}\left\{\ln \frac{d}{dt} x(t)\right\} \tag{A.14}$$

The left-hand sides of Eqs. (A.12) and (A.13), which consist of the time-frequency power spectrum and the kernel function, are equated with the rate of change of the input signal in time and frequency domains respectively. To obtain the next two constraints, similar logical techniques to those used to obtain constraints 4 and 5 are employed. In this case, the rate of change of the kernel in the time and frequency domains (θ and τ respectively) must be zero so that the rate of change of the time-frequency power spectrum is equal to that of the input signal without the kernel interference. Mathematically, the required constraints to fulfill Eqs. (A.12) and (A.13) can be stated as

$$\frac{d}{d\theta}\{\Phi(\theta,\tau)\}\bigg|_{\theta=0} = 0, \forall\tau \text{ and } \Phi(0, \tau) = 1 \text{ for all } \tau. \qquad \text{(Constraint 6)}$$

Similarly, we obtain

$$\frac{d}{d\tau}\{\Phi(\theta,\tau)\}\bigg|_{\tau=0} = 0, \forall\theta \text{ and } \Phi(\theta, 0) = 1 \text{ for all } \theta. \qquad \text{(Constraint 7)}$$

The final property is the finite support property which states that the energy distribution $P(t, \omega ; \Phi)$ should be zero over the zero region(s) of the input signals in time and frequency domains. That means the followings can be stated

In the time domain, if $f(t) = 0$ for $|t| \le T_0$ then $P(t, \omega ; \Phi) = 0$ for $|t| \le T$, and similarly in the frequency domain, if $\hat{F}(\omega) = 0$ for $|\omega| \le \omega_0$ then $P(t, \omega ; \Phi) = 0$ for $|\omega| \le \omega_0$.

The finite support properties in time and frequency domains ensure that the energy distribution or the time-frequency power spectrum is finite over the defined ranges of the time variable t and the frequency variable ω. To meet the finite support properties, the kernel weighting function must be finite over the defined ranges of time or frequency. In other words, the kernel weighting functions must decay to zero outside the defined ranges of the time and frequency variables. The following constraints are therefore obtained

$$\int e^{-j\theta t} \cdot \Phi(\theta,\tau)d\theta = 0, \text{ for } |\tau| < 2|t| \qquad \text{(Constraint 8)}$$

and similarly

$$\int e^{-j\tau\omega} \cdot \Phi(\theta,\tau)\, d\tau = 0, \text{ for} |\theta| < 2|\omega| \qquad \text{(Constraint 9)}$$

where all integrals are from $-\infty$ to $+\infty$.

APPENDIX B : SIGNAL-TO-NOISE RATIO DERIVATION OF THE *GNKD*

This section presents detailed calculations of the *SNR* of the general non-unity kernel signal detector (*GNKD*). The statistics η is calculated by substituting the appropriate signals into Eq. (5.19) in which the signals $g(\cdot)$, $h(\cdot)$, $s(\cdot)$ and $f(\cdot)$ are different. To calculate the statistics η, we put $g(\cdot) = h(\cdot) = s(\cdot)$ which is the input reference signal. For the case of H_0, we have $f(\cdot) = w(n)$. For the case of H_1, $f(\cdot) = w(n) + s(\cdot)$.

The statistics of the hypothesis H_0 is given by

$$\eta_{TFR}\big|_{H_0} = M \sum_{n=-M/2}^{M/2} [A_1 \cdot B_1]$$

$$= M \sum_{n=-M/2}^{M/2} \left[\sum_{q_k=-|n|}^{|n|} s^*(q_k) \sum_{p_k} F_k \cdot w(p_k) + \sum_{q_k=-|n|}^{|n|} s^*(q_k) \cdot (-1)^{q_k} \sum_{p_k} F_k \cdot w(p_k) \cdot (-1)^{p_k} \right] \cdot$$

$$\left[\sum_{q_l=-|n|}^{|n|} s^*(q_l) \sum_{p_l} F_l \cdot s(p_l) + \sum_{q_l=-|n|}^{|n|} s^*(q_l) \cdot (-1)^{q_l} \sum_{p_l} F_l \cdot s(p_l) \cdot (-1)^{p_l} \right]$$

(B.1)

After multiplying out all the summations and products we obtain

$$\eta_{TFR}\big|_{H_0} = M \cdot \sum_{n=-M/2}^{M/2} \left(D_1 + D_2 + D_3 + D_4\right)$$

$$= M \cdot \sum_{n=-M/2}^{M/2} \left(\sum_{q_k=-|n|}^{|n|} s^*(q_k) \sum_{p_k} F_k \cdot w(p_k) \right) \cdot \left(\sum_{q_l=-|n|}^{|n|} s^*(q_l) \sum_{p_l} F_l \cdot s(p_l) \right)$$

$$+ \left(\sum_{q_k=-|n|}^{|n|} s^*(q_k) \sum_{p_k} F_k \cdot w(p_k) \right) \cdot \left(\sum_{q_l=-|n|}^{|n|} s^*(q_l) \cdot (-1)^{q_l} \sum_{p_l} F_l \cdot s(p_l) \cdot (-1)^{p_l} \right)$$

$$+ \left(\sum_{q_k=-|n|}^{|n|} s^*(q_k) \cdot (-1)^{q_k} \sum_{p_k} F_k \cdot w(p_k) \cdot (-1)^{p_k} \right) \cdot \left(\sum_{q_l=-|n|}^{|n|} s^*(q_l) \sum_{p_l} F_l \cdot s(p_l) \right)$$

$$+ \left(\sum_{q_k=-|n|}^{|n|} s^*(q_k) \cdot (-1)^{q_k} \sum_{p_k} F_k \cdot w(p_k) \cdot (-1)^{p_k} \right) \cdot \left(\sum_{q_l=-|n|}^{|n|} s^*(q_l) \cdot (-1)^{q_l} \sum_{p_l} F_l \cdot s(p_l) \cdot (-1)^{p_l} \right)$$

(B.2)

The statistics for the hypothesis H_1 is given by

$$\eta_{TFR}\big|_{H_1} = M \sum_{n=-M/2}^{M/2}$$

$$\left[\sum_{q_k=-|n|}^{|n|} s^*(q_k) \sum_{p_k} F_k \cdot \left(w(p_k) + s(p_k)\right) + \sum_{q_k} s^*(q_k) \cdot (-1)^{q_k} \sum_{p_k} F_k \cdot \left(w(p_k) + s(p_k)\right) \cdot (-1)^{p_k} \right] \cdot$$

$$\left[\sum_{q_l} s^*(q_l) \sum_{p_l} F_l \cdot s(p_l) + \sum_{q_l} s^*(q_l) \cdot (-1)^{q_l} \sum_{p_l} F_l \cdot s(p_l) \cdot (-1)^{p_l} \right]$$

$$= M \sum_{n=-M/2}^{M/2} \left(\sum_{q_k} s^*(q_k) \sum_{p_k} F_k \cdot s(p_k) + \sum_{q_k} s^*(q_k) \cdot (-1)^{q_k} \sum_{p_k} F_k \cdot s(p_k) \cdot (-1)^{p_k} + A_1 \right) \cdot B_1$$

(B.3)

where A_1 and B_1 were defined by Eq. (B.1).

The general expression of $\eta_{TFR}|_{H_1}$ is given by

$$\eta_{TFR}|_{H_1} = M \sum_{n=-M/2}^{M/2}$$

$$\left\{\left(\sum_{q_k} s^*(q_k) \sum_{p_k} F_k \cdot s(p_k)\right) \cdot B_1 + \left(\sum_{q_k} s^*(q_k) \cdot (-1)^{q_k} \sum_{p_k} F_k \cdot s(p_k) \cdot (-1)^{p_k}\right) \cdot B_1 + A_1 B_1\right\}$$

$$= M \sum_{n=-M/2}^{M/2}$$

$$\left\{\left(\sum_{q_k} s^*(q_k) \sum_{p_k} F_k \cdot s(p_k)\right) \cdot B_1 + \left(\sum_{q_k} s^*(q_k) \cdot (-1)^{q_k} \sum_{p_k} F_k \cdot s(p_k) \cdot (-1)^{p_k}\right) \cdot B_1 + \eta_{TFR}|_{H_0}\right\}$$

$$= M \sum_{n=-M/2}^{M/2} \left\{ 2C + \eta_{TFR}|_{H_0} \right\}$$

(B.4)

where

$$C = \left(\sum_{q_k} s^*(q_k) \sum_{p_k} F_k \cdot s(p_k)\right) \cdot \left(\sum_{q_l} s(q_l) \sum_{p_l} F_l \cdot s^*(p_l)\right) + \tag{B.5}$$

$$\left(\sum_{q_k} s^*(q_k) \sum_{p_k} F_k \cdot s(p_k)\right) \cdot \left(\sum_{q_l} s^*(q_l) \cdot (-1)^{q_l} \sum_{p_l} F_l \cdot s(p_l) \cdot (-1)^{p_l}\right)$$

For the special case of $p_k = p_l$, $q_l = q_{ll}$ and $q_k = q_l$, Eq. (B.5) becomes

$$C = \left(\sum_{q_k} s^*(q_k) \sum_{p_k} F_k \cdot s(p_k)\right) \cdot \left(\sum_{q_l} s(q_l) \sum_{p_l} F_l \cdot s^*(p_l)\right) +$$

$$\left(\sum_{q_k} s^*(q_k) \sum_{p_k} F_k \cdot s(p_k)\right) \cdot \left(\sum_{q_l} s^*(q_l) \cdot (-1)^{q_l} \sum_{p_l} F_l \cdot s(p_l) \cdot (-1)^{p_l}\right) \qquad \text{(B.6)}$$

$$= A_{0F} \cdot A_{0F} + B_{0F} \cdot B_{0F}$$

$$\textit{where } A_{0F} = \sum_{q_k} s^*(q_k) \cdot F_k \cdot s(q_k) \quad \textit{and} \quad B_{0F} = \sum_{q_l} s^*(q_l) \cdot F_l \cdot s(q_l) \cdot (-1)^{q_l}.$$

The constant C (given by Eq. (B.5)) is a modified version of the signal energy in which the constant product $F_k \cdot F_l$ can be considered as the square of the weighting function of the kernel function $\Phi(\theta, \tau)$ when $k \equiv l$. As will be seen later, under the special condition of $k \equiv l$, some useful manipulations can be carried out.

In the case of the Wigner-Ville kernel in which the involvement of the kernel is included in the index of the local auto-correlation function, the reference-signal energy is not altered. For the case of a non-unity kernel, the Fourier transform of the kernel is not an impulse function which makes the calculation of the time-frequency power spectrum more difficult. As a result, the signal energy will be modified and its familiar form is not obtained. Instead, the signal energy will be reduced by multiplying with the weighting function F_k. The rest of the appendix is going to calculate the *SNR* of the statistics η by evaluating all ten terms of the summation square. The effects of the weighting function will also be demonstrated.

The notations used in this appendix are as follows. The subscript "0" indicates energy terms with only one summation. These terms are usually the original signal energy or noise energy, e.g. A_0, N_0. The subscript "1" indicates there are two summations in the energy due to different k and l. If k is similar to l, then the subscript can be safely ignored. The subscript "F" indicates that the energy term is multiplied by the kernel weighting function F and "FF" means that the term is multiplied by a square of the weighting function.

The performance of a detector is judged by the value of its *SNR*. Firstly, we calculate the mean of the statistics of the two hypotheses

$$E\{\eta_{TFR}|_{H_0}\} = 0 \text{ since } E\{w(t)\} = 0;$$
$$E\{\eta_{TFR}|_{H_1}\} = 2MC \tag{B.7}$$

Since $\eta_{TFR}|_{H_1} = C + \eta_{TFR}|_{H_0}$, the variance of $\eta_{TFR}|_{H_1}$ is equal to that of $\eta_{TFR}|_{H_0}$. The variance of $\eta_{TFR}|_{H_0}$ is given by

$$Var\{\eta_{TFR}|_{H_1}\} = Var\{\eta_{TFR}|_{H_0}\} = E\{[\eta_{TFR}|_{H_0}]^2\} \tag{B.8}$$

To evaluate Eq. (B.8), we have to square $\eta_{TFR}|_{H_0}$ given in Eq. (B.2) to obtain a square summation. From Eq. (B.2) and after the squaring process, we obtain the first term of the square summation

$$M^2 \cdot \sum_{n=-M/2}^{M/2} \sum_{j=-M/2}^{M/2} D_1^2 =$$
$$M^2 \cdot \sum_{n=-M/2}^{M/2} \sum_{j=-M/2}^{M/2} \left(\sum_{q_k} s^*(q_k) \cdot \sum_{q_{k1}} s(q_{k1}) \right) \cdot \left(\sum_{p_l} F_l \cdot s(p_l) \cdot \sum_{p_{l1}} F_{jl1} \cdot s^*(p_{l1}) \right)$$
$$\cdot \left(\sum_{p_k} F_k \cdot w(p_k) \cdot \sum_{p_{k1}} F_{jk1} \cdot w(p_{k1}) \right) \cdot \left(\sum_{q_l} s^*(q_l) \cdot \sum_{q_{l1}} s(q_{l1}) \right) \tag{B.9}$$

For a complex white noise process of variance N_0, we obtain

$$E\{w(p_k) \cdot w(p_{k1})\} = N_0 \delta(p_k - p_{k1}), \text{ thus, } p_k = p_{k1} \tag{B.10}$$

The noise process in this case is modified by the energy of the weighting function F_k which results in a new noise process of variance N_{01}. Using Eq. (B.10), Eq. (B.9) can be rewritten as

The performance of a detector is judged by the value of its *SNR*. Firstly, we calculate the mean of the statistics of the two hypotheses

$$E\{\eta_{TFR}|_{H_0}\} = 0 \text{ since } E\{w(t)\} = 0;$$
$$E\{\eta_{TFR}|_{H_1}\} = 2MC \tag{B.7}$$

Since $\eta_{TFR}|_{H_1} = C + \eta_{TFR}|_{H_0}$, the variance of $\eta_{TFR}|_{H_1}$ is equal to that of $\eta_{TFR}|_{H_0}$. The variance of $\eta_{TFR}|_{H_0}$ is given by

$$Var\{\eta_{TFR}|_{H_1}\} = Var\{\eta_{TFR}|_{H_0}\} = E\{[\eta_{TFR}|_{H_0}]^2\} \tag{B.8}$$

To evaluate Eq. (B.8), we have to square $\eta_{TFR}|_{H_0}$ given in Eq. (B.2) to obtain a square summation. From Eq. (B.2) and after the squaring process, we obtain the first term of the square summation

$$M^2 \cdot \sum_{n=-M/2}^{M/2} \sum_{j=-M/2}^{M/2} D_1^2 =$$
$$M^2 \cdot \sum_{n=-M/2}^{M/2} \sum_{j=-M/2}^{M/2} \left(\sum_{q_k} s^*(q_k) \cdot \sum_{q_{k1}} s(q_{k1}) \right) \cdot \left(\sum_{p_l} F_l \cdot s(p_l) \cdot \sum_{p_{l1}} F_{jl1} \cdot s^*(p_{l1}) \right)$$
$$\cdot \left(\sum_{p_k} F_k \cdot w(p_k) \cdot \sum_{p_{k1}} F_{jk1} \cdot w(p_{k1}) \right) \cdot \left(\sum_{q_l} s^*(q_l) \cdot \sum_{q_{l1}} s(q_{l1}) \right) \tag{B.9}$$

For a complex white noise process of variance N_0, we obtain

$$E\{w(p_k) \cdot w(p_{k1})\} = N_0 \delta(p_k - p_{k1}), \text{ thus, } p_k = p_{k1} \tag{B.10}$$

The noise process in this case is modified by the energy of the weighting function F_k which results in a new noise process of variance N_{01}. Using Eq. (B.10), Eq. (B.9) can be rewritten as

$$M^2 \cdot \sum_{n=-M/2}^{M/2} \sum_{j=-M/2}^{M/2} D_1^2 = M^2 \cdot \sum_{n=-M/2}^{M/2} \sum_{j=-M/2}^{M/2} A_0 \cdot A_{01FF} \cdot N_{0FF} \cdot A_{01} \tag{B.11}$$

where

$$N_{0FF} = \sum_{p_k} F_k \cdot w(p_k) \cdot \sum_{p_{k1}} F_{jk1} \cdot w(p_{k1}),\ A_{01} = \sum_{q_l} s^*(q_l) \cdot \sum_{q_{l1}} s(q_{l1}),$$

$$A_{01FF} = \sum_{p_l} F_l \cdot s(p_l) \cdot \sum_{p_{l1}} F_{jl1} \cdot s^*(p_{l1}) \tag{B.12}$$

and A_0 was defined in Eq. (5.13).

The second term is given by

$$M^2 \cdot \sum_{n=-M/2}^{M/2} \sum_{j=-M/2}^{M/2} D_2^2$$

$$= M^2 \cdot \sum_{n=-M/2}^{M/2} \sum_{j=-M/2}^{M/2} \left(\sum_{q_k} s^*(q_k) \cdot \sum_{q_{k1}} s(q_{k1}) \right) \cdot \left(\sum_{q_l} s^*(q_l) \cdot (-1)^{q_l} \cdot \sum_{q_{l1}} s(q_{l1}) \cdot (-1)^{q_{l1}} \right)$$

$$\cdot \left(\sum_{p_k} F_k \cdot w(p_k) \cdot \sum_{p_{k1}} F_{jk1} \cdot w(p_{k1}) \right) \cdot \left(\sum_{p_l} F_l \cdot s(p_l) \cdot (-1)^{p_l} \cdot \sum_{p_{l1}} F_{l1} \cdot s(p_{l1}) \cdot (-1)^{p_{l1}} \right)$$

$$= M^2 \cdot \sum_{n=-M/2}^{M/2} \sum_{j=-M/2}^{M/2} A_0 \cdot B_{01} \cdot N_{0FF} \cdot B_{01FF}$$

$$\text{where } B_{01} = \sum_{q_l} s^*(q_l) \cdot (-1)^{q_l} \cdot \sum_{q_{l1}} s(q_{l1}) \cdot (-1)^{q_{l1}}$$

$$B_{01FF} = \sum_{p_l} F_l \cdot s(p_l) \cdot (-1)^{p_l} \cdot \sum_{p_{l1}} F_{l1} \cdot s(p_{l1}) \cdot (-1)^{p_{l1}} \tag{B.13}$$

The third term is given by

$$M^2 \cdot \sum_{n=-M/2}^{M/2} \sum_{j=-M/2}^{M/2} D_3^2$$

$$= M^2 \cdot \sum_{n=-M/2}^{M/2} \sum_{j=-M/2}^{M/2} \left(\sum_{q_k} s^*(q_k) \cdot (-1)^{q_k} \cdot \sum_{q_{k1}} s(q_{k1}) \cdot (-1)^{q_{k1}} \right) \cdot \left(\sum_{q_l} s^*(q_l) \cdot \sum_{q_{l1}} s(q_{l1}) \right)$$

$$\left(\sum_{p_k} F_k \cdot w(p_k) \cdot (-1)^{p_k} \cdot \sum_{p_{k1}} F_{jk1} \cdot w(p_{k1}) \cdot (-1)^{p_{k1}} \right) \cdot \left(\sum_{p_l} F_l \cdot s(p_l) \cdot \sum_{p_{l1}} F_{jl1} \cdot s^*(p_{l1}) \right)$$

$$= M^2 \cdot \sum_{n=-M/2}^{M/2} \sum_{j=-M/2}^{M/2} A_0 \cdot A_{01} \cdot N_{0FF} \cdot A_{01FF}$$

(B.14)

The fourth term is given by

$$M^2 \cdot \sum_{n=-M/2}^{M/2} \sum_{j=-M/2}^{M/2} D_4^2 = M^2 \cdot \sum_{n=-M/2}^{M/2} \sum_{j=-M/2}^{M/2} \left(\sum_{q_k} s^*(q_k) \cdot (-1)^{q_k} \cdot \sum_{q_{k1}} s(q_{k1}) \cdot (-1)^{q_{k1}} \right)$$

$$\cdot \left(\sum_{p_k} F_k \cdot w(p_k) \cdot (-1)^{p_k} \cdot \sum_{p_{k1}} F_{jk1} \cdot w(p_{k1}) \cdot (-1)^{p_{k1}} \right)$$

$$\cdot \left(\sum_{q_l} s^*(q_l) \cdot (-1)^{q_l} \cdot \sum_{q_{l1}} s(q_{l1}) \cdot (-1)^{q_{l1}} \right) \cdot \left(\sum_{p_l} F_l \cdot s(p_l) \cdot (-1)^{p_l} \cdot \sum_{p_{l1}} F_{jl1} \cdot s(p_{l1}) \cdot (-1)^{p_{l1}} \right)$$

$$= M^2 \cdot \sum_{n=-M/2}^{M/2} \sum_{j=-M/2}^{M/2} A_0 \cdot N_{0FF} \cdot B_{01} \cdot B_{01FF}$$

(B.15)

The fifth term is given by

$$M^2 \cdot \sum_{n=-M/2}^{M/2} \sum_{j=-M/2}^{M/2} 2D_1 D_2$$

$$= 2M^2 \cdot \sum_{n=-M/2}^{M/2} \sum_{j=-M/2}^{M/2} \left(\sum_{q_k} s^*(q_k) \sum_{p_k} F_k \cdot w(p_k) \right) \cdot \left(\sum_{q_l} s^*(q_l) \sum_{p_l} F_l \cdot s(p_l) \right)$$

$$\cdot \left(\sum_{q_k} s^*(q_k) \sum_{p_k} F_k \cdot w(p_k) \right) \cdot \left(\sum_{q_l} s^*(q_l) \cdot (-1)^{q_l} \sum_{p_l} F_l \cdot s(p_l) \cdot (-1)^{p_l} \right) \tag{B.16}$$

$$= M^2 \cdot \sum_{n=-M/2}^{M/2} \sum_{j=-M/2}^{M/2} 2A_0 \cdot N_{0FF} \cdot B_0 \cdot B_{0FF}$$

$$where \ B_{0FF} = \sum_{p_l} F_l \cdot s(p_l) F_l \cdot s(p_l) \cdot (-1)^{p_l}$$

The sixth term is given by

$$M^2 \cdot \sum_{n=-M/2}^{M/2} \sum_{j=-M/2}^{M/2} 2D_3 D_4$$

$$= 2M^2 \cdot \sum_{n=-M/2}^{M/2} \sum_{j=-M/2}^{M/2} \left(\sum_{q_k} s^*(q_k) \cdot (-1)^{q_k} \sum_{p_k} F_k \cdot w(p_k) \cdot (-1)^{p_k} \right) \cdot \left(\sum_{q_l} s^*(q_l) \sum_{p_l} F_l \cdot s(p_l) \right)$$

$$\cdot \left(\sum_{q_k} s^*(q_k) \cdot (-1)^{q_k} \sum_{p_k} F_k \cdot w(p_k) \cdot (-1)^{p_k} \right) \cdot \left(\sum_{q_l} s^*(q_l) \cdot (-1)^{q_l} \sum_{p_l} F_l \cdot s(p_l) \cdot (-1)^{p_l} \right)$$

$$= M^2 \cdot \sum_{n=-M/2}^{M/2} \sum_{j=-M/2}^{M/2} 2A_0 \cdot B_0 \cdot N_{0FF} \cdot B_{0FF}$$

(B.17)

The seventh term is given by

$$M^2 \cdot \sum_{n=-M/2}^{M/2} \sum_{j=-M/2}^{M/2} 2D_1 D_3$$

$$= 2M^2 \cdot \sum_{n=-M/2}^{M/2} \sum_{j=-M/2}^{M/2} \left(\sum_{q_k} s^*(q_k) \sum_{p_k} F_k \cdot w(p_k) \right) \cdot \left(\sum_{q_l} s^*(q_l) \sum_{p_l} F_l \cdot s(p_l) \right) \tag{B.18}$$

$$\cdot \left(\sum_{q_k} s^*(q_k) \cdot (-1)^{q_k} \sum_{p_k} F_k \cdot w(p_k) \cdot (-1)^{p_k} \right) \cdot \left(\sum_{q_l} s^*(q_l) \sum_{p_l} F_l \cdot s(p_l) \right)$$

$$= M^2 \cdot \sum_{n=-M/2}^{M/2} \sum_{j=-M/2}^{M/2} 2B_0 \cdot M_{0FF} \cdot A_0 \cdot A_{0FF}$$

The eighth term is given by

$$M^2 \cdot \sum_{n=-M/2}^{M/2} \sum_{j=-M/2}^{M/2} 2D_1 D_4$$

$$= 2M^2 \cdot \sum_{n=-M/2}^{M/2} \sum_{j=-M/2}^{M/2} \left(\sum_{q_k} s^*(q_k) \sum_{p_k} F_k \cdot w(p_k) \right) \cdot \left(\sum_{q_l} s^*(q_l) \sum_{p_l} F_l \cdot s(p_l) \right)$$

$$\left(\sum_{q_k} s^*(q_k) \cdot (-1)^{q_k} \sum_{p_k} F_k \cdot w(p_k) \cdot (-1)^{p_k} \right) \cdot \left(\sum_{q_l} s^*(q_l) \cdot (-1)^{q_l} \sum_{p_l} F_l \cdot s(p_l) \cdot (-1)^{p_l} \right)$$

$$= M^2 \cdot \sum_{n=-M/2}^{M/2} \sum_{j=-M/2}^{M/2} 2B_0 \cdot M_{0FF} \cdot B_0 \cdot B_{0FF}$$

(B.19)

The ninth term is given by

$$
\begin{aligned}
&M^2 \cdot \sum_{n=-M/2}^{M/2} \sum_{j=-M/2}^{M/2} 2D_2D_3 \\
&= 2M^2 \cdot \sum_{n=-M/2}^{M/2} \sum_{j=-M/2}^{M/2} \left(\sum_{q_k} s^*(q_k) \sum_{p_k} F_k \cdot w(p_k) \right) \cdot \left(\sum_{q_l} s^*(q_l) \cdot (-1)^{q_l} \sum_{p_l} F_l \cdot s(p_l) \cdot (-1)^{p_l} \right. \\
&\cdot \left(\sum_{q_k} s^*(q_k) \cdot (-1)^{q_k} \sum_{p_k} F_k \cdot w(p_k) \cdot (-1)^{p_k} \right) \cdot \left(\sum_{q_l} s^*(q_l) \sum_{p_l} F_l \cdot s(p_l) \right) \\
&= M^2 \cdot \sum_{n=-M/2}^{M/2} \sum_{j=-M/2}^{M/2} 2B_0 \cdot M_{0FF} \cdot B_0 \cdot B_{0FF}
\end{aligned}
$$

(B.20)

The tenth term is given by

$$
\begin{aligned}
&M^2 \cdot \sum_{n=-M/2}^{M/2} \sum_{j=-M/2}^{M/2} 2D_2D_4 \\
&= 2M^2 \cdot \sum_{n=-M/2}^{M/2} \sum_{j=-M/2}^{M/2} \left(\sum_{q_k} s^*(q_k) \sum_{p_k} F_k \cdot w(p_k) \right) \cdot \left(\sum_{q_l} s^*(q_l) \cdot (-1)^{q_l} \sum_{p_l} F_l \cdot s(p_l) \cdot (-1)^{p_l} \right) \\
&\cdot \left(\sum_{q_k} s^*(q_k) \cdot (-1)^{q_k} \sum_{p_k} F_k \cdot w(p_k) \cdot (-1)^{p_k} \right) \cdot \left(\sum_{q_l} s^*(q_l) \cdot (-1)^{q_l} \sum_{p_l} F_l \cdot s(p_l) \cdot (-1)^{p_l} \right) \\
&= M^2 \cdot \sum_{n=-M/2}^{M/2} \sum_{j=-M/2}^{M/2} 2B_0 \cdot M_{0FF} \cdot A_0 \cdot A_{0FF}
\end{aligned}
$$

(B.21)

By summing the ten terms given by Eqs. (B.9)–(B.20), the variance of $\eta_{TFR}|_{H_1}$ is given by Eq. (B.22)

$$
\begin{aligned}
Var\{\eta_{TFR}|_{H_0}\} = M^2 \cdot \sum_{n=-M/2}^{M/2} \sum_{j=-M/2}^{M/2} & (A_0 \cdot A_{01FF} \cdot N_{0FF} \cdot A_{01}) + (A_0 \cdot B_{01} \cdot N_{0FF} \cdot B_{01FF}) \\
& + (A_0 \cdot A_{01} \cdot N_{0FF} \cdot A_{01FF}) + (A_0 \cdot N_{0FF} \cdot B_{01} \cdot B_{01FF}) + (2A_0 \cdot N_{0FF} \cdot B_0 \cdot B_{0FF}) \\
& + (2A_0 \cdot B_0 \cdot N_{0FF} \cdot B_{0FF}) + (2B_0 \cdot M_{0FF} \cdot A_0 \cdot A_{0FF}) + (2B_0 \cdot M_{0FF} \cdot B_0 \cdot B_{0FF}) \\
& + (2B_0 \cdot M_{0FF} \cdot B_0 \cdot B_{0FF}) + (2B_0 \cdot M_{0FF} \cdot A_0 \cdot A_{0FF}) \\
= M^2 \cdot \sum_{n=-M/2}^{M/2} \sum_{j=-M/2}^{M/2} & 2[(A_0 \cdot A_{01FF} \cdot N_{0FF} \cdot A_{01}) + (A_0 \cdot B_{01} \cdot N_{0FF} \cdot B_{01FF})] \\
& + 2[2(A_0 \cdot N_{0FF} \cdot B_0 \cdot B_{0FF}) + 2(B_0 \cdot M_{0FF} \cdot A_0 \cdot A_{0FF}) + 2(B_0 \cdot M_{0FF} \cdot B_0 \cdot B_{0FF})]
\end{aligned}
\tag{B.22}
$$

where

$$
\begin{aligned}
& A_{01FF} = \sum_{p_l} F_l \cdot s(p_l) \cdot \sum_{p_{l1}} F_{jl1} \cdot s^*(p_{l1}); \; A_{0FF} = \sum_{p_l} F_l \cdot s(p_l) \cdot F_{jl} \cdot s^*(p_l), \\
& B_{0FF} = \sum_{p_l} F_l \cdot s(p_l) F_l \cdot s(p_l) \cdot (-1)^{p_l}; \; B_{01} = \sum_{q_l} s^*(q_l) \cdot (-1)^{q_l} \cdot \sum_{q_{l1}} s(q_{l1}) \cdot (-1)^{q_{l1}}, \\
& B_{01FF} = \sum_{p_l} F_l \cdot s(p_l) \cdot (-1)^{p_l} \cdot \sum_{p_{l1}} F_{l1} \cdot s(p_{l1}) \cdot (-1)^{p_{l1}}; \; A_{01} = \sum_{q_l} s^*(q_l) \cdot \sum_{q_{l1}} s(q_{l1}), \\
& M_{0FF} = \sum_{p_k} F_k \cdot w(p_k) \cdot F_k \cdot w(p_k) \cdot (-1)^{p_k}; \; N_{0FF} = \sum_{p_k} F_k \cdot w(p_k) \cdot F_k \cdot w(p_k)
\end{aligned}
\tag{B.23}
$$

For the special case of $p_k = p_l$ and $q_k = q_l$, $q_l = q_{l1}$, Eq. (B.23) becomes

$$A_{01FF} = A_{0FF};\ A_{0FF} = \sum_{p_l} F_l \cdot s(p_l) \cdot F_{jl} \cdot s^*(p_l),$$

$$B_{0FF} = \sum_{p_l} F_l \cdot s(p_l) F_l \cdot s(p_l) \cdot (-1)^{p_l};\ B_{01} = A_{01} = A_0;\ B_{01FF} = B_{0FF}, \tag{B.24}$$

$$M_{0FF} = \sum_{p_k} F_k \cdot w(p_k) \cdot F_k \cdot w(p_k) \cdot (-1)^{p_k};\ N_{0FF} = \sum_{p_k} F_k \cdot w(p_k) \cdot F_k \cdot w(p_k)$$

The *SNR* of the *GNKD* for the general case is then given by

$$SNR_{GNKD} = \frac{2 \cdot M \cdot C}{\sqrt{Var\{\eta_{TFR}|_{H_0}\}}}, \text{ where } Var(\eta_{TFR}|_{H_1}) \text{ is given by Eq. (B.22).} \tag{B.25}$$

Eq. (B.22) consists of ten terms with the involvement of the fourth power of the weighting functions F_k which reduces the values of the terms significantly. The nominator of Eq. (B.25) is a function of the second power of the weighting function which is much larger than its fourth power counterpart in the denominator. The denominator of Eq. (B.25) needs to be as small as possible to maximize the *SNR*.

For the special case of $p_k = p_l$ and $q_k = q_l$ given by Eq. (B.24), Eq. (B.22) becomes,

$$Var\{\eta_{TFR}|_{H_0}\} = M^2 \cdot \sum_{n=-M/2}^{M/2} \sum_{j=-M/2}^{M/2} 2[(A_0 \cdot A_{01FF} \cdot N_{0FF} \cdot A_{01}) + (A_0 \cdot B_{01} \cdot N_{0FF} \cdot B_{01FF})]$$

$$+ 2[2(A_0 \cdot N_{0FF} \cdot B_0 \cdot B_{0FF}) + 2(B_0 \cdot M_{0FF} \cdot A_0 \cdot A_{0FF}) + 2(B_0 \cdot M_{0FF} \cdot B_0 \cdot B_{0FF})]$$

$$= M^2 \cdot \sum_{n=-M/2}^{M/2} \sum_{j=-M/2}^{M/2} 2[(A_0 \cdot A_{0FF} \cdot N_{0FF} \cdot A_0) + (A_0 \cdot A_0 \cdot N_{0FF} \cdot B_{0FF})]$$

$$+ 2[2(A_0 \cdot N_{0FF} \cdot B_0 \cdot B_{0FF}) + 2(B_0 \cdot M_{0FF} \cdot A_0 \cdot A_{0FF}) + 2(B_0 \cdot M_{0FF} \cdot B_0 \cdot B_{0FF})]$$

$$= 2M^2 \cdot A_0^2 \cdot N_{0FF} \cdot A_{0FF}$$

$$\sum_{n=-M/2}^{M/2} \sum_{j=-M/2}^{M/2} \left\{1 + \frac{B_{0FF}}{A_{0FF}} + 2\left[\frac{B_0}{A_0} \cdot \frac{B_{0FF}}{A_{0FF}} + \frac{B_0}{A_0} \cdot \frac{M_{0FF}}{N_{0FF}} + \left(\frac{B_0}{A_0}\right)^2 \cdot \frac{M_{0FF}}{N_{0FF}} \cdot \frac{B_{0FF}}{A_{0FF}}\right]\right\}$$

$$= 2M^2 \cdot A_0^2 \cdot N_{0FF} \cdot A_{0FF} \sum_{n=-M/2}^{M/2} \sum_{j=-M/2}^{M/2} \left\{1 + \frac{B_0}{A_0} + 2\left[\left(\frac{B_0}{A_0}\right)^2 + \frac{B_0}{A_0} \cdot \frac{M_0}{N_0} + \left(\frac{B_0}{A_0}\right)^3 \cdot \frac{M_0}{N_0}\right]\right\}$$

(B.26)

where $X_1 = \frac{B_0}{A_0}$ is the ratio between the energy difference of the even and odd parts of the reference signal $s(\cdot)$ to its energy A_0, $X_2 = \frac{M_0}{N_0}$ is the ratio of the energy difference between the even and odd parts of the noise process to its energy N_0, N_{0FF} and A_{0FF} are the modified noise energy and signal energy respectively scaled down by the square of the weighting function F_k.

The SNR_{GNKD} of the non-unity kernel signal detector in the special case is thus given by

$$SNR_{GNKD} = \frac{\sqrt{2} \cdot (A_{0F}^2 + B_{0F}^2)}{A_0 \sqrt{N_{0FF} \cdot A_{0FF} \cdot \left\{1 + \frac{B_0}{A_0} + 2\left[\left(\frac{B_0}{A_0}\right)^2 + \frac{B_0}{A_0} \cdot \frac{M_0}{N_0} + \frac{M_0}{N_0} \cdot \left(\frac{B_0}{A_0}\right)^3\right]\right\}}} \tag{B.27}$$

REFERENCES

[1] Vinod Chandran and Charles Pezeshki, "Bispectral and Trispectral Characterization of Transition to Chaos in the Duffing Oscillator," *International Journal of Bifurcation and Chaos*, 1993, **3**, No. 3, pp. 551-557.

[2] J. M. Lipton and K. P. Dabke, "Bispectra of non-linear systems," *Report from the Department of Electrical and Computer Systems Engineering, Faculty of Engineering, Monash University, Clayton VIC. 3168*, 1994

[3] Personal communication with Dr. Justin M. Lipton in 1996.

[4] J. M. Lipton, *et al.*, "Use of the Bispectrum to Analyse Properties of the Human Electrocardiograph," *Australasian Physical & Engineering Sciences in Medicine*, **21**, March 1998, pp. 1-11.

[5] Leon Cohen, "Instantaneous "Anything"," *ICASSP*, 1993, **4**, pp. 105-108.

[6] Leon Cohen, *Time-Frequency Analysis*, pp. 136-289, Prentice Hall, Englewood Cliffs, 1995.

[7] Leon Cohen, "On a Fundamental Property of the Wigner Distribution," *IEEE Transactions on Acoustics, Speech and Signal Processing*, April 1987, **ASSP-35**, No. 4, pp. 559-561.

[8] Leon Cohen, "Time-Frequency Distributions—A Review," *Proceedings of the IEEE*, July 1989, **77**, No. 7, pp. 941-981.

[9] Leon Cohen, "Introduction: A Primer on Time-Frequency Analysis," in *Time-Frequency Signal Analysis: Methods and Applications*, B. Boashash Eds., John Wiley, pp. 3-43, 1992.

[10] Chester H. Page, "Instantaneous Power Spectra," *Journal of Applied Physics*, January 1952, **23**, No. 1, pp. 103-106.

[11] August W. Rihaczek, "Signal Energy Distribution in Time and Frequency," *IEEE Transactions on Information Theory*, May 1968, **IT-14**, No. 3, pp. 369-374.

[12] L. B. White and B. Boashash, "On Estimating the Instantaneous Frequency of a Gaussian Random Signal by use of the Wigner-Ville Distribution," *IEEE Transactions on Acoustics, Speech and Signal Processing*, March 1988, **36**, No. 3, pp. 417-420.

[13] T. A. C. M. Claasen and W. F. G. Mecklenbrauker, "The Wigner Distribution—A Tool for Time-Frequency Signal Analysis. Part I: Discrete-Time Signals," *Philips Journal of Research*, 1980, **35**, No. 3, pp. 217-250.

[14] T. A. C. M. Claasen and W. F. G. Mecklenbrauker, "The Wigner Distribution—A Tool for Time-Frequency Signal Analysis. Part III: Relations with Other Time-Frequency Signal Transformations," *Philips Journal of Research*, 1980, **35**, No. 6, pp. 372-389.

[15] T. A. C. M. Claasen and W. F. G. Mecklenbrauker, "The Wigner Distribution—A Tool for Time-Frequency Signal Analysis. Part II: Continuous-Time Signals," *Philips Journal of Research*, 1980, **35**, No. 4-5, pp. 276-300.

[16] Cornelis P. Janse and Arie J. M. Kaizer, "Time-Frequency Distributions of Loudspeakers: The Application of the Wigner Distribution," *Journal of the Audio Engineering Society*, April 1983, **31**, No. 4, pp. 198-222.

[17] Werner Krattenthaler and Franz Hlawatsch, "Time-Frequency Design and Processing of Signal Via Smoothed Wigner Distributions," *IEEE Transactions on Signal Processing*, 1993, **41**, No. 1, pp. 278-287.

[18] B. V. K. Vijaya Kumar and C. Carroll, "Pattern Recognition Using Wigner Distribution Function," *Proceedings of the Tenth International Optical Computing Conference*, 1983, pp. 130-135.

[19] Javier R. Fonollosa and Chrysostomos L. Nikias, "Wigner Polyspectra: Higher-Order Spectra in Time Varying Signal Processing," *ICASSP*, **5**, 1991, pp. 3085-3088.

[20] Javier R. Fonollosa and Chrysostomos L. Nikias, "Transient Signal Detection using the Wigner Bispectrum," *Asilomar Conference on Signals, Systems & Computers*, **2**, 1991, pp. 1087-1092.

[21] Javier R. Fonollosa and Chrysostomos L. Nikias, "Wigner Higher Order Moment Spectra: Definition, Properties, Computation and Application to Transient Signal Analysis," *IEEE Transactions on Signal Processing*, January 1993, **41**, No. 1, pp. 245-265.

[22] Boualem Boashash and Gordon Frazer, "Time-Varying Higher-Order Spectra, Generalised Wigner-Ville Distribution and The Analysis of Underwater Acoustic Data," *ICASSP*, 1992, pp. 193-196.

[23] J. C. Cardoso, P. J. Fish, and M. G. Ruano, "Parallel Implementation of a Choi-Williams TFD for Doppler Signal Analysis," *Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Biomedical Engineering towards the year 2000 and Beyond*, NJ, USA, **3**, 1998, pp. 1490-92.

[24] C. Griffin, "A Comparison Study on the Wigner and Choi-Williams Distributions for Detection," *ICASSP*, **2**, 1991, pp. 1485-1488.

[25] Antonia Papandreou and G. Faye Boudreauz-Bartels, "Distributions for Time-Frequency Analysis: A Generalization of Choi-Williams and the Butterworth Distribution," *ICASSP*, NewYork, USA, **5**, 1992, pp. 181-192.

[26] A. Papandreou, G. F. Boudreaux-Bartels, and S. M. Kay, "Detection and Estimation of Generalized Chirps Using Time-Frequency Representations," *Asilomar Conference on Signals, Systems and Computers*, **1**, 1995, pp. 50-54.

[27] Antonio H. Costa and G. Faye Boudreaux-Bartels, "Design of Time-Frequency Representations Using a Multiform, Tiltable Exponential Kernel," *IEEE Transactions on Signal Processing*, October 1995, **43**, No. 10, pp. 2283-2301.

[28] Hyung-ILL Choi and William J. Williams, "Improved Time-Frequency Representation of Multicomponent Signals Using Exponential Kernels," *IEEE Transactions on Acoustics, Speech and Signal Processing*, June 1989, **37**, No. 6, pp. 862-871.

[29] Ljubisa Stankovic and Veselin Ivanovic, "Further Results on the Minimum Variance Time-Frequency Distribution Kernels," *IEEE Transactions on Signal Processing*, June 1997, **45**, No. 6, pp. 1650-1655.

[30] Moeness G. Amin, "Minimum Variance Time-Frequency Kernels for Signals in Additive Noise," *IEEE Transactions on Signal Processing*, September 1996, **44**, No. 9, pp. 2352-2356.

[31] B. V. K. Vijaya Kumar and C. W. Carroll, "Effects of Sampling on Signal Detection Using the Cross-Wigner Distribution Function," *Applied Optics*, November 1984, **23**, No. 22, pp. 4090-4994.

[32] B. V. K. Vijaya Kumar and Christopher W. Carroll, "Performance of Wigner Distribution Function Based Detection Methods," *Optical Engineering*, December 1984, **23**, No. 6, pp. 732-737.

[33] B. Ph. van Milligen, *et al.*, "Wavelet Bicoherence: A New Turbulence Analysis Tool," *Physics of Plasmas*, August 1995, **2**, No. 8, pp. 3017-3032.

[34] B. Ph. van Milligen, C. Hidalgo, and E. Sanchez, "Nonlinear Phenomena and Intermittency in Plasma Turbulence," *Physical Review Letters*, January 1995, **74**, No. 3, pp. 395-398.

[35] Marie Farge, Nicholas Kevlahan, Valerie Perrier, and Eric Goirand, "Wavelets and Turbulence," *Proceedings of the IEEE*, April 1996, **84**, No. 4, pp. 639-669.

[36] Patrick Flandrin, "Some Features of Time-Frequency Representations of Multicomponent Signals," *Proceedings IEEE ICASSP*, Sandiego, CA, **3**, 1984, pp. 41.B.4.1-41.B.4.4.

[37] Leon Cohen, "Generalized Phase-Space Distribution Functions," *Journal of Mathematical Physics*, 1966, **7**, pp. 781-786.

[38] Soo-Chang Pei and Er-Jung Tsai, "New Time-Frequency distribution," *IEEE International Symposium on Circuits and Systems*, NewYork, USA, **1**, 1993, pp. 204-207.

[39] Yunxin Zhao, LEs E. Atlas, and Robert J. Marks II, "The Use of Cone-Shaped Kernels for Generalized Time-Frequency Representations of Nonstationary Signals," *IEEE Transactions on Acoustics, Speech and Signal Processing*, July 1990, **38**, No. 7, pp. 1084-1091.

[40] Personal communication with Dr. P. Flandrin in 1999.

[41] B. P. Lathi, *Signals, Systems and Communication*, pp. 14-152 and 515-531, John Wiley and Sons, 1965.

[42] Javier Rodriguez Fonollosa and Chrysostomos L. Nikias, "Analysis of Transient Signals Using Higher-Order Time-Frequency Distributions," *ICASSP*, 1992, pp. 197-200.

[43] Javier R. Fonollosa and Chrysostomos L. Nikias, "Analysis of finite-energy signals using higher-order moments and spectra-based time-frequency distributions," *Signal Processing*, 1994, **36**, No. 3, pp. 315-328.

[44] Fritz Oberhettinger, *Fourier Transforms of Distributions and Their Inverses—A Collection of Tables*, pp. 17-36, Academic Press, 1973.

[45] I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Series and Products*, . ›. 92-1058, Academic Press, 1980.

[46] Douglas L. Jones and Richard G. Baraniuk, "An Adaptive Optimal-Kernel Time-Frequency Representation," *ICASSP*, NewYork, USA, 4, 1993, pp. 109-112.

[47] F. Auger, P. Flandrin, P. Goncalves, and O. Lemoine, *Time-Frequency Toolbox For Use with MATLAB*, pp. 1-50, Centre National de la Recherche Scientifique (CNRS), 1995-1996.

[48] Jechang Jeong and William J. Williams, "Kernel Design for Reduced Interference Distributions," *IEEE Transactions on Acoustics, Speech and Signal Processing*, February 1992, **40**, No. 2, pp. 402-412.

[49] LJubisa Stankovic, "Auto-Term Representation by the Reduced Interference Distributions: A Procedure for Kernel Design," *IEEE Transactions on Signal Processing*, June 1996, **44**, No. 6, pp. 1557-1563.

[50] Steven B. Hearon and Moiness G. Amin, "Minimum-Variance Time-Frequency Distribution Kernels," *IEEE Transactions on Signal Processing*, May 1995, **43**, No. 5, pp. 1258-1262.

[51] C. S. Burrus, R. A. Gopinath, and H. Guo, *Introduction To Wavelets and Wavelet Transforms: A Primer*, pp. 10-50, Prentice-Hall, Inc., USA, 1998.

[52] Ingrid Daubechies, *Ten Lectures on Wavelets*, pp. 1-16, The Society for Industrial and Applied Mathematics, Philadelphia, 1992.

[53] A. Grossmann, R. Kronland-Martinet, and J. Morlet, "Reading and Understanding Continuous Wavelet Transforms," in *Wavelets: Time Frequency Methods and Phase Space*, J.M. Combes, A. Grossmann, and P. Tchamitchian Eds., Springer-Verlag, pp. 2-20, 1989.

[54] P. G. Lemarie-Rieusset, "More Regular Wavelets," *Applied and Computational Harmonic Analysis*, 1998, **5**, No. 1, pp. 92-105.

[55] O. Rioul and M. Vetterli, "Wavelets and Signal Processing," *IEEE Signal Processing Magazine*, Oct. 1991, **8**, No. 4, pp. 14-38.

[56] G. Strang, "Wavelets," *American Scientist*, June 1994, **82**, No. 3, pp. 250-255.

[57] G. Strang, "Wavelets and Dilation Equations: A Brief Introduction," *SIAM Review*, Dec. 1989, **31**, No. 4, pp. 614-627.

[58] G. Strang and T. Nguyen, *Wavelets and Filter Banks*, pp. 200-270, MA: Wellesley-Cambridge Press, 1997.

[59] G. Strang and V. Strela, "Orthogonal Multiwavelets With Vanishing Moments," *Optical Engineering*, July 1994, **33**, No. 7, pp. 2104-2107.

[60] P. G. Lemarie-Rieusset, "On the Existence of Compactly Supported Dual Wavelets," *Applied and Computational Harmonic Analysis*, 1997, **3**, No. 1, pp. 117-118.

[61] G. Strang, "The Optimal Coefficients in Daubechies Wavelets," *Physica D*, 1992, **60**, No. 1-4, pp. 239-244.

[62] J. Shen and G. Strang, "Asymptotics of Daubechies Filters, Scaling Functions and Wavelets," *Applied and Computational Harmonic Analysis*, 1998, **5**, No. 3, pp. 312-331.

[63] B. Jawerth and W. Sweldens, "An Overview of Wavelet Based Multiresolution Analyses," *SIAM Review*, Sep. 1994, **36**, No. 3, pp. 377-412.

[64] C. E. Heil and D. F. Walnut, "Continuous and Discrete Wavelet Transforms," *SIAM Review*, Dec. 1989, **31**, No. 4, pp. 628-666.

[65] M. Holschneider, *Wavelets: An Analysis Tool*, pp. 17-53, 271-341, Oxford: Clarendon Press, 1995.

[66] D. Jordan and R. W. Miksad, "Implementation of the Continuous Wavelet Transform for Digital Time Series Analysis," *Review of Scientific Instruments*, March 1997, **68**, No. 3, pp. 1484-1494.

[67] Michel Misiti, Yves Misiti, G. Oppenheim, and Jean-Michel Poggi, *Wavelet Toolbox for Use with MATLAB—A User's Guide*, pp. 6.2-6.114, The MathWorks, Inc., MA, Chapter 6, 1997.

[68] R. N. McDonough and A. D. Whalen, *Detection of Signals in Noise*, pp. 152-242, 340-438, Academic Press, 1995.

[69] B. Celler, G. Y. C. Chung, and C. Phillips, "ECG Analysis and Processing Using Wavelets and Other Methods," *Biomedical Eneineering—Applications, Basis and Communications*, April 1997, **9**, No. 2, pp. 81-90.

[70] Ian Kaminskyj and A. Mterka, "Automatic Source Identification of Monophonic Musical Instrument Sounds," *IEEE International Conference on Neural Networks*, Perth, Western Australia, 1995, pp. 189-194.

[71] Ian Kaminskyj and A. Materka, "Automatic Monophonic Musical Instrument Sound Identification System," *Proc. ACMA Conference*, Brisbane, Australia, 1996, pp. 41-46.

[72] Ian Kaminskyj, "Multidimensional Scaling Analysis of Musical Instrument Sounds' Spectra," *Proc. ACMA Conference*, Wellington, NewZealand, 1999, pp. 36-39.

[73] Moness G. Amin, "Time-Frequency Spectrum Analysis and Estimation for Non-Stationary Random Processes," in *Time-Frequency Signal Analysis: Methods and Applications*, B. Boashash Eds., Longman Cheshire, pp. 208-232, 1992.

[74] H. L. VanTrees, *Detection, Estimation and Modulation Theory: Part 1*, pp. 30-40, 247-260, John Wiley & Sons, New York, 1968.

[75] N. M. Marinovich, "Detection of Non-Stationary Signals in Colored Noise via Time-Frequency Subspaces," in *Time-Frequency Signal Analysis: Methods and Applications*, B. Boashash Eds., John Wiley, pp. 305-323, 1992.

[76] A. M. Sayeed and D. L. Jones, "Optimal Detection Using Bilinear Time-Frequency and Time-Scale Representations," *IEEE Transactions on Signal Processing*, Dec. 1995, **43**, No. 12, pp. 2872-2883.

[77] P. Flandrin, "A Time-Frequency Formulation of Optimum Detection," *IEEE Transactions on Acoustics, Speech and Signal Processing*, September 1988, **36**, No. 9, pp. 1377-1384.

[78] J. E. Moyal, "Quantum Mechanics as a Statistical Theory," *Proceedings of Cambridge Philosophical Society*, 1949, No. 45, pp. 99-124.

[79] B. V. K. V. Kumar and C. Carroll, "Loss of Optimality in Cross Correlators," *Journal of Optical Society of America A—Optics Image Science*, April 1984, **1**, No. 4, pp. 392-397.

[80] Antonia Papandreou, Franz Hlawatsch, and G. Faye Boudreaux-Bartels, "The Hyperbolic Class of Quadratic Time-Frequency Representations; Part I: Constant-Q Warping, the Hyperbolic Paradigm, Properties and Members," *IEEE Transactions on Signal Processing*, December 1993, **41**, No. 12, pp. 3425-3444.

[81] G. F. Boudreaux and T. W. Parks, "Signal Estimation Using Modified Wigner Distributions," *ICASSP*, 1984, pp. 22.3.1-22.3.4.

[82] F. Hlawatsch, A. Papandreou, and G. F. Boudreaux-Bartels, "Regularity and Unitarity of Affine and Hyperbolic Time-Frequency Representations," *ICASSP*, NewYork, **3**, 1993, pp. 245-248.

[83] B. Boashash and P. O'Shea, "Signal Detection Using Time-Frequency Analysis," in *Time-Frequency Signal Analysis: Methods and Applications*, B. Boashash Eds., Longman Cheshire, pp. 279-304, 1992.

[84] A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing*, pp. , Prentice Hall, Englewood Cliffs, New Jersey, 1975.

[85] C. B. Rorabaugh, *DSP Primer*, pp. 469-488, McGraw-Hill, 1999.

[86] Chrysostomos L. Nikias and Jerry M. Mendel, "Signal Processing with Higher-Order Spectra," *IEEE Signal Processing Magazine*, 1993, July, **10**, No. 3, pp. 10-37.

[87] Chrysosotomos L. Nikias and Mysore R. Raghuveer, "Bispectrum Estimation: A Digital Signal Processing Framework," *Proceedings of the IEEE*, July, 1987, **75**, No. 7, pp. 869-891.

[88] W. B. Collis, P. R. White, and J. K. Hammond, "Higher-Order Spectra: The Bispectrum and Trispectrum," *Mechanical Systems and Signal Processing*, May 1998, **12**, No. 3, pp. 375-394.

[89] Kee Bong Kim and Soo Yong Kim, "Chaos from van der Pol-Duffing Oscillator: Bispectral Analysis," *Journal of the Physical Society of Japan*, July 1996, **65**, No. 7, pp. 2323-2332.

[90] Y. C. Kim and E. J. Powers, "Digital Bispectral Analysis and ITs Applications to Nonlinear Wave Interactions," *IEEE Transactions on Plasma Science*, 1979, June, **PS-7**, No. 2, pp. 120-131.

[91] B. A. Jubran, M. N. Hamdan, and N. H. Shabaneh, "Wavelet and Chaos Analysis of Flow Induced Vibration of a Single Cylinder in Cross-Flow," *International Journal of Engineering Science*, June 1998, **36**, No. 7-8, pp. 843-864.

[92] L. E. Atlas, P. J. Loughlin, and J. W. Pitton, "Signal Analysis with Cone Kernel Time-Frequency Distributions and their Application to Speech," in *Time-Frequency Signal Analysis: Methods and Applications*, B. Boashash Eds., Longman Cheshire, pp. 375-388, 1992.

[93] Bouchaib Radi and Jean-Francois Estrade, "Adaptive parallelization techniques in global weather models," *Parallel Computing*, Sept. 1998, **24**, No. 9-10, pp. 1167-1175.

[94] Gary E. Christensen, "MIMD vs. SIMD parallel processing: A case study in 3D medical image registration," *Parallel Computing*, September 1998, **24**, No. 9-10, pp. 1369-1383.

[95] J. M. MacLaren and J. M. Bull, "Lessons Learned when Comparing Shared Memory and Message Passing Codes on Three Modern Parallel Architectures," *International Conference and Exhibition Proceedings on High-Performance Computing and Networking*, Berlin, Germany, 1998, pp. 337-346.

[96] Edward A. Carmona and Charles L. Matson, "Performance of a Parallel Bispectrum Estimation Code," *Proceedings of SPIE on Advanced Signal Processing Algorithms, Architectures and Implementations II, Franklin T. Luk; Ed.*, Dec. 1991, **1566**, pp. 329-340.

[97] John N. Kalamatianos and Elias S. Manolakos, "Parallel Computation of Higher Order Moments on the MASPAR-1 Machine," *International Conference on Acoustics, Speech and Signal Processing*, New York, USA, **3**, 1995, pp. 1832-1835.

[98] Yushu Chen and Andrew Y. T. Leung, *Bifurcation and Chaos in Engineering*, pp. 265-305, Springer-Verlag, 1998.

[99] W. J. Staszewski and K. Worden, "Wavelet Analysis of Time-Series: Coherent Structures, Chaos and Noise," *International of Bifurcation and Chaos*, 1999, **9**, No. 3, pp. 455-471.

[100] R. M. Wilcox, "Exponential Operators and Parameter Differentiation in Quantum Physics," *Journal of Mathematical Physics*, April 1967, **8**, No. 4, pp. 962-982.

[101] SiliconGraphic course note on Parallel Computing.

[102] B. Zhang and S. Sato, "A Time-Frequency Distribution of Cohen's Class With a Compound Kernel and Its Applications to Speech Signal Processing," *IEEE Transactions on Signal Processing*, Jan. 1994, **42**, No. 1, pp. 54-64.

[103] C. Delfs, F. Jondral, "Classification of piano sounds using time-frequency signal analysis," *ICASSP*, Los Alamitos, CA, USA, **3**, 1997, pp. 2093-6.

[104] K. N. Hamdy, A. H. Tewfik, Ting Chen, S. Takagi, "Time-scale modification of audio signals with combined harmonic and wavelet representations," *ICASSP*, Los Alamitos, CA, USA, **1**, 1997, pp. 439-42.

[105] P. De Gersem, B. De Moor, M. Moonen, "Applications of the continuous wavelet transform in the processing of musical signals," *1997 13th International Conference of Digital Signal Processing Proceedings*, NY, USA, **2**, 1997, pp. 563-6.

[106] H. H. Fletcher, "Inharmonicity, nonlinearity and music," *Physicist*, Sept. 2000, **37**, No. 5, pp. 171-5.

[107] J. Jeong, M. K. Joung, S. Y. Kim, "Quantification of emotion by nonlinear analysis of the chaotic dynamics of electroencephalograms during perception of 1/f music," *Biological Cybernetics*, Mar. 1998, **78**, No. 3, pp. 217-25.

[108] N. Birbaumer, W. Lutzenberger, H. Rau, C. Braun, G. Mayer-Kress, "Perception of music and dimensional complexity of brain activity," *International Journal of Bifurcation & Chaos in Applied Sciences & Engineering*, Feb. 1996, **6**, No. 2, pp. 267-78.

[109] M. Morando, G. Lorgavi, C. Martini, "Local prediction in musical time series," *International Journal of Modelling & Simulation*, 1996, **16**, No. 1, pp. 1-5.

[110] A. S. Dmitriev, A. I. Panas, S. O. Starkov, "Experiments on speech and music signals transmission using chaos," *International Journal of Bifurcation & Chaos in Applied Sciences & Engineering*, Aug. 1995, **5**, No. 4, pp. 1249-54.

[111] T. Lambrou, P. Kudumakis, R. Speller, M. Sandler, A. Linney, " Classification of audio signals using statistical features on time and wavelet transform domains," *ICASSP*, NY, USA, **6**, 1998, pp. 3621-4.

[112] G. Olmo, F. Dovis, P. Benotto, C. Calosso, P. Passaro, "Instrument-independent analysis of music by means of the continuous wavelet transform," Denver, Colorado, **3813**, July 1999, pp. 716-26.

[113] R. Sussman, M. Karsh, "Analysis and resynthesis of musical instrument sounds using energy separation," *ICASSP*, 1996, **2**, pp. 997-999.

[114] O. M. Nielsen and M. Hegland, "Parallel performance of fast wavelet transforms," *International Journal of High Speed Computing*, Mar. 2000, **11**, No. 1, pp. 55-74.

[115] A. Smyk and M Tudruj, "Inter-process communication for parallel computations of wavelet transforms on Hitachi SR2201 supercomputer," *Proceedings International Conference on Parallel Computing in Electrical Engineering*, Aug. 2000, pp. 248-52.

[116] Z. Yong-Hong, M. Xiang-Jie, H. Li-Jun and L. Xiao-Mei, "Parallel algorithms for discrete transforms and wavelet transforms with their applications," *Journal of National University of Defense Technology*, April 2000, **22**, No. 2, pp. 41-5.

[117] L. R. C. Suzuki, J. R. Reid, T. J. Burns, G. B. Lamont, S. K. Rogers, "Parallel computation of 3K wavelets," *Proceedings of the Scalable High-Performance Computing Conference*, Knoxville, TN, USA, May 1994, pp. 454-61.

[118] M. M. Pic, H. Essafi, "Wavelet transform on Connection Machine and SYMPATI 2," *International Journal of Modern Physics C-Physics & Computers*, Feb. 1993, **4**, No. 1, pp. 97-103.

[119] M. F. L. Pereira, L. V. Koenigkan, P. E. Cruvinel, "Parallel DSP architecture for reconstruction of tomographic images using wavelets techniques," *Proceedings XIV Brazilian Symposium on Computer Graphics and Image Processing*, Los Alamitos, CA, USA, Oct. 2001, pp. 384.

[120] S. Hungenahally and J. You, "Parallel wavelet transform over distributed computer network for real-time applications," *Real-Time Imaging*, Oct. 2000, **6**, No. 5, pp. 375-89.

[121] L. Yang, M. Misra, "Coarse-grained parallel algorithms for multi-dimensional wavelet transforms," *Journal of Supercomputing*, Jan.-Feb. 1998, **12**, No. 1-2, pp. 99-118.

[122] F. Marino, V. Piuri and E. E. Swartzlander Jr, "A parallel implementation of the 2-D discrete wavelet transform without interprocessor communication," *IEEE Transactions on Signal Processing*, Nov. 1999, **47**, No. 11, pp. 3179-84.

[123] X. Juan and W. Maohua, "Parallel algorithm of wavelet transformation in image processing," *Mini-Micro Systems*, Jan. 1999, **20**, No. 1, pp. 29-32.

[124] H. J. Lee, J. C. Liu, A. K. Chan and C. K. Chui, "Parallel implementation of wavelet decomposition/reconstruction algorithms," *Proceedings of Spie – the International Society for Optical Engineering*, Orlando FL, USA, April 1994, **2242**, pp. 248-59.

[125] H. Szu, "Why the Soliton Wavelet Transform Is Useful for Nonlinear Dynamics Phenomena," *Visual Information Processing*, **1705**, pp. 280-288, 1992.

[126] H. Szu, "Why Adaptive Wavelet Transform?" *SPIE-93*, Orlando, **1961**, 1993.

[127] K. N. Le, K. P. Dabke and G. K. Egan, "The Hyperbolic Time-Frequency Power Spectrum," *Report from the Department of Electrical and Computer Systems Engineering, Faculty of Engineering, Monash University, Clayton, VIC. 3168*, 2000.

PUBLISHED PAPERS

The following papers have been published in international conferences and journals.

K. N. Le, G. K. Egan and K. P. Dabke, "Parallel Computation of the F [illegible] n," *ISSPA 99*, Brisbane, Australia, 1, Aug. 1999, pp. 251-254.

K. N. Le, K. P. Dabke and G. K. Egan, "The Hyperb[illegible] *SPIE AeroSense Proceedings: Wavelet Applications VIII*, April 2[illegible], Florida, USA, pp. 411-22.

K. N. Le, G. K. Egan and K. P. Dabke, "Parallel Computation of the Time-Frequency Power Spectrum," *ICII 2001 Proceedings – Conference C*, Beijing, China, Nov. 2001, pp. 322-28.

K. N. Le, K. P. Dabke and G. K. Egan, "Signal Detection Using Non-Unity Kernel Time-Frequency Distributions," *Optical Engineering*, Dec. 2001, 40, No. 12, pp. 2866-77.

The measured speedup factor of the time-frequency power spectral parallel calculation process is evidently almost linear. The speedup factor increases linearly as there are no "humps" or sudden "jumps" in the speedup factor. This shows that the time-frequency power spectrum can be calculated efficiently using parallel computing.

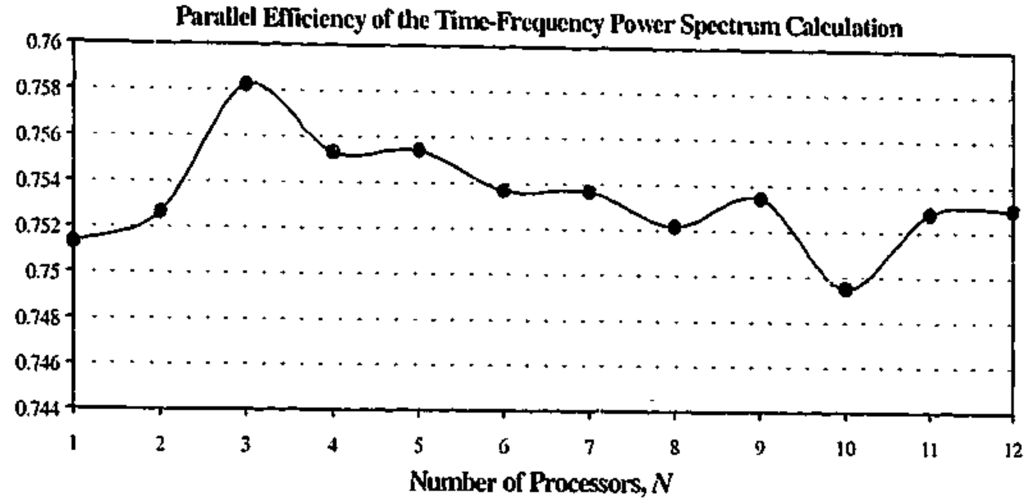

Figure 1.2.2: Parallel efficiency of the time-frequency power spectral calculation process, the loop size is 256 samples

1.3 Discussion

Conditional executions within the main parallel loop unbalance the load among the processors as values of the running variables n, η and τ vary and thus lower the parallel efficiency of the process. Consider a fragment of the code given below

```
index1 = n + η + τ;
index2 = n + η - τ;
   if ((index1 >= M) | (index2 < 0)) temp = (float) 0.0;
   Task 1
   else temp = (float) WFCW*ff[index1]*ff[index2];
   Task 2
```

The above code is located in the inner-most loop of the main parallel loop. Thus, the conditional execution will occur repeatedly many times. Since the number of nested `for` loops of the bispectrum process is less than that of the time-frequency power spectrum process, this explains why the parallel efficiency of the latter is lower than that of the former. As values of the temporary running variables `index1` and `index2` change, so does that of the `temp` variable which results in two unbalanced tasks as can be seen in the above code. For large values of n, η and τ which are close to the loop size M, the load is less computational-heavy compared to that for the case of small values n, η and τ. Thus, processors that had been assigned large-order tasks will take less time to complete these tasks and therefore they will be re-employed to finish any remaining iterations of the main parallel loop.

Parallel efficiency might be improved by increasing the loop size M which increases the ratio of useful work to total work whereas the amount of wasted work remains almost unchanged. However, at the same time the parallel overhead might be increased. In other words, if the loop size M is large compared to the total number of processors N in the system (which is 12 in this case), the parallel efficiency might be improved. For the time-frequency power spectrum case, the number of iterations for the main parallel loop is 256, which is a large number compared to the number of processors of the system (12 independent processors). Thus, the waiting time of any processor will be considerably shorter than its useful time. For example, if the number of processors is $N = 5$, then the number of iterations that each processor has to complete is 256 `div` 5 = 51 and there is 256 mod 5 = 1 iteration left. After the 5 processors finish their 51-iteration work then the one that finishes its tasks earliest has to finish the remaining iteration and other processors (four in this case) have to wait until the remaining iteration is completed. The waiting time associated with one iteration is clearly much less than that of completing 51 iterations. This explains why the parallel efficiency of the time-frequency power spectral calculation process is more stable than that of the bispectral calculation process.

However, the time-frequency power spectrum calculation process suffers from a larger number of conditional executions compared to that of the *direct* bispectrum calculation process which lowers the parallel efficiency. The parallel efficiency of the time-frequency power spectrum calculation process is higher than that of the *indirect* method bispectrum calculation for some values of N since the parallel efficiency of the latter depends strongly on the number of conditional executions in the code and the number of processors used to execute the parallel program (hence the unstability) as can be seen in Figure 1.3.1.

The parallel efficiency of the time-frequency power spectrum calculation process attains its minimum of 74.9% with $N = 10$ and its maximum value of 75.8% with $N = 3$. Most of the time, the efficiency remains almost unchanged around the average value of 75.2% for other values of the number of processors N.

If the program is run serially, then the time it would take to finish it is approximately 2583 seconds or 43 minutes for a small loop size of 256. The parallel program would take less than 15 minutes to complete since the efficiency is higher than 75%. If the loop size further increases, the efficiency of parallel techniques will be increased. The following table summarises the advantages and disadvantages of parallel calculation of the bispectrum and the time-frequency power spectrum.

Table 1.3.1: Comparison of the parallel bispectrum and time-frequency power spectrum calculation processes

Common features		
Parallel efficiency can be improved by increasing the loop size, parallel codes are flexible and can be easily modified for other types of parallel architecture. Data dependency can be resolved effectively.		
	Advantages	**Disadvantages**
Bispectrum	Higher efficiency due to efficient inner loop (*direct* method) and a small loop size.	Unstable efficiency (*indirect* method) due to a small number of loop size of 10. Performance strongly depends on the number of processors used to run the program in parallel.
Time-frequency power spectrum	Stable efficiency due to a large number of iterations of 256 or larger. Performance is uniform as N varies.	Lower efficiency due to a larger number of conditional executions.

Overall, although the parallel efficiency of the time-frequency power spectral process (75.2%) is not that high compared to that of the parallel bispectral calculation process (about 90% or higher), it has been shown that the time-frequency power spectrum can be calculated efficiently by using parallel computing. In addition, parallel computing is a useful tool to solve large and lengthy signal processing tasks. The parallel efficiency of the time-frequency power spectral calculation process might be improved by increasing the loop size. The measured speedup factor was obtained by using 256-sample `for` loops. Since there are 4 nested `for` loops, the number of iterations would be $(256)^4$ or approximately 4.3×10^9 iterations. Because the `stack_size` and `memory_size` quotas on the parallel computer on which all experiments were run were limited, larger loop size (for example 1024 or larger) could not be performed.

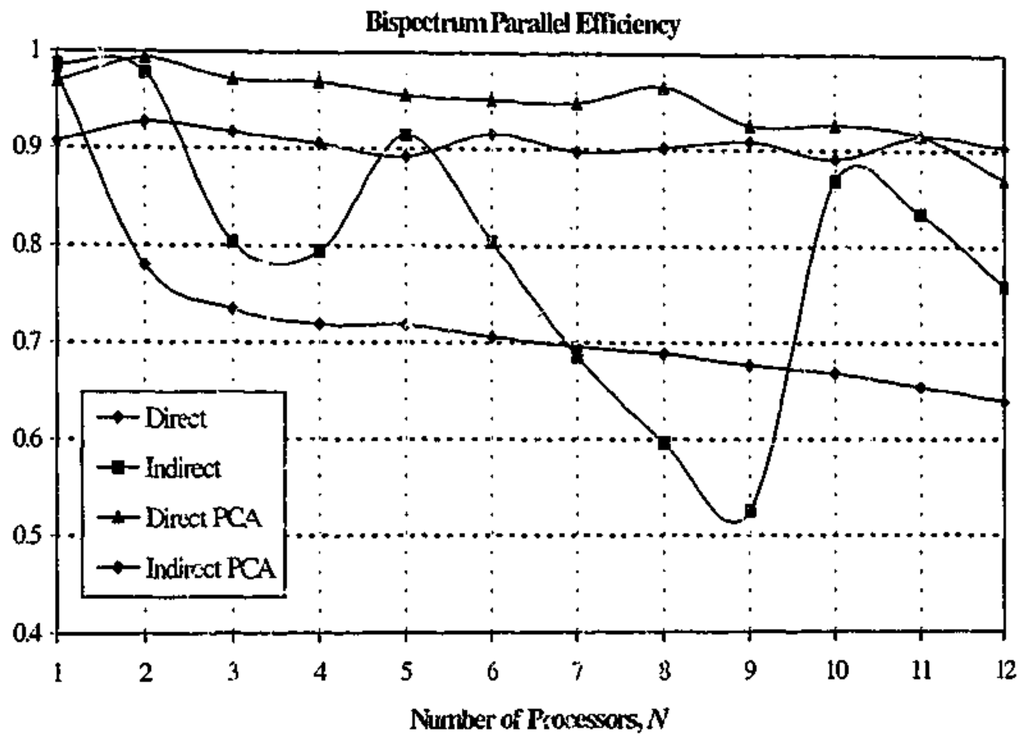

Figure 1.3.1: Parallel efficiency of the bispectrum calculation process [24]

1.4 Conclusion

For the time-frequency power spectral parallel calculation process, near linear-speedup factor has been observed with the minimum efficiency of 75.1% with $N = 1$. The maximum efficiency of 75.6% was obtained when $N = 3$ at the speedup factor of 2.3. With $N = 12$, the efficiency was 75.3% and the speedup factor was 9.03. The average efficiency was approximately 75%. It has been observed that the PCA method could not provide suitable parallel solution to the program hence only the semi-automatic method was employed using the coarse-grained method for the time-frequency power spectrum calculation process. From the results obtained, it appears that the time-frequency power spectrum can be calculated efficiently by using parallel computing. However, the efficiency of the process is not high compared to the bispectral calculation process due to a large number of conditional executions.

References

[1] Leon Cohen, "Introduction: A Primer on Time-Frequency Analysis," in *Time-Frequency Signal Analysis: Methods and Applications*, B. Boashash Eds., John Wiley, pp. 3-43, 1992.

[2] Leon Cohen, *Time-Frequency Analysis*, pp. 136-289, Prentice Hall, Englewood Cliffs, 1995.

[3] J. M. Lipton and K. P. Dabke, "Bispectra of non-linear systems," *Report from the Department of Electrical and Computer Systems Engineering, Faculty of Engineering, Monash University, Clayton VIC. 3168*, 1994.

[4] Personal communication with Dr. Justin M. Lipton in 1996.

[5] J. M. Lipton, *et al.*, "Use of the Bispectrum to Analyse Properties of the Human Electrocardiograph," *Australasian Physical & Engineering Sciences in Medicine*, **21**, March 1998, pp. 1-11.

[6] Javier R. Fonollosa and Chrysostomos L. Nikias, "Wigner Higher Order Moment Spectra: Definition, Properties, Computation and Application to Transient Signal Analysis," *IEEE Transactions on Signal Processing*, January 1993, **41**, No. 1, pp. 245-265.

[7] W. B. Collis, P. R. White, and J. K. Hammond, "Higher-Order Spectra: The Bispectrum and Trispectrum," *Mechanical Systems and Signal Processing*, May 1998, **12**, No. 3, pp. 375-394.

[8] Vinod Chandran and Charles Pezeshki, "Bispectral and Trispectral Characterization of Transition to Chaos in the Duffing Oscillator," *International Journal of Bifurcation and Chaos*, 1993, **3**, No. 3, pp. 551-557.

[9] David R. Brillinger, "Some Basic Aspects and Uses of Higher-Order Spectra," *Signal Processing*, 1994, **36**, pp. 239-249.

[10] August W. Rihaczek, "Signal Energy Distribution in Time and Frequency," *IEEE Transactions on Information Theory*, May 1968, **IT-14**, No. 3, pp. 369-374.

[11] Leon Cohen, "Generalized Phase-Space Distribution Functions," *Journal of Mathematical Physics*, 1966, **7**, pp. 781-786.

[12] Leon Cohen, "Time-Frequency Distributions—A Review," *Proceedings of the IEEE*, July 1989, **77**, No. 7, pp. 941-981.

[13] Leon Cohen and Theodore E. Posch, "Generalized Ambiguity Functions," *IPASSP*, NewYork, USA, **3**, 1985, pp. 1033-1036.

[14] T. A. C. M. Claasen and W. F. G. Mecklenbrauker, "The Wigner Distribution—A Tool for Time-Frequency Signal Analysis. Part III: Relations with Other Time-Frequency Signal Transformations," *Philips Journal of Research*, 1980, **35**, No. 6, pp. 372-389.

[15] A. Uhl, "Vector and Parallel Wavelet Transforms for the Analysis of Time-Varying Signals," *Proceedings of the Seventh SIAM Conference on Parallel Processing for Scientific Computing*, Philadelphia, **1**, 1995, pp. 9-14.

[16] O. N. Moller, "Parallel Wavelet Transforms," *Large Scale Scientific and Industrial Problems*, Berlin, **1**, 1998, pp. 385-389.

[17] S. Hungenahally and I. You, "Parallel Wavelet Transform Over Distributed Computer Network for Real-Time Applications," *Real-Time Imaging*, Oct. 2000, **6**, No. 5, pp. 375-389.

[18] M. Feil, R. Kutil, and A. Uhl, "Parallel Wavelet Transforms on Multiprocessors," *5th International Euro-Par Conference*, Berlin, **3**, 1999, pp. 1013-1017.

[19] W. Cai and Xian-He Sun, "Adaptive Wavelet ADI Method: Application and Parallelization," *International Workshop on Parallel Processing*, New Jersey, **2**, 2000, pp. 547-550.

[20] A. M. Rassau, G. Alagoda, and K. Eshraghian, "Massively Parallel Wavelet Based Video Codec for an Intelligent-Pixel Mobile Multimedia Communicator," *ISSPA '99*, Brisbane, **2**, 1999, pp. 793-795.

[21] P. Moravie, H. Essafi, and M. Pic, "Parallel Wavelet Transform Algorithm for Image Compression," *SPIE International Society of Optical Engineering*, USA, **2488**, 1995, pp. 112-122.

[22] A. Andrews, "Parallel Time-Frequency Analysis," *ICASSP*, UK, **4**, 1989, pp. 2234-2237.

[23] J. C. Cardoso, P. J. Fish, and M. G. Ruano, "Parallel Implementation of a Choi-Williams TFD for Doppler Signal Analysis," *Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Biomedical Engineering towards the year 2000 and Beyond*, NJ, USA, 3, 1998, pp. 1490-92.

[24] Khoa N. Le, Greg K. Egan, and Kishor P. Dabke, "Parallel Computation of the Bispectrum," *ISSPA '99*, Brisbane, Australia, **1**, August 1999, pp. 251-254.

[25] Edward A. Carmona and Charles L. Matson, "Performance of a Parallel Bispectrum Estimation Code," *Proceedings of SPIE on Advanced Signal Processing Algorithms, Architectures and Implementations II, Franklin T. Luk; Ed.*, Dec. 1991, **1566**, pp. 329-340.

Signal detection using time-frequency distributions with nonunity kernels

Khoa Nguyen Le
Kishor P. Dabke
Gregory K. Egan
Monash University
Department of Electrical and Computer Systems Engineering
Clayton Campus
Melbourne, Australia
E-mail: Khoa.Le@eng.monash.edu.au

Abstract. A new technique is proposed to solve the simple binary signal-detection problem using a nonunity kernel time-frequency signal detector (GNKD). The GNKD is based on a Cohen time-frequency power spectrum, employing nonunity kernels only. This class of signal detectors includes the Choi-Williams detector (CWWD) and the recently proposed hyperbolic detector (HyD). This work extends the work done by Kumar and Carroll, who investigated the cross unity-kernel Wigner-Ville detector (CWD), which is a special case of the GNKD class. The discrete Moyal's formula for the nonunity kernel time-frequency distribution is derived. The performance of the GNKD is then compared to that of the CWD and the cross-correlator (CORR) detectors by calculating the signal-to-noise ratio (SNR) and the loss factor Q. The GNKD is shown to be better than both the CWD and the CORR with improvement in the SNR by a factor of $\sqrt{2}$. The HyD can improve the SNR by about 18% compared to the CWWD. Detection of some practical nonstationary signals is also investigated to exemplify the proposed method. [DOI: 10.1117/1.1417498]

Subject terms: Wigner-Ville detector; Moyal's formula; hyperbolic kernel; Choi-Williams kernel; signal-to-noise ratio; Cohen time-frequency power spectrum.

Paper 200411 received Oct. 19, 2000; revised manuscript received June 19, 2001; accepted for publication June 20, 2001.

1 Introduction

Detection of known and deterministic signals in the presence of noise is a classical problem that has been studied intensively in the literature.[1,2] To solve this problem, the signals and the additive noise are assumed to be stationary or wide-sense stationary and zero-mean processes. The matched-filter technique has been shown to be the most effective method to detect signals in this case. However, if the signal is nonstationary, i.e., its power spectrum varies with time or the signal is not known beforehand, then the classical method using the matched-filter technique is limited. Nonstationary signals in practice include radar, sonar signals, image matching,[3,4] and so on. For such nonstationary signals, time-frequency signal detectors need to be employed so that the signals can be detected effectively.

One typical time-frequency detector is the Wigner-Ville unity-kernel detector, which can be used to solve the simple binary detection problem.[5,6] There are two reasons that the Wigner-Ville time-frequency detector is popular. First, the Wigner-Ville distribution is simple and easy to implement, and it provides perfect frequency concentration in the time-frequency plane.[7] Second, originally used in quantum mechanics,[8] Moyal's formula, which is required for calculation of the SNR, is available for the Wigner-Ville distribution. The noise, which is assumed to be complex, wide-sense stationary, can be of two common types that are usually encountered in practice, namely white and colored noise. Using the Wigner-Ville unity kernel detector, detection of nonstationary signals in white noise was done by Flandrin[7] and in colored noise by Marinovich.[3] Both these researchers used a method to detect signals by estimating a statistical function η, which is then compared with a threshold value.[5-7,9-11] If η is greater than the threshold, then the signal is said to be present; otherwise, the signal is not present.

The nonunity kernel time-frequency signal detectors form a class of detectors of which the Wigner-Ville unity-kernel signal detector is one special case. This class of detectors employs Cohen's time-frequency distributions with different kernel functions. Each kernel corresponds to a unique distribution and hence to a unique signal detector. The kernel function strongly influences the performance of the detector in terms of SNR and the higher the SNR, the better the performance of the signal detector. The simplest nonunity kernel of Cohen time-frequency class is the Rihaczek kernel, $\Phi_{\text{Rihaczek}}(\theta,\tau)=\exp(j\theta\tau/2)$. The Choi-Williams kernel signal detector (CWWD) can be considered as the most useful and popular detector due to the effectiveness of the Choi-Williams kernel in suppressing cross terms and its robustness in noisy conditions. A different class of signal detectors is the bilinear signal detectors, in which the nonstationary structure of the signal is exploited to ensure the best match of the signal to the detector's filter.[4] Another class of distribution associated with a detector is the quadratic class of time-frequency power spectrum called the hyperbolic class, which was first proposed by Papandreou and Bartels.[12] Signal detection using this particular class is examined in Refs. 13–15 using the

…d of estimating the statistical function η.

…unity kernel signal detectors have been studied in some detail in the literature, in particular, detectors using the Rihaczek and Choi-Williams kernels. A comparison of the Wigner-Ville and Rihaczek distributions has been done in Ref. 7 in which the Wigner-Ville distribution was found to be more suitable than the Rihaczek distribution in terms of signal detection and preservation of the inner product of Moyal's formula. The Wigner-Ville detector was compared to the Choi-Williams detector[9] for the case of the Doppler target-return signal using the same method presented in Ref. 7. In Ref. 9, the reverberation ratio SRR was estimated instead of the SNR due to the specific requirements of the application of calculating the target return.

Although the Choi-Williams time-frequency distribution has been used to detect the Doppler signal, the statistical function η of GNKD has not been derived. It should also be noted that to estimate the SNR of a time-frequency detector, Moyal's formula for the corresponding time-frequency distribution of the detector is required. While Moyal's formula has been derived for the case of the Wigner-Ville time-frequency distribution of unity kernel only, this formula has not been derived for a nonunity kernel time-frequency distribution. We derive this formula for nonunity kernel time-frequency distribution and then apply it to the statistical function η to calculate the SNR of a detector. Thus, derivation of Moyal's formula is an important step before any performance calculation of a time-frequency detector is carried out. Furthermore, using Moyal's formula is the only method currently available to estimate the SNR of time-frequency detectors, since the process involves multiple products of the corresponding time-frequency distributions that the detector is based on. If Moyal's formula for a particular class of time-frequency detectors, i.e., Moyal's formula for the corresponding time-frequency distributions, does not exist, then it is not possible to estimate the performance of the detector class.

We aim to achieve three goals. First, to derive the prerequisite Moyal's formula for nonunity kernel detectors. This formula can be used for any nonunity kernel detector if a new kernel function and hence its corresponding time-frequency distribution are available. Secondly, the hyperbolic detector (HyD) and Choi-Williams detector (CWWD) are compared so that the effectiveness of the hyperbolic kernel over the Choi-Williams kernels can be clearly identified. Thirdly, the ability of nonunity kernel detectors in detecting practical signals such as ECG, music, and speech is examined in detail.

Section 2 briefly defines the binary signal detection problem and outlines the general expression of the SNR. Section 3.2 derives Moyal's formula in detail for the nonunity kernel time-frequency distribution. In Sec. 3.3, the general detailed expression of the SNR of the GNKD is given using Moyal's formula. The relative performance of the HyD and CWWD is compared by using the geometrical features of the hyperbolic and Choi-Williams weighting functions. Section 4 calculates the SNR by using Moyal's formula from Sec. 3.2 and compares the loss factor Q of three signal detectors, namely, GNKD, CWD, and the cross-correlator detector (CORR). The value of the energy ratio X_1, which plays an important role in determining the performance of a signal detector, is estimated in Sec. 4.4 for a number of signals, including a sinusoid $\sin(t)$, an exponential $\exp(-t)$, exponentially decaying sinusoid $\sin(t)\cdot\exp(-t)$, chirped $\cos(Ct^2)$, the ECG, and speech including all the vowels and the "sh"-sound signals. These signals are used to test the performance of the GNKD, CWD, and CORR. Overall, the GNKD provides a substantially improved SNR for most practical signals compared to the CWD and CORR. The HyD performs better than the popular CWWD with a larger SNR.

2 Binary Detection Problem

The binary detection problem can be understood as a problem of determining the presence of a nonstationary signal $s(t)$ in the presence of the stationary, white, zero mean and complex noise $w(t)$, given the received noisy signal $f(t)$. The signal energy and the noise variance are assumed to be A_0 and N_0, respectively. These parameters are used to estimate the detector's performance by estimating its SNR.

Since the signal is nonstationary, the classical method employed for stationary and known signals cannot be used. Instead, time-frequency signal detectors have to be employed to detect the presence of the nonstationary and unknown signals, which are corrupted by channel noise and other noise sources. It is assumed that it is not possible to separate frequency power spectra of the signal and the noisy received signal $f(t)$, and also that the signal is completely masked by the noise $w(t)$. The two hypotheses for detecting the signal that need to be considered are given in Eq. (1)

$$H_0: f(t)=\omega(t) \quad \text{and} \quad H_1: f(t)=s(t)+\omega(t), \tag{1}$$

in which H_0 means that the signal $s(t)$ is not present and H_1 means the signal is present. The reference signal $s(t)$ is assumed to be unknown, nonstationary, and could be of random type.

The hypotheses are then examined and the main goal is to decide which one of them is likely to hold. This is done by forming a statistics η using the received noisy signal $f(t)$ and the reference signal $s(t)$. The hypotheses are then decided by comparing the statistics η with a threshold value. If η is greater than the threshold, the signal is said to be present. Otherwise, the signal is not present.[3,4,7] The performance of a particular statistics η is determined by estimating its SNR. The SNR of a statistical function η for random variables, which is equivalent to the likelihood ratio, is given by[5,6]

$$\mathrm{SNR}=\frac{|E\{\eta|_{H_1}\}-E\{\eta|_{H_0}\}|}{(\frac{1}{2}\mathrm{Var}\{\eta|_{H_1}\}+\mathrm{Var}\{\eta|_{H_0}\})^{1/2}}, \tag{2}$$

where $E(\cdot)$ and $\mathrm{Var}(\cdot)$ denote the expectation and variance operations on the statistical function η under the hypotheses H_0 and H_1. The SNR of the matched filter or CORR can be found by using the general formula [Eq. (2)], which will be shown in Sec. 4.1. The next section derives Moyal's formula for the general nonunity kernel time-frequency distributions based on the same Moyal's formula for the unity-kernel Wigner-Ville time-frequency distribution.

3 Derivation of the Discrete Moyal's Formula for a General Time-Frequency Distribution

3.1 *Discrete Moyal's Formula for the Wigner-Ville Time-Frequency Distribution*

To successfully estimate the SNR of a time-frequency detector, Moyal's formula of a particular time-frequency distribution must be known. The discrete Moyal's formula for the Wigner-Ville time-frequency distribution has been derived by Moyal and forms the basis for deriving Moyal's formula for a general time-frequency distribution, which is vital in estimating the SNR of a detector using a nonunity kernel time-frequency distribution.

The general time-frequency distribution is denoted as $\mathrm{TFR}(\omega,t)$ in continuous form, or $\mathrm{TFR}(m,n)$ in discrete form, with ω and t the frequency and time variables respectively, and m,n the discrete frequency and time variables, respectively. The general time-frequency distribution is derived by Cohen [16,17] and given in Eq. (3)

$$\mathrm{TFR}(t,\omega)=\frac{1}{4\pi^2}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\underbrace{[\exp-j\theta(t-u)]\cdot\Phi(\theta,\tau)}_{F(t-u,\tau)}\cdot\exp(-j\tau\omega)\cdot R_{t,1}(t,\tau)du\,d\tau\,d\theta, \tag{3}$$

where $F(t-u,\tau)$ is the weighting function, which is the 1-D Fourier transform of the kernel function $\Phi(\theta,\tau)$, $u=t+\tau/2$, and the autocorrelation function $R_{t,1}(t,\tau)=x(u+\tau/2)\cdot x^*(u-\tau/2)$. The Choi-Williams kernel is given by $\Phi_{\mathrm{CW}}(\theta,\tau)=\exp(-\theta^2\tau^2/\sigma)$ (Ref. 18) and the hyperbolic or first-order hyperbolic kernel is given by $\Phi_{\mathrm{Hy}}(\theta,\tau)=[\mathrm{sech}(\beta\theta\tau)]^n$, where $n=1$. Two other kernels that can be used for signal detection are the third-order hyperbolic kernel, $\Phi_{\mathrm{CubHy}}(\theta,\tau)=[\mathrm{sech}(\beta\theta\tau)]^n$, (where $n=3$), and the Choi-Williams-Butterworth (CWB) kernel $\Phi_{\mathrm{CWB}}(\theta,\tau)=\exp(-\theta^2\tau^2/\sigma)/\theta^2\tau^2+1$. It should be noted that the $\Phi_{\mathrm{CWB}}(\theta,\tau)$ kernel satisfies admissibility constraints[17–19] and has not been reported in the literature.

The weighting functions F_{CW}, F_{Hy}, F_{CubHy}, and F_{CWB} of the Choi-Williams, hyperbolic, cubic hyperbolic, and Choi-Williams-Butterworth kernels are given by Eqs. (4)–(7), respectively

$$F_{\mathrm{CW}}=\frac{\sigma\sqrt{\pi}}{\tau}\exp\left(-\frac{\sigma(t-u)^2}{4\tau^2}\right) \tag{4}$$

$$F_{\mathrm{Hy}}=\frac{\pi}{\beta\tau}\mathrm{sech}\left(\frac{\pi(t-u)}{2\beta\tau}\right) \tag{5}$$

$$F_{\mathrm{CubHy}}=\frac{\pi(\beta^2\tau^2+u^2)}{2\beta^3\tau^3}\mathrm{sech}\left(\frac{\pi(t-u)}{2\beta\tau}\right) \tag{6}$$

$$F_{\mathrm{CWB}}=\frac{\pi\exp(1/\sigma)}{2\tau}\cdot\left\{\exp\left(-\frac{t-u}{\tau}\right)\cdot\mathrm{Erfc}\left(\frac{1}{\sigma}-\sqrt{\sigma}\cdot\frac{t-u}{2\tau}\right)+\exp\left(\frac{t-u}{\tau}\right)\cdot\mathrm{Erfc}\left(\frac{1}{\sigma}+\sqrt{\sigma}\cdot\frac{t-u}{2\tau}\right)\right\}. \tag{7}$$

The Wigner-Ville time-frequency distribution in continuous form is given by substituting $\Phi(\theta,\tau)=1$, which is a unity kernel,[16,17,20–22] into Eq. (3) to obtain

$$W(t,\omega)=\int x\left(t+\frac{\tau}{2}\right)\cdot x^*\left(t-\frac{\tau}{2}\right)\cdot\exp(-j\omega\tau)d\tau, \tag{8}$$

where the range of integration is from $-\infty$ to $+\infty$ unless otherwise stated.

In discrete form, the Wigner-Ville distribution of two signals $f(\cdot)$ and $s(\cdot)$ is given by Eq. (9)

$$W_{fs}(n,m)=2\cdot\sum_{\tau_k=-(M/2-|n|)}^{M/2-|n|}\{f(n+\tau_k)\cdot s^*(n-\tau_k)\cdot\exp[-j2\pi km/(M+1)]\}, \tag{9}$$

where m,n are the discrete time and frequency variables, τ_k is the lag parameter, and M is the number of data samples. More detailed background on the Wigner-Ville time-frequency distribution can be found in Refs. 20–22.

The continuous Moyal's formula for the Wigner-Ville time-frequency distribution derived by Moyal in 1949[8] is given by Eq. (10)[5,6,10,23]

$$B_{\mathrm{WV}}=\int_\omega\int_t W_{fg}(t,\omega)\cdot W_{hs}^*(t,\omega)dt\,d\omega=\left\{\int f(t)\cdot g^*(t)dt\right\}\cdot\left\{\int h^*(t)\cdot s(t)dt\right\}, \tag{10}$$

which is a product of two energy terms of the four functions, i.e., the inner product has been reserved for the Wigner-Ville time-frequency distribution.[7] As is seen later, the discrete Moyal's formula for the Wigner-Ville and for the nonunity kernel time-frequency distributions are more complicated with the involvement of the odd and even samples of the signal in the time and frequency domains.

The discrete Moyal's formula of the Wigner-Ville distribution is given by Eq. (11)

$$\beta_{\mathrm{WV}}=\sum_{n=-M/2}^{M/2}\sum_{m=0}^{M-1}W_{fg}(n,m)\cdot W_{hs}^*(n,m). \tag{11}$$

The derivation of the discrete Moyal's formula [Eq. (12)] is given in detail in Ref. 5 and is repeated here as

$$\beta_{\mathrm{WV}}=2M\cdot\left[\sum_{u=-M/2}^{M/2}f(u)\cdot h^*(u)\right]\cdot\left[\sum_{\nu=-M/2}^{M/2}g^*(\nu)\cdot s(\nu)\right]+2M\cdot\left[\sum_{u=-M/2}^{M/2}(-1)^u f(u)\cdot h^*(u)\right]\cdot\left[\sum_{\nu=-M/2}^{M/2}(-1)^\nu g^*(\nu)\cdot s(\nu)\right]. \tag{12}$$

To apply the discrete Moyal's formula to find the SNR of the GNKD, the following identities are applied to Eq. (12): $g(\cdot)\equiv h(\cdot)\equiv s(\cdot)$. The following section derives the discrete Moyal's formula for the GNKD.

3.2 Derivation of the Discrete Moyal's Formula for the GNKD

The discrete Moyal's formula for the Wigner-Ville distribution was given in the previous section. This section extends Moyal's formula for the general Cohen nonunity kernel time-frequency distribution. Given the reference signal $s(t)$ with the energy A_0 and the white noise, zero mean process $w(t)$ of variance N_0, the problem we have to solve is to determine the existence of the reference signal in noisy conditions. The signal energy A_0 and the absolute energy difference B_0 between the even and odd samples of the signal $s(t)$ are defined by

$$A_0=\sum_{k=-M/2}^{M/2}|s(k)|^2$$

and

$$B_0=\left|\sum_{k_{\rm even}=-M/2}^{M/2}|s(k_{\rm even})|^2-\sum_{k_{\rm odd}=-M/2}^{M/2}|s(k_{\rm odd})|^2\right|. \quad (13)$$

The energy and energy difference of the noise $w(t)$ are similarly defined by Eq. (14)

$$N_0=\sum_{k=-M/2}^{M/s}|w(k)|^2$$

and

$$M_0=\left|\sum_{k_{\rm even}=-M/2}^{M/2}|w(k_{\rm even})|^2-\sum_{k_{\rm odd}=-M/2}^{M/2}|w(k_{\rm odd})|^2\right|. \quad (14)$$

From Eqs. (13) and (14), the dimensionless energy ratios of the signal $s(t)$ and the noise $w(t)$ are defined as $X_1=B_0/A_0$ and $X_2=M_0/N_0$, respectively. It is evident that the ratios X_1 and X_2 are $0\leq X_1,X_2\leq 1$, since $A_0\geq B_0\geq 0$ and $N_0\geq M_0\geq 0$. Generally, the values B_0 could be in the range of $-A_0\leq B_0\leq A_0$, however, in this work, only the positive half of B_0 is considered due to its usefulness and convenience in practical situations. The physical meaning of X_1 is discussed in detail in Secs. 3.3 and 4.3.

The discrete form of the general time-frequency distribution is also given by Eq. (11) but with $W(n,m)$ replaced by $\mathrm{TFR}(n,m)$ as shown in Eq. (15)

$$\beta_{\rm GNKD}=\sum_{n=-M/2}^{M/2}\sum_{m=0}^{M-1}\mathrm{TFR}_{fg}(n,m)\cdot\mathrm{TFR}^*_{hs}(n,m). \quad (15)$$

The discrete form of the general time-frequency distribution with a nonunity kernel is given by

$$\mathrm{TFR}(n,m)=2\sum_{\tau=-L}^{L}\sum_{u=-M/2}^{M/2}f(u+\tau)\cdot g^*(u-\tau)\cdot F(n-u,\tau)\cdot\exp(-j2\pi m\tau/M), \quad (16)$$

where $L=M/2-|n|$, $F(n-u,\tau)$ is the 1-D Fourier transform of the kernel functions $\Phi(\theta,\tau)$, M is the length of the input discrete signal, and m and n are the discrete time and frequency variables, respectively.

The discrete Moyal's formula for a nonunity kernel distribution is obtained by taking a product of two discrete $\mathrm{TFR}(n,m)$ given in Eq. (17)

$$\beta_{\rm GNKD}=4\sum_{n=-M/2}^{M/2}\left[\sum_{m=0}^{M-1}\exp[j2\pi\cdot(\tau_l-\tau_k)\cdot m/M]\right]\cdot\left\{\left[\sum_{\tau_k=-L}^{L}\sum_{u_k=-M/2}^{M/2}F\cdot f(u_k+\tau_k)\cdot g^*(u_k-\tau_k)\right]\cdot\left[\sum_{\tau_l=-L}^{L}\sum_{u_l=-M/2}^{M/2}F\cdot h^*(u_l+\tau_l)\cdot s(u_l-\tau_l)\right]\right\}, \quad (17)$$

where F is the weighting function of the kernel.

The summation with respect to m in Eq. (17) can be replaced by $M\cdot\delta(\tau_l-\tau_k)$ (Refs. 5 and 24), which results in $\tau_l=\tau_k=\tau$ so that the impulse function exists. After putting $p_k=u_k+\tau$ and $q_k=u_k-\tau$, and similarly $p_l=u_l+\tau$ and $q_l=u_l-\tau$, Eq. (17) can be rewritten as

$$\beta_{\rm GNKD}=4M\sum_{n=-M/2}^{M/2}\{A\cdot B\}=4M\sum_{n=-M/2}^{M/2}\left\{\left[\sum_{q_k=-|n|}^{|n|}\sum_{p_k=-M+|n|}^{M-|n|}F_{pk}\cdot f(p_k)\cdot g^*(q_k)\right]\cdot\left[\sum_{q_l=|n|}^{|n|+M}\sum_{p_l=-M+|n|}^{M+|n|}F_{pl}\cdot h^*(p_l)\cdot s(q_l)\right]\right\}, \quad (18)$$

where A and B correspond appropriately to the square-bracketed terms in Eq. (18).

From these, we also obtain $P_k=p_k+q_k=2u_k$ and $P_l=p_l+q_l=2ul$, which are even numbers. Thus, to allow the summation over the specified range given in Eq. (18), the factors $1/2[1+(-1)^{p_k+q_k}]$ and $1/2[1+(-1)^{p_l+q_l}]$ are inserted into the expressions A and B in Eq. (18), respectively, without affecting the value of the expression, since the inserted factors are unity in value. After multiplying, separating, and rearranging the variables appropriately, we obtain

$$
\begin{aligned}
\beta_{\mathrm{GNKD}} &= 4M \sum_{n=-M/2}^{M/2} [A \cdot B] \\
&= 4M \sum_{n=-M/2}^{M/2} \left[\sum_{q_k=-|n|}^{|n|} g^*(q_k) \sum_{p_k} F_{pk} \cdot f(p_k) \right. \\
&\quad \left. + \sum_{q_k=-|n|}^{|n|} g^*(q_k) \cdot (-1)^{q_k} \sum_{p_k} F_{p_k} \cdot f(p_k) \cdot (-1)^{p_k} \right] \\
&\quad \cdot \left[\sum_{q_l=-|n|}^{|n|} h^*(q_l) \sum_{p_l} F_{p_l} \cdot s(p_l) + \sum_{q_l=-|n|}^{|n|} h^*(ql) \right. \\
&\quad \left. \cdot (-1)^{q_l} \sum_{pl} F_{p_l} \cdot s(p_l) \cdot (-1)^{p_l} \right],
\end{aligned} \tag{19}
$$

where A and B correspond to the square-bracketed items.

Eq. (19) is the final form of the discrete Moyal's formula of the general time-frequency power spectrum with a nonunity kernel. The next section gives the calculations of the SNR of the statistics of the hypotheses H_0 and H_1 for the nonunity kernel general case by using Eq. (19).

3.3 SNR Calculation of the GNKD and Performance Comparison of Different Nonunity Kernel Detectors

Having obtained Moyal's formula for the general time-frequency distribution, detailed derivation of the SNR of the GNKD can be made by employing Eq. (2). The mean and variance of the statistical function η are given by

$$E\{\eta_{\mathrm{TFR}}|_{H_0}\} = 0, \text{ since } E\{w(t)\} = 0$$

$$E\{\eta_{\mathrm{TFR}}|_{H_1}\} = 2MC$$

where M is the length of the input data samples, and

$$
\begin{aligned}
C &= \left(\sum_{q_k} s^*(q_k) \sum_{p_k} F_k \cdot s(p_k) \right) \cdot \left(\sum_{q_l} s(q_l) \sum_{p_l} F_l \cdot s^*(p_l) \right) \\
&\quad + \left(\sum_{q_k} s^*(q_k) \sum_{p_k} F_k \cdot s(p_k) \right) \\
&\quad \cdot \left(\sum_{q_l} s^*(q_l) \cdot (-1)^{q_l} \sum_{p_l} F_l \cdot s(p_l) \cdot (-1)^{p_l} \right).
\end{aligned} \tag{20}
$$

Under the special conditions $p_k = p_l$, $q_l = q_{l1}$, and $q_k = q_l$, the term given by Eq. (20) becomes a constant $C = (A_{0F})^2 + (B_{0F})^2$, where

$$A_{0F} = \sum_{q_k} s^*(q_k) F_k \cdot s(q_k)$$

and

$$B_{0F} = \sum_{q_l} s^*(q_l) \cdot F_l \cdot s(q_l) \cdot s(q_l) \cdot (-1)^{q_l}.$$

Since the

$$\eta_{\mathrm{TFR}}|_{H_1} = C + \eta_{\mathrm{TFR}}|_{H_0},$$

the variance of $\eta_{\mathrm{TFR}}|_{H_1}$ is equal to that of $\eta_{\mathrm{TFR}}|_{H_0}$. The variance of the statistical functions $\eta_{\mathrm{TFR}}|_{H_1}$ is given by

$$
\begin{aligned}
\mathrm{Var}\{\eta_{\mathrm{TFR}}|_{H_0}\} &= \mathrm{Var}\{\eta_{\mathrm{TFR}}|_{H_1}\} = E\{[\eta_{\mathrm{TFR}}|_{H_1}]^2\} \\
&= M^2 \cdot \sum_{n=-M/2}^{M/2} \sum_{j=-M/2}^{M/2} (A_0 \cdot A_{01FF} \cdot N_{0FF} \cdot A_{01}) \\
&\quad + (A_0 \cdot B_{01} \cdot N_{0FF} \cdot B_{01FF}) + (A_0 \cdot A_{01} \cdot N_{0FF} \\
&\quad \cdot A_{01FF}) + (A_0 \cdot N_{0FF} \cdot B_{01} \cdot B_{01FF}) + (2A_0 \\
&\quad \cdot N_{0FF} \cdot B_0 \cdot B_{0FF}) + (2A_0 \cdot B_0 \cdot N_{0FF} \cdot B_{0FF}) \\
&\quad + (2B_0 \cdot M_{0FF} \cdot A_0 \cdot A_{0FF}) + (2B_0 \cdot M_{0FF} \\
&\quad \cdot B_0 \cdot B_{0FF}) + (2B_0 \cdot M_{0FF} \cdot B_0 \cdot B_{0FF}) \\
&\quad + (2B_0 \cdot M_{0FF} \cdot A_0 \cdot A_{0FF}) \\
&= M^2 \cdot \sum_{n=-M/2}^{M/2} \sum_{j=-M/2}^{M/2} 2[(A_0 \cdot A_{01FF} \cdot N_{0FF} \\
&\quad \cdot A_{01}) + (A_0 \cdot B_{01} \cdot B_{0FF} \cdot B_{01FF})] + 2[2(A_0 \\
&\quad \cdot N_{0FF} \cdot B_0 \cdot B_{0FF}) + 2(B_0 \cdot M_{0FF} \cdot A_0 \cdot A_{0FF}) \\
&\quad + 2(B_0 \cdot M_{0FF} \cdot B_0 \cdot B_{0FF})]
\end{aligned} \tag{21}
$$

where

$$
\begin{aligned}
A_{01FF} &= \sum_{p_l} F_l \cdot s(p_l) \cdot \sum_{p_{l1}} F_{jl1} \cdot s^*(p_{l1}), \\
A_{0FF} &= \sum_{p_l} F_l \cdot s(p_l) \cdot F_{jl} \cdot s^*(p_l), \\
B_{0FF} &= \sum_{p_l} F_l \cdot s(p_l) F_l \cdot s(p_l) \cdot (-1)^{p_l}, \\
B_{01} &= \sum_{q_l} s^*(q_l) \cdot (-1)^{q_l} \cdot \sum_{q_{l1}} s(q_{l1}) \cdot (-1)^{q_{l1}}, \\
B_{01FF} &= \sum_{p_l} F_l \cdot s(p_l) \cdot (-1)^{p_l} \cdot \sum_{p_{l1}} F_{l1} \cdot s(p_{l1}) \cdot (-1)^{p_{l1}}, \\
A_{01} &= \sum_{q_l} s^*(q_l) \cdot \sum_{q_{l1}} s(q_{l1}), \\
M_{0FF} &= \sum_{p_k} F_k \cdot w(p_k) \cdot F_k \cdot w(p_k) \cdot (-1)^{p_k}, \\
N_{0FF} &= \sum_{p_k} F_k \cdot w(p_k) \cdot F_k \cdot w(p_k).
\end{aligned} \tag{22}
$$

The SNR of the GNKD is given by

$$\mathrm{SNR_{GNKD}}=\frac{E\{\eta_{\mathrm{TFR}}|_{H_1}\}}{\mathrm{Var}\{\eta_{\mathrm{TFR}}|_{H_1}\}}=\frac{2MC}{\mathrm{Var}\{\eta_{\mathrm{TFR}}|_{H_1}\}^{1/2}}. \tag{23}$$

The $\mathrm{SNR_{GNKD}}$ for the special case of $p_k=p_l$, $q_k=q_l$ and $q_l=q_{ll}$ is given by

$$\mathrm{SNR_{GNKD}}=\frac{\sqrt{2}\cdot(A_0^2+B_0^2)}{A_0\left(N_0\cdot A_0\cdot\left\{1+\frac{B_0}{A_0}+2\left[\left(\frac{B_0}{A_0}\right)^2+\frac{B_0}{A_0}\cdot\frac{M_0}{N_0}+\frac{M_0}{N_0}+\frac{M_0}{N_0}\cdot\left(\frac{B_0}{A_0}\right)^3\right]\right\}\right)^{1/2}}. \tag{24}$$

It is worth repeating that $X_1=B_0/A_0$ and $X_2=M_0/N_0$, which were defined by Eqs. (13) and (14) in Sec. 3.2, are ratios of the absolute energy difference between the even and odd samples of a signal to its total energy of the input signal and noise, respectively. The ratio X_1 can be estimated by using simulation at different sampling intervals. It is shown later in Sec. 4.4 that the sampling rate can affect the value of X_1, which in turn will affect the performance of the signal detector.

The physical meanings of the energy ratio X_1 can be understood as the ratio of the bandwidth [the Difference_Energy given by Eq. (25)] to the total energy of the input signal. As is shown later in the figures in Sec. 4, the smaller the value of X_1 the higher the signal detector performance in detecting a particular signal. In addition, satisfactory performance can be achieved by having the value of X_1 close to 1.0, provided that X_2 is small (Sec. 4.3). However, the latter scenario is not applicable to situations in which the X_2 ratio of the noise is large. The energy of the input signal can be expressed in terms of the even and odd energy of the input signal terms

$$\begin{aligned}&\text{Total\_Energy}\\&\quad=\text{Even\_Energy}+\text{Odd\_Energy}\\&\quad=\frac{1}{2\pi}\int_{-\infty}^{+\infty}|\hat{F}(\omega)|^2d\omega=\int_{-\infty}^{+\infty}f^2(t)dt\end{aligned}$$

$$\begin{aligned}&\text{Difference\_Energy}\\&\quad=|\text{Even\_Energy}-\text{Odd\_Energy}|\sim\text{Signal\_Bandwidth}\end{aligned}$$

$$\text{Absolute\_Energy\_Ratio}=\frac{\text{Difference\_Energy}}{\text{Total\_Energy}}, \tag{25}$$

where $\hat{F}(\omega)$ is the Fourier transform of the input signal $f(t)$.

Theoretically, the constant signal $(-\infty\leqslant t\leqslant+\infty)$, which according to Eq. (25) has zero bandwidth, is most effectively detected since there is no energy difference between the even and odd samples of the signal. The Fourier transform or the energy density of the constant signal is a single impulse $\delta(\omega)$ located at the origin. This impulse is regarded as a perfect way to concentrate the energy in the frequency domain, since there is no smearing of energy in the frequency domain. The bandwidth of a constant signal is zero, since there is no width in the frequency domain for an impulse. In the case of periodic sinusoidal signals, the Fourier transforms of the functions $\sin(\omega_0 t)$ and $\cos(\omega_0 t)$ are impulses located at frequencies $\pm\omega_0$. These impulses concentrate the energy of the input signal perfectly in the frequency domain, and their bandwidths are effectively zero. Thus, it can be concluded that signals that have zero Fourier-frequency-domain bandwidth, such as the constant and periodic sinusoid signals, are most effectively detected using a time-frequency signal detector. Simulation results in Sec. 4.4 show that the periodic sinusoid has a zero-valued X_1, which is consistent with the theoretical prediction of Eq. (25).

These two cases validate Eq. (25) in which the absolute energy difference between the even and odd samples of an input digital signal is directly proportional to its Fourier frequency domain bandwidth. Other types of signals including exponential transient $\exp(-t)$, chirped signals $\cos(C\cdot t^2)$, exponentially decaying sinusoid $\sin(\omega_0)\cdot\exp(-t)$, and so on have nonzero bandwidths, which result in larger energy ratio X_1. Hence, detecting signals with wider bandwidths is more difficult than detecting the ones with narrower bandwidths. Real signals such as the ECG, speech signals, and so on are reported to be effectively detected using a time-frequency signal detector, as is seen in Sec. 4.4.

The fact that wide-band signals are more difficult to detect than narrow-band signals can also be explained by looking at the problem from the filter point of view. If the signal is wide-band, it is more likely to be contaminated by other signals such as noise, or different types of signals that have been sent at the same time in the same channel in the passband of the system. The role of the filter is to extract the passband of the detected signal. If the passband contains not just the signal but a mixture of two or more signals, it is more difficult to detect the signal.

From Eqs. (20) and (21), it is evident that the SNR of a signal detector is proportional to the volume under the surface of the weighted signal, i.e., a product of signal $s(t)$ and the weighting function of the kernel, and inversely proportional to the volume of the weighted-signal variance. Thus, if the volume under the surface of the weighted signal is larger than that of the weighted-signal variance, the SNR of the corresponding detector is high. Furthermore, it has been found that the hyperbolic kernel is more robust[19,25,26] than the Choi-Williams kernel. Thus, the HyD provides a smaller variance than that of the CWWD for well chosen values of β, as can be seen in Fig. 1 in which the volume under the surface of the weighted-signal variance of the two kernels is displayed. It is clear from Fig. 1

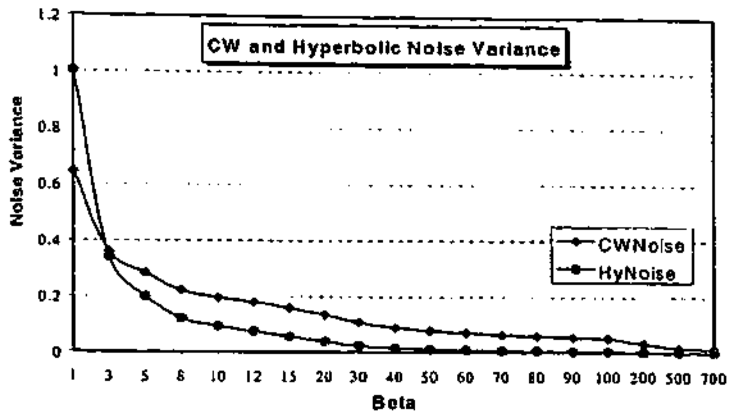

Fig. 1 Noise variance of the hyperbolic and Choi-Williams time-frequency signal detectors.

that the SNR of the HyD is better than that of the CWWD due to having smaller noise variance.

Table 1 gives the volume under the surface of the weighting function of four kernels: hyperbolic, cubic hyperbolic (the third order of the hyperbolic kernel), Choi-Williams, and Choi-Williams-Butterworth kernel (a product of the Choi-Williams and Butterworth kernels[27]). The volumes under the surface of the weighted-signal variance of the four kernels for some typical values β are listed in Table 2. It is evident from Table 2 and Fig. 1 that for different values of the control parameter β, a different volume under the surface is obtained. Thus the control parameter of a kernel plays an important role in determining the performance of the corresponding signal detector.

The volume under the surface of the weighted-signal variance is directly proportional to the variance of the time-frequency signal detector. The smaller this volume, the better is the performance of a particular time-frequency signal detector. From Figs. 2 and 1, it is evident that for $\beta \leq 3$, the performance of the hyperbolic time-frequency signal detector is worse than that of the Choi-Williams due to larger noise variance or larger volume under the surface of the weighted-signal variance. For $\beta > 3$, and typically $\beta = 5$, the HyD provides a larger SNR than that of the CWWD.

The performance of the GNKD in terms of SNR is dependent on the volume under the surface of the weighted signal and its variance. The loss factor Q of the GNKD over the Wigner-Ville unity-kernel signal detector, i.e., a ratio of SNR_{GNKD} to SNR_{CWR}, is given by

$$Q(\text{GNKD/CWD}) = \frac{(V_{\text{GNKD}})^2}{[4(V_{\text{GNKD}})^2 \cdot SV_{\text{GNKD}} + 12(SV_{\text{GNKD}})^2]^{1/2}}, \tag{26}$$

where V_{GNKD} and (SV_{GNKD}) are the volume under the surface of the weighted signal and its variance, respectively. Equation (26) can be used to estimate the improvement factor for each different nonunity kernel time-frequency signal detector. Using the data provided by Tables 1 and 2, the improvement factor Q of the HyD and CWWD are calculated and given in Table 3. It should be noted that the minimum lower bound value of $\sqrt{2}$ (about 3.01 dB) is obtained by employing the special case as stated by Eq. (24).

To measure the relative performance of the HyD and CWWD, the ratio of their SNRs (the Q factor or loss factor) is formed as

Table 1 Volume under the surface of the Choi-Williams, hyperbolic, cubic hyperbolic, and CW-Butterworth [Eqs. (4)–(7), respectively] weighting functions.

	Volume under the surface of the weighting function			
β	Hyperbolic kernel	Cubic hyperbolic kernel	Choi-Williams (CW) kernel	CW-Butterworth kernel
0.1	12.08	12.04	12.014	11.91
1	11.997	11.88	11.98	11.78
5	9.9	7.58	11.02	10.78
10	7.39	4.98	9.953	9.787
20	4.884	2.98	8.63	8.53
50	2.44	1.33	6.78	6.74
100	1.3173	0.68	5.47	5.45

Table 2 Volume under the surface of the variance of the weighting function.

β	Volume under the surface of the variance of the weighting function			
	Hyperbolic kernel	Cubic hyperbolic kernel	Choi-Williams (CW) kernel	CW-Butterworth kernel
0.1	5.25	2.766	1.02	0.36
1	0.51	0.274	0.322	0.234
5	0.1	0.05	0.1425	0.13
10	0.046	0.02	0.0985	0.094
20	0.0195	0.008	0.067	0.065
50	0.0054	0.0018	0.0387	0.0038
100	0.00175	0.000494	0.025	0.025

$$Q(\mathrm{HyD/CWWD}) = \frac{\mathrm{SNR_{HyD}}}{\mathrm{SNR_{CWWD}}} = \frac{C_{\mathrm{HyD}}}{C_{\mathrm{CWWD}}} \cdot \left(\frac{\mathrm{Var}\{\eta_{\mathrm{CWWD}}|_{H_0}\}}{\mathrm{Var}\{\eta_{\mathrm{HyD}}|_{H_0}\}} \right)^{1/2} \approx \left(\frac{V_{\mathrm{HyD}}}{V_{\mathrm{CWWD}}} \right)^2 \cdot \frac{SV_{\mathrm{CWWD}}}{SV_{\mathrm{HyD}}}, \tag{27}$$

where V_{HyD} and V_{CWWD} are the volumes under the surface of the weighted signal and $SV_{\mathrm{HyD}}, SV_{\mathrm{CWWD}}$ are the volumes under the surface of the weighted-signal variance of the hyperbolic and Choi-Williams kernel, respectively.

The relative performance of the HyD to that of the CWWD is displayed graphically in Fig. 2. From Eq. (27), for the case of the hyperbolic and Choi-Williams time-frequency signal detectors, the CWWD is more effective than the HyD by a factor of about 1.6 (60%) for $\beta=1$. However, for $\beta=5$, the HyD yields a larger SNR than that of the Choi-Williams by a factor of 1.15 (15%). For $\beta=10$, the performance of the HyD is approximately 1.18 (18%) times better than the CWWD in terms of the SNR. As β further increases, the performance of the hyperbolic degrades gradually even though at $\beta \geqslant 500$ the performance is slightly improved. This is due to an unequal rate of change of the volume under the surface of the weighted signal and that of the weighted-signal variance.

From Fig. 2, it can be suggested that the HyD is better than the CWWD in terms of SNR over the typical range of the control parameter β of $3 \leqslant \beta \leqslant 10$. Outside this range, the CWWD outperforms the HyD. For large values of β ($\beta \geqslant 500$) the HyD might provide a large SNR, which is mainly due to the relatively large value of the volume under the surface of the weighted-signal variance. It should be noted that large values of β are not applicable in practice, since the hyperbolic weighting function collapses (in shape) into a near-flat function with a very small volume under the surface. This shape of the weighting function indicates that the kernel is not stable under these specific conditions of large β (small σ for the Choi-Williams kernel) and should not be employed as a time-frequency kernel. In contrast, the Choi-Williams weighting function retains its original shape for very small values of σ by having a finite volume under the surface. This makes the Choi-Williams kernel more stable than the hyperbolic kernel over extreme values of the control parameters β and σ.

4 Performance Comparison of Some Time-Frequency Signal Detectors

Section 3.3 derived an expression of the SNR of the GNKD in detail and analyzed the physical meanings of the energy ratio X_1. Relative performance of the HyD and CWWD

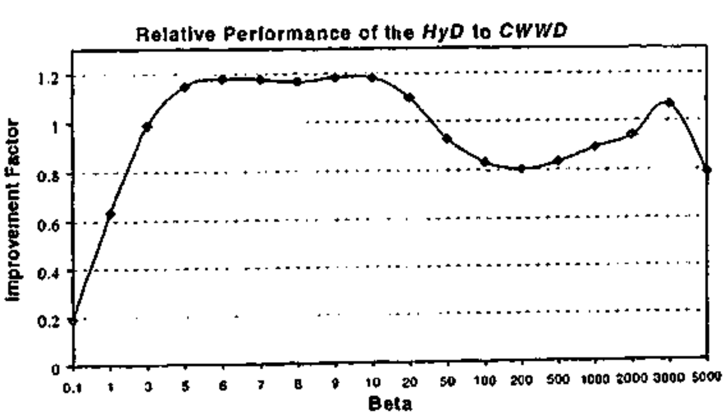

Fig. 2 The Q factor [Eq. (27)] of the HyD and CWWD as a function of the kernel control parameter $\beta = 1/\sigma$. The useful range of β is evidently from 3 to 10.

Table 3 Improvement factors Q of the HyD and CWWD, $3 \leq \beta \leq 10$.

Signal detector	Improvement factor Q
HyD	$22.5 \leq Q_{HyD} \leq 24.8$ dB
CWWD	$22.554 \leq Q_{CWWD} \leq 23.9$ dB

was measured successfully based on geometrical characteristics of the hyperbolic and Choi-Williams kernels, respectively. In this section, performance of three signal detectors, namely, CORR, CWD, and GNKD, will be estimated as a function of $X_1 = B_0/A_0$ and $X_2 = M_0/N_0$ under general cases and special cases. The SNR expression of the GNKD derived in Sec. 3.3 is employed to determine its performance. The SNR expressions of the CWD and CORR have been given in the literature and will be used to compare their performance with that of the GNKD.

4.1 Performance of the Cross-Correlator Signal Detector

The performance of the cross-correlator method, known as the matched filter method, is considered as the best method in binary signal detection, since it provides the best SNR.[2,5] The statistical function η is given by

$$\eta_{CORR} = \int_t f(t) \cdot s^*(t) dt, \quad \text{where} \quad -\infty \leq t \leq +\infty. \tag{28}$$

The SNR of the cross-correlator detector is given by[6]

$$\mathrm{SNR}_{CORR} = \sqrt{\frac{A_0}{N_0}}, \tag{29}$$

where A_0 and N_0 are the signal energy and noise variance, respectively.

The SNR of the CORR is not affected by the energy difference between the even and odd samples of the signal (X_1) as in the case for the CWD with a unity-kernel function, as is discussed in the next section. For the case of a nonunity kernel signal detector, the effects of the absolute energy difference between the even and odd samples of the digital input signal and noise (X_2) are included as is shown in Sec. 4.3.

4.2 Performance of the CWD

The performance of the CWD was studied by Kumar and Carroll for both of the continuous and discrete cases.[5,6,10,11] The SNR of the Wigner-Ville-based signal detection method is given by Eq. (30)

$$\mathrm{SNR}_{CWD} = \sqrt{\frac{A_0}{N_0}} \cdot \frac{1 + \left(\frac{B_0}{A_0}\right)^2}{\left[1 + 3 \cdot \left(\frac{B_0}{A_0}\right)^2\right]^{1/2}}. \tag{30}$$

This SNR_{CWD} of the Wigner-Ville time-frequency signal detector is clearly smaller than the SNR_{CORR} of the cross-correlator detector given by Eq. (29) due to the effects of the ratio $X_1 = B_0/A_0$.

4.3 Performance of the Nonunity-Kernel Signal Detector

The performance of the GNKD was briefly estimated in Sec. 3.3. In this section, performance under special and general cases, such as for small values of X_1 and X_2, is discussed. Relative performance of the GNKD, CORR, and CWD is also estimated by taking ratios of SNRs to form the loss factor Q. The larger the value of the Q factor, the better the performance of the relevant signal detector.

From Eq. (24), it is evident that the SNR of the detection system depends on X_1 and X_2, which clearly shows the effects of the noise process X_2 on the performance of the system. It should be noted that in the case of the Wigner-Ville distribution employing a unity kernel, the effects of the ratio X_2 are not apparent.[5] In addition, the effects of X_1 range from the first order to the third order, as shown by Eq. (24). The noise ratio X_2 is of the first order only.

If $X_2 = M_0/N_0$ is very small, i.e., the noise energy difference is most evenly distributed between its even and odd samples or the noise bandwidth is small, then Eq. (24) becomes

$$\mathrm{SNR}_{GNKD} = \frac{\sqrt{2} \cdot (A_{0F}^2 + B_{0F}^2)}{A_0 \left(N_{0FF} \cdot A_{0FF} \cdot \left\{ 1 + \frac{B_0}{A_0} + 2\left(\frac{B_0}{A_0}\right)^2 \right\} \right)^{1/2}}. \tag{31}$$

After separating the kernel's weighting function, we obtain

$$\mathrm{SNR}_{GNKD} = \sqrt{\frac{A_0}{N_0}} \cdot \frac{\sqrt{2} \cdot \left[1 + \left(\frac{B_0}{A_0}\right)^2\right]}{\left[1 + \frac{B_0}{A_0} + 2 \cdot \left(\frac{B_0}{A_0}\right)^2\right]^{1/2}}, \tag{32}$$

where X_2 is very small.

If $X_2 = M_0/N_0$ is not small, then SNR_{GNKD} will be further reduced and the performance of the GNKD is degraded. It is also important to note that the SNR of the GNKD has been calculated under the special conditions of $p_k = p_l$, $q_k = q_l$ and $q_l = q_{l1}$ in Sec. 3.3. This means that only the autoterms of the summations are included and interaction between the autoterms are ignored. The performance in this case can be considered as the lower limit performance of the detector. For the general case, the SNR of the GNKD, which will be improved, is given by Eq. (33)

$$\mathrm{SNR}_{GNKD} = \sqrt{\frac{A_0}{N_0}} \cdot \frac{\sqrt{2} \cdot \left[1 + \left(\frac{B_0}{A_0}\right)^2\right]}{\left\{1 + \frac{B_0}{A_0} + 2\left[\left(\frac{B_0}{A_0}\right)^2 + \frac{M_0}{N_0} \cdot \left(\frac{B_0}{A_0} + \left(\frac{B_0}{A_0}\right)^3\right)\right]\right\}^{1/2}}. \tag{33}$$

The 3-D graphical presentation of the normalized SNR of the GNKD, SNR_{GNKD}, as a function of $X_1 = B_0/A_0$ and $X_2 = M_0/N_0$ is displayed in Fig. 3. It should be noted again that for the case of the nonunity kernel time-frequency sig-

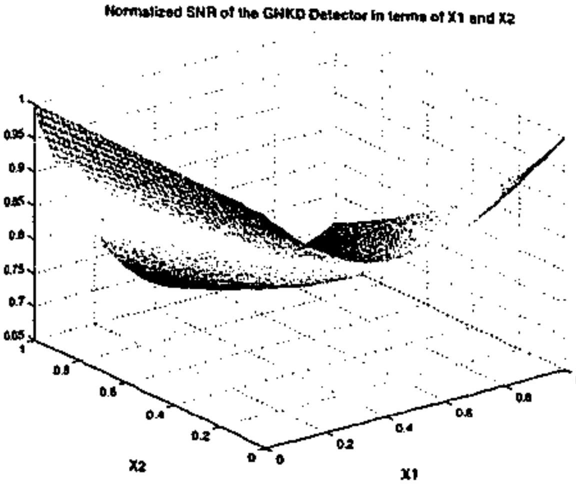

Fig. 3 Normalized SNR_{GNKD} of the GNKD [Eq. (33)] as a function of X_1 and X_2. Maximum performance is obtained by having small values of X_1 of $0.0 \leq X_1 \leq 0.2$ or large values of $0.9 \leq X_1$ and $X_2 \leq 0.2$.

nal detector, the effects of noise are taken into account, which reduces the performance of the detector. From Fig. 3, the minimum SNR_{GNKD} is -0.6602 dB at ($X_1 = 0.63$, $X_2 = 1$), i.e., when the energy difference of the even and odd samples of the noise is equal to its energy. Figures 4 and 5 show the absolute and normalized loss factors of three systems, GNKD/CORR, GNKD/CWD, and CWD/CORR, respectively, as a function of X_1. The absolute plot of the SNR_{GNKD} in Fig. 4 has the same shape as that of Fig. 5 except that its maximum value is $\sqrt{2} \approx 1.414$, i.e., the SNR_{GNKD} is improved by a factor of $\sqrt{2}$ or about 41.4%. From these figures, detailed comparisons of the three signal detectors are shown clearly in Table 4.

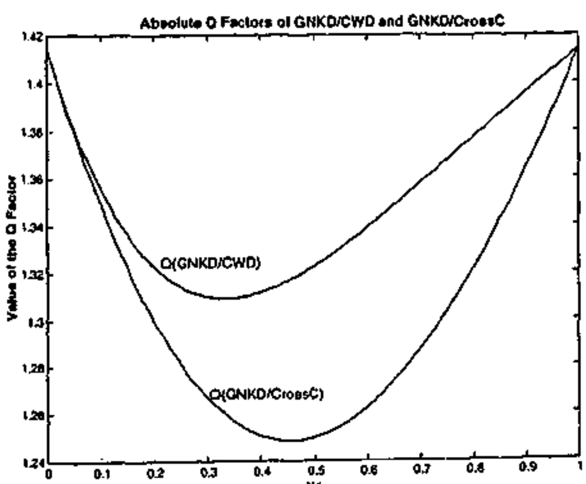

Fig. 4 Loss factor Q(GNKD/CWD) [Eq. (34)] and Q (GNKD/CORR) [Eq. (35)] as a function of X_1. The typical range of X_1 is $0.0 \leq X_1 \leq 0.1$ or $0.9 \leq X_1 \leq 1.0$ when X_2 is small.

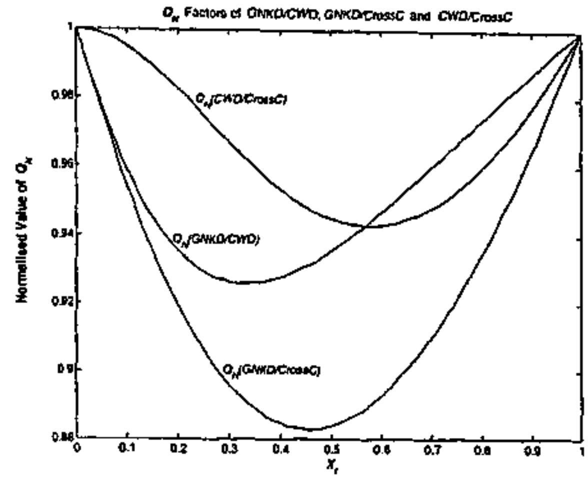

Fig. 5 Normalized Q_N (GNKD/CWD) [Eq. (34)], Q_N (GNKD/CORR) [Eq. (35)] and Q_N (CWD/CORR) [ratio of Eq. (30) to Eq. (29)] as a function of X_1. For the CWD, the typical range of X_1 can be extended to $0.0 \leq X_1 \leq 0.3$. The maximum value of each Q factor was used as the normalization factor.

The Q factor or loss factor of the GNKD and the CWD is then given by the ratio of Eq. (32) to Eq. (30) (when X_2 is small)

$$Q(\text{GNKD/CWD}) = \frac{SNR_{GNKD}}{SNR_{CWD}} = \sqrt{2} \cdot \left[\frac{1 + 3 \cdot \left(\frac{B_0}{A_0}\right)^2}{1 + \frac{B_0}{A_0} + 2 \cdot \left(\frac{B_0}{A_0}\right)^2} \right]^{1/2} = \sqrt{2} \cdot \left(\frac{1 + 3 \cdot X_1^2}{1 + X_1 + 2 \cdot X_1^2} \right)^{1/2}. \quad (34)$$

When X_2 is small, the loss factor of the GNKD and the classical CORR is given by

Table 4 Worst performance ratio of the GNKD to the CWD [Eq. (34)], CWD to the CORR [ratio of Eq. (30) to Eq. (29)] and the GNKD to the CORR [Eq (35)] as a function of X_1 as seen from Fig. 5. The best performance is obtained at $Q_N = 1$ and the corresponding SNR=0 dB.

	Worst Performance		
Normalized loss factor	$X_1 = \frac{B_0}{A_0}$	Q_N	SNR (dB)
GNKD/CWD	0.35	0.9258	−0.67
CWD/Cross-correlator	0.6	0.9428	−0.5
GNKD/Cross-correlator	0.45	0.8829	−1.08

Table 5 The best and worst cases in detection of some popular signals using the GNKD in terms of normalized SNR_{GNKD} in Fig. 3 when X_2 is small.

Signal	Worst case			Best case		
	$\Delta t(ms)$	X_1	SNR (dB)	$\Delta t(ms)$	X_1	SNR (dB)
$\sin(2\pi\times 50n)$	5.00	1.00	0.75	2.00	0.00	1.00
$\exp(-n)$	500	0.462	0.65	10	0.01	0.98
$\exp(-n)\cdot\sin(n)$	600	0.14	0.935	100	0.04	0.975
$\cos[2\pi\cdot n\cdot 0.125\cdot(n/M)]$	800	0.008	$\approx$0.99	0.1	5.0×10^{-7}	$\approx$1.00
ECG (averaged over 12 channels)	The sampling intervals for these cases are fixed. There are no worst or best cases for these signals.			1.00	4.75×10^{-6}	1.00
Speech[28] (vowels a, e, o, u, i, and the sound "sh")				—	3.7×10^{-4}	1.00

$$Q(\text{GNKD/CORR})=\frac{\text{SNR}_{\text{GNKD}}}{\text{SNR}_{\text{CORR}}}=\frac{\sqrt{2}\cdot\left[1+\left(\frac{B_0}{A_0}\right)^2\right]}{\left[1+\frac{B_0}{A_0}+2\cdot\left(\frac{B_0}{A_0}\right)^2\right]^{1/2}}$$

$$=\frac{\sqrt{2}\cdot(1+X_1^2)}{(1+X_1+2\cdot X_1^2)^{1/2}}. \quad (35)$$

It was found in Sec. 3.3 that the ratio of the bandwidth to the total energy of the signal [Eqs. (25) and (13)], X_1, determines the performance of a time-frequency signal detector. Decreasing the bandwidth and increasing the energy of the signal lowers the ratio and leads to better performance. For good performance, a typical range of X_1 is $0.0\leqslant X_1\leqslant 0.2$ or $0.8\leqslant X_1\leqslant 1.0$ if and only if $0.0\leqslant X_2\leqslant 0.2$, as can be seen in Fig. 3. Thus, any value of X_1 in the range of $0.2<X_1<0.8$ will lower the SNR of the detector considerably and should not be used. The performance of the GNKD has been estimated and compared with other detectors, such as the CWD and the CORR. The next section studies the effects of sampling on the performance of a time-frequency signal detector using typical waveforms in practice.

4.4 *Some Typical Examples*

In Secs. 4.1 through 4.3, the performance of the CORR, CWD, and the GNKD were estimated theoretically by using the discrete Moyal's formula derived in Sec. 3.2. In this section, the experimental aspects of detection performance and the effects of sampling on the input signal are examined. Moreover, particular attention is given to how the energy ratio X_1 varies with different values of the sampling interval (Δt). Since X_1 is the energy ratio of the even and odd samples of the digital input signal, its value depends strongly on the type of the signal and the sampling interval Δt. Some typical and popular signals in practice are examined, such as sinusoid at 50 Hz $[\sin(2\pi\times 50t)]$, decaying exponential $\exp(-t)$, exponentially decaying sinusoid $\sin(t)\cdot\exp(-t)$, chirped $\cos(C\cdot t^2)$, ECG, and speech.

As was mentioned in Sec. 3.3, for digital input signals, the sampling interval does affect the value of the energy ratio X_1 of the signal. A number of waveforms have been digitized at different sampling rates and the experimental results are summarized in Table 5. The sampling interval should be small enough to enable small values of X_1. In this case, the sampling frequency is set to be at about four times larger than the critical Nyquist frequency of the input signal. It is important to emphasize that for periodic signals, the signal interval should be chosen long enough so that X_1 can be estimated correctly. From Table 5, it appears that sinusoidal signals can be detected efficiently using the GNKD because of the low value of X_1 in the best case scenario. The transient signal $\exp(-t)$ has a large X_1, which can have low SNR if the sampling interval $\Delta t\geqslant 0.5$ s. The exponentially decaying sinusoidal signal has the worst X_1 of 0.14 at $\Delta t=0.6$ s, with the corresponding SNR=0.935, as can be seen in Table 5. The ECG and speech signals appear to have small X_1, which might suggest that these signals can be detected successfully using the GNKD. The nonstationary chirp signals can be detected very efficiently by using the GNKD, with the worst and best SNRs very close at 0.99 and 1.00, respectively. From Fig. 5, it is clear that signals with $X_1\leqslant 0.1$ yields $Q_N\geqslant 0.95$, which corresponds to satisfactory SNR.

Based on the performance of the general time-frequency signal detector, it is evident that stationary signals such as sinusoids can be detected effectively. However, there are other simpler and equally effective methods for doing this. Thus, time-frequency signal detectors can be employed to detect stationary signals. Detecting nonstationary signals such as decaying exponential $\exp(-t)$, chirp, and exponentially decaying sinusoid $\sin(t)\cdot\exp(-t)$ signals is dependent on the sampling interval used to sample the signal. If the sampling interval Δt is small enough, the detection process will be most effective. This is consistent with Nyquist sampling theorem.

5 Conclusions

We report on some contributions in the field of time-frequency signal detection.

First, the discrete Moyal's formula has been derived for the general case in which the kernel function is not a unity kernel. The performance of the general nonunity kernel signal detector GNKD has been examined by using the discrete Moyal's formula to obtain the SNR of the statistical function η. It has been shown that the GNKD performs better than the cross Wigner-Ville detector CWD by increasing the loss factor and the SNR by a minimum factor of $\sqrt{2}$. The performance of the correlator detector CORR has also been examined and compared with that of the CWD and GNKD. It has been found that the hyperbolic detector HyD and Choi-Williams detector CWWD can improve the SNR over the CWD by a factor Q in the range of $22.5\,\mathrm{dB} \leq Q_{\mathrm{Hy}} \leq 24.8\,\mathrm{dB}$ and $22.5\,\mathrm{dB} \leq Q_{\mathrm{CW}} \leq 23.99\,\mathrm{dB}$, respectively, over the typical range of $3 \leq \beta \leq 10$.

Second, a new signal detector, the hyperbolic time-frequency signal detector, has been investigated. The new detector has been proven to be better than the famous CWWD and CWD by improving the SNR by 18% with the range of the control parameter β being $3 \leq \beta \leq 10$.

Third, the performance of time-frequency signal detectors using a number of typical signals has been examined. It has been shown that the sampling interval used for sampling the input signal can affect the performance of a time-frequency signal detector by varying the energy ratio $X_1 = B_0/A_0$. It has been observed by simulation that sinusoidal and chirped signals can be efficiently detected with satisfactory SNR. Transient signals can be detected efficiently using a suitable sampling interval. Physiological signals such as the ECG and speech can be detected successfully with the normalized SNR in the approximate range of 0.99 to 1.00.

References

1. R. N. McDonough and A. D. Whalen, *Detection of Signals in Noise*, pp. 152–242, 340–438, Academic Press, New York (1995).
2. H. L. VanTrees, *Detection, Estimation and Modulation Theory: Part I*, pp. 30–40, 247–260, John Wiley & Sons, New York (1968).
3. N. M. Marinovich, "Detection of non-stationary signals in colored noise via time-frequency subspaces," in *Time-Frequency Signal Analysis: Methods and Applications*, B. Boashash Eds., pp. 305–323, John Wiley & Sons (1992).
4. A. M. Sayeed and D. L. Jones, "Optimal detection using bilinear time-frequency and time-scale representations," *IEEE Trans. Signal Process.* 43(12), 2872–2883 (1995).
5. B. V. K. Vijaya Kumar and C. W. Carroll, "Effects of sampling on signal detection using the cross-Wigner distribution function," *Appl. Opt.* 23(22), 4090–4994 (1984).
6. B. V. K. Vijaya Kumar and W. Carroll, "Performance of Wigner distribution function based detection methods," *Opt. Eng.* 23(6), 732–737 (1984).
7. P. Flandrin, "A time-frequency formulation of optimum detection," *IEEE Trans. Acoust., Speech, Signal Process.* 36(9), 1377–1384 (1988).
8. J. E. Moyal, "Quantum mechanics as a statistical theory," *Proc. Cambridge Philosophical Soc.* No. 45, pp. 99–124 (1949).
9. C. Griffin, "A comparison study on the Wigner and Choi-Williams distributions for detection," *ICASSP* 2, 1485–1488 (1991).
10. B. V. K. Vijaya Kumar and C. Carroll, "Pattern recognition using Wigner distribution function," *Proc. Tenth Intl. Opt. Computing Conf.* pp. 130–135 (1983).
11. B. V. K. V. Kumar and C. Carroll, "Loss of optimality in cross correlators," *J. Opt. Soc. Am. A* 1(4), 392–397 (1984).
12. A. Papandreou, F. Hlawatsch, and G. F. Boudreaux-Bartels, "The hyperbolic class of quadratic time-frequency representations: Part 1: Constant-Q warping, the hyperbolic paradigm, properties and members," *IEEE Trans. Signal Process.* 41(12), 3425–3444 (1993).
13. A. Papandreou, G. F. Boudreaux-Bartels, and S. M. Kay, "Detection and estimation of generalized chirps using time-frequency representations," *Asilomar Conf. Sign. Syst. Computers* 1, 50–54 (1995).
14. G. F. Boudreaux and T. W. Parks, "Signal estimation using modified Wigner distributions," *ICASSP* 22.3.1–22.3.4 (1984).
15. F. Hlawatsch, A. Papandreou, and G. F. Boudreaux-Bartels, "Regularity and unitarity of affine and hyperbolic time-frequency representations," *ICASSP* 3, 245–248 (1993).
16. Leon Cohen, *Time-Frequency Analysis*, pp. 136–289, Prentice Hall, Englewood Cliffs, NJ (1995).
17. Leon Cohen, "Time-Frequency Distributions—A Review," *Proc. IEEE* 77(7), 941–981 (1989).
18. H. I. Choi and W. J. Williams, "Improved time-frequency representation of multicomponent signals using exponential kernels," *IEEE Trans. Acoust., Speech, Signal Process.* 37(6), 862–871 (1989).
19. Khoa N. Le, Kishor P. Dabke, and Gregory K. Egan, "The hyperbolic kernel for time-frequency power spectrum," submitted for publication.
20. T. A. C. M. Classen and W. F. G. Mecklenbrauker, "The Wigner distribution—a tool for time-frequency signal analysis. Part 3: Relations with other time-frequency signal transformations," *Philips J. Res.* 35(6), 372–389 (1980).
21. T. A. C. M. Classen and W. F. G. Mecklenbrauker, "The Wigner distribution—a tool for time-frequency signal analysis. Part 1: Discrete-time signals," *Philips J. Res.* 35(3), 217–250 (1980).
22. T. A. C. M. Classen and W. F. G. Mecklenbrauker, "The Wigner distribution—a tool for time-frequency signal analysis. Part 2: Continuous-time signals," *Philips J. Res.* 35(4–5), 276–300 (1980).
23. B. Boashash and P. O'Shea, "Signal detection using time-frequency analysis," in *Time-Frequency Signal Analysis: Methods and Applications*, B. Boashash Eds., pp. 279–304, Longman Cheshire, Melbourne (1992).
24. A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing*, pp. 110–120, Prentice Hall, Englewood Cliffs, NJ (1975).
25. L. Stankovic and V. Ivanovic, "Further results on the minimum variance time-frequency distribution kernels," *IEEE Trans. Signal Process.* 45(6), 1650–1655 (1997).
26. M. G. Amin, "Minimum variance time-frequency kernels for signals in additive noise," *IEEE Trans. Signal Process.* 44(9), 2352–2356 (1996).
27. A. Papandreou and G. F. Boudreauz-Bartels, "Distributions for time-frequency analysis: A generalization of Choi-Williams and the Butterworth distribution," *ICASSP*, 5, 181–192 (1992).
28. C. B. Rorabaugh, *DSP Primer*, pp. 469–488, McGraw-Hill, New York (1999).

Proceedings of the Fifth International Symposium on Signal Processing and its Applications ISSPA'99, 1, Brisbane, Australia

PARALLEL COMPUTATION OF THE BISPECTRUM

K. N. Le, G. K. Egan and K. P. Dabke
Department of Electrical and Computer Systems Engineering, Monash University, Clayton Campus, Wellington Road, Melbourne, VIC. 3168, Australia.
Email: {khoa.le, greg.egan, kishor.dabke}@eng.monash.edu.au

ABSTRACT

This paper investigates the effectiveness of parallel computing in calculation of the bispectrum. The bispectrum is estimated by using two different methods namely *direct* and *indirect*. The *direct* method employs 1-D FFT algorithms and the *indirect* method employs the 2-D FFT algorithm to estimate the bispectrum. Both methods have been implemented using 2 different parallel programming techniques: semi-automatic and fully automatic using the Power C Analyzer (PCA). The Silicon Graphics Power Challenge Multiprocessor System (with 12 CPUs) is used to run the parallel codes. Near-linear speedup was observed by employing both techniques. Overall, the *maximum* speedup of 10.84 at $N = 12$ can be achieved for the *direct* method and of 8.7 at $N = 10$ for the *indirect* method using the semi-automatic parallel technique. For the PCA fully parallel technique, the *maximum* measured speedup for the *direct* and *indirect* methods are 10.07 at $N = 11$ and 7.67 at $N = 12$ respectively.

1 INTRODUCTION

Parallel programming and parallel machines have been studied and used extensively in the last few years mainly for predicting weather pattern [1] and in image processing, [2,3]. However, to the best of our knowledge, parallel programming techniques have not been widely used in the field of higher-order statistics and higher-order spectra. Recently, there were two relevant papers applying parallel computing in estimating the bispectrum. The first paper [4], which was published in 1991, reported the performance of an 8-CPU shared-memory CRAY Y-MP machine and 1024-CPU distributed-memory nCUBE machine in calculation of the bispectrum. In particular, the speed-up factor was measured and compared for different machine configurations. Near super-linear speedup was obtained. The second paper [5] proposed an algorithm to estimate higher-order moments using the MASPAR-1 machine, which is a SIMD (Single Instruction Multiple Data) machine.

This paper focuses on the effectiveness of the Silicon Graphics Power Challenge Multiprocessor shared-memory MIMD Machine (HOTBLACK)[1] in calculation of the bispectrum. Each CPU can be considered as an independent PC within the system with separate local memory and cache. To program the system effectively, it is important to arrange the loop parameters and data structure inside the program so that they are suitable for the specific configuration of a particular system. This is the most difficult part of parallel programming in which the programmer must understand the configuration of the particular machine.

The bispectrum [6,7] is estimated using the *direct* method as in Eq. 1

$$B(f_1,f_2) = X(f_1)\cdot X(f_2)\cdot X^*(f_1+f_2) \tag{1}$$

where $X(f)$ is the 1-D FFT of a given discrete series $x(n)$ of M samples and $X^*(\cdot)$ is the complex conjugate of $X(\cdot)$. For more information on the bispectrum, the interested reader should consult references [6, 7].

The *indirect* method uses the 2-D FFT of the tricorrelation function $R_{xxx}(\tau_1,\tau_2)$ abbreviated as R_x;

$$B(f_1,f_2) = 2\text{-}\mathrm{D\ DFFT}\{R_{xxx}(\tau_1,\tau_2)\} \tag{2}$$

where $R_x = \sum_{n=0}^{M-1} x(n)\cdot x(n+\tau_1)\cdot x(n+\tau_2)$ and $\tau_1,\tau_2 = 0,1,2,\ldots,M-1$.

[1] HOTBLACK is a local name of the machine at Monash University.

Sequential C programs were written first based on Eqs. 1 & 2. Then the semi-automatic and fully automatic parallel programs were constructed based on the sequential versions. Semi-automatic programs are obtained by inserting `#pragma` *directives* into the sequential program at appropriate points. This technique is based on the coarse-grained method whereas the PCA method is based on the fine-grain method[2]. Also, arrays and loop parameters of the sequential program are controlled so that they can be accessed independently by different CPUs to avoid data dependency. The fully automatic option is activated by running the `pca` flag of the Power C compiler.

2 EXPERIMENTS AND RESULTS

There are 12 parallel programs with different `numthread(N)` to run on 12 different processors on the system. `numthread(N)` is a parallel *directive* from Silicon Graphics that allows the program to be executed in parallel using N independent CPUs. For example, if N = 3 then the program will be executed in parallel using only 3 CPUs on the system. To ensure efficient compilation, the programs are submitted into a batch queue on the system to obtain more `CPU_time`, `memory_use` and `stack_data_size` quota. Four script files have been written to run the programs under the UNIX operating system.

The speedup factor is estimated as

$$\text{Speedup} = \frac{\text{Sequential\_Time}}{\text{Parallel\_Time}} \tag{3}$$

where the Sequential_Time is the real CPU time used to run the sequential source code and the Parallel_Time is the real CPU time of the slowest thread in a parallel program.

The parallelling efficiency of a parallel program can be estimated as

$$\text{Efficiency} = \frac{\text{Measured speedup}}{\text{Ideal speedup}} \tag{4}$$

Theoretically, the ideal speed-up or super-linear speedup of a program is defined as N if the parallel program is run using N CPUs [8]. The measured speedup is defined as in Eq. 3. Practically, the measured speedup is less than the super-linear speed-up due to parallel overhead.

To ensure consistency between parallel and sequential programming techniques, the output files are compared and it has been observed that they are identical. If the files are not identical, data dependency must have occurred in the sequential source code. The difference between fine-grained and coarse-grained parallel techniques is for the former, only small repeated loops are paralleled and thus results in more than one parallel loop in a parallel program This technique is often used by Power C Compiler. The latter technique parallels loop(s) (usually done manually) that have the largest work in the program. Usually, there is only one largest repeated loop in the program.

For comparison, the size of each segment of the *direct* method is 2048 data points which is twice that of the *indirect* method of 1024 data points (since 1024 data-point segments are still not large enough for the *direct* method, 2048-data-point segments are used instead, also longer segment size up to 10,240 points can be used). From simulation results, it has been observed that the serial program of the *direct* method took approximately 23 seconds to run compared to 587 seconds running time of the *indirect* method although the segments are half as long. Thus the *direct* method is more efficient than the *indirect* method in terms of computing efficiency.

As N is increased, the amount of parallel overhead increases due to synchronization and waiting time of slave processors. However, near super-linear speedup is obtained using the semi- and fully automatic parallel techniques for the *direct* method as seen in Fig. 1. For the *direct* method using PCA parallel technique and for large values of N (for instance, at $N = 12$), the speed-up factor starts to decrease which illustrates the limitation of the fine-grained paralleling method: large amount of parallel overhead for large values of N which lowers the performance (the Direct_PCA and

[2] The idea of these methods will be explained later.

Indirect_PCA curves in Fig. 1). For other values of N less than 12, the PCA method provides better speed-up factor which indicates that the fine-grained method is more suitable for the *direct* method than the coarse-grained method. If the segment size is increased further, near linear-speedup might not be obtained due to long waiting time as explained for the *indirect* method in the following section.

For the *indirect* method using the semi-automatic parallel technique, near super-linear speedup is observed only with some specific number of CPUs which is a multiple of the loop size of 10. That means when N = 1, 2, 5 or 10, near super-linear speed-up will be obtained. For other values of N, since the work associated with each iteration of the loop is large, there will be "unemployed" processors waiting for other processors to complete the tasks. For example, if $N = 6$, all six CPUs will be assigned to the first six iterations of the loop. After finishing the 6 iterations, four of the six CPUs will be used to complete the remaining 4 iterations and two CPUs have to wait ("spin") until the iterations are finished. Since the associated work of each iteration is large, this results in long waiting time and thus the performance of the parallel program will be lowered. This is illustrated for cases of N = 6, 7, 8 and 9 in Fig. 1. Hence, if N is not a factor of the loop size and if the work of each iteration is large, increasing N will increase parallel overhead and constrain the speed-up factor. Since the PCA parallel technique employs the fine-grained method (several small repeated loops will be paralleled instead of the largest repeated loop), parallel overhead will be increased substantially as N is increased, i.e., parallel overhead of using N CPUs will be N times larger than using 1 CPU.

For the *indirect* method applying the PCA parallel technique, the speed-up factor is increased linearly although with lower values compared to the case of semi-automatic parallel technique due to high parallel overhead in several small parallel loops. However, the performance of the PCA method is predictable. From Fig. 1, in contrast with the *direct* method, the coarse-grained method is more appropriate for the *indirect* method since better speed-up factor is achieved. However, to obtain the best performance, the number of CPUs, N, used to run a parallel program applying the *indirect* method must be chosen to be a factor of the loop size.

The measured speed-up factor and paralleling efficiency of the semi- and fully automatic parallel techniques are plotted against the number of processors, N, in Figs. 1 & 2 respectively. In these figures, Direct and Indirect are the speed-up curves using the semi-automatic parallel technique for the *direct* and *indirect* methods respectively. Table 1 compares the *maximum* speed-up factor of the two parallel techniques. Direct PCA and Indirect PCA are speed-up curves obtained using the PCA fully parallel technique for the *direct* and *indirect* methods respectively.

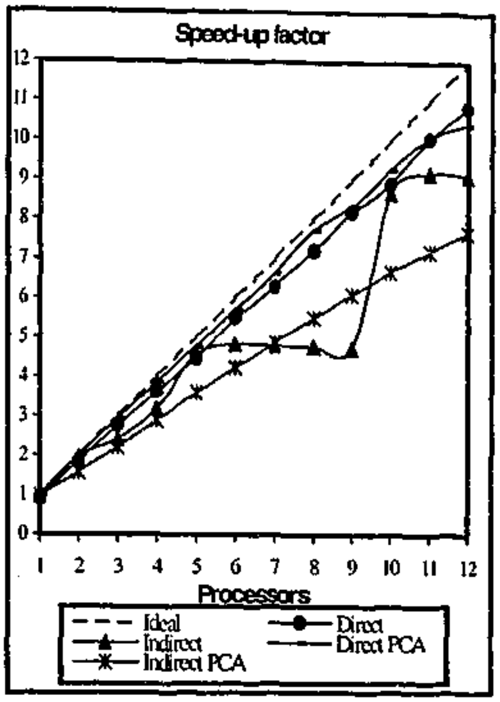

Figure 1: Measured speedup factor for the *direct* and *indirect* methods using the semi- and fully automatic parallel techniques.

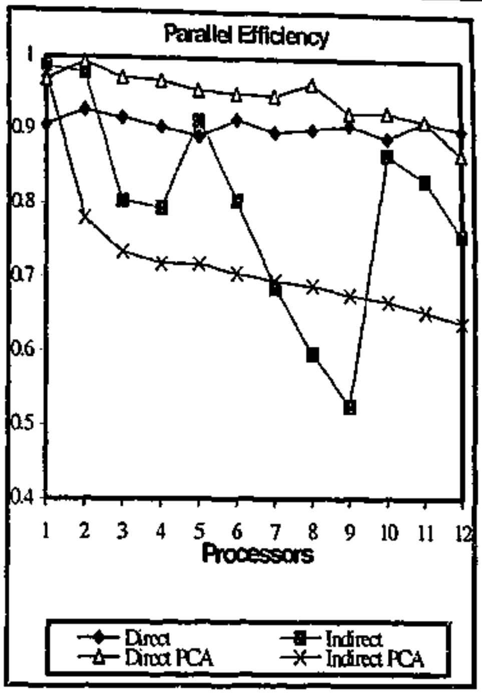

Figure 2: Parallel efficiency for the *direct* and *indirect* methods using the semi- and fully automatic (PCA) parallel techniques.

Table 1: Maximum speedup comparison of semi-automatic and PCA parallel programming techniques.

Method	Semi-automatic	PCA
Direct	10.84 at N=12	10.44 at N=11
Indirect	9.17at N=11	7.67 at N=12

3 CONCLUSIONS

Near linear-speedup was achieved using the semi- and fully automatic parallel techniques for the *direct* method. For the *indirect* method, the amount of overhead gradually increases when $N \geq 6$ due to specific loop structure of the serial program, however, for $N \leq 5$ or $N = 10$, near linear-speedup was observed using the semi-automatic parallel technique. Thus it can be concluded that the *direct* method is more suitable for parallel programming than the *indirect* method. The PCA technique can be used to achieve the speedup factor of 7.67 at $N = 12$ (for the *indirect* method). However, the PCA (Power C Analyzer) method suffers from high parallel overhead for large values of N ($N \geq 12$) since a PCA parallel program (employing the fine-grained parallel method) contains several small parallel loops inside. Further research can be done by applying parallel computing to higher-order spectra such as the trispectrum.

REFERENCES

[1] Bouchaib Radi and Jean-Francois Estrade, "Adaptive parallelization techniques in global weather models," *Parallel Computing*, Sept. 1998, **24**, No. 9-10, pp. 1167-1175.

[2] J. M. MacLaren and J. M. Bull, "Lessons Learned when Comparing Shared Memory and Message Passing Codes on Three Modern Parallel Architectures," in: *International Conference and Exhibition Proceedings on High-Performance Computing and Networking*, Berlin, Germany, 1998, Springer Verlag, pp. 337-346.

[3] Gary E. Christensen, "MIMD vs. SIMD parallel processing: A case study in 3D medical image registration," *Parallel Computing*, Sept. 1998, **24**, No. 9-10, pp. 1369-1383.

[4] Edward A. Carmona and Charles L. Matson, "Performance of a parallel bispectrum estimation code," in: *Proceedings of SPIE on Advanced Signal Processing Algorithms, Architectures and Implementations II*, Franklin T. Luk; Ed., Dec. 1991, **1566**, pp. 329-340.

[5] John N. Kalamatianos and Elias S. Manolakos, "Parallel Computation of Higher Order Moments on the MASPAR-1 Machine," in: *International Conference on Acoustics, Speech and Signal Processing*, New York, USA, 1995, Signal Processing Society IEEE, pp. 1832-1835.

[6] Chrysosotomos L. Nikias and Mysore R. Raghuveer, "Bispectrum Estimation: A Digital Signal Processing Framework," *Proceedings of the IEEE*, July 1987, **75**, No. 7, pp. 869-891.

[7] Chrysostomos L. Nikias and Jerry M. Mendel, "Signal Processing with Higher-Order Spectra," *IEEE Signal Processing Magazine*, 1993, July, **10**, No. 3, pp. 10-37.

[8] "Course Notes on Parallel Computing, Power Learn: Parallel Programming on Silicon Graphics Multiprocessor Systems," Silicon Graphics Computer System, Inc., March, 1995.

Proceedings of SPIE on Wavelet Applications VIII, 4391, AeroSense Conference 2001, Orlando, Florida, USA

THE HYPERBOLIC WAVELET FUNCTION

Khoa N. Le, K. P. Dabke and G. K. Egan
Department of Electrical and Computer Systems Engineering, Clayton, Melbourne, Australia
Email: {Khoa.Le, Kishor.Dabke, Greg.Egan}@eng.monash.edu.au

Abstract

A survey of known wavelet groups is listed and properties of the symmetrical first-order hyperbolic wavelet function are studied. This new wavelet is the negative second derivative function of the hyperbolic kernel function, $[sech(\beta\theta)]^n$ where n = 1, 3, 5,... and n = 1 corresponds to the first-order hyperbolic kernel, which was recently proposed by the authors as a useful kernel for studying time-frequency power spectrum. Members of the "crude" wavelet group, which includes the hyperbolic, Mexican-hat (Choi-Williams) and Morlet wavelets, are compared in terms of band-peak frequency, aliasing effects, scale limit, scale resolution and the total number of computed scales. The hyperbolic wavelet appears to have the finest scale resolution for well-chosen values of $\beta \le 0.5$ and the Morlet wavelet seems to have the largest total number of scales.

Keywords: Mexican-hat wavelet, Morlet wavelet, hyperbolic wavelet, scale resolution, aliasing, wavelet transform.

1 INTRODUCTION

Study of wavelet functions and wavelet transform was done many years ago, starting with the simplest wavelet system, the Haar wavelet [1]. There is a strong connection between wavelet transform and time-frequency power spectrum since both of these techniques view the energy density of the signal in two dimensions, time and frequency. One of the most popular wavelets is the Mexican-hat wavelet, which is negative second derivative function of the Gaussian pulse or Choi-Williams kernel, $\Phi(\theta, \tau) = e^{-\theta^2\tau^2/\sigma}$, where σ is the kernel control parameter [2]. This is one typical case which shows that there exists a link between time-frequency kernels and wavelet functions.

By taking negative second derivative of the kernel, it is possible to generate a variety of different wavelet functions, $\psi(t)$. However, these wavelet functions need to satisfy the admissibility condition of odd symmetry, i.e.,

$$\int_{-\infty}^{+\infty} \psi(t)dt = 0 \tag{1.1}$$

Over the years, a large number of wavelet systems have been proposed and studied extensively, starting with the Haar wavelet system [1, 3] proposed in 1910. In the 1980s, a number of excellent wavelet systems were proposed such as the Daubechies wavelet system [3], the Meyer wavelet system [4] and the Mallat wavelet system [5-10]. These wavelets provide excellent features such as orthogonality, bi-orthogonality, a number of vanishing moments, existence of the scaling function, continuous and discrete transform of the wavelet function, which have received considerable attention from mathematicians.

The purpose of this paper is to investigate the new hyperbolic wavelet function and some of its properties. The central theme of this paper is to establish the permanent link between time-frequency kernels and wavelet functions so that more wavelet functions can be discovered by taking the negative second derivative functions of the corresponding kernels and also new kernels can be found by using known wavelet functions. Only the Morlet wavelet was studied in detail in [11]. The main contribution of this paper is to study and compare properties of the hyperbolic, Choi-Williams and Morlet wavelets in terms of scale resolution, band-peak frequency, aliasing effects, total number of computed scales and scale limit. Symmetrical wavelets with explicit expressions are particularly focused in this paper. Other types of wavelets will not be considered. The Mexican-hat (Choi-Williams) and Morlet wavelets [12, 13] are considered as "crude" wavelets [14] with explicit expressions and symmetrical properties. The hyperbolic wavelet appears to belong to the same wavelet group. However, properties of the hyperbolic wavelet need to be studied before drawing any conclusions about the wavelet itself. Some useful overview papers that summarise the developments in the field of wavelet can be found in [6, 8, 10, 15, 16].

2 WAVELET SYSTEMS AND THE "CRUDE" WAVELET GROUP

There are many interesting wavelet systems that have been proposed and studied such as Daubechies, Mallat, Meyer, Morlet and so on whose wavelet systems are named after the investigators and proponents. These wavelets have been studied extensively and their many interesting and useful properties can be found in [1, 3, 4, 6, 7, 9-11, 15-19]. Wavelet functions have been classified into four different types [14]

1. Type 1 (orthogonal with *FIR* filtering): the wavelet is orthogonal and the *FIR* filter of the wavelet exists. This class includes the Daubechies, Coiflets and Symlets wavelets.
2. Type 2 (biorthogonal with *FIR* filtering): the wavelet is bi-orthogonal and the *FIR* filter of the wavelet exists. The BiorSplines wavelet belongs to this class.
3. Type 3 (orthogonal with scale function): the wavelet and scale functions exist, the *FIR* filter does not exist, however. The Meyer wavelet is a typical member of this class.
4. Type 4 (*FIR* filter and scaling function do not exist): this class has been considered as a "crude" wavelet class since the *FIR* filter and the wavelet's scaling function do not exist. However, the support range of wavelets in this class can be identified as the time-base interval (T in Section 2.1). The wavelets in this class are usually symmetrical and have explicit expressions. As already noted, this paper deals with this particular class of wavelets. The hyperbolic, Choi-Williams (Mexican-hat) and Morlet wavelets belong to this class.

Unlike the Daubechies wavelet family, the Mexican-hat and Morlet wavelets have explicit expressions and are odd symmetrical about the origin. By having explicit expressions, the Morlet and Choi-Williams wavelets are considered as "crude" wavelet systems in which the scaling function has been proven to be non-existent [14]. The Mexican-hat or Choi-Williams wavelet, given in Eq. (2.1) is found by taking the negative second derivative of the Choi-Williams kernel [2] as given by

$$\psi_{CW}(t) = \frac{2}{\sigma} \cdot exp\left(-\tau^2/\sigma\right)\cdot\left(-1+2\tau^2/\sigma\right) \tag{2.1}$$

The hyperbolic wavelet can be considered to be in the same group as the Mexican-hat and Morlet wavelets since they are all symmetrical and have explicit expressions. Their frequency representations are given by

$$F\{\psi_{CW}(t)\} = \hat{\psi}_{CW}(\omega) = \sqrt{\pi\sigma}\cdot\omega^2\cdot exp\left(-\sigma\omega^2/4\right), \text{ and} \tag{2.2}$$

$$F\{\psi_{Morlet}(t)\} = \hat{\psi}_{Morlet}(\omega) = \sqrt{\pi\sigma}\cdot exp\left(-\sigma(\omega-\omega_\psi)^2/4\right) \tag{2.3}$$

where the symbol $F\{\cdot\}$ denotes the Fourier transform operation of the function $\{\cdot\}$ and CW stands for Choi-Williams

The hyperbolic wavelet function is generated by taking negative second derivative of the 2-variable hyperbolic kernel, $\Phi(\theta) = sech(\beta\theta)$, which was proposed by the authors recently. The hyperbolic wavelet function $\psi_{Hy}(\theta)$ is given by

$$\psi_{Hy}(\theta) = (-1)\cdot n\beta^2\cdot[sech(\beta\theta)]^n\cdot\{n-(n+1)\cdot[sech(\beta\theta)]^2\} \tag{2.4}$$

For $n = 1$, the frequency domain representation of the first-order hyperbolic wavelet function is given by

$$F\{\psi(\theta)\} = \hat{\psi}_{Hy}(\omega) = \frac{\pi\omega^2}{\beta}\cdot sech\left(\pi\omega/2\beta\right) \tag{2.5}$$

The hyperbolic wavelet function will be shown by simulation in Section 2.1.1 to satisfy the odd-symmetry condition imposed by Eq. (1.1), i.e., the area under the curve of the second derivative of the hyperbolic kernel is zero.

Other wavelets in the same family with different scales can be obtained by using the translation and dilation relationship or multi-resolution relationship [1]

$$\psi_{a,b}(t)=\frac{1}{\sqrt{a}}\cdot\psi_{a,b}\left(\frac{t-b}{a}\right) \tag{2.6}$$

where a is the scale index and b is the translation or time index of the wavelet. The mother wavelet corresponds to $a = 1$ and $b = 0$.

For each value of the scale index a, there is one unique corresponding wavelet function which can be considered as a band pass filter. In the frequency domain, the multi-resolution relationship becomes [11]

$$\hat{\psi}_{a,b}(\omega)=\sqrt{a}\cdot\psi_{a,b}(a\omega)\cdot exp(-i\omega b) \tag{2.7}$$

where $\hat{\psi}_{a,b}(\omega)$ is the Fourier transform of the wavelet function $\psi_{a,b}(t)$.

In this section, explicit expressions of the Choi-Williams (Mexican-hat) (Eqs. (2.1)–(2.2)), Morlet wavelets (Eq. (2.3)) and the hyperbolic (Eqs. (2.4) and (2.5)) have been given in both time and frequency domain. It is necessary to examine some important properties of these wavelets by estimating the number of sampling points for the wavelet, aliasing effects, maximum possible scale that can be supported by the wavelet system and the scale resolution. These properties will be studied in detail in Section 2.1.

2.1 Properties of the Mexican-Hat (Choi-Williams), Morlet and Hyperbolic Wavelets

From an engineering point of view, to study properties of a wavelet function, it is important to investigate the scale resolution, maximum scale used in wavelet transform, the sampling of the wavelet and its relation to the time sampling of the time input signal and the aliasing effects. The main reason that sampling of a wavelet function is of concern is that digital signal processing is practical and important. In addition, the input waveform is usually a discrete set of samples from a continuous process. This section examines the above mentioned properties in detail. Firstly, some preliminary parameters of the Choi-Williams, Morlet and hyperbolic wavelets are estimated.

2.1.1 Fundamental Parameters

The Morlet wavelet was studied and used to investigate transition to turbulence in [11] by Jordan, Miksad and Powers in which formulas of the admissibility constant C_ψ , the first moments in time t_0 and frequency domain ω_0, the time variance σ_t and frequency variance σ_ω are given in detail. The parameters of the hyperbolic, Mexican-hat and Morlet wavelets are calculated for $\beta = 0.5$ and given in Table 2.1. For various values of the control parameters σ and β, from simulation, it can be concluded that the Morlet and hyperbolic wavelets satisfy the admissibility condition imposed by Eq. (1.1) by having a very-small area under the curve. The error in this case for both wavelets is always less than 1 part in a million.

Table 2.1: Fundamental parameters of the hyperbolic, Choi-Williams and Morlet wavelets for $\beta = 0.5$

Wavelet	Parameter values, $\beta = 0.5$				
	C_ψ	T_0	σ_t	ω_0	σ_ω
Choi-Williams	1.785	0.0	1.245	1.47	1.08
Morlet	1.58	0.0	0.6656	5.0	2.36
Hyperbolic	0.15	0.0	0.62	0.817	0.213

The larger the values of σ_t and σ_ω are, the less time- and frequency-support the corresponding wavelet has respectively.

2.1.2 Dimensional Expressions and Band-Peak Frequency

It is assumed that the dimensional sampling time interval of the input data series of length M is $\Delta t'$ and the non-dimensional sampling time of the wavelet, whose time base interval is from $-T$ to T, is Δt, where the symbol " ′ "* indicates a dimensional quantity [11]. Let N be the number of samples that should be sampled for the wavelet function. To calculate the non-dimensional time base of the wavelet function, we have to map the sampling time interval of the input waveform to that of the wavelet, i.e., $[-T, T] \leftrightarrow [0, N(\Delta t')]$. The wavelet time base is therefore given by

$$t = \frac{2T}{N(\Delta t')} \cdot t' \tag{2.8}$$

The expression for the non-dimensional frequency f is obtained by taking the inverse of Eq. (2.8) yielding

$$f = \frac{N(\Delta t')}{2T} \cdot f' \text{ or } \omega = \frac{\pi \cdot N(\Delta t')}{T} \cdot \omega' \tag{2.9}$$

The dimensional frequency expressions of the Choi-Williams, Morlet and hyperbolic wavelets are given by Eqs. (2.10), (2.11) and (2.12)

$$\hat{\psi}_{a,b'}^{CW}(f') = \sqrt{a\pi\sigma} \cdot exp\left(-\frac{iN\pi f'(\Delta t')b'}{T}\right)\left(\frac{N\pi f'(\Delta t')}{T}\right)^2 \cdot exp\left[-\frac{\sigma}{4}\cdot\left(\frac{aN\pi f'(\Delta t')}{T}\right)^2\right] \tag{2.10}$$

$$\hat{\psi}_{a,b'}^{Morlet}(f') = \sqrt{\frac{a}{2\pi}} \cdot exp\left(-\frac{iN\pi f'(\Delta t')b'}{T}\right) \cdot exp\left[-\frac{1}{\sigma}\cdot\left(\frac{a\pi f' N(\Delta t')}{T} - \omega_\psi\right)^2\right] \tag{2.11}$$

$$\hat{\psi}_{a,b'}^{Hy}(f') = \frac{\pi\sqrt{a}}{\beta} \cdot exp\left(-\frac{iN\pi f'(\Delta t')b'}{T}\right)\cdot\left(\frac{\pi f' N(\Delta t')}{T}\right)^2 \cdot sech\left(\frac{\pi \cdot a \cdot \pi f' N(\Delta t')}{2\beta \cdot T}\right) \tag{2.12}$$

where typically, $5.0 \leq \omega_\psi \leq 6.0$ rad/s to ensure that the condition imposed by Eq. (1.1) is met. Throughout this paper, $\omega_\psi = 5.0$ rad/s for the Morlet wavelet. The dimensional quantity b' is similarly defined via Eqs. (2.8) and (2.9).

The band-peak frequency, f'_p, is the frequency at which the wavelet filter has the maximum value. To estimate the band-peak frequency, the first derivative of the real-part dimensional frequency expressions of the wavelets is obtained. Since the real parts of the first derivative functions are exponential functions, the second derivative functions are not required. For the Morlet wavelet, to maximise $\hat{\psi}_{a,b'}^{Morlet}(f')$ (given by Eq. (2.11)), the exponent of the exponential term is made to be zero which yields [11]

$$f'_{p(Morlet)} = \frac{T_{Morlet}\,\omega_\psi}{\pi a N_{Morlet}(\Delta t')} = \frac{\omega_\psi}{aC_{Morlet}}, \text{ where } C_{Morlet} = N\pi(\Delta t')/T. \tag{2.13}$$

The band-peak frequency of the Choi-Williams wavelet is similarly obtained as

$$f'_{p(CW)} = \frac{2}{aC_{CW}\sqrt{\sigma}} = \frac{2T}{aN\pi(\Delta t')\sqrt{\sigma}} \tag{2.14}$$

where σ is the kernel control parameter of the Choi-Williams kernel.

The band-peak frequency of the hyperbolic wavelet is similarly given by

* Throughout this paper, some MATLAB graphs use a "*" instead of a " ′ " to indicate dimensional quantities and $\psi_T(f')$ is equivalent to $\hat{\psi}(f')$ for convenience.

$$f'_{p(Hy)} = \frac{4\beta}{a\pi C_{Hy}} = \frac{4\beta T_{Hy}}{aN_{Hy}\pi^2(\Delta t')} \tag{2.15}$$

2.1.3 Aliasing Effects

To avoid aliasing effects, in sampling the wavelet non-dimensionally and in sampling the input time series dimensionally, the Nyquist criterion must be satisfied. The Nyquist frequency of the input time series with the sampling time ($\Delta t'$) can be given by

$$f'_{Ny} = \frac{1}{2(\Delta t')}, \text{ where } (\Delta t') \text{ is the dimensional sampling time of the input series.} \tag{2.16}$$

To avoid aliasing in the mother wavelet, the overlapping fraction α of two adjacent wavelet filters at different scales of a wavelet system must be prescribed so that it is less than a threshold value. This fraction can be defined as an absolute ratio of the magnitude of the wavelet at the frequency $f'_{overlapp}$ at which α is sufficiently small to the magnitude of the wavelet at the band-peak frequency f'_p (Eq. (2.17)). At the time that two adjacent wavelet filters overlap, to recover the input time signal and to avoid aliasing of the wavelet filters, the overlapping frequency must be at least equal to the Nyquist frequency f'_{Ny}, i.e., $f'_{overlapp} = f'_{Ny}$. The mathematical expression of the ratio α is therefore given by Eq. (2.17) and graphical representation of α is seen in Figure 2.1.

$$\alpha = \left| \frac{\hat{\psi}_{a=1,b'}(f'_{Ny})}{\hat{\psi}_{a=1,b'}(f'_p)} \right| \tag{2.17}$$

If α is known beforehand, then it is possible to estimate the number of sampling points for the wavelet system. Jordan, Miksad and Powers [11] calculated the required number of sampling points N_{Morlet} for the Morlet wavelet system for a typical case of $\sigma = 2$. The number of sampling points of the Morlet wavelet function N_{Morlet} for a general value of σ is given by

$$N_{Morlet} = \frac{2T}{\pi}\cdot\left(\omega_\psi + \sqrt{-\sigma\ln\alpha}\right), \text{ where } \omega_\psi = 5.0 \text{ rad/s} \tag{2.18}$$

By using Eq. (2.16) for f'_{Ny}, Eq. (2.15) for f'_p and Eq. (2.12) for the expression of $\hat{\psi}^{Hy}_{a=1,b'}(f')$ we obtain the minimum number of sampling data points N_{Hy} for the hyperbolic wavelet. The number of sampling points N_{Hy} is found by a graphical method by plotting the graphs of two functions f_1 and f_2 given by the following equation

$$f_1 = f_2 \text{ where } f_1 = \ln\left[\frac{\alpha\beta^2 T_{Hy}^2}{5.73a^2}\cdot\left(exp\left(N\cdot\frac{a\pi^2}{2\beta T_{Hy}}\right)+1\right)\right] \text{ and } f_2 = \ln(2) + N\cdot\frac{a\pi^2}{4\beta T_{Hy}} \tag{2.19}$$

Eq. (2.19) yields a good estimate of N_{Hy} and therefore is used throughout this paper. Similarly, by using Eq. (2.10) for the expression of $\hat{\psi}^{CW}_{a=1,b'}(f')$, Eq. (2.16) for f'_{Ny}, Eq. (2.15) for f'_p and after performing mathematical manipulations, the sufficient number of sampling points N_{CW} for the Choi-Williams wavelet is given by the relation

$$0.617\sigma\cdot\left(\frac{N_{CW}}{T_{CW}}\right)^2 = \ln\left[\frac{1.67\sigma}{\alpha}\cdot\left(\frac{N_{CW}}{T_{CW}}\right)^2\right] \tag{2.20}$$

2.1.4 Scale Limit

The maximum possible scale using in a wavelet system is determined based on the number of wrapped-around points or end-points since these points do not provide useful information. It has been observed that the number of end points is proportional to the scale a [11]. That means if the scale increases to a certain value, the number of end points will dominate the estimated wavelet transform coefficients.

From [11], the number of wrap-around points at one end is a function of the scale a and can be given approximately by

$$N_{wrap}(a) \approx \frac{a(N-1)}{2} \tag{2.21}$$

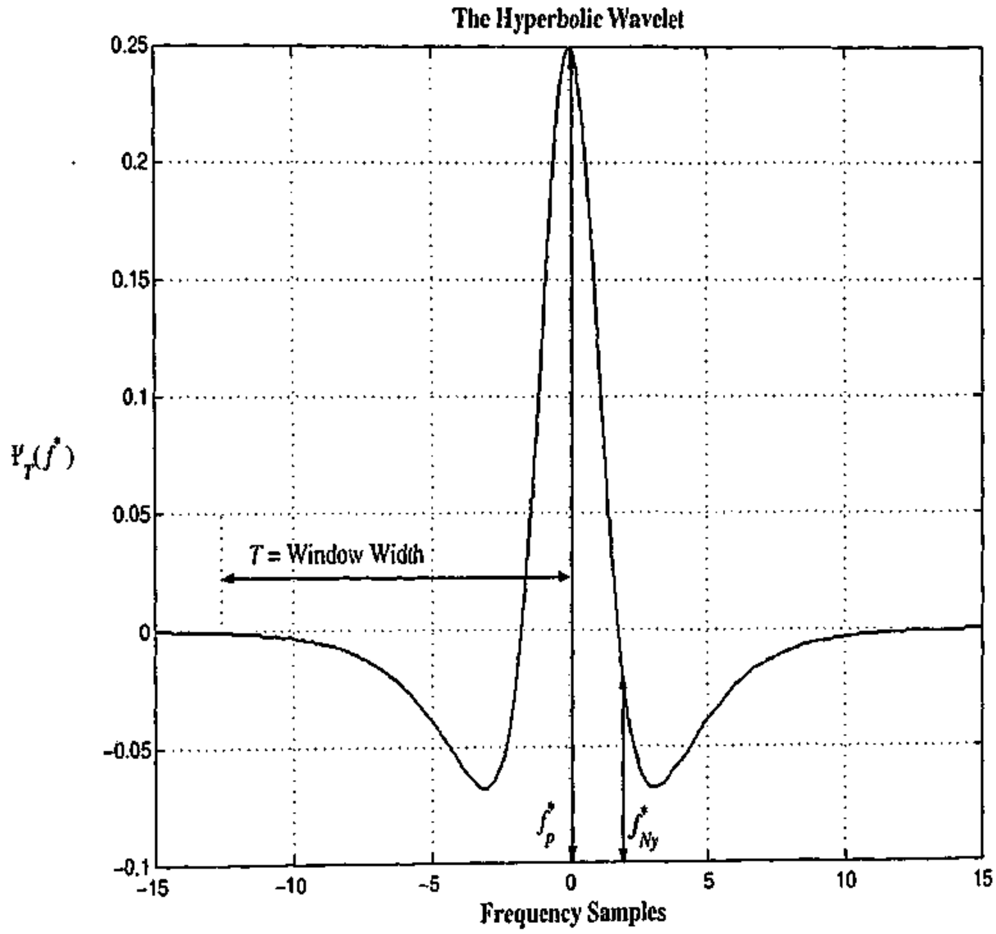

Figure 2.1: Graphical representation of α and window width T of the hyperbolic wavelet for $\beta = 0.5$

To estimate the largest scale of a wavelet system, introduce η as the fraction that the number of wavelet coefficients being affected by the number of wrap-around points N_{wrap} and $M = 2^m$ as the number of input data points into the wavelet system. Then

$$\frac{a_{\max}(N-1)/2}{2^m} \le \frac{\eta}{2}, \tag{2.22}$$

where N is the number of sampling points of the wavelets which was discussed in Section 2.1.3.

To speed up the calculation process of the wavelet transform coefficients, M should be a power of 2. The fraction $\eta = \frac{1}{3}$ was used by Jordan, Miksad and Powers for the estimation of the largest scale. For both ends and from Eq. (2.22), the maximum scale a_{max} is given by

$$a_{\max} \le \frac{2^m \cdot \eta}{N-1} \tag{2.23}$$

The number of input sampling points M can be estimated from the maximum scale a_{max} using Eq. (2.24)

$$M = 2^m \ge \frac{(N-1)\cdot a_{\max}}{\eta} \quad \text{or} \quad m \ge 1.443 \cdot \ln\left(\frac{(N-1)\cdot a_{\max}}{\eta}\right) \tag{2.24}$$

where η is a ratio of the number of wrap around points at the largest scale to the total number of points in the time series.

The maximum scale is inversely proportional to N, the number of sampling points of a wavelet function, as seen from Eq. (2.23). The next section calculates the scale resolution.

2.1.5 Scale Resolution

The scale resolution constant ω_d is defined as the distance between two band-peak frequencies of the two adjacent wavelet filters [11]. The finer the scale resolution ω_d is, the smaller the resolution constant. This distance between two adjacent band-peak frequencies can be determined by specifying a decay constant λ which has the mathematical form given by Eq. (2.25)

$$\lambda = \frac{\hat{\psi}(a\omega_p)}{\hat{\psi}(a\omega_p + \omega_d)} \tag{2.25}$$

where ω_d is the scale resolution constant and $\hat{\psi}(\omega)$ is the frequency expression of the wavelet function given by Eqs. (2.2), (2.3) and (2.5) for the Morlet, Choi-Williams and hyperbolic wavelets respectively. In most practical systems, the scale resolution constant must be small to capture rapid changes in the energy density of the input waveform, which is usually non-stationary in cases of turbulence and chaos [11], ECG [20], music signal [21-23] or random processes [24]. It is important to note that in Eq. (2.25), the frequency quantities are non-dimensional, thus appropriate conversion of the variables must be used to obtain the correct answer.

As the scale a increases, the scale resolution constant decreases since the frequency in a wavelet system is inversely proportional to the scale [1, 3]. If j is the index of an instant scale that is going to be used in a wavelet system, then we have the following relationship

$$\omega_{p(j+1)} - \omega_{p(j)} = -\frac{\omega_d}{a_j} \tag{2.26}$$

where a_j is the j^{th} scale of the wavelet system and $\omega_{p(j+1)}$ is the band-peak frequency at the $(j+1)^{th}$ scale.

The scale resolution of the Morlet can be obtained analytically. For the Choi-Williams and hyperbolic wavelets, the approximate scale resolutions are estimated by eliminating the third- and higher-order terms in the time series of $\ln(1-x)$, where x is a function of the scale resolution ω_d. The main reason that the third-order terms are ignored is that the scale resolution ω_d is expected to be less than 1. In addition, for these two particular wavelets, the third-order constants are quite small that they can be safely ignored without making large differences in value of the final answer. For the Morlet wavelet function, the *exact* scale resolution constant ω_d^{Morlet} is found to be

$$\omega_d^{Morlet} = \sqrt{-\sigma \cdot \ln \lambda}\,, \text{ where } \lambda < 1 \tag{2.27}$$

The approximate scale resolution constant of the hyperbolic wavelet system ω_d^{Hy} is given by Eq. (2.28)

$$\omega_d^{Hy} \approx \frac{4\beta\cdot\sqrt{-\ln\lambda}}{\pi} = \frac{4\sqrt{-\ln\lambda}}{\pi\sigma}\text{, where } \lambda < 1 \tag{2.28}$$

The approximate scale resolution of the Choi-Williams or Mexican-hat wavelet system ω_d^{CW} is given by

$$\omega_d^{CW} \approx \sqrt{-\frac{2\cdot\ln\lambda}{\sigma}} = \sqrt{-2\beta\cdot\ln\lambda}\text{, where } \lambda < 1 \tag{2.29}$$

From Eqs. (2.27)–(2.29), it is evident that the scale resolution ω_d of the three wavelet systems are independent of the sampling interval (Δt) which makes the wavelets unique. These equations are obtained analytically or with practical approximations. The following table lists values of the scale resolution of the three wavelets for different values of β

Table 2.2: The scale resolutions of the Morlet, hyperbolic and CW wavelets for $\beta = 0.5$, 1 and 2

β	ω_d^{Morlet}	ω_d^{Hy}	ω_d^{CW}
0.5	0.459	**0.2066**	0.3246
1	0.3246	**0.4133**	0.459
2	0.459	**0.8266**	0.649

Thus, over the typical range of the hyperbolic control parameter $0.5 < \beta < 2$, the hyperbolic wavelet appears to have a fine scale resolution constant as shown. For $\beta = 2$, the scale resolution is four times larger than for the case of $\beta = 0.5$ which suggests that for $\beta \geq 2$, the hyperbolic wavevlet will have coarse scale resolution.

To obtain the largest number of scales that can be utilised in a wavelet system (provided that the scale resolution is known), it is convenient to take the first band-peak frequency to be the reference frequency. The subsequent band-peak frequencies are obtained by dividing the reference band-peak frequency by the scale that corresponds to the particular band-peak frequency, i.e., $\omega_{(p)j} = \omega_{(p)1}/a_j$. Using this relation and Eq. (2.26) one can obtain [11]

$$\frac{\omega_{(p)1}}{a_{j+1}} - \frac{\omega_{(p)1}}{a_j} = -\frac{\omega_d}{a_j} \tag{2.30}$$

The minus sign on the right hand side of Eq. (2.30) is employed to ensure that the total number of scales j_{max} is a positive number (Eq. (2.33) and Section 2.1.6) without affecting the correctness of the equation.

The recursive relationship of the scale a is then given by

$$a_{j+1} = \left(\frac{\omega_{(p)1}}{\omega_{(p)1} + \omega_d}\right)a_j = \kappa a_j\text{, where } \kappa = \frac{\omega_{(p)1}}{\omega_{(p)1} - \omega_d} \tag{2.31}$$

The first band-peak frequencies (corresponding to $a = 1$ for the mother wavelet) of the Morlet, Choi-Williams and hyperbolic wavelets can be estimated by using Eqs. (2.13)–(2.15) respectively. From Eq. (2.31), it is evident that previous scales are dependent on the present scale. This relationship can be understood via the constant κ, which is a function of the peak frequency of the first scale wavelet (mother wavelet) and the scale resolution constant ω_d. As the scale a becomes larger, the width of the corresponding wavelet becomes smaller and smaller. The wavelet width as a function of the scale a is a function of the mother wavelet peak frequency $\omega_{p(1)}$ and the scale resolution constant ω_d. Assuming that $a_1 = 1$ ($j = 1$ as the starting point), Eq. (2.31) can be rewritten to find the number of scales that are required in a wavelet system under certain conditions [11]

$$a_j \approx \kappa^{j-1}, \text{ where } \kappa \text{ was defined by Eq. (2.31).} \tag{2.32}$$

From Eq. (2.32), one can obtain an expression for the total number of required scales j_{max}

$$j_{max} = \frac{\ln(a_{j\max})}{\ln \kappa} + 1 \tag{2.33}$$

By using the maximum value of a_{jmax} given by Eq. (2.23), the total number of scales j_{max} required for a wavelet system can be obtained. For each wavelet system, the number of sampling points of the mother wavelet is different and so are the band-peak frequency, resolution, a_{max} and the total number of required scales. To gain more practical insight into the three wavelet systems, Section 2.1.6 provides one typical example in which the necessary parameters and important properties (which have been estimated and discussed throughout Sections 2.1.1–2.1.5) of a wavelet system under some special conditions are estimated.

2.1.6 Remarks and One Typical Example

One practical example was used in [11] in which the transition to turbulence in a subsonic wake was investigated using the Morlet wavelet transform. The major conclusions about the behaviour of the subsonic wake were made in [11] and will not be repeated here. This section compares the three wavelet systems namely Morlet, Choi-Williams and hyperbolic in terms of band-peak frequency, maximum scales, aliasing, resolution and the total number of scales used in this particular application. For the three wavelets Morlet, Choi-Williams and hyperbolic, the value of $\beta = 1/\sigma = 0.5$ is used throughout this section.

The sampling interval of the input time series was $(\Delta t') = 0.2\ ms$. The aliasing parameter is chosen to be $\alpha = 0.01$ (1%) so that only 1% of the mother wavelet is overlapped. From Table 2.1, the one-sided length of the hyperbolic, Choi-Williams and Morlet mother wavelets are $T_{Hy} \approx 10$, $T_{CW} \approx 5$ and $T_{Morlet} \approx 3$ respectively. The values of the required number of sampling points of the mother wavelets are hence found by using Eqs. (2.18)–(2.20). From Eqs. (2.19) and (2.20), the approximate number of sampling points of the hyperbolic and Choi-Williams wavelets are $N_{Hy} = 9$ and $N_{CW} = 13$ respectively.

The band-peak frequencies are obtained by employing Eqs. (2.13)–(2.15) for the Morlet, Choi-Williams and hyperbolic wavelets respectively. Since the band-peak frequency can be down to about 30 Hz [11], the maximum scales of each wavelet can be found. From Eqs. (2.24), the required number of data points in each wavelet system with $\eta = \frac{1}{3}$ can be calculated. For $\beta = 0.5$ and $\sigma = 2$, the scale resolution of each wavelet system is estimated next using Eqs. (2.27)–(2.29). It should be noted that the number of input sampling points can be varied by changing the value of η to provide satisfactory solutions to a particular application or problem (Eq. (2.24)). However, η should be kept small so that aliasing effects can be avoided effectively.

The total number of scales that can be computed is directly proportional to the scale resolution ω_d. By employing Eq. (2.33) the total number of scales of each wavelet system can be worked out. Table 2.3 summarises values of important parameters for the hyperbolic, Choi-Williams and Morlet wavelets that have been estimated in this section.

Table 2.3: Parameter comparison of the hyperbolic, Choi-Williams (Mexican-hat) and Morlet wavelets for $\beta = 1/\sigma = 0.5$

Wavelet	T	N	a_{max}	ω_d	j_{max}
Morlet	3	17	49	0.459	42
Choi-Williams	5	13	29	0.3246	14
Hyperbolic	10	9	38	0.2066	11

From Table 2.3, the hyperbolic wavelet appears to have the finest scale resolution, however, the total number of scales j_{max} of the hyperbolic wavelet is smaller compared to those of the Mexican-hat and

Morlet wavelets. This suggests that the hyperbolic wavelet is most suitable for signals which do not have a wide frequency spectrum to resolve. In addition, with a fine scale resolution, the hyperbolic wavelet can be used as an instant energy-monitoring tool over a number of short time intervals. This feature is most suitable for non-stationary signals whose energy intensity changes rapidly with time. Figure 2.2 and Figure 2.3 illustrate these points using the speech signal of the vowel "e".

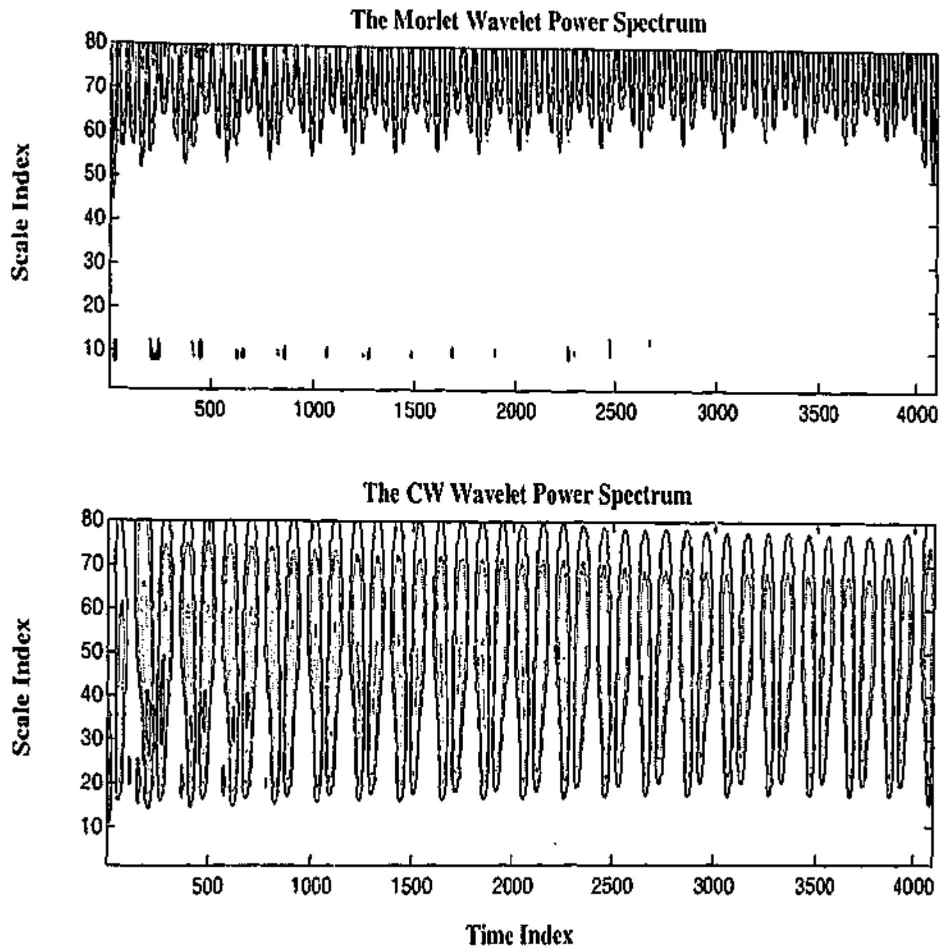

Figure 2.2: Contour plot of the Morlet and Choi-Williams (Mexican-hat) wavelet power spectra

As can be seen, the Morlet wavelet does not reveal energy components in the scale range of 20 to 40, whereas the hyperbolic and CW wavelets do. Moreover, the hyperbolic wavelet can monitor components at very high frequencies which correspond to scales as low as 10. The Choi-Williams wavelet can be considered as a mixture of the Morlet and hyperbolic wavelets. From Figure 2.3, the hyperbolic wavelet power spectrum is different from those seen in Figure 2.2 in which a certain degree of slope is present. This suggests that the hyperbolic wavelet cannot be used to examine signals that have broad spectra (as explained earlier) since part of the energy pattern is masked by the slope. This is another trade-off of the hyperbolic wavelet in having a fine scale resolution and a small total number of scales. The Choi-Williams wavelet seems to be the most suitable wavelet in this case. However, the hyperbolic wavelet does reveal the energy pattern continuously at almost every scale from 10 to 80, whereas the Morlet (only for scales greater than 40) and CW (for scales greater than 20) wavelets do not. Thus, it can be suggested that having a fine scale resolution makes the wavelet more effective in revealing the energy pattern of the input signal but simultaneously limits the ability in displaying the full energy pattern as illustrated in Figure 2.2 and Figure 2.3 due to a non-90^0 slope. Therefore, it is not always advantageous to have a fine scale resolution and a small total number of scales. Instead, there should be a balance between the scale resolution and the total number of scales as for the Choi-Williams wavelet.

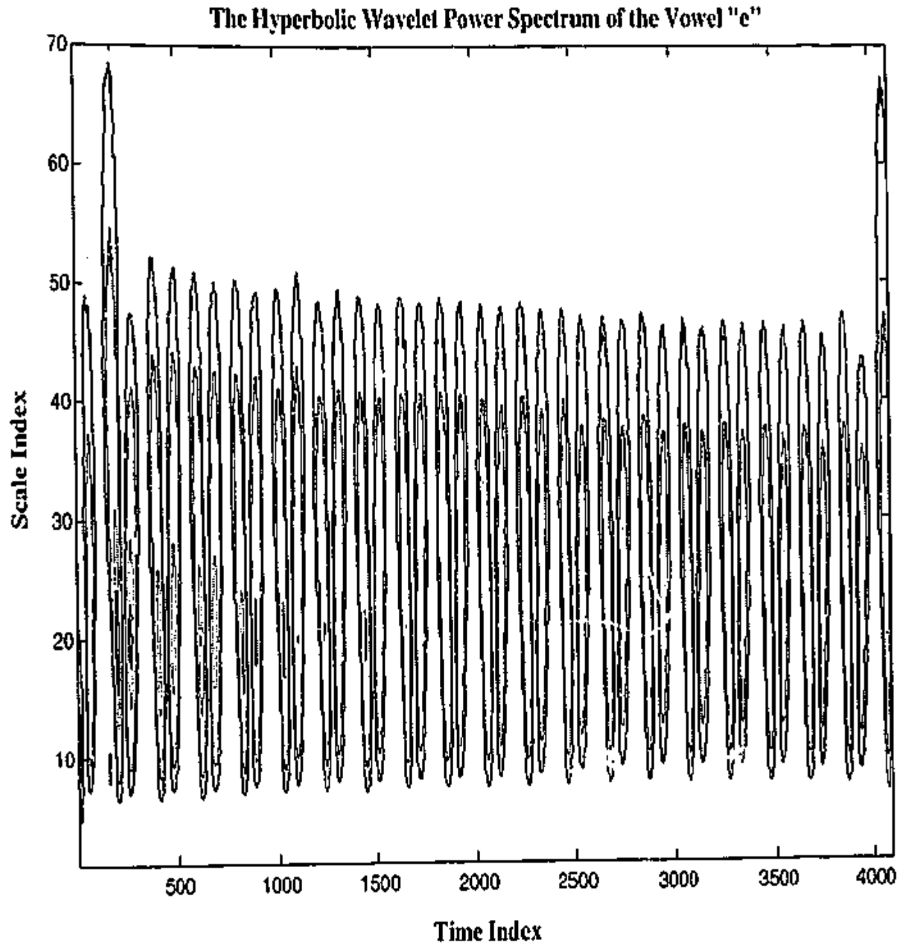

Figure 2.3: Contour plot of the hyperbolic wavelet power spectrum

2.2 Conclusion

The hyperbolic, Choi-Williams and Morlet wavelets have been compared in detail in terms of scale resolution, scale limit and aliasing effects. The hyperbolic wavelet appears to have the finest scale resolution for well-chosen values of β. The Choi-Williams and Morlet wavelets appears to be suitable for broad-spectrum signals, whilst the hyperbolic wavelet is applicable for signals whose spectra vary rapidly with time, i.e., non-stationary signals. By having a fine scale resolution, the hyperbolic wavelet might be run in parallel on independent processors to monitor the energy intensity of different segments of a discrete input signal. This is one major advantage of the hyperbolic over the Morlet and Choi-Williams wavelets. The major disadvantage of the hyperbolic wavelet is that its total number of scales is small which might be difficult when examining broad-spectrum signals. Moreover, there exists a non-90^0 slope which limits the effectiveness of the hyperbolic wavelet in displaying energy patterns. Higher-order hyperbolic wavelets can be investigated in future work so that more useful applications can be found.

References

[1] C. S. Burrus, R. A. Gopinath, and H. Guo, *Introduction To Wavelets and Wavelet Transforms: A Primer*, pp. 10-50, New Jersey: Prentice-Hall, Inc., 1998.

[2] Hyung-ILL Choi and W. J. Williams, "Improved Time-Frequency Representation of Multicomponent Signals Using Exponential Kernels," *IEEE Transactions on Acoustics, Speech and Signal Processing*, June 1989, **37**, No. 6, pp. 862-871.

[3] I. Daubechies, *Ten Lectures on Wavelets*, pp. 1-16, Philadelphia: The Society for Industrial and Applied Mathematics, 1992.

[4] A. Grossmann, R. Kronland-Martinet, and J. Morlet, "Reading and Understanding Continuous Wavelet Transforms," in *Wavelets: Time Frequency Methods and Phase Space*, J.M. Combes, A. Grossmann, and P. Tchamitchian Eds., Springer-Verlag, pp. 2-20, 1989.

[5] P. G. Lemarie-Rieusset, "More Regular Wavelets," *Applied and Computational Harmonic Analysis*, 1998, **5**, No. 1, pp. 92-105.

[6] O. Rioul and M. Vetterli, "Wavelets and Signal Processing," *IEEE Signal Processing Magazine*, Oct. 1991, **8**, No. 4, pp. 14-38.

[7] G. Strang, "Wavelets," *American Scientist*, June 1994, **82**, No. 3, pp. 250-255.

[8] G. Strang, "Wavelets and Dilation Equations: A Brief Introduction," *SIAM Review*, Dec. 1989, 31, No. 4, pp. 614-627.

[9] G Strang and T. Nguyen, *Wavelets and Filter Banks*, pp. 200-270, MA: Wellesley-Cambridge Press, 1997.

[10] G. Strang and V. Strela, "Orthogonal Multiwavelets With Vanishing Moments," *Optical Engineering*, July 1994, **33**, No. 7, pp. 2104-2107.

[11] D. Jordan and R. W. Miksad, "Implementation of the Continuous Wavelet Transform for Digital Time Series Analysis," *Review of Scientific Instruments*, March 1997, **68**, No. 3, pp. 1484-1494.

[12] Shyh-Jier Huang and Cheng-Tao Hsieh, "Application of Morlet Wavelets to Supervise Power System Disturbances," *IEEE Transactions on Power Delivery*, Jan. 1999, **14**, No. 1, pp. 235-243.

[13] Shyh-Jier Huang and Cheng-Tao Hsieh, "High-Impedance Fault Detection Utilising a Morlet Wavelet Transform Approach," *IEEE Transactions on Power Delivery*, Oct. 1999, **14**, No. 4, pp. 1401-1410.

[14] M. Misiti, Y. Misiti, G. Oppenheim, and Jean-Michel Poggi, *Wavelet Toolbox for Use with MATLAB—A User's Guide*, pp. 6.2-6.114, MA: The MathWorks, Inc., Chapter 6, 1997.

[15] B. Jawerth and W. Sweldens, "An Overview of Wavelet Based Multiresolution Analyses," *SIAM Review*, Sep. 1994, **36**, No. 3, pp. 377-412.

[16] C. E. Heil and D. F. Walnut, "Continuous and Discrete Wavelet Transforms," *SIAM Review*, Dec. 1989, **31**, No. 4, pp. 628-666.

[17] M. Holschneider, *Wavelets: An Analysis Tool*, pp. 17-53, 271-341, Oxford: Clarendon Press, 1995.

[18] P. G. Lemarie-Rieusset, "On the Existence of Compactly Supported Dual Wavelets," *Applied and Computational Harmonic Analysis*, 1997, **3**, No. 1, pp. 117-118.

[19] G. Strang, "The Optimal Coefficients in Daubechies Wavelets," *Physica D*, 1992, **60**, No. 1-4, pp. 239-244.

[20] B. Celler, G. Y. C. Chung, and C. Phillips, "ECG Analysis and Processing Using Wavelets and Other Methods," *Biomedical Eneineering—Applications, Basis and Communications*, April 1997, **9**, No. 2, pp. 81-90.

[21] Ian Kaminskyj and A. Materka, "Automatic Source Identification of Monophonic Musical Instrument Sounds," *IEEE International Conference on Neural Networks*, Perth, Western Australia, 1995, pp. 189-194.

[22] Ian Kaminskyj and A. Materka, "Automatic Monophonic Musical Instrument Sound Identification System," *Proc. ACMA Conference*, Brisbane, Australia, 1996, pp. 41-46.

[23] Ian Kaminskyj, "Multidimensional Scaling Analysis of Musical Instrument Sounds' Spectra," *Proc. ACMA Conference*, Wellington, NewZealand, 1999, pp. 36-39.

[24] Moness G. Amin, "Time-Frequency Spectrum Analysis and Estimation for Non-Stationary Random Processes," in *Time-Frequency Signal Analysis: Methods and Applications*, B. Boashash Eds., Longman Cheshire, pp. 208-232, 1992.

PARALLEL COMPUTATION OF THE TIME-FREQUENCY POWER SPECTRUM: ANALYSIS AND COMPARISON TO THE BISPECTRUM

Khoa N. Le, Gregory K. Egan and Kishor P. Dabke
Department of Electrical and Computer Systems Engineering,
Monash University, Clayton campus, Melbourne, Australia
Email: {Khoa.Le, Greg.Egan, Kishor.Dabke}@eng.monash.edu.au

Abstract

Experiments of large data sets are computationally expensive. Signal processing analysis on a single CPU leads to unacceptably long execution times. The paper presents initial experiments on calculating the time-frequency power spectrum using the coarse-grained parallel programming technique. Experimental speedup factors are given and discussed. The measured speedup factor of the time-frequency power spectrum parallel calculation process is sub-linear which indicates that the time-frequency power spectrum is a suitable application for parallel programming. The parallel efficiency is acceptable with the lowest value of 75.1% occurring at $N = 10$. The maximum speedup factor of 9.1 is obtained when $N = 12$ at 75.3% of efficiency.

1.1 Introduction

Time-frequency power spectrum is a useful tool to study non-stationary signals in detection and signal behaviour. In practice, non-stationary signals are often encountered in speech, sound, under-water signal analysis, non-linear response, random processes, ECG, complex exponential signals, chirped signals and so on [1, 2]. For wide-sense stationary signals, the bispectrum has been extensively used to extract vital information about the signal. Therefore, the signal behaviour and characteristics can be viewed systematically [3-9]. Since Fourier transform is the basis of the bispectrum, for non-stationary signals, it is not accurate to use the Fourier transform (due to its infinite time support) but a joint time-frequency distribution must be developed to capture the changes in energy density with frequency of these signals. Time-frequency power spectrum is one such tool which was proposed and studied by Rihaczek [10], Cohen [1, 11-13] and others. The general formula for the time-frequency power spectrum is [2, 12]

$$P(t,\omega)=\frac{1}{4\pi^2}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\underbrace{\left[e^{-j\theta(t-u)}\cdot\Phi(\theta,\tau)\right]}_{W(t-u,\tau)}\cdot e^{-j\tau\omega}\cdot R_{t,1}(t,\tau)\,du\,d\tau\,d\theta \tag{1.1.1}$$

where $x(n)$ is the discrete-time input signal, $\Phi(\theta, \tau)$ is the kernel function, $W(t-u, \tau)$ is the weighting function of the kernel $\Phi(\theta, \tau)$ and $R_{t,1}(t,\tau)=x\left(u+\frac{\tau}{2}\right)\cdot x^*\left(u-\frac{\tau}{2}\right)$ is the local auto-correlation function, where $u = t+\frac{\tau}{2}$.

The discrete version of Eq. (1.1.1) is given by Eq. (1.1.2)

$$DTF(n,k) = 2\sum_{\tau=-M}^{M} e^{-j2\pi k\tau/M} \sum_{\mu=-M}^{M} W(\mu,\tau)\cdot f(n+\mu+\tau)\cdot f^{*}(n+\mu-\tau) \tag{1.1.2}$$

where $W(\mu, \tau)$ is the weighting function as defined by Eq. (1.1.1), M is the number of input samples and $f(\cdot)$ is the discrete-time input signal. The weighting function is an arbitrary function that satisfies a number of admissibility constraints [14].

The discrete version of the time-frequency power spectrum is used to construct the serial program (and hence the parallel program) to run on a parallel machine. As can be seen in Eq. (1.1.2), there are four running variables of n, k, τ and μ which result in four nested `for` loops in the program. Moreover, there are conditional executions that are employed to evaluate the local auto-correlation function.

Although it is a very useful tool in analysing non-stationary signals, the time-frequency power spectrum is complicated and expensive to compute. This is due to the fact that instant values of the auto-correlation function must be calculated so that rapid changes of the input signal can be viewed at any time instant. The smaller the auto-correlation time window is, the more expensive the calculation. This is the main reason that time-frequency power spectrum has not been popular or widely used compared to the Fourier transform, the power spectrum and the bispectrum. One solution to the above problem is to reduce the calculation process time of the time-frequency power spectrum by using a parallel computer.

The wavelet transform, which is another form of the time-frequency power spectrum, has received considerable attention in recent years with proposals on applying parallel computing to compute the wavelet transform [15-21]. Some authors have reported on methods of developing ways to compute the time-frequency power spectrum concurrently [22, 23]. This indicates that time-frequency power spectrum could be a suitable candidate for parallel computing. Even though there are proposals to calculate time-frequency tasks in parallel, no hands-on experiments have been performed and reported in the literature. This motivates the work reported in this paper.

In this paper, a MIMD 12-processor Silicon Graphics Power Challenge parallel machine is utilised to perform experiments in estimating the time-frequency power spectrum. From this, recommendations can be made whether the time-frequency power spectrum is suitable for parallel computing. The main aim of this paper is to show that the time-frequency power spectrum is a suitable application for parallel programming as is the case for the bispectrum. Moreover, this paper hopes to make the time-frequency power spectrum a more popular signal-processing tool by reducing its calculation time using a parallel computer. The basic background of this paper is the time-frequency power spectrum and the bispectrum with some knowledge on parallel computing being employed as a new exploring tool to do the calculations efficiently.

The coarse-grained parallel technique is used for the semi-automatic and the fine-grained parallel technique is employed for the full-automatic PCA method as was applied in the case of the bispectrum calculation process. In this paper, however, only the semi-automatic parallel method is employed due to inefficient of coding of the C annotator (full-automatic PCA method). An examination of the parallel code showed that the compiler mistook the inner loops as the most efficient loop for parallel programming. In other words, it employed the fine-grained parallel method inefficiently which resulted in large parallel overhead and thus lowered the speedup factor. For the semi-automatic parallel method, the structure of the parallel program is constructed manually so that the coarse-grained parallel method can be employed more effectively to give better speedup factors and parallel efficiency.

The key factors that affect the performance of parallel programs are parallel load (the load among the processors should be evenly balanced), parallel overhead (the amount of communication among the processors should be minimised) and data dependency in the parallel loop. There are four nested `for` loops in the program which require a large amount of computation. If the number of input samples (or loop size) is greater than 256 ($M \geq 256$), the number of iterations of the program is in the order of 10^9 or even larger. The coarse-grained parallel technique is utilised in the outer-most loop of the program by dividing it into smaller tasks. Each small task has 3 nested `for` loops, which can be executed independently by independent processors (CPU) of the systems thus preventing data dependency among the processors. Moreover, in using the coarse-grained parallel technique, the parallel overhead is minimised since the number of parallel loops is only one in this case. Since the main parallel loop and the smaller loops inside it are identical, the load division among the processors is equal which maximises the efficiency of the system.

1.2 Experimental Results

The parallel speedup factor and parallel efficiency [24, 25] are estimated by using Eq. (1.2.1)

$$\text{Speedup\_Factor} = \frac{\text{Sequential\_Time}}{\text{Parallel\_Time}}, \qquad (1.2.1)$$

$$\text{Parallel\_Efficiency} = \frac{\text{Measured\_Speedup}}{\text{Ideal\_Speedup}}$$

The parallel speedup factors of the time-frequency power spectral calculation and bispectrum processes are shown in Figure 1.2.1 which also shows the ideal speedup factor. The parallel efficiency of the time-frequency power spectrum calculation is displayed in Figure 1.2.2.

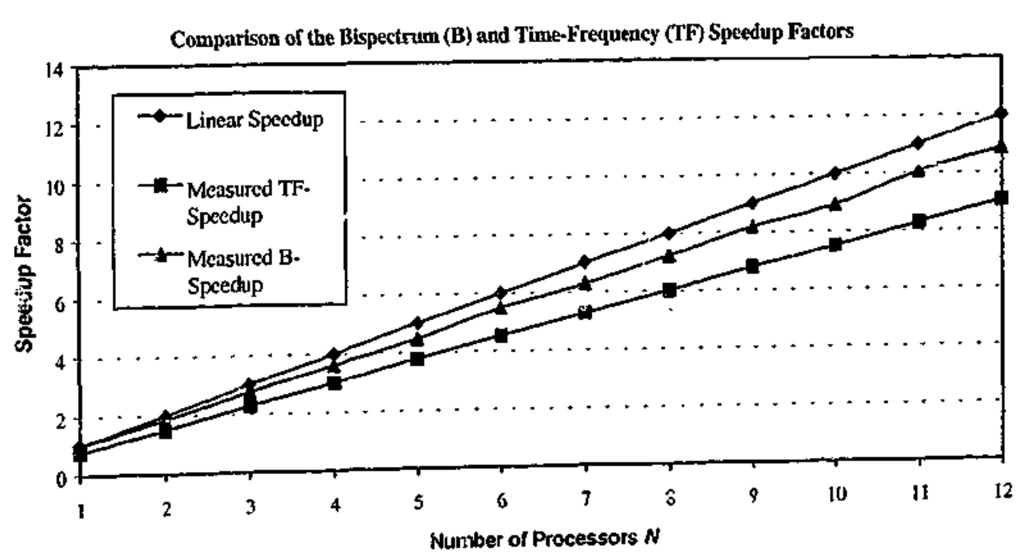

Figure 1.2.1: Measured speedup factor of the time-frequency power spectrum (the loop size is 256 samples) and bispectrum calculation processes