DIGITAL SIGNAL PROCESSING AND DISPLAY OF LUNG SOUNDS

Pasika

DIGITAL SIGNAL PROCESSING AND DISPLAY OF LUNG SOUNDS

by

Hugh Joseph Christopher Pasika

A Thesis
Submitted to the School of Graduate Studies
in Partial Fulfillment of the Requirements
for the Degree
Master of Engineering
McMaster University
April 1993

Master of Engineering (1993) (Electrical Engineering)		McMaster University Hamilton, Ontario, Canada
Title	:	Digital Signal Processing and Display of Lung Sounds
Author	:	Hugh Joseph Christopher Pasika, B. Eng McMaster University
Supervisor	:	Dr. L. David Pengelly
Number of Pages	:	viii; 134

Abstract

Presented here is an examination of the issues surrounding the analysis of lung sounds and their display. The project is aimed at providing a visual representation of the information that a physician gleans from auscultation of the lungs. Such a tool would be of benefit to those who are hearing impaired and also in teaching auscultation. A second goal is to provide a tool that will allow the examination and quantification of lung sounds thus permitting linkage between the acoustic events and their physical causes.

The project is divided into two tasks. The first is the isolation of the wheezes and crackles; the second is their display.

The isolation problem is difficult due to the variance in the frequency characteristics of the sounds; wheezes may appear anywhere in a two thousand hertz band and crackles also display a varying spectrum. The difficulty in separation is further compounded by the spectral overlap of the two. These problems preclude any 'simple' filter solution. In order to separate the sounds, filtering methods based on exploiting the statistical differences namely the stationarity of the wheeze and non-stationarity of the crackle are utilized. Of the several methods attempted, the most promising was the Adaptive Line Enhancement process when driven by the Least Mean Squares adaptive

algorithm.

An important criteria for being able to display the sounds was to access their temporal information. Accomplishing this with the standard short time Fourier transform precludes adequate resolution to identify the frequency characteristics of crackles. Display of the crackle information was facilitated by the use of high resolution time-frequency methods based on Cohen's Class of time-frequency representations. These methods are able to simultaneously provide high time and frequency resolution. A method for automatic adjustment of the parameters involved in the process was developed in order to yield the best display possible.

Dedication

To my parents:

Their love and support have enabled me to realize many goals and dream of many more for the future.

Acknowledgements

This work would not have been possible without the willingness and guidance of Dr. Pengelly. For these I am grateful.

I would like to thank Dr. T. Bhattacharya, Dr. G. Jones, Norm Nantel, John Smerek and Andrew Ukrainec who in the true spirit of collegiality were always willing to lend assistance and provide a humbling injection of ideas.

In my years at McMaster, I have met many. Several have had a profound affect on my psyche and educated me in ways other than academic. I present their initials here so that late in my life, should our paths never cross again, I can still remember them, the good times shared and the lessons they have taught me.

Chronologically they are: MC, RC, SB, JP, SH, JO, CC, CM, PM - Thank you

Table of Contents

Section 1 - Background and Purpose

1.0 Introduction

2.0 Lung Sounds

3.0 Heart Sounds

4.0 Literature

5.0 Rationale

1.0 Introduction

The objective of this thesis is to enable viewing of the auscultative information given by lungs. The end product is not intended as a 'black box' type of analyzer that produces a diagnosis but rather as a mechanism that will better aid the physician in forming a diagnosis. Currently, this information is used much the same as phonocardiography (heart sounds) in only assessing the functional state of the organ (Des90, Ran88). This thesis tackles the creation of a system to separate and quantify lung sounds with the hope that further study will allow correlation between the physical processes of lung sound generation and the acoustic parameters of spectral and temporal characteristics.

Human interpretation of lung sounds is an imprecise procedure. Large variability can exist in the manner in which two observers perceive the same sound. Perhaps most notably, age plays a large role in determining one's ability to hear. The problem of individual interpretation is compounded by the physical characteristics of the hearing process which are ill suited to auscultation.

Human hearing mechanisms are unable to determine a split of less than 20 milliseconds between two sounds (Fei71). In addition, logarithmic compression of sound intensity produces a scale where an acoustic event must contain ten times the energy of another in order to be perceived as twice as loud (Guy87). While this property allows the human ear an extremely large dynamic range, it can make subtle changes between sounds difficult to appreciate. As well, the ear's frequency sensitivity follows a logarithmic scale in the 0-1000 hertz range. This results in sounds with a higher pitch

being perceived as louder than a sound with the same energy but at a lower frequency. Since lung sounds occur between 60 and 2000 Hz (Fle90) and heart sounds between 20 and 600 Hz (Ran88) this factor is extremely important in auscultation. Clearly, the human ear is not the best mechanism for discriminating in auscultation.

Technological revolution has brought many marvellous methods to the clinical environment. Computing power has enabled methods such as magnetic resonance imaging (MRI) and computed tomography (CT) to becomes staples of modern medicine. Applied herein to the analysis of lung sounds is the strength of digital signal processing that has enabled these other techniques to flourish. The overall goal is to provide methods that will move lung sound analysis out of the realm of descriptive terms such as noisy, moist, quiet, dry, clicking, bubbly and slurpy (Des90) and into the arena where it is possible to quantify and classify the sounds by their physical characteristics.

In this thesis, lung sound analysis is presented as a two-fold problem. The first to overcome is the separation of wheeze and crackle information. The second is to display the information in a meaningful manner that will support quantification of the sounds and provide a method for better understanding of the factors affecting lung sound genesis.

2.0 Lung Sounds

The topic of lung sounds seldom occupies more than a handful of pages in any medical text. Confusion and controversy surround their genesis, utility and even

nomenclature. The following text is a reflection of the current status of lung sound understanding and application in the clinical setting.

2.1 Physiology of the Respiratory System

On its way to the lungs, air must first pass through the trachea. Reinforcing hyaline cartilage rings prevent collapse of this tube when the intra-tracheal pressure falls below the external pressure (Gra89). The trachea bifurcates into the bronchi; one bronchus entering each lung. Bronchi further divide producing progressively smaller tubes with fewer and fewer cartilage rings. These tubes are named bronchioles when their diameter reaches the millimetre range.

Subdivision continues until up to 23 levels of branching from the trachea have occurred. The pathways end with an alveolar duct that opens into alveolar sacs lined with bubble-like alveoli. These alveoli are covered by a capillary network through which gas exchange takes place. Because of the need to facilitate diffusion of gases, the alveolar and capillary walls are extremely thin. Pulmonary surfactant is necessary for the prevention of alveolar collapse resulting from the surface tension of the thin alveoli. This surfactant is secreted from within the alveoli.

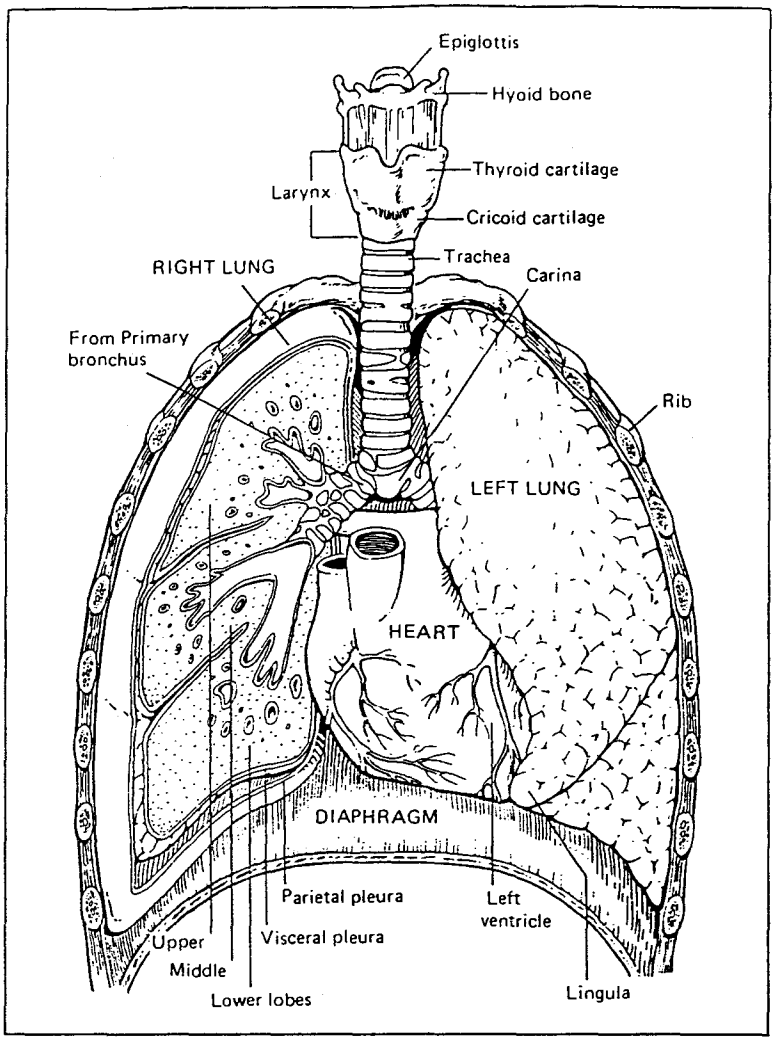
Each lung is covered by a double layer of lubricated membrane. These pleural membranes provide a frictionless environment for the healthy lung as it expands and contracts during respiration. The inner membrane, the visceral pleura, is in contact with

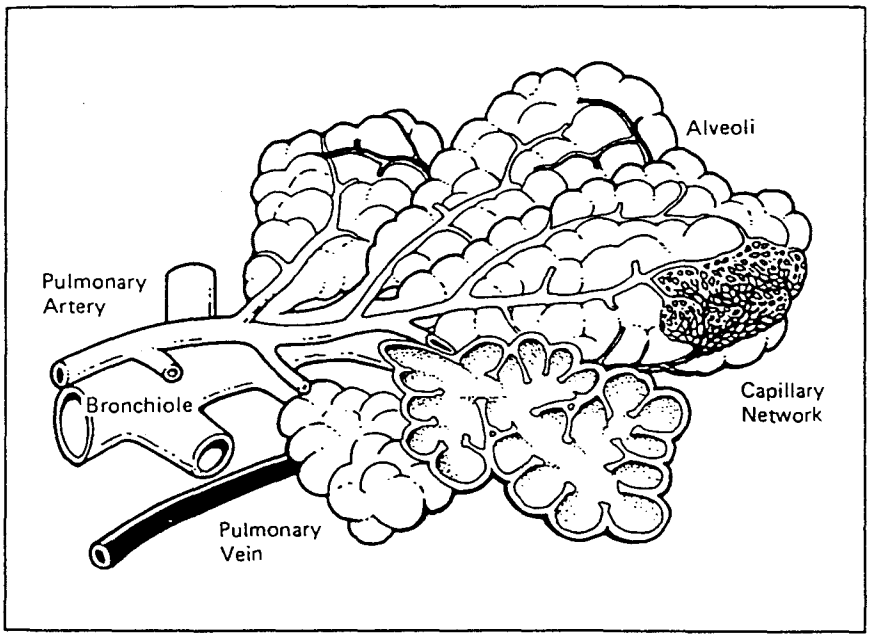
Figure 2.1a *Gross Anatomical Relationships in the Thoracic Cavity*

The respiratory pathway is shown starting with the trachea and subdividing into the bronchi. Further sub-division occurs (up to 23 times) but is not illustrated. The pleural membranes are also shown. (Mar87)

Figure 2.1b *Alveolar Structure*

A view of the bubble like alveoli found at the terminal points of the respiratory pathway. Their thin walls are covered with a network of capillaries. (Mar87)


Epiglottis
Hyoid bone
Thyroid cartilage
Larynx
Cricoid cartilage
Trachea
RIGHT LUNG
Carina
From Primary bronchus
Rib
LEFT LUNG
HEART
DIAPHRAGM
Parietal pleura
Left ventricle
Upper
Visceral pleura
Middle
Lower lobes
Lingula


Alveoli
Pulmonary Artery
Bronchiole
Capillary Network
Pulmonary Vein

the lung surface making its way between the lung's lobes. The parietal pleura rests against the rib cage and diaphragm. Figure 2.1a shows the lungs, pleural membranes and and respiratory pathway. A view of the alveolar structure is presented in Figure 2.1b. This presentation is an extremely brief overview of the macroscopic components that generate lung sounds. Greater detail may be found in any one of many references (Ber90, Gra89, Guy87).

2.1.2 The Mechanics of Breathing

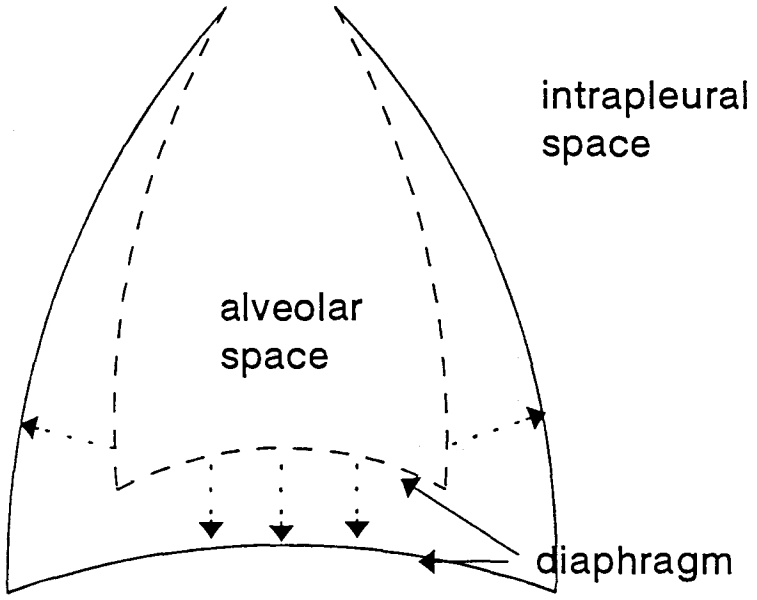
The lung's elastic nature is due to the presence of elastin and collagen fibres working in conjunction with the surface tension of the fluid lining the alveoli (Guy87). In order to explain the process of respiration, the lung can be simply divided into two areas; the alveolar space (inside the lung) and the intra-pleural space (the area outside the lung but still inside the thoracic cavity). Inspiration occurs when the diaphragm moves downward and the intercostal muscles expand the circumference of the rib cage. This causes a decreased intra-pleural pressure followed by a decrease in the intra-pulmonary pressure. The end result is an influx of air. While maintaining the intra-pleural pressure lower than the intra-pulmonary pressure, the pressure in the alveolar space will keep the lung in a state of inflation. Expiration is a passive exercise relying on the elasticity of the tissues to return the lung to a resting position when the intra-pleural pressure is returned to the relaxed value. Figure 2.2 illustrates the expansion process.

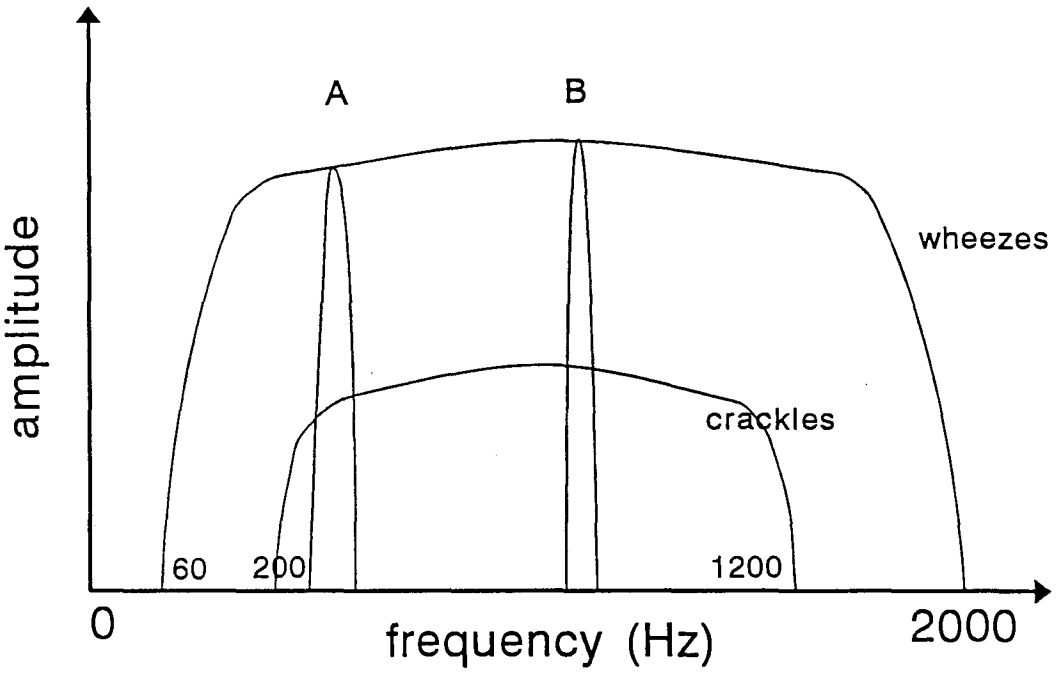
Figure 2.2 *The Mechanics of Breathing.*

A decrease in the intrapleural pressure is caused by downward movement of the diaphragm and an upward and outward expansion of the rib cage by the inter-costal muscles. When this pressure becomes lower than the that of the alveolar space, the lung inflates. Expiration occurs when the diaphragm and inter-costals are relaxed. The elastic properties in the lung and surrounding structures cause the lungs to empty. This is a passive process.

Figure 2.3 *Spectral Range of Wheezes and Crackles.*

The spectral range of crackles is shown to be approximately 200 to 1200 Hz. An individual crackle spreads its energy over this entire range. Wheezes occur as sharp spikes in the frequency domain. The spikes at A and B show wheezes at approximately 300 Hz and 800 Hz. Because the wheezes are narrow band and their energy is concentrated, they tend to overpower the crackles in the frequency domain.


intrapleural
space
alveolar
space
diaphragm


A
B
wheezes
crackles
amplitude
60
200
1200
0
frequency (Hz)
2000

2.2 Lung Sounds of Significance in Clinical Diagnosis

Originating with Laennec, râles and crepitations were the popular terms for lung sounds circa 1900. These were then sub-divided by accompaniment of descriptive terms such as sonorous or sibilant. As more classifications and different descriptions were added, the nomenclature became very confusing until Robertson and Coope proposed the currently accepted terminology in 1957 (For78). Lung sounds used in diagnosis are now referred to as crackles, wheezes and breath sounds. Breath sounds are due to normal healthy breathing. Wheezes and crackles fall under the category of adventitious (abnormal) lung sounds. Throughout this thesis, the sounds will be referred to in this manner. Sub-divisions describing different types of wheezes and crackles will be outlined in their respective sections. Several other sounds exist and are explained later.

2.2.1 The Wheeze

Wheezes are continuous semi-stationary sounds that may be identified visually as a series of identical waveforms (Kos90). They are often referred to as musical sounds since their waveforms are periodic and stationary in nature. These waveforms can be differentiated into a few well defined, harmonically related frequencies. Wheeze genesis of these sounds has been proven to result from air rushing past a point where the walls of the passageway are almost in contact. The driving condition for the following

phenomena is the instability of the airway walls. As the velocity of the air increases through the narrowed passage, a vacuum is created. This draws the walls closer together. When they come in contact, air flow is halted and the wall's elastic recoil pulls them apart. This cycle is repeated and the speed with which it occurs determines the frequency of the wheeze (For78). Wheezes may also be caused under pathologic conditions such as the presence of tumours or other foreign objects in the airways (Fle90). These objects form a point of closure and are set into vibration in the same manner as for a point of closure formed by the narrowing of walls.

Monophonic wheezes occur if only one point of closure exists. Polyphonic wheezes occur when multiple points of closure are present. Characteristically, wheezes occupy the frequency range of 60 to 2000 Hz (Fle90, For78, Kos90).

2.2.2 The Crackle

Two types of crackles exist. Necessary for the generation of the first type, the fine crackle, is a closed compartment and a significant pressure differential between the inside of this compartment and its outside environment. The removal of the closure results in a rapid pressure equalization. The transition is so sharp and rapid that two to three rapidly fading irregular oscillations are generated. These oscillations decay exponentially and completely expire within a few milliseconds (For78). Fine crackles are high pitched. Their timing is reproducible from respiratory cycle to respiratory cycle

Table 2.1 *Characteristics of Crackles for Several Conditions*, (Cla81). The table details variations in crackles that help differentiate between lung conditions.

Characteristic of crackle	Obstructive chronic bronchitis	Bronchiectasis	Fibrosing alveolitis
Timing of inspiratory crackles	Present the early phase of inspiration	Present in the early and mid-phase of inspiration	Present in the latter phase of inspiration
Number of crackles	Always few	Usually moderate	Can be profuse
Effect of cough	No change	Temporarily reduced	No change
Effect of position	No change	No change	Modified or abolished
Intensity	Faint	Loud	Moderately Loud
Pitch	Low pitched	Low pitched	High pitched
Expiratory crackles	May be present	Typically present	May be present
Transmission through chest cavity	Transmitted	Transmitted	Not transmitted

and cannot be abolished by coughing (Des90). Another characteristic is the localization to their area of genesis. The second type of crackle, coarse crackles, originate from free fluid in the airways. Their spectral characteristics are a function of fluid viscosity (Ols78). They are low pitched, continuous over the entire respiratory cycle, audible over entire lung and vary greatly with posture or coughing (Des90).

Crackles may be heard in both inspiration and expiration though they tend to occur most commonly during inspiration. The combination of timing, distribution and transmission of the crackles identify distinct patterns particular to certain disorders.

Difficulty in auscultation occurs due to the spectral overlap of wheezes and crackles. Figure 2.3 shows the ranges for these sounds. Wheezes are narrow band and can spring up anywhere in the 60-2000 Hz range. Crackles are wide band and occupy most of the 200-1200 Hz range. Due to their wide band nature, their spectral energy is spread thinly.

2.2.3 Other Sounds

When the pleural surface is roughened by fibrin deposits or other infiltrations, the normal silent smooth sliding motion becomes a series of short jerks. The pulling motion sets the chest wall into oscillation often resulting in a continuous sound produced in much the same manner as the drawing of a bow across violin strings. Otherwise, a crackly rubbing sound is heard (Des90). Pleural friction rub is often accompanied by sharp pain

localized to the region rubbing.

2.2.4 Pathophysiological Significance of Lung Sounds

Though simple in nature, adventitious sounds reveal a great deal about the pathology of a patient's condition. The following details some of the acoustic characteristics of lung disease.

The nature of a wheeze is indicative of the pathology involved, as the following examples demonstrate (Ogi90). A fixed monophonic wheeze is typical of cicatricial stenosis (formation of new tissue that narrows the airway) or an intra-bronchial tumour. The pitch is related to mass and varies over a very narrow range. This variation is determined by air flow velocity. Random monophonic wheezes are characteristic of widespread airflow obstruction as in asthma or bronchitis. The expiratory polyphonic wheeze is also characteristic of widespread airflow obstruction but when accompanied by inspiratory wheezes is a sure sign of emphysema. Pulmonary fibrosis is characterised by wheezes late in inspiration. These are accompanied by inspiratory crackles. Sequential wheezes (squawks) are caused by the opening of areas of the lung previously apposed. This is typical of pulmonary fibrosis.

Table 2.1 outlines the timing, number, effect of posture, acoustic characteristics and transmission of crackles for chronic bronchitis (narrowing of airways through inflammation), bronchiectasis (dilation of bronchi due to damage of the elastic and

muscular layers) and fibrosing alveolitis (stiffening of alveoli due to increase in fibrin content) (Nat81).

From this discussion, it becomes apparent that different kinds of pathology are associated with different physical properties of the airways and their contents, thus giving rise to individual crackle and wheeze signatures.

2.2.5 Sound Transmission Through the Chest

The quality of transmission depends on many factors since the sound must pass through several tissues and tissue interfaces. Factors affecting transmission must be kept in mind during auscultation.

Experiments on excised lungs have shown that sound energy is lost at the interface of the lung and chest wall. This loss is due to the mismatch of acoustic impedance and becomes more pronounced with increasing obesity (For78b). Lung consolidation, while on first thought would seem to attenuate sound energy, actually enhances sound transmission (Des90). Fluid collecting in the pleural cavity, attenuates low frequencies and if enough collects, a complete acoustic block may occur (For78b). By comparing lung sounds heard at the mouth and passed through the chest, it has been determined that the lung and associated tissues act as a low-pass filter. The associated high frequency loss accounts for the purity of wheezes heard through the chest when compared to those heard at the mouth.

3.0 Heart Sounds

The analysis of heart sounds is plagued by many of the same problems as the analysis of lung sounds. At the forefront of this list is the ongoing debate as to the exact genesis of the sounds (Ran88). The signal analysis problem is similar particularly when compared to the analysis of crackles. However, the phonocardiogram has undergone a great deal more study. This is probably because the genesis of heart sounds is better understood than that of lung sounds and it is known that changes in heart function will produce abberations in acoustic properties before other symptoms are detectable such as modification of the electrocardiogram (Ran88).

Frequency domain analysis has been performed by many different methods. Bandpass filter banks (Ado70), pole-zero modelling (Joo84) and linear prediction (Iwa79, Nan84) have all been attempted with varying degrees of success. The method most utilized is the fast fourier transform which has been used for several decades (Ado70, Fro74, Yog76a, Yog76b, Pin79). Many of these studies rely on pattern recognition techniques to classify the sounds. For reasons revealed in section 7, it would be interesting to revisit these studies as new methods of time-frequency analysis provide resolution not previously available.

Study in the time domain has primarily focused on the analysis of the phonocardiogram's envelope (Kar75, Sar76). These studies have been particularly useful in identifying murmurs between the first and second heart sounds.

Despite the great deal of research in phonocardiograms, the methods have only

been able to assist in the assement of the functional state of the heart and classify the condition. The procedures will not become a reliable part of the diagnostic and monitoring process until acoustic events can be correlated with specific mechanical events and methods of quantifying these relationships exist.

4.0 Previous Work in Lung Sound Analysis

Little improvement in auscultative techniques has occurred since the invention of the stethoscope by Laennec in 1885. In recent years, with the development of high powered computing techniques, there has been a renewed interest in auscultation. This interest has resulted in several attempts to better recount auditory information and to provide computer diagnosis. The literature can be categorized as follows; frequency domain, time domain, time frequency representations and signal pre-processing.

4.1 Work in the Frequency Domain

Early attempts at providing a system for diagnosis have centered around the fast Fourier transform (FFT) algorithms. This algorithm reduces a sampled waveform into its spectral components. The most common approach involves sampling of the lung sounds to yield a time series, breaking this data into blocks, performing the fast Fourier

transform and averaging the results from each block. Each point of this average spectrum is considered a feature. Identification of features that have the greatest variance between conditions is performed by principal component analysis (PCA). Urquhart used 200 millisecond overlapping windows and PCA to reduce the spectrum to only two features for comparison (Urq81). This process resulted in dimensionless graphs that did show clustering but no absolute boundaries were discernable.

The inability of this approach to produce tight clusters is partially due to the great variability in the methods of sound genesis. For example, the wide frequency band that wheezes can occupy depends on the mass and rigidity of the vibrating structure or the pitch of crackles is related to the amount of fluid and viscosity of fluid in the lungs. In addition, large variances in soma-type or lung structure can filter the same original lung sound to produce two dissimilar spectra. Flow rate must also be monitored closely as it can play a profound effect in shifting the overall lung sound spectra. Lessard and Wong have shown that an increase in flow rate raises the mean frequency of the overall lung sound spectrum (Les80).

One severe disadvantage in studies of this nature is the complete loss of timing information. As seen previously, the position of an adventitious sound in the respiratory cycle is of clinical relevance. In the above instance, this information is completely ignored as the FFT is taken over the entire respiratory cycle providing only the total frequency content.

While the concept of computer diagnosis using auscultation alone is appealing, it is not very practical. The 'black box' approach taken by the methods above

prevents the inclusion of many other variables in making the diagnosis.

4.2 Work in the Time Domain

Recently, there has been an interest in time series analysis methods. Manual examination of the crackle waveform has led to measures that would be useful in distinguishing between fine and coarse crackles. These characteristics are the initial deflection width (IDW), the time required for the first two cycles (2CD), the time required for the first 1/4 cycle of crackles and the time required for the first 9/4 cycle of crackles. Mean values of these characteristics for patients with pulmonary fibrosis (fine crackles) and chronic bronchitis (coarse crackles) were determined and showed distinct differences (Mun91). Hoevers characterized crackles based on IDW and 2CD as well as measurements based around the largest deflection in the crackle (Hoe90). These types of studies are able to classify the crackles with a high degree of precision. Unfortunately these processes require manual identification of crackles and often manual examination of the waveform resulting in time consuming and tedious work.

More recently, investigation into the separation of crackles from the rest of the lung sound has been attempted with a non-linear prediction error filter (Ono89). Based on the standard prediction error filter, the method produces two outputs; stationary and non-stationary. If the prediction error exceeds a preset level, the signal is shunted to the non-stationary output through a non-linear function. The resulting gap in the stationary

output is filled in with the prediction. The work has been furthered with the addition of a filter that identifies non-stationarities with widths comparable to those of crackles (Ara91).

4.3 Time-Frequency Representations

Pasterkamp has devised a display that simultaneously displays a spectrogram of breath sounds, flow information, the raw lung sound signal and the electro-cardiogram (Pas89). The presentation of the ECG allows one to visually compensate for the effects of the spectral overlap between heart and breath sounds. This system works in approximately 1.5 times real-time. Since a 1024 point FFT is utilized in forming the spectrogram, significant temporal smearing of the crackles occurs. This is more clearly demonstrated in section 6.1.1.1.

Other methods have utilized image analysis techniques on the spectrogram. Edge detection has been applied in order to identify the wide band crackles (Cha83).

4.4 Signal Pre-Processing

A significant problem to be managed is the spectral overlap between heart and lung sounds. The problem is quite serious since the phonocardiogram originates from the

same physical area as the lung sounds. This has been addressed by Iyer through the use of adaptive filter techniques (Iye86). Using the electrocardiogram's QRS structure as a timing reference for the first heart sound and a delayed version of it for the second, phonocardiogram reductions in the order of 70% have been reported.

5.0 Rationale

This body of research was launched by the need of a hearing impaired medical student to 'hear' the diagnostic sounds of the lungs. The only method available was to watch the acoustic waveform on a portable oscilloscope. Though this one problem provides an immediate application for a lung sound analysis device, it is not the only one. A unit built incorporating the algorithms developed here would also be an excellent teaching tool allowing a solid visual link with the acoustic phenomena of lung disease. Visual representations are often much easier to comprehend and could reduce the learning curve as a great deal of effort is required to 'train' the ear and associated processes for auscultation. Also, as one ages and the characteristics of their hearing change, such a device could still permit the use of auscultative information. Lack of quantitative assessment due to the individual perception of sound may be overcome by the use of a standardized method of analysis. One example of a useful application was cited by a physician during the course of this research. It is well known that chronic lung disease and heart failure produce different types of crackles. In order to determine whether treatment is having a positive effect, it is necessary to determine if the number of

crackles due to the heart condition are decreasing. No methodology exists for determining the ratio of these lung condition induced and heart condition induced crackles; save a subjective judgement.

This thesis sees the application of several current sophisticated signal processing techniques. Although they have their origins in the field of communication where they have seen wide application, they are also well matched to lung sound analysis. The first of these techniques; adaptive filtering; exhibits its power by allowing filtering with no *a priori* information about the signal. In the case of lung sound wheezes, these stationary waveforms can appear anywhere in a 2000 Hz frequency band and will fluctuate slightly. Adaptive filters are able to detect the appearance of such signals, lock onto them and track their evolution. In terms of displaying lung sound information, it will be shown that high resolution time frequency methods are a vast improvement over the standard Fourier transform allowing highly non-stationary crackles to be examined with detail never before realized. These techniques are particularly useful in the biologic environment since noise is usually high. Detection of transient signals at a signal to noise ratio of -6 dB has been reported with these methods (Boa90).

Perhaps the greatest accomplishment would be to re-introduce auscultation as a diagnostic tool. In order for this to be accomplished, the acoustic phenomena need an inexorable link to the physical processes creating them. With the high resolution now available, information that was previously hidden is now displayed providing a point of departure for such an investigation. The phonocardiogram could also benefit from this type of analysis.

Section 2 - Theory of Methods

6.0 Adaptive Techniques

The agenda of adaptive filter theory is to define mechanisms that monitor a filter's output and update the tap weights accordingly. This provides a time-varying, self-adjusting system that is able to operate in an optimum sense despite time-variations in filter criteria. These criteria must often be averaged and a compromise reached when designing a fixed parameter system.

This discussion is limited to adaptive filters based on the single input transversal filter. It is often referred to as a finite impulse response (FIR). In these filters, the impulse response is only as long as the filter order. Filters with feedback loops are able to generate impulse responses that extend to infinity and hence are called infinite impulse response filters (IIR). The block diagram for a FIR filter is given in figure 6.1a where the block operations denoted 'z^{-1}' are unit delay operators. These delays hold the value of the previous sample presented at their input. Multiplications are performed at the tap weights, 'w_{mk}' resulting in the output of this filter being a linear combination of the inputs. The subscript '$_k$' indicates that the values are calculated for time k. The number of tap weights defines the filter order, 'M'.

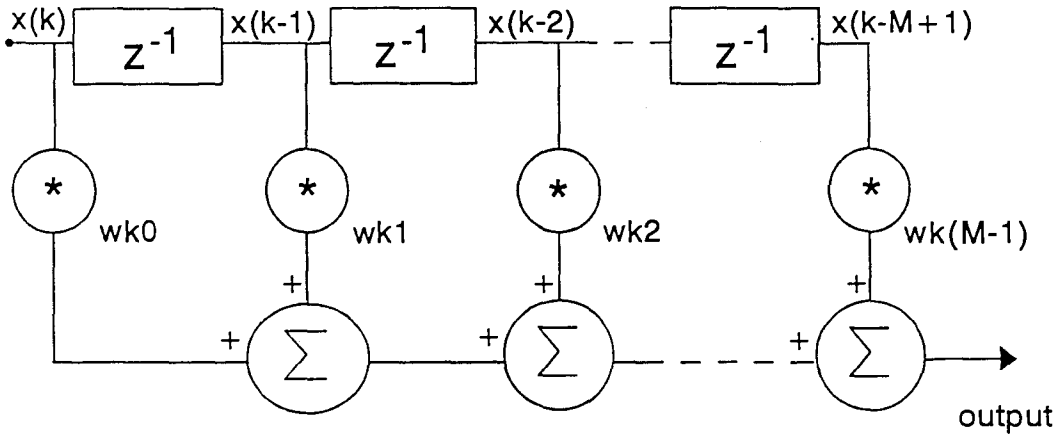
$$y_k = \sum_{m=0}^{M-1} w_{mk}\, x_{k-m} \tag{6.1}$$

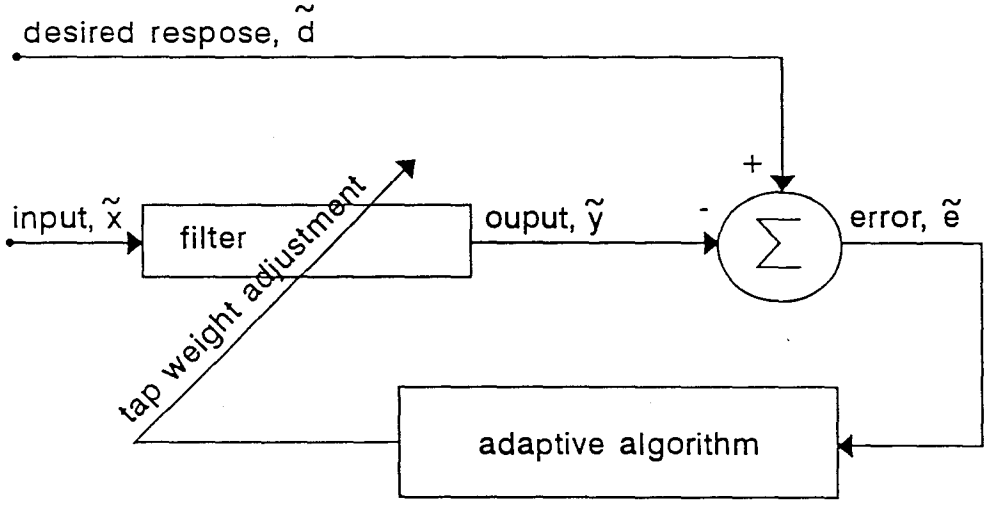
From the purpose of this discussion on filters, the following notation will be

Figure 6.1a, 6.1b *Adaptive Filter Building Blocks.*

Figure 6.1a Transversal Filter. The number of delay elements (z^{-1}) is the filter order. The output is a linear combination of the vector held in tapped delay lines.

Figure 6.1b The Adaptive Filter as a Closed Loop System. Adaptation is performed adjusting the filter weights in order to minimize the error between the desired response and the input.


x(k)
z-1
x(k-1)
z-1
x(k-2)
z-1
x(k-M+1)
*
wk0
*
wk1
*
wk2
*
wk(M-1)
+
+
+
+
Σ
Σ
Σ
output


desired respose, d̃
+
input, x̃
filter
ouput, ỹ
-
Σ
error, ẽ
tap weight adjustment
adaptive algorithm

adhered to:

1) scalars will appear as lower case letters
2) matrices will be shown as upper case letters
3) vectors will be denoted by lower case letters with one of the following overstruck

$\bar{}$ will only be used to annotate vectors formed by filter elements. For example, $\bar{x}$ is the signal segment currently in the filter. Its composition is $[x_k, x_{k-1}, x_{k-2}, \ldots, x_{k-M+1}]$ where M is the order of the filter. Similarly, $\bar{w}$ is the vector made of the tap weights.

$\tilde{}$ indicates a signal vector, as an example, $\tilde{x}$ is the signal itself in its entirety and $\tilde{y}$ is its filtered version.

$\check{}$ indicates an estimated vector, usually an estimation of its optimal value

$\hat{}$ indicates that a vector is optimal with respect to some criteria

Equation 6.1 may be more conveniently expressed in vector notation where the superscript 'T' denotes the matrix transpose operation. The transpose operator is used instead of the hermitian operator (transposition and complex conjugation) since the data to be used in conjunction with the adaptive algorithms presented will be real making complex conjugation redundant.

$$y_k = \bar{x}_k^T \bar{w}_k = \bar{w}_k^T \bar{x}_k \tag{6.2}$$

Filter output characteristics are determined by the values of the tap weights. Generally speaking, the greater the filter order, the better will be the filter's performance since more information is used in forming the output.

Adaptive systems may be conceived as a closed loop process with two functional blocks and two inputs. Typical adaptive system architecture is illustrated in figure 6.1b. The signal input, $\tilde{x}$, goes directly to the filter. The second input required is the desired output, $\tilde{d}$ which acts as a reference in determining what the filter output, $\tilde{y}$, should approximate. A subtraction of the filter output from the desired vector yields the error vector $\tilde{e}$. This vector is a measure of 'how close' the signal vector is to approximating the desired response and when fed to the adaptive algorithm, causes a filter adjustment that minimizes some criterion. The usual criterion is the average power (mean-square value) of the error signal. This is referred to as the mean square error (MSE) or simply ξ. Unfortunately one requires a fairly accurate reference vector, $\tilde{d}$, as the degree of success in filtering depends crucially on it.

A concept that helps in visualizing the process of adaptation is the error performance surface (EPS). The EPS function plots ξ against the filter weights. Optimal performance is achieved when the adaptive algorithm locates the surface's minimum. For a filter of order of two, the error performance surface takes the form of a bowl with the optimum value for ξ and the optimum filter weights being given by the coordinates at the bottom of the bowl. The geometry of the bowl is constant for a stationary environment. However, if the environment is non-stationary, the bowl's shape will change in time and the adaptation must track its minimum. Adaptive filter theory is concerned with generating methods to locate the minimum of the error performance surface and follow its movements.

6.1 Adaptive Algorithms

Before algorithmic development takes place, a mathematic description of the EPS surface is required. Filter output is very simply calculated. For a particular point in time,

$$e_k = d_k - y_k = d_k - \bar{w}_k^T \bar{x}_k = d_k - \bar{x}_k^T \bar{w}_k \tag{6.3}$$

Squaring this value yields,

$$e_k^2 = d_k^2 + \bar{w}_k^T \bar{x}_k \; \bar{x}_k^T \bar{w}_k - 2 d_k \bar{x}_k^T \bar{w}_k \tag{6.4}$$

Assuming stationarity, the expected values are,

$$E[e_k^2] = E[d_k^2] + \bar{w}_k^T \; E[\bar{x}_k \bar{x}_k^T] \bar{w}_k - 2 \; E[d_k \bar{x}_k^T] \bar{w}_k \tag{6.5}$$

It is more convenient to express equation 6.5 in matrix notation where the value of $E[\bar{x} \; \bar{x}^T]$ is referred to as the input correlation matrix (R). The upper limit on the summation, 'n' is the length of the signal vector $\tilde{x}$.

$$R = E[\bar{x}_k \bar{x}_k] = \sum_{k=1}^{n} \bar{x}_k \; \bar{x}_k^T \tag{6.6}$$

$$R = \begin{bmatrix} r(0) & r(1) & \cdots & r(M-1) \\ r(-1) & r(0) & \cdots & r(M-2) \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ r(-M+1) & r(-M+2) & \cdots & r(0) \end{bmatrix} \tag{6.7}$$

This square matrix is Toeplitz meaning that the values on the main diagonal are equal as are the elements on any diagonal parallel to the main diagonal. Another property of the correlation matrix is that it displays conjugate symmetry about the main diagonal (Wid85). A simple interpretation of the correlation matrix is that it tells 'how much' a data value in the future will depend on one in the past. The first diagonal above the main diagonal relates the degree of a sample's dependency upon the previous one for its value. For the second diagonal, the same relationship is expressed with the exception that one is now looking two samples into the future. As an example, the correlation matrix for white noise is the identity matrix; zero everywhere except along the main diagonal. This is due to the complete unpredictability of white noise.

By similar argument to the formation of R, the expectation $E[d\bar{x}]$ is the cross correlation between the desired response and the vector held by the tapped delay line and is denoted $\bar{p}$. Now the mean square error is more succinctly expressed as,

$$\xi = E[e_k^2] = E[d_k^2] + \bar{w}_k^T R \bar{w}_k - 2\bar{p}^T \bar{w}_k \tag{6.8}$$

Adaptive algorithms focus on the use of the gradient to seek the minimum of the EPS. Designated by the symbol $\bar{\nabla}$, the gradient is the partial derivative of the mean-square error taken with respect to the weight vector.

$$\bar{\nabla}(\xi) = \frac{\partial \xi}{\partial \bar{w}} = \left[\frac{\partial \xi}{\partial w_0} \quad \frac{\partial \xi}{\partial w_1} \quad \frac{\partial \xi}{\partial w_2} \quad \cdots\cdots \quad \frac{\partial \xi}{\partial w_b} \right] \tag{6.9}$$

Differentiating the expansion of equation 6.8 with respect to the weight vector yields equation 6.10. Setting this equation to 0 and solving for $\bar{w}$ yields the optimal weight vector ($\hat{w}$) that ensures minimum mean square error.

$$\bar{\nabla} = 2R\bar{w} - 2\bar{p} \tag{6.10}$$

$$\bar{\nabla} = 0 = 2R\hat{w} - 2\bar{p} \tag{6.11}$$

$$\hat{w} = R^{-1}\,\bar{p} \tag{6.12}$$

Equation 6.12 is referred to as the Wiener-Hopf equation.

Unfortunately, in real world applications, difficulty arises in deriving the EPS through explicit measurement of the correlation matrix and also in R's inversion. This inversion is extremely time consuming particularly if the order of the filter is high (Hay91). Because of this, the analytic approach outlined above is not useful for our purposes. It is possible, however, to determine the location of points on the EPS through monitoring of ξ over time. In this manner, adjustments can be made to the weight vector moving it in the direction of zero gradient. This is the method of steepest descent (Wid85) which is defined mathematically as follows.

$$\bar{w}_{k+1} = \bar{w}_k + \mu(-\bar{\nabla}_k) \tag{6.13}$$

The negative sign ensures that the movement on the EPS proceeds in a 'downhill'

direction. Since the true gradient is not available, an estimate of it, $\check{\nabla}$, must be used. This estimate is usually noisy and a parameter to control the size of the weight adjustments is used to combat the imprecise nature of the gradient estimation (Wid85). This parameter is called a 'forgetting factor', and is abbreviated, μ. The value of μ is also directly related to the algorithm's rate of convergence to $\hat{w}$ and to the stability of the algorithm. Attention now turns to methods of finding practical methods of generating $\hat{w}$.

6.1.1 The Least Mean Square Algorithm

The simplest and computationally least expensive method of updating the adaptive filter's weight vector is the least mean square (LMS) algorithm. The assumption at the heart of this algorithm is the use e_k^2 as a crude but effective estimate of the mean square error (Wid75).

$$\check{\nabla}_k = \begin{bmatrix} \frac{\partial e_k^2}{\partial w_0} \\ \cdot \\ \cdot \\ \cdot \\ \frac{\partial e_k^2}{\partial w_l} \end{bmatrix} = 2e_k \begin{bmatrix} \frac{\partial e_k}{\partial w_0} \\ \cdot \\ \cdot \\ \cdot \\ \frac{\partial e_k}{\partial w_l} \end{bmatrix} = -2\, e_k \bar{x}_k \qquad (6.14)$$

Substitution of equation 6.14 into the equation for updating the filter weights by the method of steepest descent yields the LMS algorithm.

$$\bar{w}_{k+1} = \bar{w}_k + 2\mu\, e_k\, \bar{x}_k \tag{6.15}$$

Convergence has been proven for the LMS algorithm (Wid85). The constraint placed on μ is that it must lie between zero and the reciprocal of the largest eigenvalue of the correlation matrix.

$$0 < \mu < \frac{1}{\lambda_{max}\ of\ R} \tag{6.16}$$

However, this is only true for the case of a stationary uncorrelated input. For conditions other than these, analysis of convergence is extremely complex.

6.1.2 The Recursive Least Squares Algorithm

One of the most powerful methods of weight vector updating is the Recursive Least Squares (RLS) algorithm. It is an extension of equation 6.12 where an explicit solution to the Wiener-Hopf equation is derived. The correlation matrix (R) and the cross-correlation vector ($\overline{p}$) are generated recursively at every point in the filtering process. To begin the development, the Wiener-Hopf equation is rewritten as,

$$\Phi_k \hat{w}_k = \bar{\theta}_k \tag{6.17}$$

where Φ_k is the approximated correlation matrix at time k formed by adding the new information stored in the unit delay operators. Φ_k is given by,

$$\Phi_k = \sum_{i=1}^{k} \lambda^{k-i} \bar{x}_i \bar{x}_i^T \tag{6.18}$$

Because the of exponential weighting of each contribution to the overall correlation matrix, the algorithm is able to track statistical changes in the data. By similar argument, $\bar{\theta}_k$, the cross-correlation vector is,

$$\bar{\theta}_k = \sum_{i=1}^{k} \lambda^{k-i} \bar{x}_i d_i^T \tag{6.19}$$

Expanding these two equations to elucidate the recursive nature of the updating procedure, equation 6.18 becomes,

$$\Phi_k = \lambda \left[\sum_{i=1}^{k-1} \lambda^{k-1-i} \bar{u}_k \bar{u}_k^T \right] + \bar{u}_k \bar{u}_k^T \tag{6.20}$$

$$\Phi_k = \lambda \Phi_{k-1} + \bar{x}_k \bar{x}_k^T \tag{6.21}$$

and equation 6.19 becomes,

$$\bar{\theta}_k = \lambda\, \bar{\theta}_{k-1} + \bar{x}_k\, \bar{d}_k^T \tag{6.22}$$

With an estimate of the data's correlation matrix and cross correlation vector, the solution to the Weiner-Hopf equation is the next step. In order to avoid the costly and sensitive inversion of Φ_k to solve for the optimum weight vector, $\hat{w}$, the matrix inversion lemma is utilized. This powerful relationship was discovered in 1950 by Woodbury (Hay91) and will permit the calculation of the least squares estimate of the weight vector, $\check{w}$. The lemma is completely represented in the following two equations.

$$\begin{aligned} A &= B^{-1} + CD^{-1}C^H \\ A^{-1} &= B - BC(D + C^HBC)^{-1}\, C^HB \end{aligned} \tag{6.23}$$

By making the following substitutions

$$\begin{aligned} A &= \Phi_k \\ B^{-1} &= \lambda\, \Phi_{k-1} \\ C &= \bar{x}_k \\ D &= 1 \end{aligned} \tag{6.24}$$

and letting

$$P_k = \Phi_k^{-1} \tag{6.25}$$

the RLS algorithm is realized by evaluating the following equations at each point in the filtering process.

$$\begin{aligned} m_k &= \frac{\lambda^{-1} P_{k-1} \bar{x}_k}{1 + \lambda^{-1} \bar{x}_k^T P_{k-1} \bar{x}_k} \\ \alpha_k &= d_k - \check{w}_{k-1}^T \bar{x}_k \\ \check{w}_k &= \check{w}_{k-1} + \bar{m}_k \alpha_k \\ P_k &= \lambda^{-1} P_{k-1} - \lambda^{-1} \bar{m}_k \bar{x}_k^T P_{k-1} \end{aligned} \tag{6.26}$$

The "memory" of the filter is determined by the "forgetting factor", λ. For the special case where $\lambda=1$, the filter is said to have infinite memory. In practice, λ is extremely sensitive to change and should be kept close to a value of one. For the conditions required for initiation of the algorithm and analysis of convergence, consult Hay91.

6.1.2 Comparison of the LMS and RLS Algorithms

The LMS and RLS are at opposite ends of the spectrum in terms of cost and performance. Cost can easily be ascertained by comparing the number of operations required to perform each filtering step. These are summarized in the table below (Table 6.1) (Hay91).

Algorithm	Multiplications	Divisions	Additions Subtractions
LMS	$2M + 1$	0	$2M$
RLS	$2M^2 + 7M + 5$	$M^2 + 4M + 3$	$2M^2 + 6M + 4$

Table 6.1 *Operations per Iteration as a Function of Filter Order.* The table illustrates the difference in the computational complexity of the RLS and LMS algorithms. Clearly, the RLS method requires a great deal more computation. M is the filter's order.

The RLS converges an order of magnitude faster than the LMS (Hay91). However, in a non-stationary environment, the LMS algorithm often exhibits superior tracking performance (Hay91).

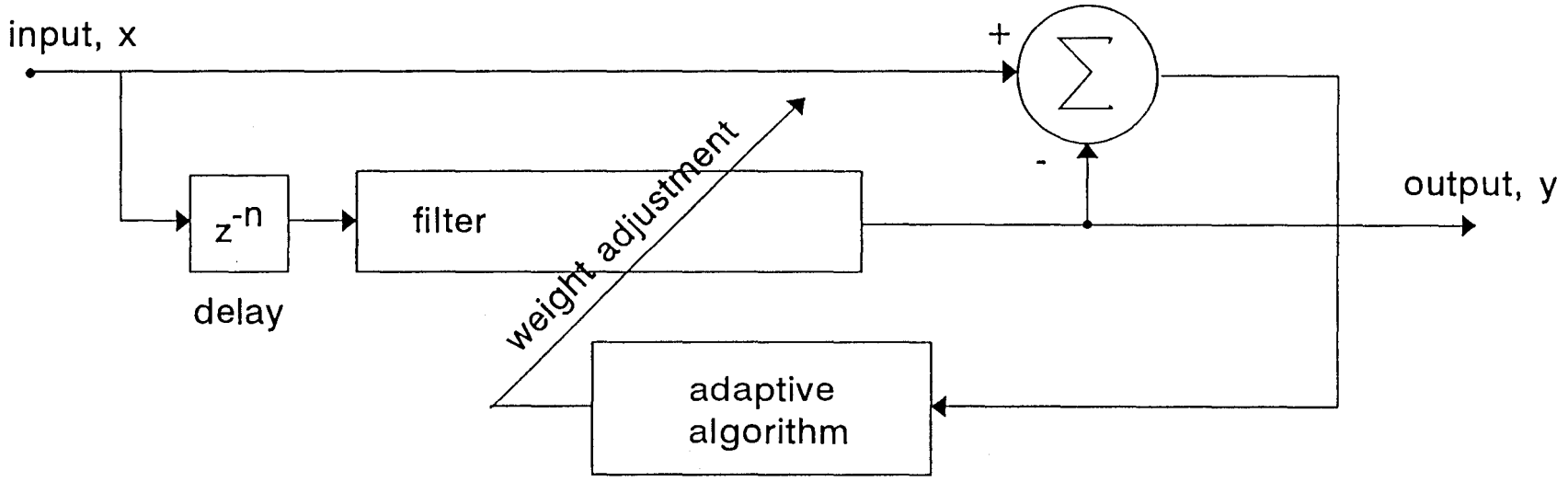
6.2 Adaptive Line Enhancer

The adaptive line enhancer was devised to deal with the classic problem of the detection of a low level sine wave in noise (Wid75). It works particularly well in

Figure 6.2 *Block Diagram of the Adaptive Line Enhancer.*

Adaptation is accomplished by minimizing the error between the signal ($\tilde{x}$) and a time-delayed filtered version it itself ($\tilde{y}$). This process retains signal components that result in a high correlation between the signal and its delayed version. The optimal filter impulse response is the matched filter response for the common signal components.

Figure 6.2


input, x
+
-
output, y
z^{-n}
delay
filter
weight adjustment
adaptive
algorithm

isolating sinusoids of narrow bandwidth. A block diagram is present in figure 6.2. At the heart of the adaptive line enhancer (ALE) is the comparison of the input signal, $\tilde{x}$, with a time delayed, filtered version of itself, $\tilde{y}$. The time-delayed filtered signal is the ALE's output which is forced to approximate the input. This is performed by minimization of the difference, $\tilde{e}$, between $\tilde{x}$ and $\tilde{y}$. In performing this minimization, the filter extracts signal components that are common between the $\tilde{x}$ and its delayed version.

In addition to the choice of adaptive algorithm, the value selected for the delay controls the ALE's performance characteristics. This "decorrelation" parameter controls 'how far back' the algorithm looks for correlation. For the purpose of wheeze and crackle separation, the wheezes will appear at the stationary output and the crackles at the non-stationary output. The wheezes are not necessarily of constant pitch. This creates a problem since if the delay is too long, the there will not be any significant correlation between the signal and its delayed version and no common components will be identified.

Once the ALE's steady state has been reached, the ideal impulse response for the filter portion is the equivalent of the matched filter response for the sinusoid under detection (Wid85).

7.0 Time-Frequency Representations

The time domain and frequency domain are two alternate ways of looking at a signal. In the strictest sense, time information is not accessible in the frequency domain and vice-versa. A joint time-frequency representation is much more revealing as it allows determination of frequency content at a specific time. These time-frequency representations will be referred to as TFRs and notated mathematically as $T_x(t,f)$ the subscript '$_x$' denoting the signal under analysis. Time frequency representations are displayed in the time-frequency plane as a three dimensional figure. Along the x-axis is the variable time. Frequency is plotted along the y-axis and the energy for a particular frequency at a particular time is plotted on the z-axis. Unfortunately, achieving a high resolution TFR is difficult and requires a considerable amount of computation. This section explores the fundamentals responsible for two of the most currently favoured methods of time-frequency analysis; linear and quadratic. These categories differ greatly in the manner in which they treat multi-component signals.

Discussion will be exclusive to the discrete forms of the representations. Formulae for the continuous cases may be found in many of the references cited. With discreteness comes the problem of resolution. It is ideal to have as many divisions in time and frequency as possible in the TF plane. However the different representations support varied levels of resolution and their computational cost is directly related to their degree of resolution.

In this section, definitions are given in conjunction with graphic examples.

Strengths and weaknesses associated with each method will also be considered and demonstrated.

7.1 Linear Time-Frequency Representations

The linear class of TFRs satisfy the linear superposition principle (equation 7.1). This principle states that the TFR of a multi-component signal is nothing more than a linear combination of the individual TFRs of each component.

$$\begin{aligned} x(t) &= c_1 x_1(t) + c_2 x_2(t) \\ &\Downarrow \\ T_x(t,f) &= c_1 T_{x_1}(t,f) + c_2 T_{x_2}(t,f) \end{aligned} \tag{7.1}$$

This property, deemed 'finite support', is extremely desirable since it states that time-frequency elements occur only when the signal is present or the frequency components comprising the signal actually exist. This is not the case in all time-frequency representations as will be seen when quadratic TFRs are examined in section 7.2.

7.1.1 The Short Time Fourier Transform

The short time Fourier transform (STFT) belongs to the linear class of time-frequency representations. It stems from the regular Fourier transform whose definition is given in equation 7.2.

$$X(f) = \int x(t)\, e^{-2j\pi ft}\, dt \tag{7.2}$$

Here, the magnitude of X(f) determines the signal's global frequency content. This results in a spectrum that is the summation of all spectral components present during the signal. While this may suit certain applications, such a representation does not allow time localization of frequency components and is primarily useful only for the analysis of stationary signals. Temporal information is contained in the phase spectrum, the argument of X(f), but it is difficult to interpret in this form (Hla92). The STFT was designed to induce time localization through the use of multiplicative windows. These windows break the signal into smaller segments on the assumption that these smaller units will be stationary.

$$STFT_x(t,f) = \int_{t'} [\, x(t')\gamma^*(t' - t)]\, e^{-j2\pi ft'}\, dt' \tag{7.3}$$

The mathematical definition of the STFT (equation 7.3), corresponds to the Fourier transform of the signal x(t′) multiplied by a shifted "analysis window", $\gamma^*(t'-t)$,

Figure 7.1a-b *Implementation of Short Time Fourier Transform.*

Figure 7.1a Window Placement in the Short Time Fourier Transform. Windows move along the signal zeroing terms outside of their boundaries. The windows are 64 samples in length.

Figure 7.1b Time Frequency Representation by the STFT. The time-frequency representation is created by fast fourier transform of the windowed segments. The results are stacked in matrix and displayed as a three dimensional plot. The resolution is set at 16 bins on the frequency axis and 10 bins on the time axis.

Figure 7.1a

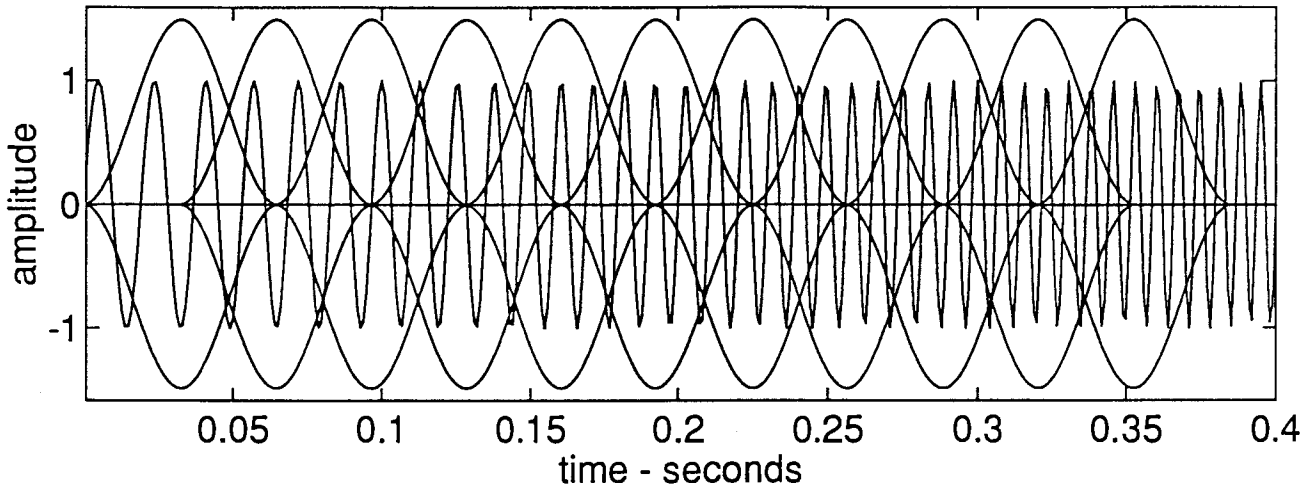
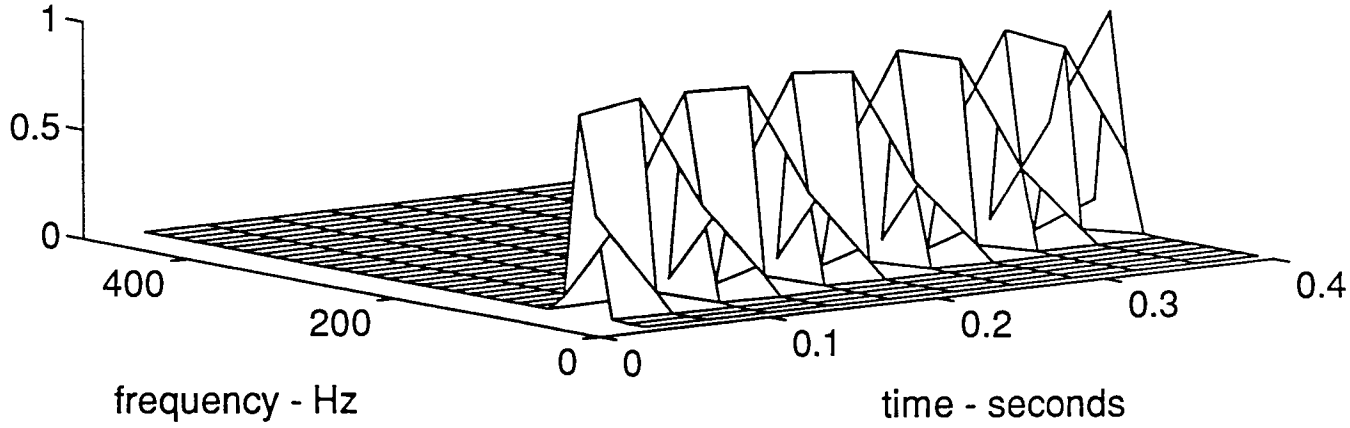


Figure 7.1b



that is centered on t. The window zeros the signal outside of the window and provides smooth edges that will prevent occurrence of Gibb's phenomena (Opp75). The resulting signal segment is then Fourier transformed producing a local spectrum. Localization of individual spectral components in time is restricted by the length of the window γ. This results in a binning process where the time-frequency plane is broken into cells.

The STFT may also be expressed in the frequency domain.

$$STFT_x(t,f) = e^{-j2\pi tf} \int_{f'} X(f')\ \Gamma^*(f'-f)\ e^{j2\pi tf'}\ df' \qquad (7.4)$$

$$e^{-j2\pi tf} \text{ is a phase factor}$$

Equation 7.4 may be interpreted as passing the signal through a bank of bandpass filters, each filter having the frequency response $\Gamma^*(f'-f)$ and being centered around the frequency of analysis, f (Hla92). Since the function Γ, is the Fourier transform of the window function γ, the filter's bandwidth is proportional to the window's width. In order for the STFT to produce a high degree of frequency resolution, a narrow filter band and hence, long window are required. This resolution/window length relationship is the reciprocal for time variable so that good time resolution results in poor frequency resolution. Because the STFT at time t is the spectrum of the signal x(t′) pre-windowed by $\gamma^*(t'-t)$, all spectral features in that window are collected in the time-frequency cell centered on time t in the STFT. This means that superior time resolution demands a short window, $\gamma^*(t')$. In all cases of STFT use a compromise between

temporal and spectral resolution must be made since the analysis window cannot support high time and frequency resolution simultaneously.

For maximum resolution and accuracy to occur, the window length should be optimized so that its length is related to the period of the input series (Har91). This requires *a priori* information that is difficult to obtain in a multi-component, non-stationary, environment.

7.1.1.1 Discrete Realization of the STFT

The STFT is extremely easy to implement in discrete time. The general method is to utilize windows that overlap by 50% although Harris has derived that an overlap of 75% is optimal (Har91). Many standard analysis windows exist. The more common ones are the hanning, hamming and cosine windows which have all been well studied and their properties examined in detail (Opp75). Fast Fourier transforming each windowed data segment is the next step. The resulting vectors are then stacked in a matrix where rows index the frequency bins and columns index the temporal center of each window. The values at the intersection of these indices are the energy levels for the energy of that frequency bin over the window's duration. The process is illustrated in figure 7.1a where the sliding analysis windows are superimposed on the signal to be analyzed. Taking the FFT of the resulting segments and displaying them yields figure 7.1b.

Temporal and spectral smearing due to the time/frequency resolution trade-off are

Figure 7.2a-b *Spectral and Temporal Resolution of the Short Time Fourier Transform.*

Figure 7.2a Crackle Waveform. The signal exhibits three crackles. They occur at $t \in \{0.35, 0.06, 0.07\}$. The signal is approximately 500 samples long.

Figure 7.2b STFT using a window of 64 samples (0.016 seconds). The first crackle is almost non-observable. The other two are identifiable though the resolution is fairly poor. The two spikes fuse together about half-way up their sides. If the windows were larger, the peaks would completely fuse.

Figure 7.2c STFT using a window of 32 samples (0.008 seconds). The crackles have become separated and the first crackle begins to emerge due to the increased temporal resolution over figure 7.2b. The frequency resolution has decreased significantly.

Figure 7.2d STFT using a window of 16 samples (0.004 seconds). Frequency resolution is now extremely poor.

Figure 7.2a

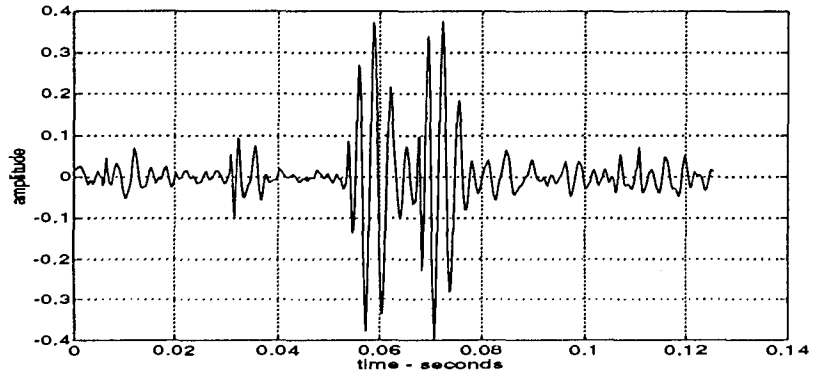


Figure 7.2b

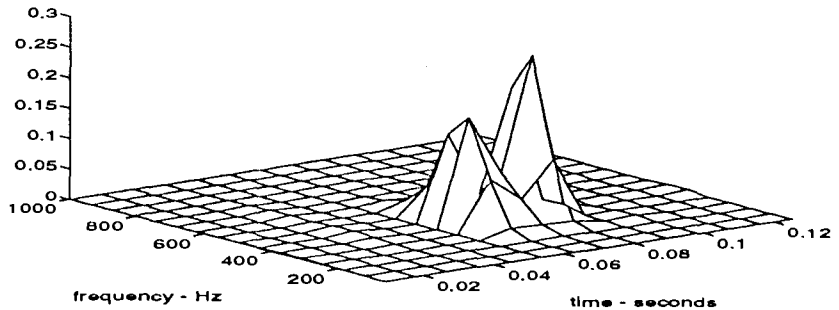


Figure 7.2c

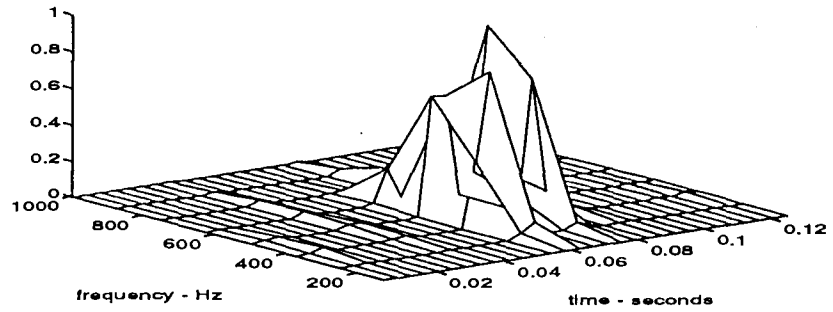
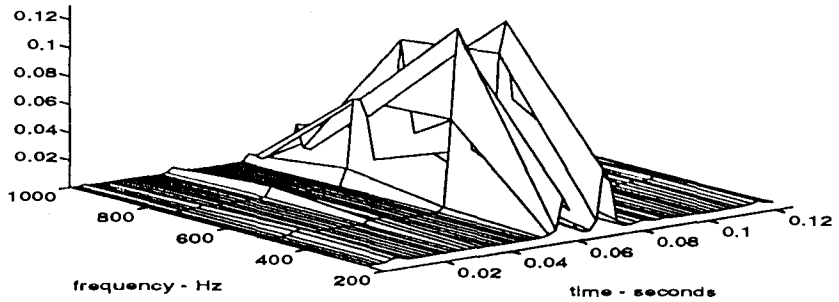


Figure 7.2d



demonstrated in figures 7.2a-d. Figure 7.2a shows the signal to be analyzed. The next three figures show the STFT calculated for three different sized windows. Time-frequency representations with resolution in time and frequency of 13x16 (fig. 7.2b), 29x8 (fig. 7.2c) and 60x4 (fig. 7.2d) result. Of note is the increased temporal resolution and localization in time of the crackles and the slight emergence in figure 7.2d of the small crackle at 0.03 seconds as the window's size decrease. This increase in spectral resolution is accompanied by a concomitant decrease in frequency resolution.

7.2 Quadratic Time Frequency Representations

The quadratic class is thought to be a more realistic approach for generating time-frequency representations. This stems from the definition of a signal's power and spectral density, both of which are squared functions. One class of quadratic TFRs has recently undergone a great deal of investigation. These are the so-called "energetic TFRs". In this group, the signal's energy is given by integration over the entire time frequency plane. In order for this to be possible, the energetic class must satisfy certain mathematical conditions; the two most important being the time and frequency marginals. The marginals are formed by integrating out variables selectively so that values for $|x(t)|^2$ and $|X(f)|^2$ are given. These relationships are expressed in equations 7.5-7.7 (Coh89). The integration of all frequencies at a particular time yields the instantaneous energy.

$$\int T_x(t,f)\, df = |s(t)^2| \tag{7.5}$$

Integrating for one frequency over the signal's duration results in the energy density spectrum $|X(f)^2|$.

$$\int T_x(t,f)\, dt = |S(f)^2| \tag{7.6}$$

Total energy may be derived by integration over the entire frequency-time plane.

$$E = \int T_x(t,f)\, dfdt = \int |x(t)|^2\, dt = \int |X(f)|^2\, df \tag{7.7}$$

Like the linear class of TFRs, the quadratic class also satisfies a superposition principle. Unfortunately, due to the quadratic nature of the TFR, this leads to a more complex representation. The two component signal defined in equation 7.8 results in a time frequency representation that contains terms other than those corresponding directly in time to the components x_1 and x_2. These additional components appear in the second line of equation 7.9.

$$x(t) = a_1x_1(t) + a_2x_2(t) \tag{7.8}$$

$$\begin{aligned} T_x(t,f) = {} & |a_1|^2T_{x_1}(t,f) + |a_2|^2T_{x_2}(t,f) \\ & + a_1a_2^*T_{x_1,x_2}(t,f) + a_2a_1^*T_{x_2,x_1}(t,f) \end{aligned} \tag{7.9}$$

The $T_{x1}(t,f)$ and $T_{x2}(t,f)$ terms result directly from the signal terms and are referred to as the auto terms. The $T_{x1,x2}(t,f)$ and $T_{x2,x1}(t,f)$ terms are referred to as the cross terms, interference terms or simply, ITs. By extension of the principle leading to equation 7.9, it becomes obvious that the cross component problem is augmented as the number of signal components grows. This situation brings to light several other desirable properties that time frequency representations should have.

These other properties are well summarized by Hlawatsch (Hla92). Previously mentioned in section 7.1 was the concept of finite support. This stated that no time-frequency components should exist when the signal or spectral components of the signal are not actually present. The cross terms violate this property. Although cross terms can provide useful information and are of benefit in some applications, they generally confuse the representation particularly when an accurate display of the signal is required. In addition, intuitively, the ideal TFR should be non-negative everywhere though there exists a severe lack of understanding of negative frequency terms (Coh91).

An alternative to looking at the signal in the time-frequency domain is to move the signal into the dual correlative domain (DCD). The DCD combines the concepts of spectral and time correlation forming a joint time-frequency correlation function that represents all possible combinations of time and frequency shifts (Jon92). These representations will be denoted as $D_x(\tau, \upsilon)$. Axes of the dual correlative domain are referred to as lag, for temporal correlation, and Doppler for frequency correlation. Typically, the units on the graph are normalized so that the lag is expressed in seconds and the Doppler in hertz. Actual values are easily computed if one knows the sampling

frequency. The dual correlative function displays conjugate symmetry about the origin (Hla92). Conversion of the signal's representation from the time-frequency domain to the DCD is effected by a 2-dimensional Fourier transform.

$$D_x(\tau,\nu) = \iint T_x(t, f)\, e^{-j2\pi(\nu t - \tau f)}\, dt\, df \tag{7.10}$$

$$T_x(t, f) = \iint D_x(\tau, \nu)\, e^{-j2\pi(\tau f - \nu t)}\, d\tau\, d\nu \tag{7.11}$$

The utility of displaying a signal in the DCD lies in its ability to identify which TFR terms are auto terms and which are cross terms. Elimination of the cross terms in this domain followed by a transformation back to the time-frequency plane effects a cleaning of the TFR leaving only the auto terms.

When the TFR is moved to the DCD via Fourier transform, correlation in the time domain is mapped onto the Doppler axis (υ) and correlation in the frequency domain is mapped onto the lag axis (τ). More importantly, the function describing the signal terms cluster around the origin. Functions describing the interference terms occur distal to the origin. Interference term elimination can be effected by designing a function that will retain components located close to the origin and zero those further away from it.

7.2.1 Cross Term Removal

The general approach to cross term suppression may be likened to the band pass filtering of a time series. Figure 7.3a shows a time series made of two sets of sinusoids. One group is clustered around 250 Hz, the other is clustered around 400 Hz. The spectrum of this signal is shown in figure 7.3b. In order to retain only the group at 250 Hz, the group at 400 Hz could be zeroed with a boxcar function centered on 250 Hz and the inverse Fourier transform would yield a time series whose signal components are only those originally around 250 hertz.

The same type of process can be made to work on the TFR. Figure 7.3c shows a quadratic time-frequency representation of two sinusoid pulses; one at 150 Hz and the other at 400 Hz. The structure between the two signal terms is the interference term. When this TFR is moved to the dual correlative domain (fig 7.3d), multiplication by a mask that passes the structure at the origin and eliminates the interference term structures at $(\tau,\upsilon) \in \{(10,.2), (-10,-.2)\}$ will result in a TFR devoid of interference terms when the function is transformed back to the time-frequency domain. Such a function corresponds to a low-pass filter and forms the basis for an entire family of time frequency representations named Cohen's Class.

7.2.2 Location of Cross Terms

Figure 7.3a-7.3d *Interference Term Suppression through Filtering of the Quadratic Time Frequency Representation. A Comparison to 1-Dimensional Filtering.*

Figure 7.3a Time series made of sinusoids which cluster around either 250 Hz or 400 Hz.

Figure 7.3b Spectrum of the signal in fig. 7.3a. Filtering to retain the sinusoids around 250 Hz may be accomplished by zeroing the components centered on 400 Hz then converting back to the time domain.

Figure 7.3c Quadratic TFR of two sinusoid pulses (150 Hz, 400 Hz) occurring .01 seconds apart. The structure centered on 275 hz .025 second is the interference term.

Figure 7.3d Dual correlative domain representation of TFR in figure c. Note how the mask will zero the ITs located around (-10,-0.2) and (10,0.2). Conversion back to the time-frequency domain will yield a TFR with only signal terms.

Figure 7.3a

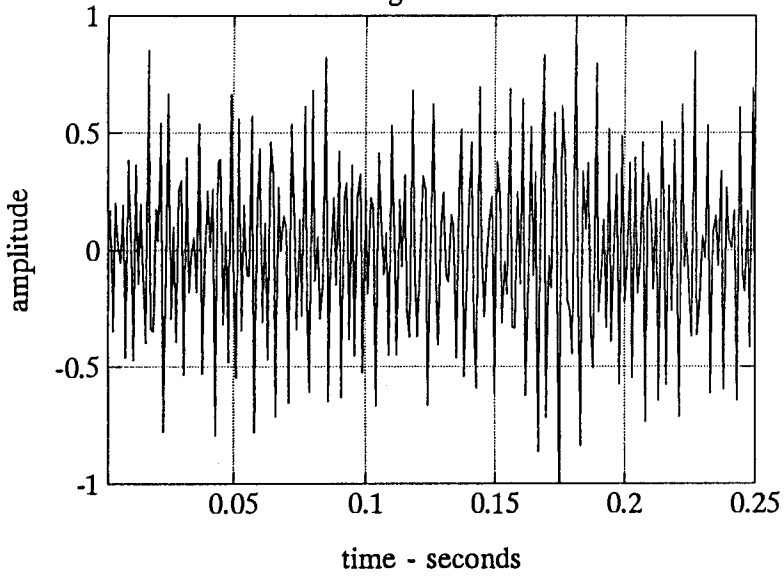


Figure 7.3b

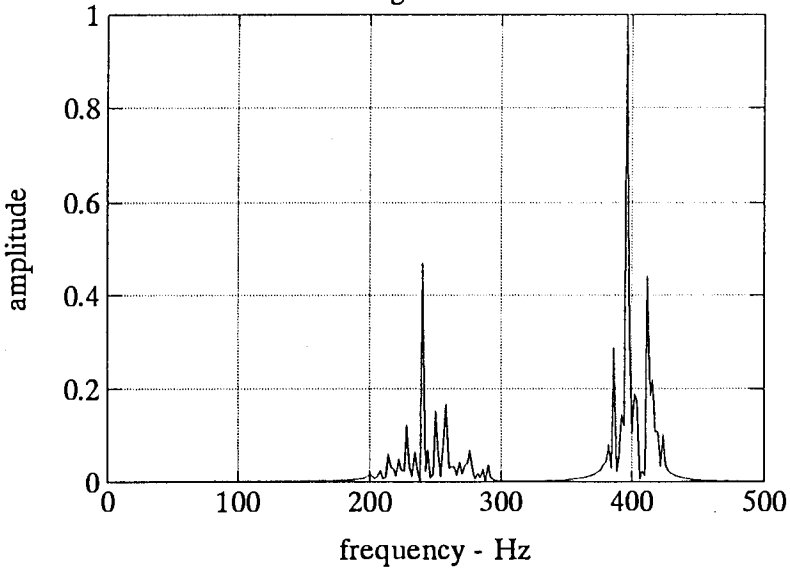


Figure 7.3c

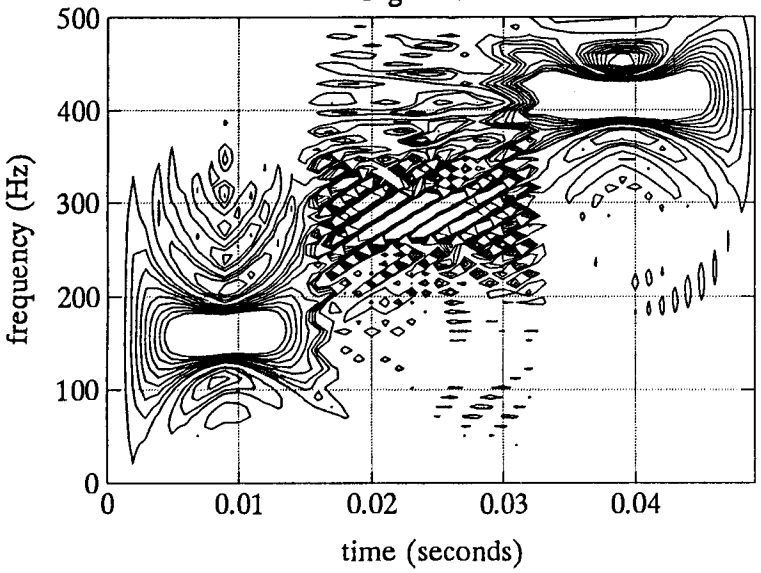
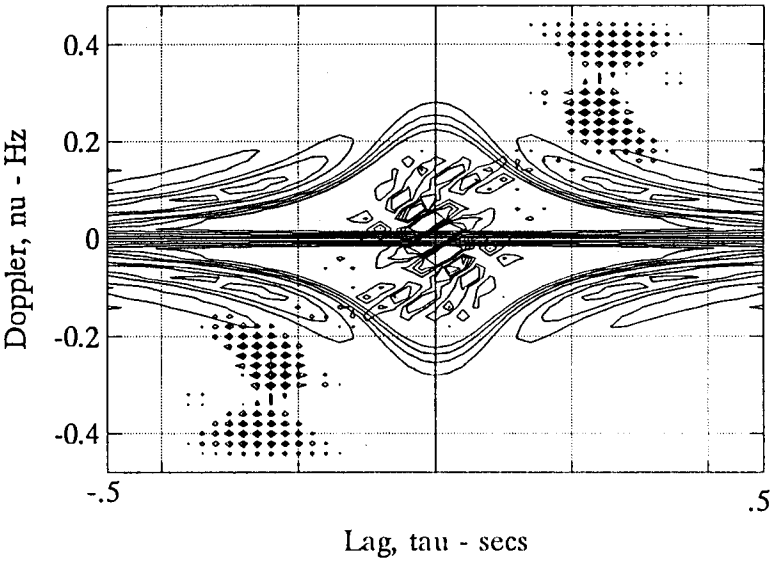


figure 7.3d



A general understanding of cross term location in the DCD is necessary for determining the type of masking function to use when zeroing ITs. For the purpose of illustrating their genesis and placement, the Wigner-Ville Distribution (WVD) will be used. The WVD is further studied in section 7.2.4.1 but for this discussion it suffices to know that it is the simplest member of the quadratic TFR family. A consequence of this fact is the WVD's propensity for cross terms. Many of the TFRs currently undergoing study are based on the WVD. Thus, a firm understanding of WVD cross-terms and their position in the corresponding DCD representation, the Ambiguity Function (AF), is the precursor to the design of time-frequency representations that exhibit a reduction in interference terms.

Hlawatsch has shown that cross terms are formed by every pairing of components in the signal. He has further shown that cross terms are formed by the interference of the TFR signal terms themselves and that these cross terms are located halfway in time and frequency between the pair of interfering terms when in the time frequency plane. At these co-ordinates an oscillatory term twice the amplitude of the original terms occurs. Its frequency is proportional to the distance between the two interfering terms and its direction of oscillation is perpendicular to the line joining the two interfering terms (Hla84).

When the 2-dimensional Fourier transform is invoked to move the TFR into the dual correlative domain, the non oscillating signal terms contribute to the DC portion of the spectrum which maps into the origin of the AF (τ=0, υ=0). The DCD terms that correspond to the TFRs oscillating interference terms occur distal to the origin by an

amount proportional to their oscillatory frequency in the time-frequency domain. In the time frequency domain, interference term oscillation frequency is proportional to the separation of the signal components. The closer the components, the lower the frequency of oscillation and the closer the ITs will be to the origin in the AF.

The following three presentations form an intuitively satisfying explanation of cross terms and their appearance in the dual correlative domain. A strict mathematical approach may be found in Hla84. In each example, the time series to be analyzed is bi-component resulting in a TFR with two signal terms and one cross term. Each presentation consists of three plots. The first is the signal undergoing analysis, the second is the Wigner-Ville Distribution of the signal and the third plot is the signal's Ambiguity Function. The cases are as follows.

Signal components are:

case 1	well separated in time but not frequency (two 15 ms 150 Hz tone bursts separated by 20 ms)
case 2	well separated in frequency but not time (50 Hz and 400 Hz simultaneous 20 ms tone bursts)
case 3	well separated in both time and frequency (15 ms 150 Hz and 20 ms 400 Hz tone bursts separated by 15 ms)

Case one is illustrated in figures 7.4a-c. Since the oscillation in the WVD occurs only in the direction of the frequency axis, the resulting Ambiguity Function shows the interference terms being mapped parallel to the tau axis specifically, on the plane $\tau=0$ since there is no oscillation in the time direction in the WVD.

Case two is illustrated in figures 7.5a-c. Oscillation of the cross term is parallel to the time axis in the WVD resulting in the interference term placement in the AF along

Figures 7.4a-c *Cross Term Location for Bi-Component Signal Separated in Time but not in Frequency*

7.4a The signal under analysis.
7.4b The signal's Wigner-Ville Distribution.
7.4c The signal's Ambiguity Function.

Figure 7.4a

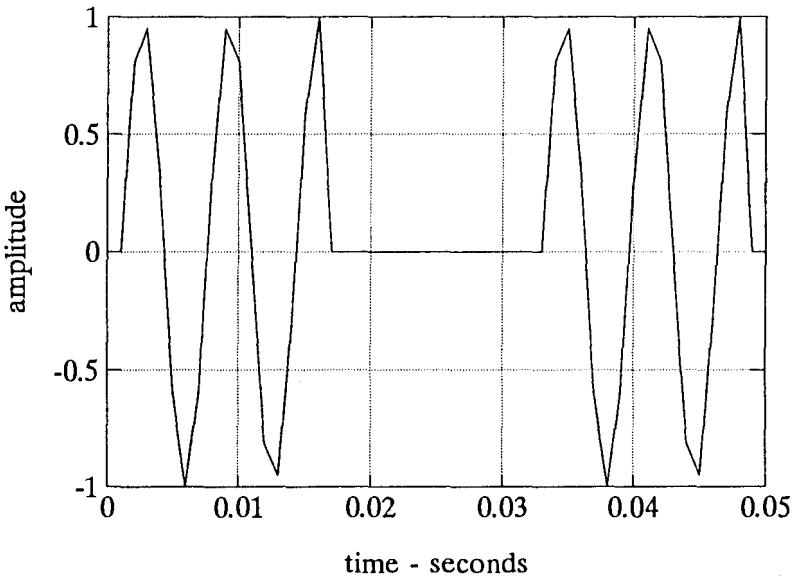


Figure 7.4b

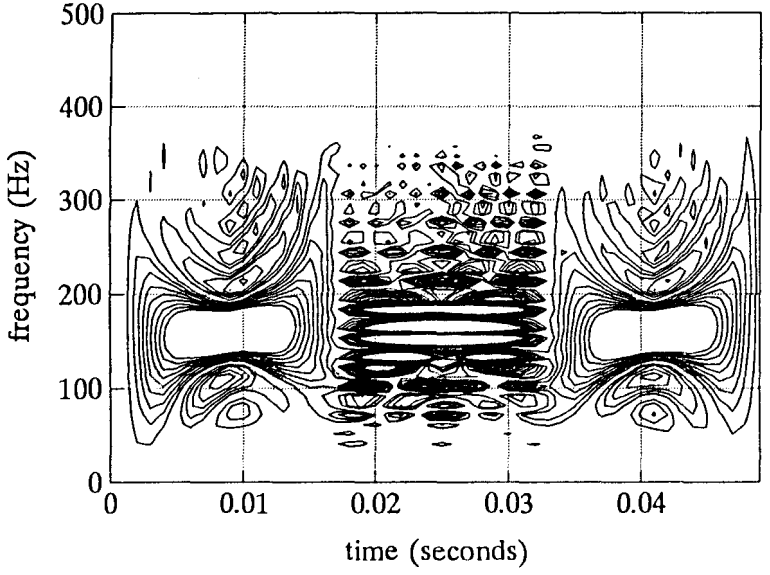
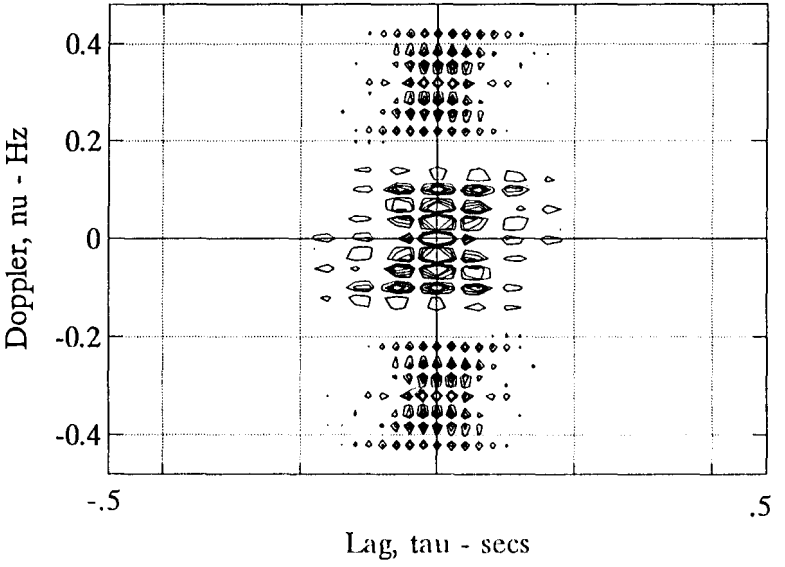


Figure 7.4c



Figures 7.5a-c *Cross Term Location for Bi-Component Signal Separated in Frequency but not in Time*

7.5a The signal under analysis.
7.5b The signal's Wigner-Ville Distribution.
7.5c The signal's Ambiguity Function.

Figure 7.5a

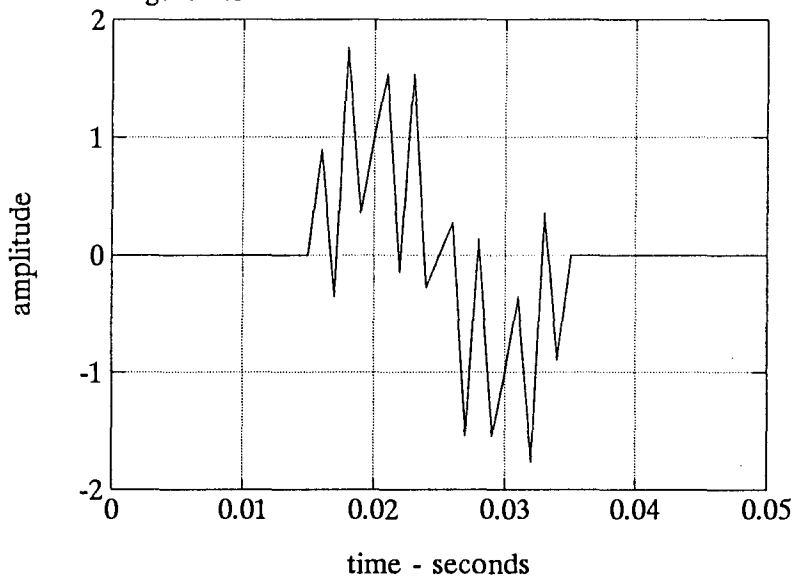


Figure 7.5b

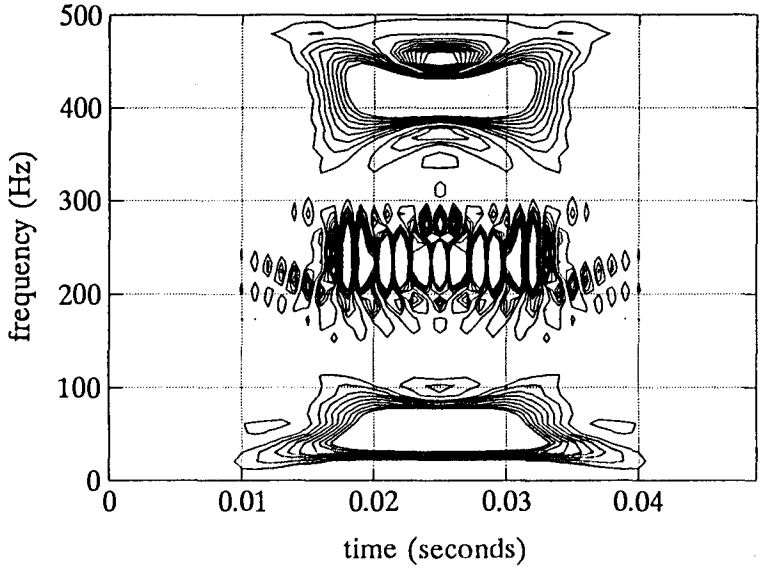


Figure 7.5c

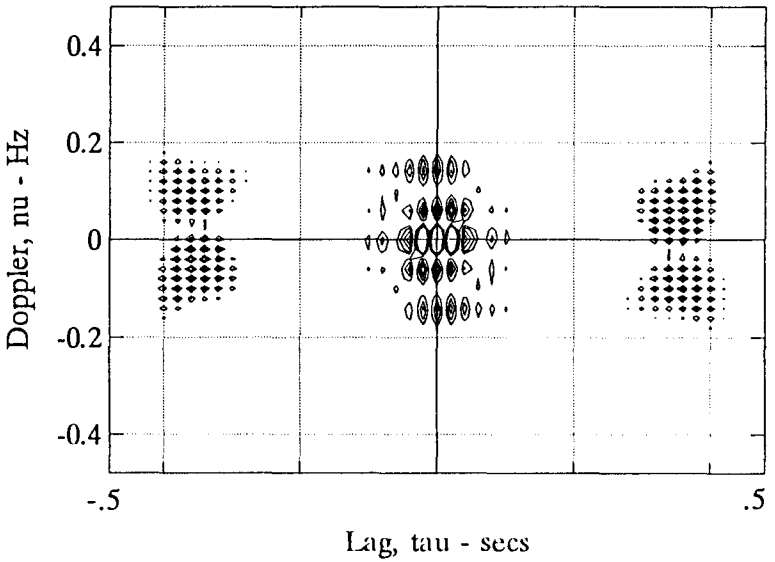


Figure 7.6a-c *Cross Term Location for Bi-Component Signal not Separated in Time or Frequency*

7.6a The signal under analysis.
7.6b The signal's Wigner-Ville Distribution.
7.6c The signal's Ambiguity Function.

Figure 7.6a

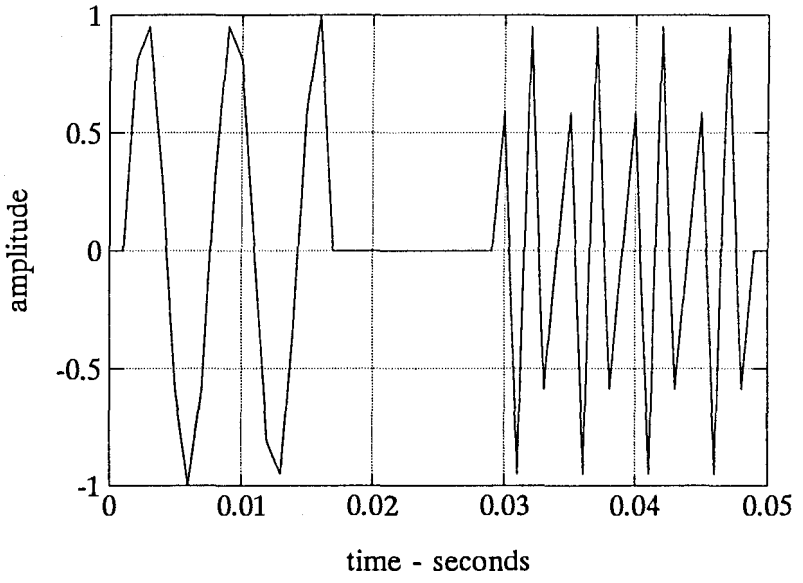


Figure 7.6b

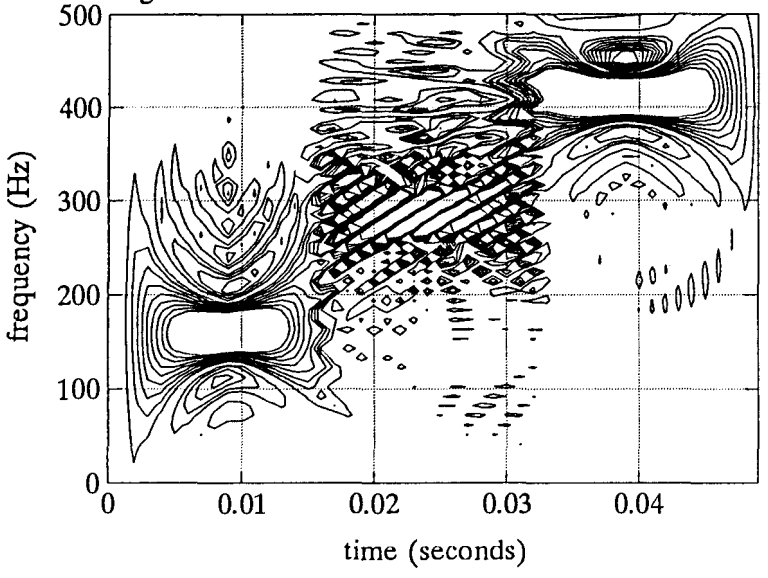
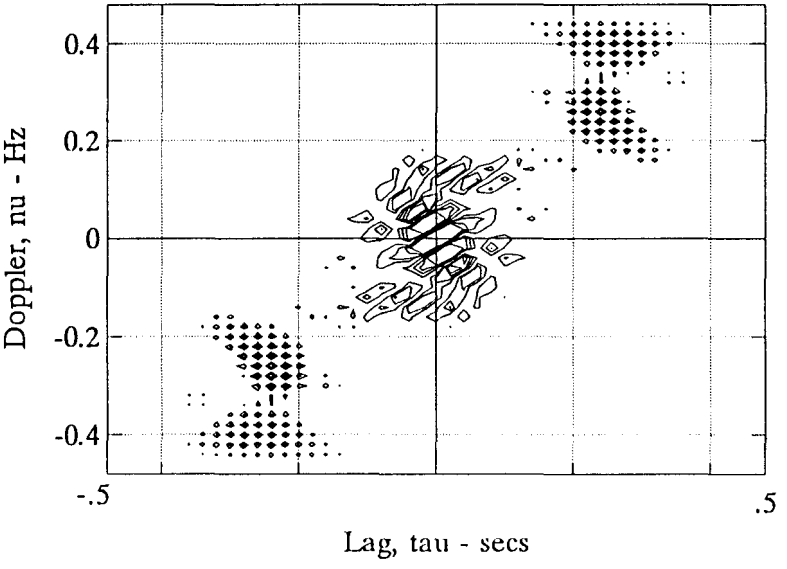


Figure 7.6c



the plane $\upsilon=0$.

Case three is illustrated in figures 7.6a-c. Due to the difference in both time and frequency of the pulses the ITs in the WVD are not perpendicular to either axis and therefore do not lie along the co-ordinate axes in the dual correlative domain.

7.2.3 The Analytic Signal

When data are sampled, they are generally real valued. However, analysis by the methods presented here is more conveniently performed on a signal's complex representation.

The spectrum of a real valued signal possesses symmetry about the origin resulting in a spectrum of negative frequencies that mirrors that of the positive frequencies (Opp75). These negative frequencies have no physical realization and will generate problems that would otherwise not exist. To eliminate these problems, a signal containing the same positive frequency information but without the negative frequencies is constructed. This process utilizes an imaginary component to cancel the negative spectrum and the result is referred to as the analytic signal. Cohen cites three reasons for using the analytic signal (Coh91). They are as follows:

1) If the real signal were used, cross terms would evolve from the reaction between the positive and negative frequencies. Use of the analytic signal reduces the complexity of the resulting TFR.

2) Use of the analytic signal facilitates easy calculation of the instantaneous frequency as the integral of the first moment of the distribution. This relationship does not hold with a real signal.

3) Use of the analytic signal results in the elimination of the requirement to sample at twice the Nyquist rate.

The analytic signal may be formed by first calculating the spectrum of the real signal, discarding the negative frequency components and then utilizing an inverse Fourier transform to put the signal back into the time domain. While this procedure is intuitively satisfying, it is not used due to the difficulty in dealing with the phase information and the amount of computation required. In practice, the analytic signal is calculated using the Hilbert transform. The Hilbert transform $\hat{x}$,is defined as follows.

$$\hat{x}(t) = \frac{1}{\pi} \int \frac{x(\tau)}{t - \tau} d\tau \tag{7.12}$$

Denoted x_+, the analytic signal is the real signal plus an imaginary component that is the Hilbert transform of the real signal.

$$x_+(t) = x(t) + j\hat{x}(t) \tag{7.13}$$

In the communication field, the analytic signal is referred to as the pre-envelope. The name arises from the masking of the negative frequencies. The Hilbert transform possesses many useful properties which are well documented (Hay78).

7.2.4 Generalized Time-Frequency Representation (GTFR)

Cohen (Coh85) has summarized a large group of time frequency representations under one equation. This equation (7.14) defines a group of shift invariant TFRs of the energetic class. Shift invariance is defined as the property whereby a shift in time or frequency in the signal results in a corresponding shift in time or frequency in the time-frequency domain. Cohen's class is capable of describing the energy density of a signal simultaneously in time and frequency with a high degree of resolution. The most prominent member of Cohen's class is the Wigner Ville Distribution (Hla92) (section 7.2.2.1). At the heart of Cohen's class is the use of a 'kernel', $\psi(t,f)$. The kernel is simply a mask when in the DCD and the equivalent of a two-dimensional filter when in the time frequency plane. In the time frequency plane, Cohen's Class may be expressed as follows.

$$T_X(t,f) = \int\int\int \psi_T(t-t', f-f') \, x(t+\frac{\tau}{2}) \, x^*(t-\frac{\tau}{2}) e^{(-j2\pi f\tau)} \, d\tau dt' df' \qquad (7.14)$$

Equation 7.14 shows that a GTFR of Cohen's Class is obtained through the convolution in time and frequency of a kernel $\psi(t,f)$ with the signal's instantaneous correlation function, $(x(t+\tau/2)\, x^*(t-\tau/2))$ and a Fourier transform. An infinite number of quadratic TFRs may be derived from this equation. However, not all of them will satisfy the marginals or possess the desirable characteristics previously mentioned. It is the construction of the kernel ψ, that gives a particular GTFR its characteristics making

kernel selection based on the type of data to be worked with extremely important process.

An alternate interpretation of Cohen's class can be formed by moving the TFR into the dual correlative domain. This results in the following expression.

$$T_x(t,f) = \iint_{\tau\ \nu} \left(\Psi_T(\tau,\nu) \quad D_x(\tau,\nu) \right) e^{-j2\pi(\tau f - t\nu)}\, d\nu\, d\tau \qquad (7.15)$$

Here, the convolution in the time frequency domain now becomes a multiplication and may be envisioned in the manner presented in figure 7.3d. The kernel $\Psi(\tau, \upsilon)$ acts as a mask passing signal terms and zeroing the cross terms. Unfortunately, while filtering the TFR may significantly reduce the amplitude of interference terms, the auto-terms themselves may also be smoothed. Severe degrees of smoothing will adversely affect the information in the signal terms leading to possible misinterpretation. Also, careful attention must be paid to the marginals and the so-called 'nice' properties of the resulting TFR if it is desirable to retain them (Hla92).

7.2.4.1 The Wigner-Ville Distribution and the Ambiguity Function

The Wigner-Ville Distribution (WVD) is the simplest member of the GTFR family. It is formed by using a kernel of one ($\Psi(\tau, \upsilon) = 1$) and is expressed as equation 7.14 with the ψ term removed. Due to its characteristically large degree of interference

terms, the WVD is often confusing when analyzing multi-component signals.

The dual correlative of the WVD is the Ambiguity Function (AF) which is formed by the 2 dimensional Fourier transform of the WVD. At $\nu=0$ for all τ, the AF becomes the time-domain auto-correlation function and if $\tau=0$ for all ν, the AF becomes the frequency domain auto-correlation function.

All member's of Cohen's class stem from the WVD and are produced by smoothing it with a kernel. If one is unable to design a low-pass kernel which totally isolates the signal terms one of two things will occur. Either, all the signal terms will be passed with some interference terms, or all the ITs will be removed along with partial truncation of the auto-terms. The latter situation results in a loss of time-frequency concentration since a smoothing causes a broadening of the WVD's signal terms (Hla92).

7.2.4.2 Conic Kernel GTFR

Zhao, Atlas and Marks have utilized a cone shaped kernel in forming a GTFR of Cohen's class. The resulting distribution is commonly referred to as the ZAM transform. The ZAM transform is based on the concept of lateral inhibition. This process enhances spectral peaks by allowing a positive contribution by components at center of a neighbourhood and a negative contribution by those surrounding the area (Zha92). When a convolution between the kernel and signal spectrum occurs, spectral peaks due to the signal terms are enhanced; cross terms are suppressed.

Figure 7.7a, 7.7b *Kernels in the Dual Correlative Domain for the ZAM Distribution and the Choi-Williams Distribution.*

Figure 7.7a The ZAM Kernel. Kernel cannot be used for multi-component signals that have simultaneously occurring components as their interference components that lie along the plane $\tau=0$ and will be passed.

Figure 7.7b The Choi-Williams kernel. Kernel cannot be used for multi-component signals that have simultaneously occurring components (see ZAM case) or components that are identical as this leads to cross terms that occur along the $\upsilon=0$ plane and will be passed.

Figure 7.7a

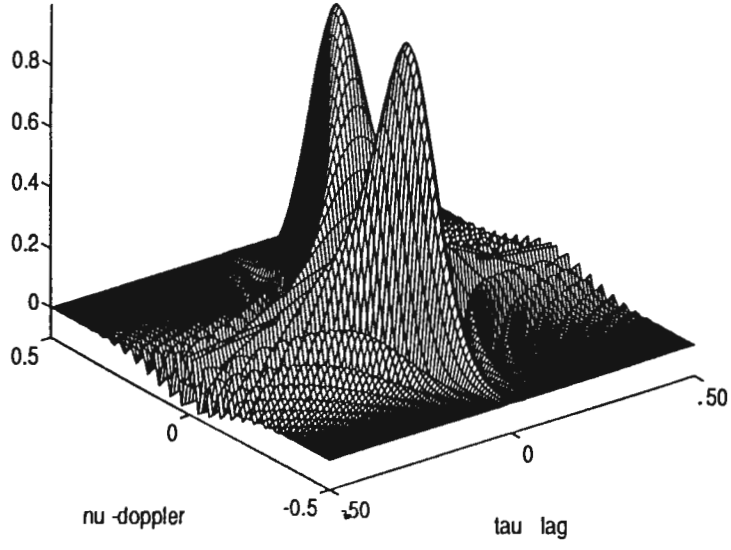
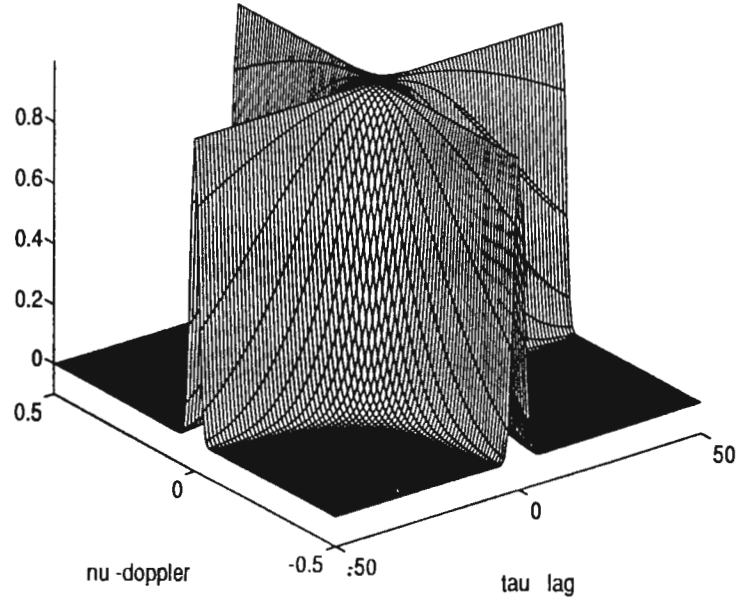


Figure 7.7b



$$g_{ZAM}(\upsilon,\tau) = g_1(\tau)\ \frac{\sin(2\pi\upsilon|\tau|/a)}{\pi\upsilon} \qquad (7.16)$$

Equation 7.16 gives the analytic form of the conic kernel when in the dual correlative domain. The function $g_1(\tau)$ is an arbitrary function generally taken to be 1. The variable a must have a value of 2 or greater to ensure finite time support (Zha92). The resulting TFR does not satisfy the marginals or result in an exclusively non-negative distribution but it does yield an easily understood TFR. The DCD representation of the ZAM transform is shown in figure 7.7a.

Because multi-component signals where two components occur simultaneously have interference terms that lie along the τ axis (figure 7.4c), the ZAM transform is not useful for cross term suppression in this case since the kernel will pass these terms.

7.2.4.3 The Choi-Williams Distribution

Choi and Williams (Cho89) have developed another popular kernel. In the DCD it is analytically defined as follows.

$$g_{CW}(\nu,\ \tau\ ) = e^{\frac{(-2\pi\tau\nu)^2}{\pi\tau\nu}} \qquad (7.17)$$

Figure 7.7b depicts this kernel in the dual correlative domain. By virtue of the same circumstance that yields the ZAM kernel ineffective for multi component signal with simultaneous components, the Choi-Williams kernel is also unusable in this situation. In addition, because a multi-component signal with identical components results in cross terms that lie on the υ axis, the Choi-Williams distribution is problematic for this situation also.

7.2.5 Parameter Selection in Kernel Design

The variables in the kernel formulae control the kernel's spread in the τ and υ directions as well as the roll-off on the sides of the kernel. Not all kernels possess individual control in these dimensions. The two kernels presented above each have only one parameter that will determine the final kernel shape. As seen previously, the proximity of the cross terms to the origin in the DCD varies according to the signal undergoing analysis. Therefore the spread and roll-off parameters must be adjusted accordingly. If the spread is too large, cross terms are retained. If it is too small, signal terms are excluded.

By determining the halfway point between the ambiguity functions origin and the nearest cross term, a cut-off point for the kernel can be identified. Substitution of the values for τ and υ into the kernel's equation and setting it to zero will allow calculation of the parameters controlling a kernel's shape. The height that the cut-off point is above

Figure 7.8a-d *Locating the Cut-off Point for Determining Kernel Parameters.*

Figure 7.8a Wigner Ville Distribution of Two Tonebursts. The Interference term is located in the center of the plot. Signal terms are located at the outside corners.

Figure 7.8b Absolute Value of the Ambiguity Function for 7.8a.

Figure 7.8c Smoothed Version of 7.8b. This is necessary so that the peak search algorithm does not falsely identify the peak of the closest cross term.

Figure 7.8d Top Half of 7.8c. This is now searched for a local maxima not located at the AF's origin. The point halfway between the origin and the maxima indicates the require cut-off point for the kernel. Substitution of the values for τ and υ at this point into the kernel's formula, setting it to zero and solving will yield the values for the parameters controlling kernel shape.

Figure 7.8a

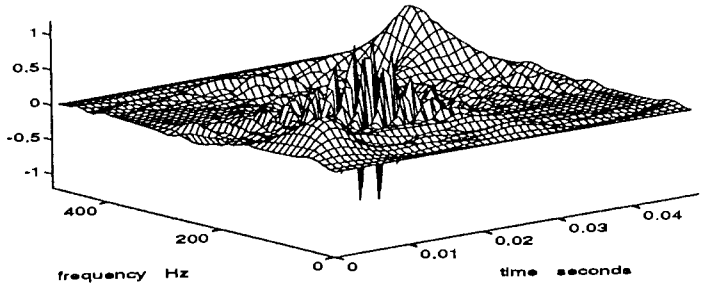


Figure 7.8b

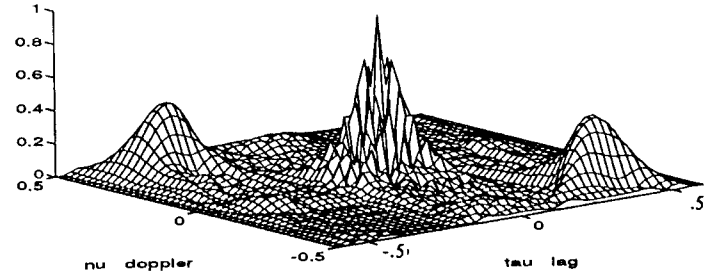


Figure 7.8c

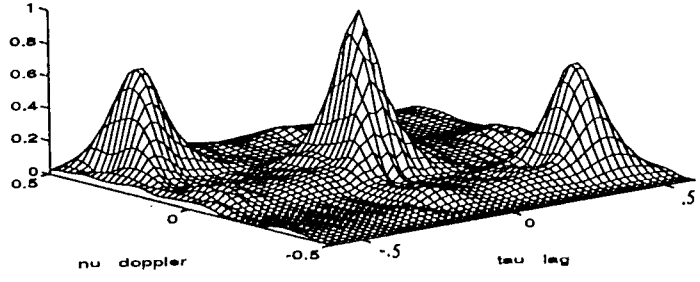
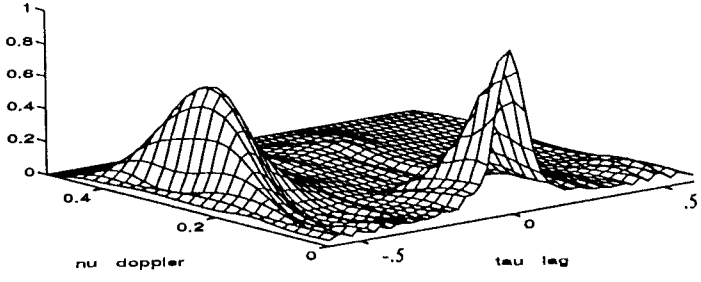


Figure 7.8d



the τv plane gives an indication as to the degree of overlap between the auto and cross terms.

A heuristic approach applied to determining this cut-off point entails taking the absolute value of the top half of the ambiguity function (permissable due to the AF's conjugate symmetry), using a 2-dimensional filter to create smooth sides on the auto and cross terms and then searching for the point closest to the origin that has eight nearest neighbours with values lower than its own. This point is the center of the closest cross term. The cut-off point is determined as the point exactly half-way between the origin and center of the point identified above. This method is appropriate since the largest component of any term, auto or cross, occurs at its center and the procedure is in essence an implementation of the minimum euclidean distance classifier (MED) commonly used in pattern recognition (DUD75). This process is demonstrated in figures 7.8a-d.

7.3 Comparison of the Spectrogram and GTFR Methods

The choice of spectrogram and generalized time frequency methods is directly dependent upon two factors; the required resolution and computational cost. While the spectrogram is incapable of providing high resolution in time and frequency simultaneously, it is incredibly inexpensive computationally.

A comparison is made in the case of analysis of a transient, non-stationary signal in figures 7.9a-d. A linear chirp only 21 samples long is the signal for analysis. The

chirp is decreasing in frequency. Displayed in figure b is the chirp's short time Fourier transform with a window length of 8 samples. Extremely little detail and resolution result. The signal term is located between the two sets of approximately parallel lines. Here, the resolution is 4 frequency bins by 4 time bins. The Wigner-Ville distribution is displayed next (fig. c) and then the ZAM distribution (fig. d). In both of these, the resolution is 21*21 bins. The resulting increase in resolution is significant as it changes from $\pm$ 62.5 Hz and $\pm$ 2.625 ms in the STFT to $\pm$ 12 Hz and $\pm$ 0.0238 ms for the GTFR.

Despite its high degree of resolution, application of the GTFR method is sometimes prevented by its extreme computational overhead. There are six steps involved in the generation of the representation. They are:

1) generate the discrete time instantaneous auto-correlation function
2) Fourier transform to yield the Wigner-Ville Distribution
3) convert the WVD to the dual correlative domain (Ambiguity Function)
4) generate the kernel to be used
5) multiply the kernel and Ambiguity Function
6) convert the result back to the time-frequency plane

In order to give an indication of the computation required for GTFR methods, the STFT and ZAM transform were calculated for a signal of 171 samples. The STFT used a window length of 32 samples and resulting in a resolution of 16x9 requiring 8216 floating point operations (flops). The ZAM transform of the same requires 31107917 flops. The increase in resolution is 203 fold.

Figure 7.9a-d *Comparison of Short Time Fourier Transform and Generalized Time Frequency Methods.*

Figure 7.9a Linear Chirp. The signal is 21 samples in length.

Figure 7.9b Short Time Fourier Transform of Chirp. Resolution is only 4*4.

Figure 7.9c Wigner Ville Distribution of Chirp. Resolution is now has been increased by a factor of 27.

Figure 7.9d ZAM transform of the Chirp. Many of the cross terms have been eliminated. Computational overhead is extremely high (see text).

Figure 7.9a

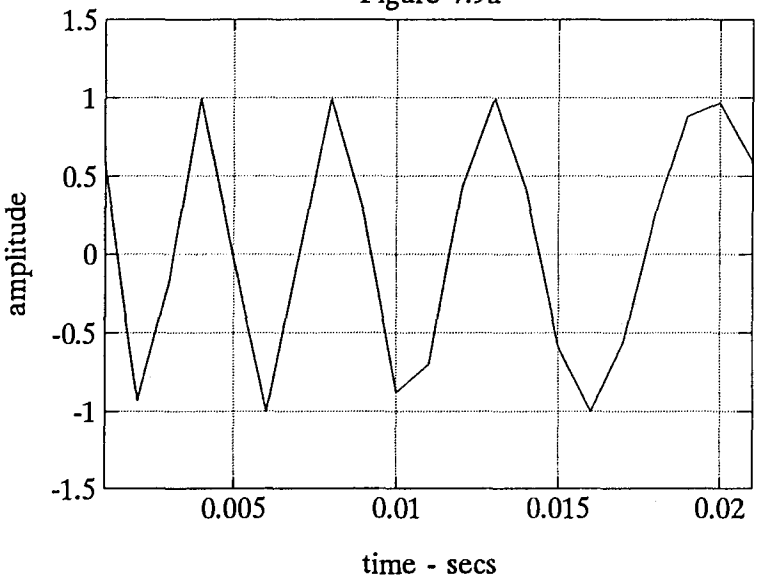


Figure 7.9b

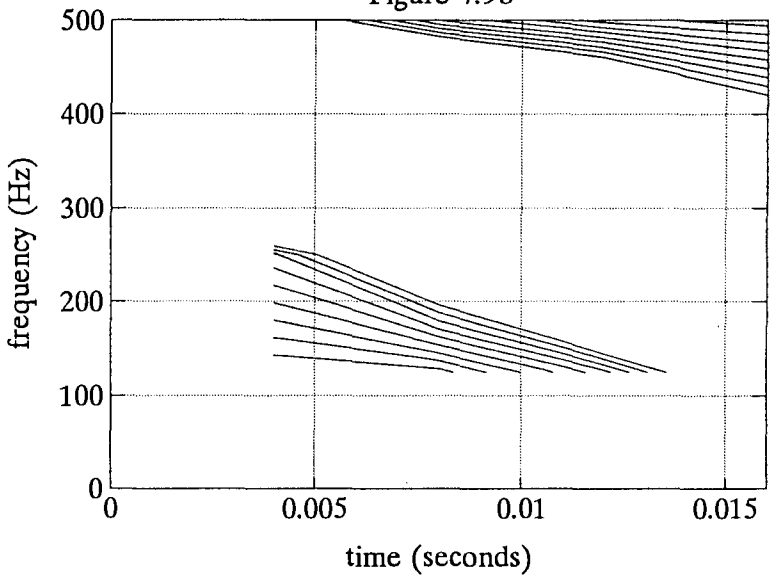


Figure 7.9c

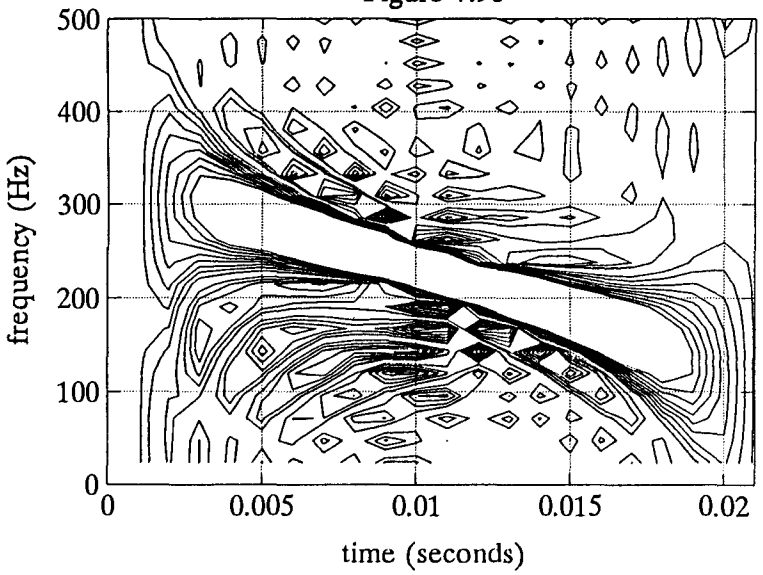
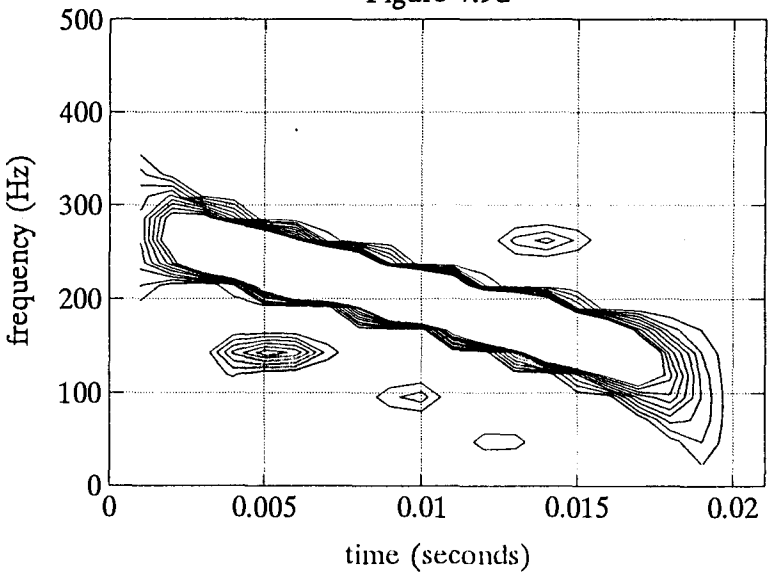


Figure 7.9d



8.0 Other Methods

Contained in this section are two methods that were utilized in the anaylsis of the lung sound signal and do not fall under the two previous categories. The first method, the self-tuning block filter uses a mechanism that not grounded in established theory and is heuristic in its approach. The second method, prediction error filtering has a rich history and has seen application in a variety of problems (Hay91). A brief description of these methods follows and each is explained in greater detail in its respective section.

The self-tuning block filter, was designed in order to combat the wheeze artifact that remained in the non-stationary output of the adaptive line enhancer. It uses the stationary output as a reference to design local filters that operate on windowed segments of the non-stationary output. These band-stop filters are centered on the frequency of the periodic components located in the stationary output of the ALE.

Prediction error (PE) filtering has been used previously in lung sound analysis (Ara91, Ono89). Here, it is developed in both fixed weight and adaptive formats. Fixed tap weight PE-filtering utilizes the signal's correlation matrix (R) to construct a filter that predicts the signal's value one sample into the future. Use of the correlation matrix results in a prediction based on the average statistics of the signal. The prediction value is compared to the actual value and the difference between the two becomes the filter's output. It is postulated that the crackle's contribution to the correlation matrix is minimal and that the PE filter's output will consist solely of the non-stationary, random crackles. Adaptive implementation of the PE filter uses time-varying weights whose values are

adjusted to minimize the prediction error.

8.1 Self Tuning Block Filter

As will be seen in the results section, the adaptive line enhancer does not perfectly isolate the stationary and non-stationary components. Some artifact of the wheeze occurs in the non-stationary output and vice-versa. The primary concern is to remove the wheeze component from the non-stationary output. The inverse problem of removing the crackles from the stationary output is not tackled.

Simple subtraction of a scaled version of the stationary waveform from the non-stationary one is precluded by the phase distortions induced by the ALE's filter. Instead, artifact removal is performed by using the ALE's stationary output as a reference in designing a moving bandstop filter for the wheeze components. Since the wheeze's pitch is subject to some fluctuation, the filter parameters cannot be constant for the entire waveform. To overcome this obstacle, local filters are constructed that operate on windowed sections of the waveform. These sections are then summed to form the filtered non-stationary output. A fundamental assumption required to be true for this technique is that the wheeze component still contains more energy than the crackle. This is a valid assumption to make since the wheeze is narrowband and has its energy much more concentrated than the crackle.

The STFT of the ALE's stationary output yields a matrix of column vectors. The

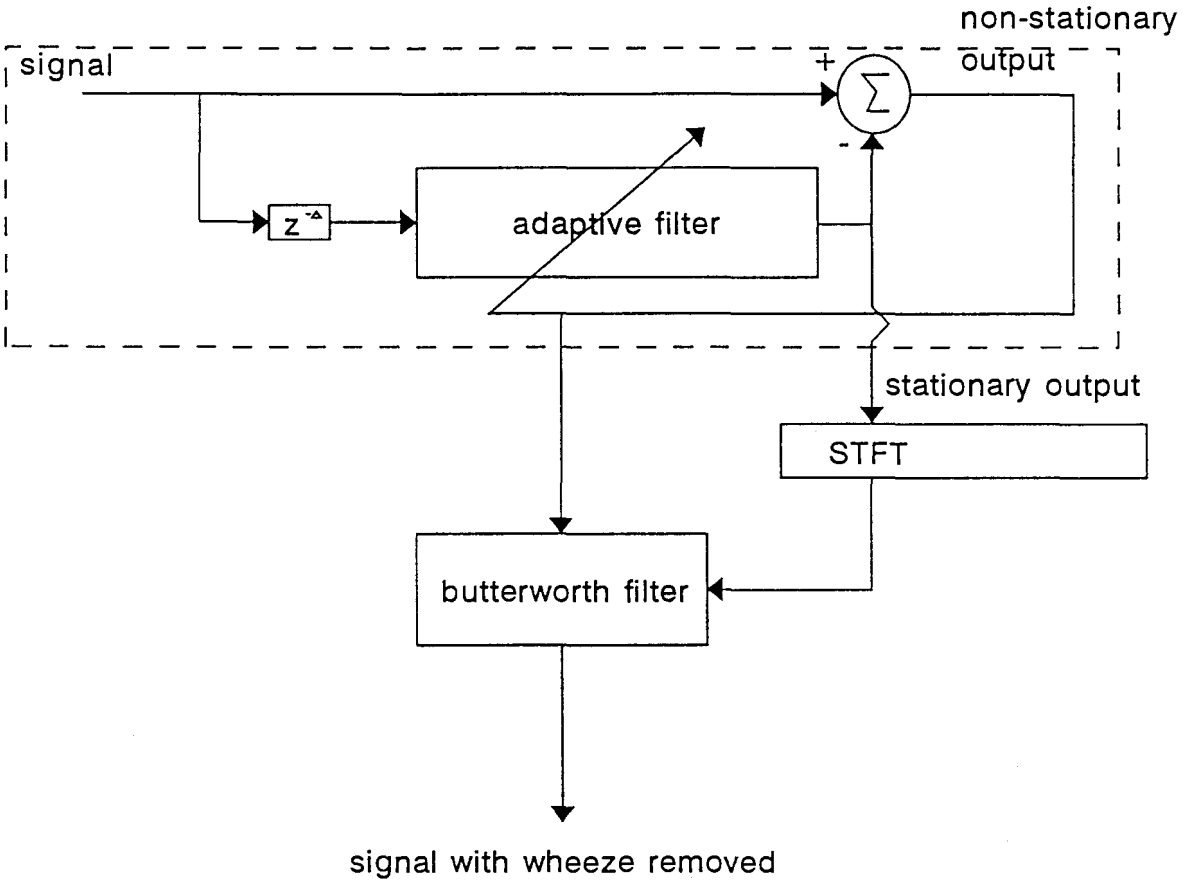


Figure 8.1 *Inner Workings of the Self-Tuning Block Filter.*

The elements enclosed by the broken line comprise the adaptive line enhancer previously discussed in section 6.2. The non-stationary output contains mostly crackle waveforms though some corruption by wheezes still persists. The stationary output is mostly wheeze information and its short time fourier transform (STFT) tells what the center frequency of the butterworth filter should be. The non-stationary output is filtered accordingly and the wheeze artifact is removed.

following procedure is performed as many times as there are columns in the STFT.

1) take a column of the STFT and perform a peak search to locate the frequency component with the highest energy
2) generate a bandstop filter with its center frequency being the peak that was identified above, bandwidth is arbitrarily set to 40 Hz
3) pass the windowed section of the non-stationary output that corresponds to the segment of the stationary yielding the peak above through the bandstop filter

When this process is complete, the filtered segments of the non-stationary output are summed in the order that they were processed to produce the filtered signal. A block disgram of the system is shown in figure 8.1.

8.2 Prediction Error Filtering

Prediction error filtering may be used to detect non-stationarities in the data. A prediction for time (k) is made using the data from time (k-M) to data (k-1). For a fixed parameter system, the predictor takes the form of the transversal filter illustrated in figure 6.1. The predicted value will be denoted f^{M}_{k} where M is the order of the predictor. Given that the filter weights are w_k where $k \in \{1, 2, .., M\}$, the prediction error at time k is given by,

$$f_k^M = u_k - \sum_{n=1}^{M} w_n^T x_{k-n} \qquad (8.1)$$

This may be expressed more succinctly as,

$$f_k^M = \sum_{n=0}^{M} a_n^T \; x_{k-n} \tag{8.2}$$

where the prediction error filter's tap weights a_k are related to the tap weights of the forward predictor by the following relation.

$$a_n = \begin{pmatrix} 1 , & n = 0 \\ -w_n , & n = 1,2,..,M \end{pmatrix} \tag{8.3}$$

A method has been devised to calculate the prediction error filter co-efficients. This is the Levin-Durbin recursion which simplifies to three equations (Hay91). At this point, another subscript is placed on the prediction error filters tap weights so that they are annotated $a_{m,k}$. The m gives the current order of the filter. The equations are applied iteratively giving an increase in the order m each time until the predictor coefficients are given for the desired order M.

$$\Gamma_m = -\frac{1}{P_{m-1}} \sum_{k=0}^{m-1} a_{m-1,k} \; r(k-m) \tag{8.4}$$

$$a_m = \begin{bmatrix} a_{m-1} \\ 0 \end{bmatrix} + \Gamma_m \begin{bmatrix} O \\ a_{m-1}^{B*} \end{bmatrix} \tag{8.5}$$

where the subscripted "B*" signifies that the vector's order is reversed. The third step is

$$P_m = P_{m-1}(1 - |\Gamma_m|^2) \tag{8.6}$$

The recursion is initiated at $P_0=r(0)$ and $a_{m,0}=1$. The values for r(m) come from the lag correlation matrix given by the top line of the correlation matrix (R) (equation 6.7). When iteration is complete. The last M-1 values of $-a_{M,k}$ give the coefficients of the forward prediction.

Rather than using a fixed parameter system, the tap weights may be determined using an adaptive algorithm. In this type of process, the PE filter is a special case of adaptive line enhancement where the decorrelation parameter is set to one. This is illustrated in figure 6.2. Filter weights are adjusted to minimize the prediction error.

Section 3 - Results and Discussion

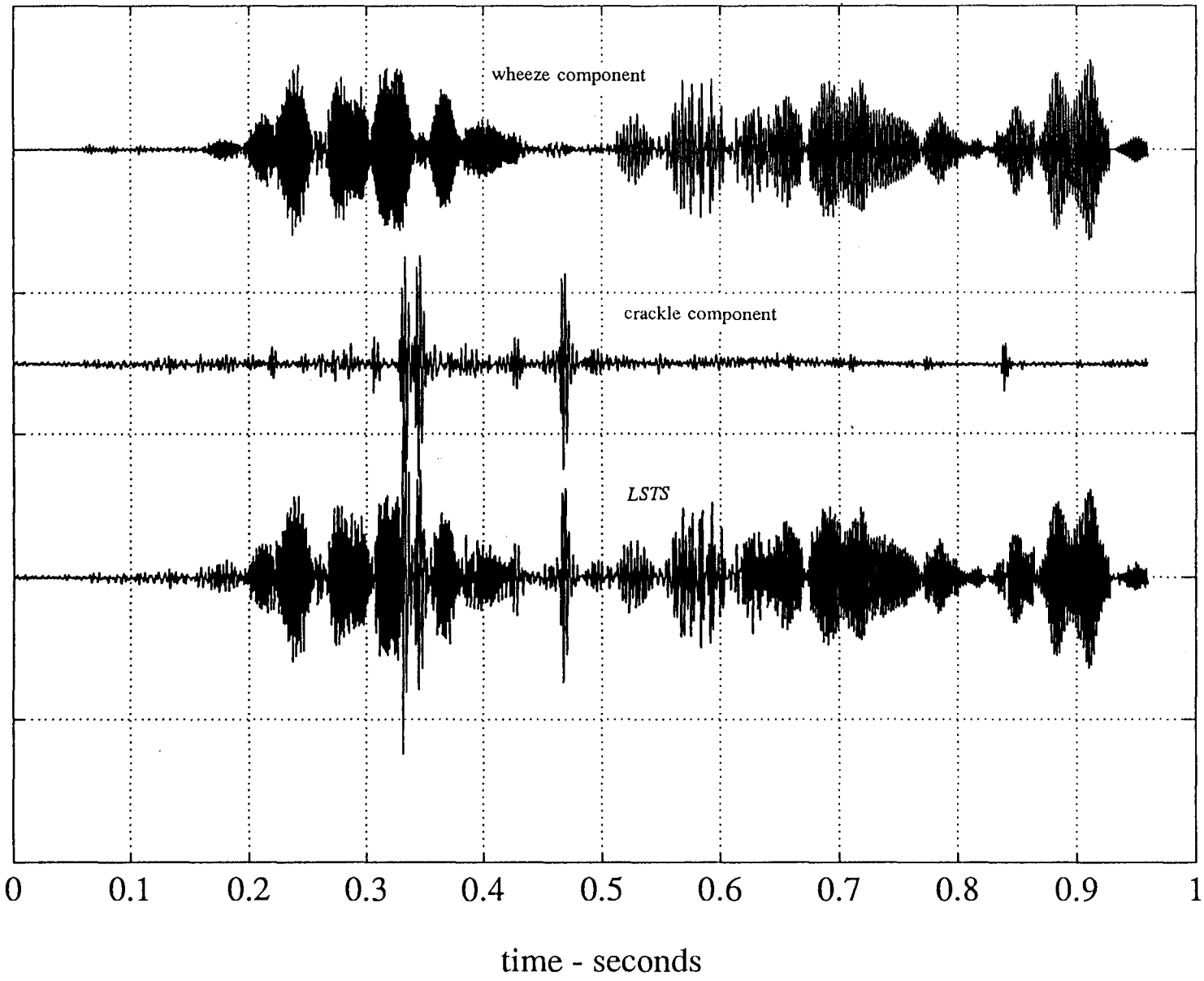
9.0 The Experiments - Introduction

Algorithms for adaptive line enhancement, prediction error filtering and generalized time frequency distributions with the Zhao Atlas and Marks and Choi-Williams kernel were developed. These were implemented on a 386 personal computer in the MATLAB programming environment. The data was derived form teaching tapes (Wil88) that were sampled at 4000 Hz.

Display of the lung sounds is not as much of a problem when crackles and wheezes are isolated. However, significant difficulty arises for the case where both crackles and wheezes are present simultaneously. In order to study the separation of these sounds, a composite waveform was generated with both crackles and wheezes as signal components. Construction of this lung sound test signal (*LSTS*) is depicted in figure 9.1. This signal was created by isolating a single monophonic wheeze from the tape and high-pass filtering it so that the noise was removed (trace 1). The addition of a waveform containing fine crackles (trace 2) made the composite signal (trace 3). The advantage realized in using this test signal is the ability to know exactly the components comprising the signal thus being able to assess the degree of separation.

An extremely important concern is the degree of distortion that a processing algorithm may impart on a signal. This is particularly important as one of the major goals of this project is to provide a method to link lung sounds with physiologic processes. Distortion of the information would inhibit formation of this link. This aspect is considered after the presentation of any filter results.

Figure 9.1 *Construction of Test Signal LSTS*



10.0 Separation

Separation of crackles from wheezes was attempted with prediction error filtering and adaptive line enhancement. The PE filter methods did not perform well and were abandoned after brief investigation. Adaptive line enhancement showed promise particularly when coupled with the self-tuning block filter. Two different adaptive algorithms, the least mean squares (LMS) and recursive least squares (RLS), were tested in driving the tap weight vector and the algorithms' associated parameters were optimized by testing the ratio of wheeze artifact energy to crackle energy. What follows are the details of these experiments.

10.1 Prediction Error Filtering

The problem of extracting the crackles from the composite signal was first tackled with prediction error filtering. Results for application to the test signal *LSTS*, using fixed tap weights and for the adaptive cases using the LMS and RLS algorithms are presented in figure 10.1a,b. Figure a shows the entire signal; pre-processing (trace 1), post PE filtering with fixed weights (trace 2), post PE filtering with the LMS (trace 3) and RLS (trace 4) algorithms. By visual inspection and comparison to the crackle component known to be in *LSTS* (see figure 9.1) , the LMS routine appears to perform best at

removing the wheeze. The enhanced performance in the LMS when compared to the RLS is attributable to the LMS's slower rate of convergence. The LMS algorithm first adjusts the filters weights to be able to predict the wheeze. When it encounters the crackle, it takes some time to adapt to the new statistics. While adjusting, the crackle is separated as prediction error. The RLS routine on the other hand, adapts too quickly and in doing so adjusts the filter to incorporate the crackle's statistics in making its predictions.

Despite the apparent success of the LMS driven PE filter, closer inspection of a segment known to contain two crackles shows significant distortion (figure 10.1b). Arguably, μ (the 'step size' or 'forgetting factor') should have been increased to allow more of the crackle to be separated as prediction error. However, adaptation of the filter weights to those required to accurately predict the wheeze would then take longer and a greater amount of wheeze would show up in the output. Of the three methods, the fixed tap weight PE filter performed the poorest. This is attributable to the rather 'loose' statistics of the signal and local variations in them. Since the statistics of the entire signal are in effect averaged and then used to generate the predictor co-efficients, the many irregularities in the signal cannot be accounted for and appear as prediction error.

10.2 Adaptive Line Enhancement

Two tests were performed on the adaptive line enhancer. The first was to confirm

its operation and suitablity for the separation problem. When suitability was confirmed the ALE was tested on *LSTS*.

10.2.1 Simulation on an Ideal Signal

Initial testing of the adaptive line enhancer was performed on a test signal that mimicked a crackle superimposed on a wheeze. The results of this test are shown in figure 10.2 for both the LMS and RLS case. Each trace shows the one tenth of a second where the simulated crackle occurred. The first trace is the segment of the input test signal which in its entirety, consisted of two seconds of a constant sine wave of 50 Hz with a 15 sample (15 ms) tone burst of 200 hertz superimposed. The next four traces show the results of the ALE process. Ideally, the 50 hertz sine wave should appear at the stationary output and the tone burst at the non-stationary output. The decorrelation parameter was set at 22 samples (22 ms) for both the RLS and LMS algorithms.

Results of the LMS tap weight adjustment are shown in traces two and three. While the stationary output is largely devoid of the tone burst, some distortion exists. In the non-stationary output which should display only the three 200 hertz cycles, distortion occurs while the filter readjusts to the 50 hertz sine wave and the filter's impulse response for the 200 Hz signal is slowly adapted out of the filter weights.

The RLS adaptation in shown in traces three and four. Performance is exceptional. The stationary output shows a sine wave with virtually no distortion, and the

non-stationary output shows perfect extraction of the three 200 hertz cycles. A small amount of distortion does occur but stabilizes very quickly.

The difference in performance can be attributed to the RLS's exploitation of the rigid statistics of the test signal. The pure sine wave yields a correlation matrix that perfectly describes the process and introduction of the tone burst is not enough to upset this framework.

Performance of the LMS routine is highly dependant on the choice for μ. If μ if made too large, the algorithm adapts quickly and forgets the statistics it was previously 'trained' on (the sine wave), and the tone burst is included in the stationary output.

10.2.2 Tests on Lung Sounds

Moving away from ideal conditions, the ALE was used to isolate a wheeze in a noisy lung sound signal taken from (Wil88). Figure 10.3a shows the short time Fourier transform of the original signal. The wheeze is easily identifiable as a long horizontal line. All other structures are noise. Post processing is shown in figure b. Here, the noise has been substantially reduced and the wheeze retained.

10.2.3 Criteria for Assessing ALE Performance

Application of the ALE was made to *LSTS*. The results of varying the parameters controlling the algorithm's performance; the decorrelation parameter, the number of taps and μ (in the LMS case) were studied to see their effect on the ALE's performance. In order to assess the filter's performance, signals were normalized to a local 'landmark'. All non-stationary output were scaled so that the crackle at approximately 0.475 seconds had a peak amplitude of one.

Performing this normalization allows filter performance to be monitored using the assumption that the contribution to the signal's energy by the crackle will be the same from signal to signal and that the only change in energy between signals will be due to the contribution of the wheeze artifact. In this sense, optimal parameters are obtained by determining values that create a minima in the energy for the non-stationary output. This is the point where wheeze artifacts are at their lowest possible level with respect to the parameter being varied. Conversely, the stationary outputs were normalized so that the wheeze structure at 0.92 seconds had a peak amplitude of one. This allowed the contribution to the stationary output made by crackle artifact to be measured.

10.2.4 Tests on *LSTS*

The ALE (LMS) was applied to *LSTS*. The first parameter to be adjusted was the

decorrelation. A minima in the non-stationary signal's energy occurred when the decorrelation parameter was set to 8. The non-stationary and stationary outputs are shown in figures 10.4 and 10.5. Careful scrutiny of figure 10.5 shows that the wheeze artifact is indeed smallest in trace 3 where the decorrelation parameter is set to 8. This is confirmed in figure 10.10a which shows a minima for the signal energy for the normalized non-stationary output when plotted against decorrelation.

This figure is the result of several factors. The first and the one that guides the baseline of the non-stationary output trace is the amount of similarity that occurs between the delayed and non-delayed signal. A maximum decorrelation of 25 was attempted. With respect to the wheeze, the signal is stationary over this period and a decorrelation anywhere in this range will have little affect on the output. The experiment confirms this as the energy changes little over this period. The local minima at 8 is related to the frequency of the wheeze. It is at this point the the filter becomes matched in the 'matched filter' sense to the wheeze and a slightly better performance is realized.

The next test of the LMS (ALE) came in studying the effect of μ on the filter's output. In figure 10.7, as μ is increased, more wheeze appears in the stationary output and impulsive type components appear. The larger step sizes taken in adjusting the tap weight vectors by using a high value for μ resulted in large filter adjustments being made at each point. These jerky modifications become very prevalent at high values of μ. They are visible between the two large crackles around 0.33 seconds in trace 5 ($\mu = 1$) in figure 10.6. Although figure 10.10b shows that the energy in the non-stationary output decreases and the energy in the stationary output increases, this is not due to an

increasing separation between the crackle and wheeze. The decreasing non-stationary energy is due to the greater speed of adaption which places more signal components, including crackles, in the stationary (common component) output.

The greater the number of tap weights used, the better were the results. This phenomena is easily attributed to the increase in information that the algorithm can use in order to adjust the weights. Figures 10.8 and 10.9 show the filter outputs and figure 10.10c shows the signal energies for an increasing number of tap weights. No substantial improvement occurs above 10 tap weights.

The *LSTS* was passed through the adaptive line enhancement procedure where the filter was controlled by the RLS algorithm. Similar results were noted as for the LMS case. The optimum decorrelation parameter was shown to be 6 which is close to the value in the LMS case (figures 10.11, 10.12, 10.15a).

No forgetting factor analysis was performed since the algorithm is extremely sensitive to adjustments to this parameter and anything less than a value of .999 resulted in a great degree of error. A similar relationship as in the LMS case was noticed for increasing the number of tap weights (figures 10.13 and 10.14).

10.2.5 Distortion

As mentioned previously one of the major considerations for the signal processing is the degree of distortion a procedure imparts on the signal. From previous observations,

it is noted that the parameter that affects the adaptive line enhancement process the greatest is the decorrelation. *LSTS* was run through the RLS and LMS ALE algorithms in order to determine the best value for the decorrelation parameter in terms of the amount of crackle distortion. It was found that the decorrelation parameter had to be at least equal to the number of samples that were required to represent the crackle otherwise, the latter part of the crackle becomes distorted.

This distortion occurs at low decorrelation since the latter part of the signal is determined to be a common component when compared to the early part of the signal and hence is shunted to the stationary output. Figure 10.16 shows several runs of the LMS and RLS ALE algorithms where the decorrelation parameter is 8, 15 and 30 samples for both. In this series, figure a shows a crackle before any processing takes place. For the other figures the title bar relays the following information. The decorrelation parameter is x in dcp=x. If the figure is the result of the RLS driven ALE algorithm '(RLS)' appears after the decorrelation parameter. Otherwise, the figure is the result of LMS driven ALE and the value for μ is given. All runs used 15 tap weights. After the filtering with the ALE, crackles should bear the same information content as in figure 10.16a. Clearly, figures 10b-f do not. Only figure 10.g-h bear resemblance. The time frequency representations for these eight signals are shown in figure 10.17a-g. As for the time series, figure h best resembles the original waveform. For further discussion of the use of time-frequency representations, see section 11.

As the RLS and LMS algorithms perform with about the same degree of distortion at a decorrelation of 30, the next criteria is the amount of attenuation of the wheeze

waveform. Figure 10.18 shows, in succession, the test waveform, the error output from the RLS and LMS routine. By visual inspection, the LMS routine performs better at isolating the crackles. Reasons for this are the same as those given for its superior performance in PE filtering.

10.3 Self-Tuning Block Filter

The self tuning block filter was applied to the results of the LMS algorithm and results are presented in figure 10.19 The five traces from top to bottom are; the test signal, stationary output of the LMS ALE, non-stationary output of the LMS ALE, the non-stationary output filtered by the self-tracking block filter process according to the information in the stationary output and finally, the original crackle waveform for comparison.

Despite the good performance of this process, some inherent difficulties exist. Since each block is processed individually by a filter with different parameters, each block and hence the resulting signal will have different phase relationships throughout. However, this may not be a problem for the individual study of crackles.

The bandwidth of the local filters was set at 40 Hz. If pursued this method would require a method of varying the filter's bandwidth in order to accommodate for variations is wheezes. This could be performed by monitoring the bandwidth of the stationary output and designing the local filters appropriately.

10.4 Confounding Factors

After viewing the results, it is felt that the methods suffered somewhat from the type of signal they were tested on. The sound taken from the tapes (Wil88) were almost certainly recorded without any monitoring of flow information. This is evidenced by the sharp drop in frequency of the monophonic wheeze in figure 10.3. Had the flow rate been kept constant, the pitch would not exhibit the large changes in frequency and or the 'choppiness' exhibited in figure 9.1. The elimination of these two problems would have facilitated better performance on behalf of the PE filter due to more stationary statistics. Performance of the ALE would also have been enhanced.

Figure 10.1a *Prediction Error Filtering of LSTS*

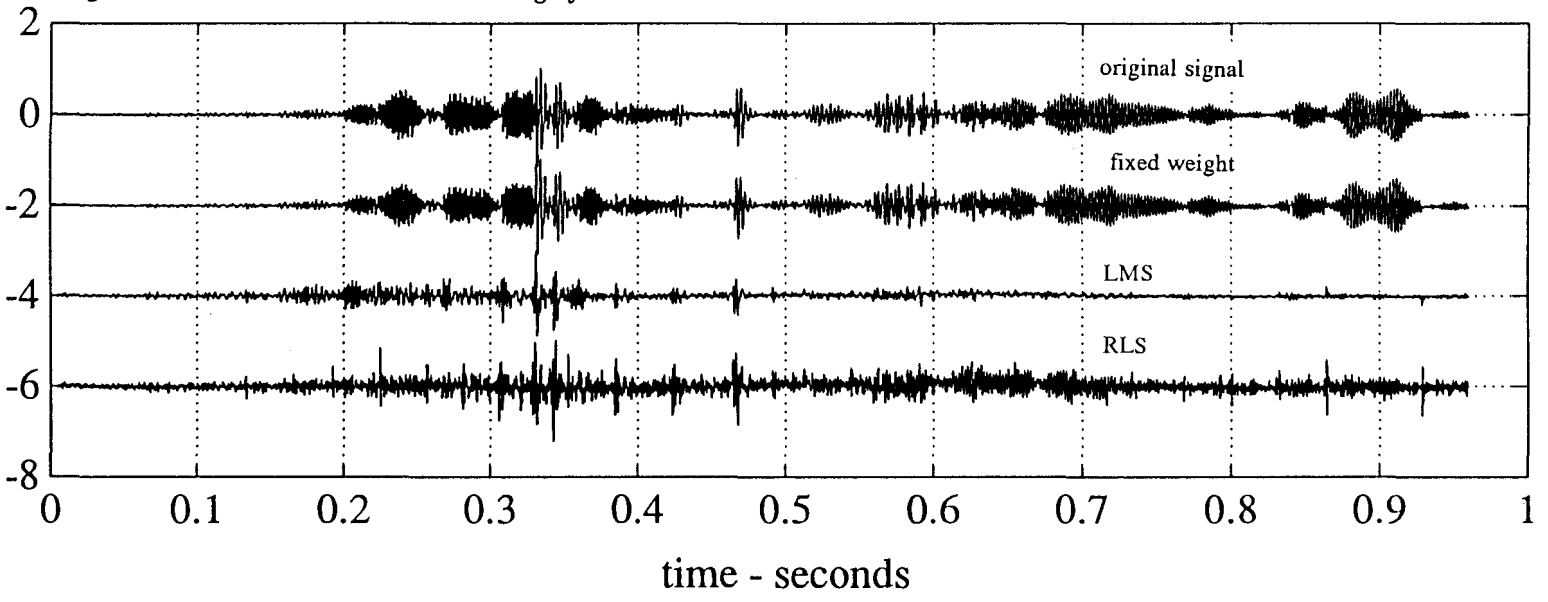


Figure 10.1b *Prediction Error Filtering of LSTS - Zoomed View of Two Crackles*

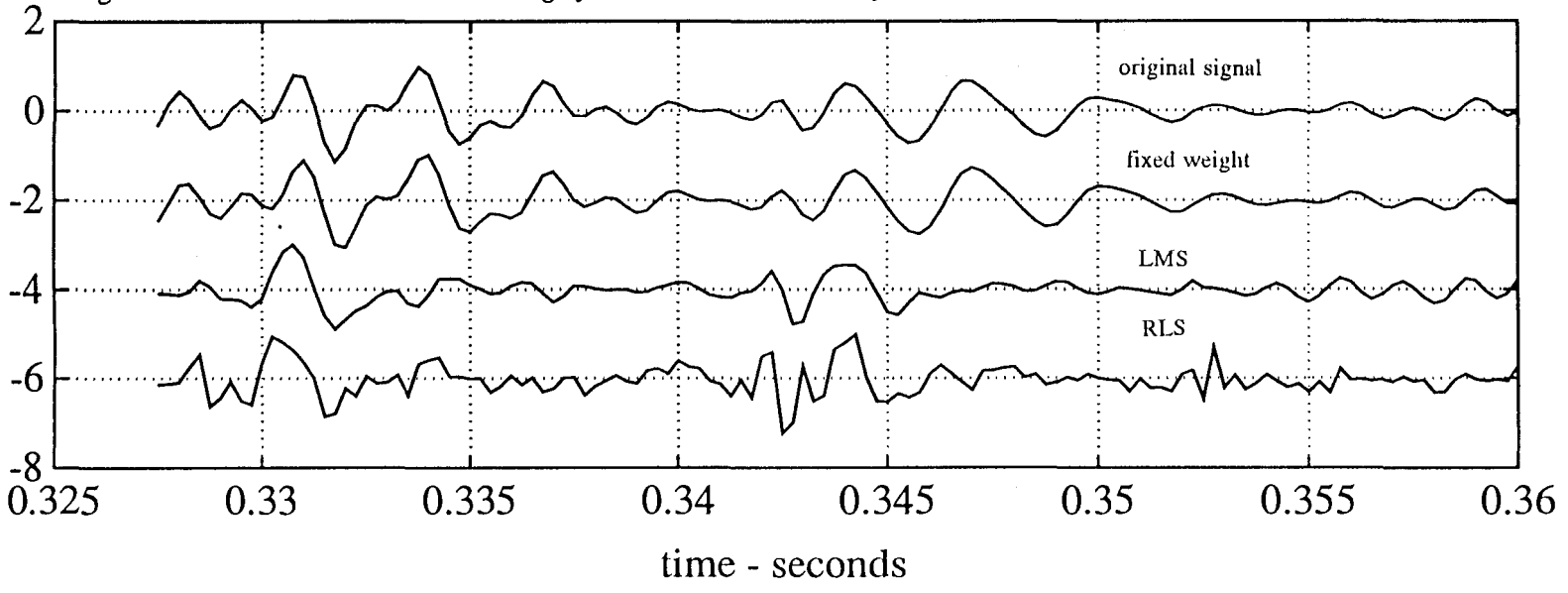


Figure 10.2 *Test of RLS and LMS ALE on Ideal Signal*

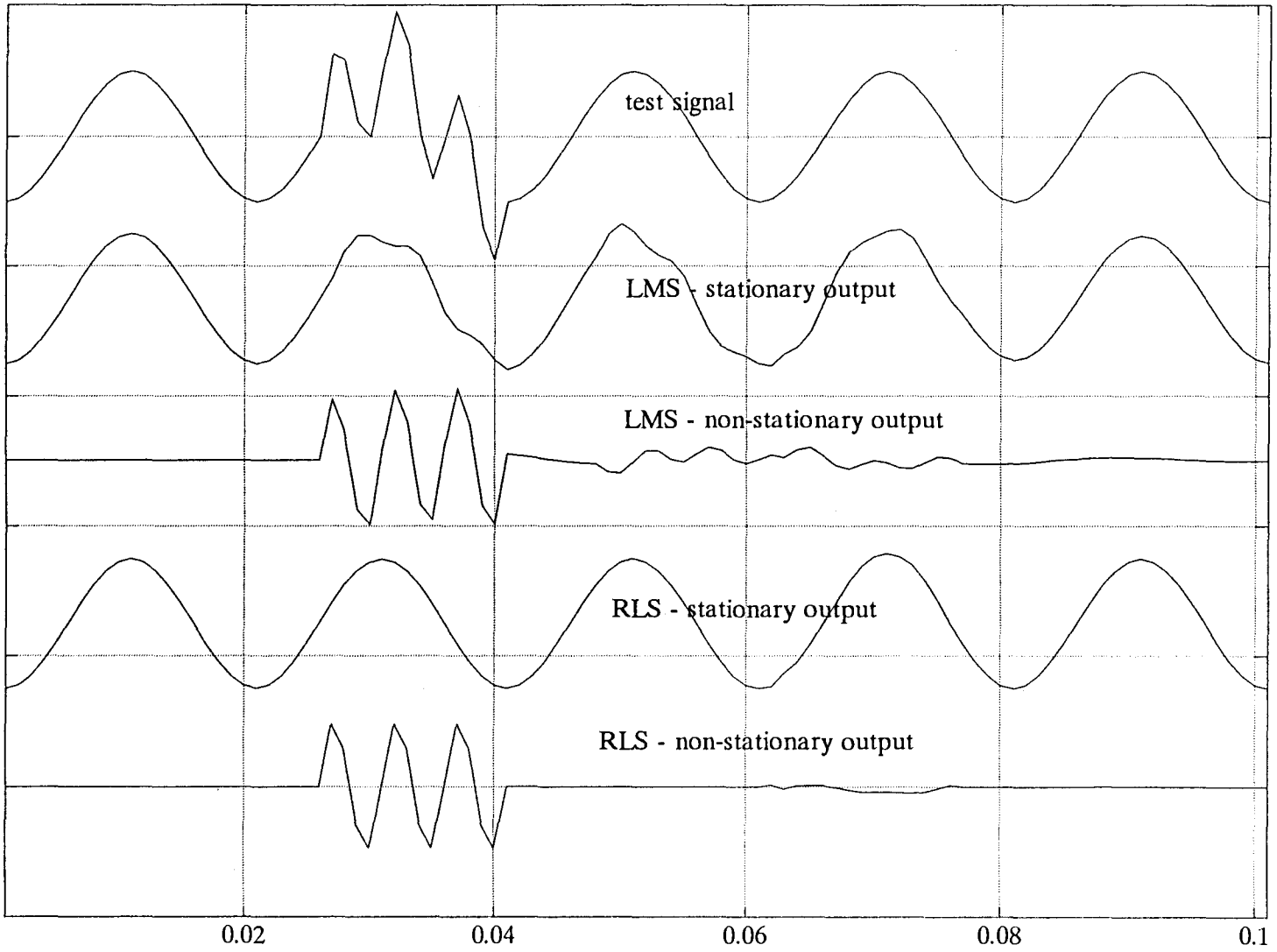


Figure 10.3a *STFT of Monophonic Wheeze*

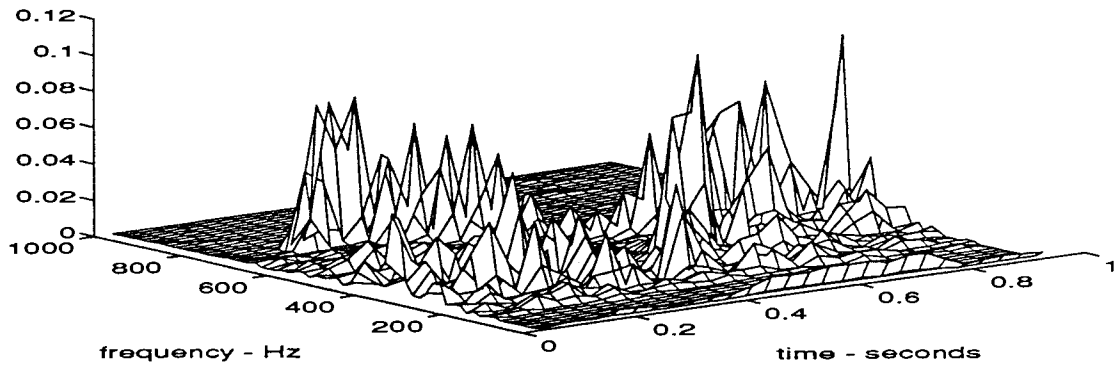


Figure 10.3b *STFT of Monophonic Wheeze after Adaptive Line Enhancement*

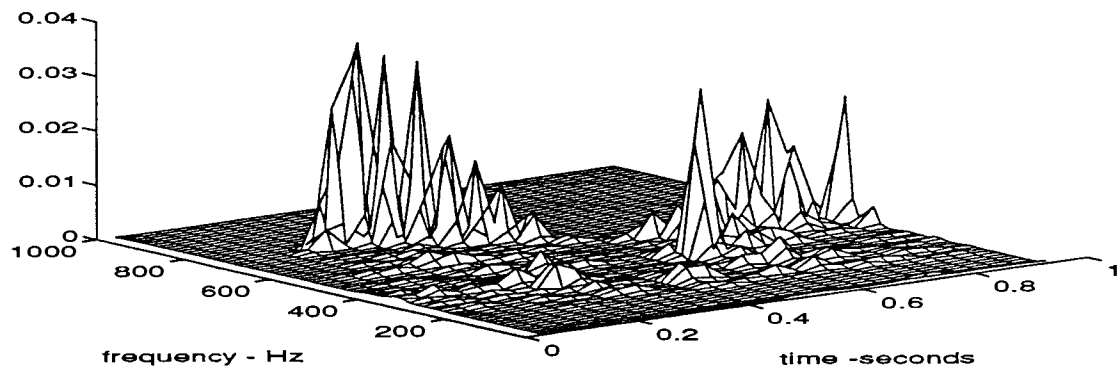


Figure 10.4 *Non-Stationary Output Variable Decorrelation (LMS)*

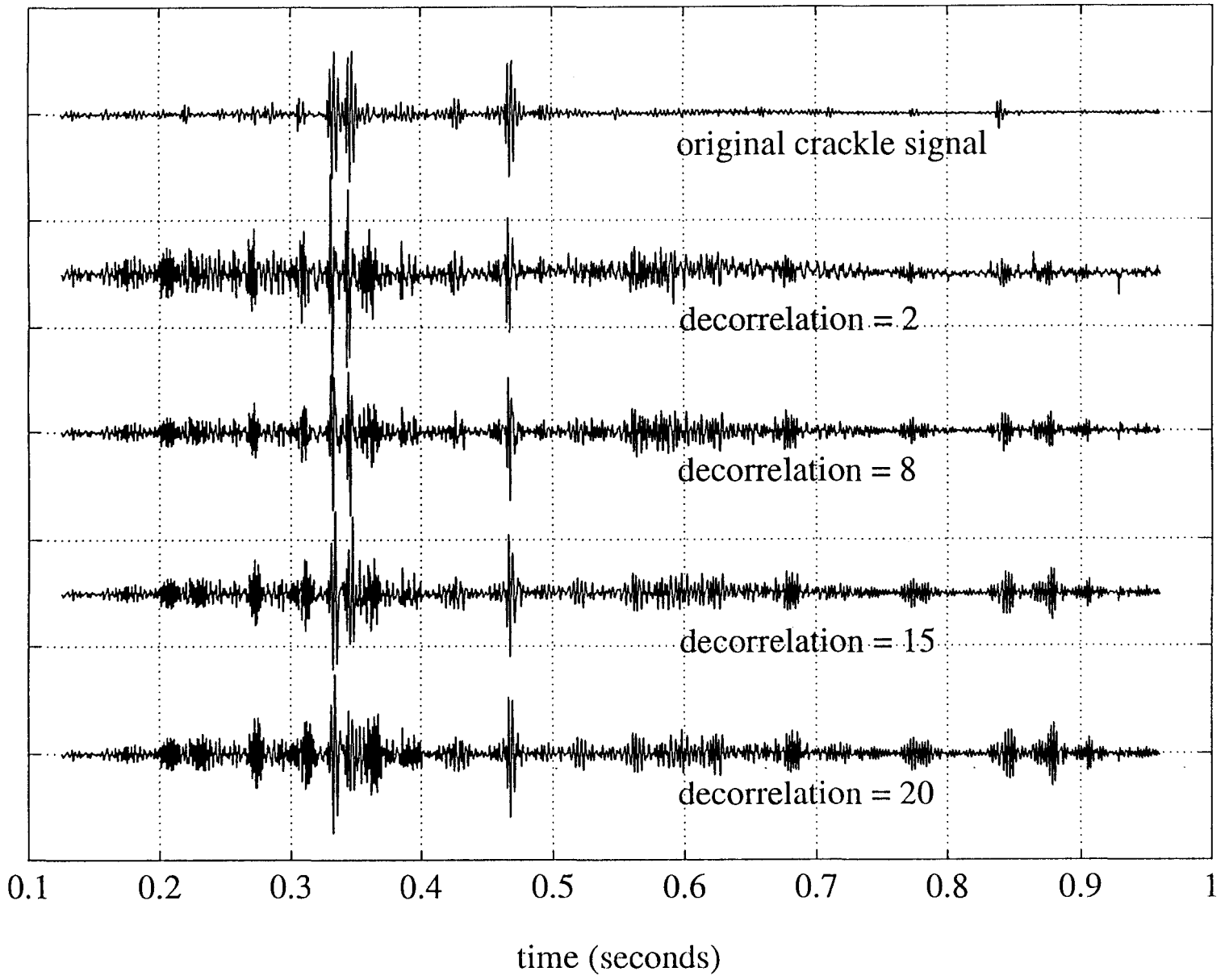


Figure 10.5 *Stationary Output Variable Decorrelation (LMS)*

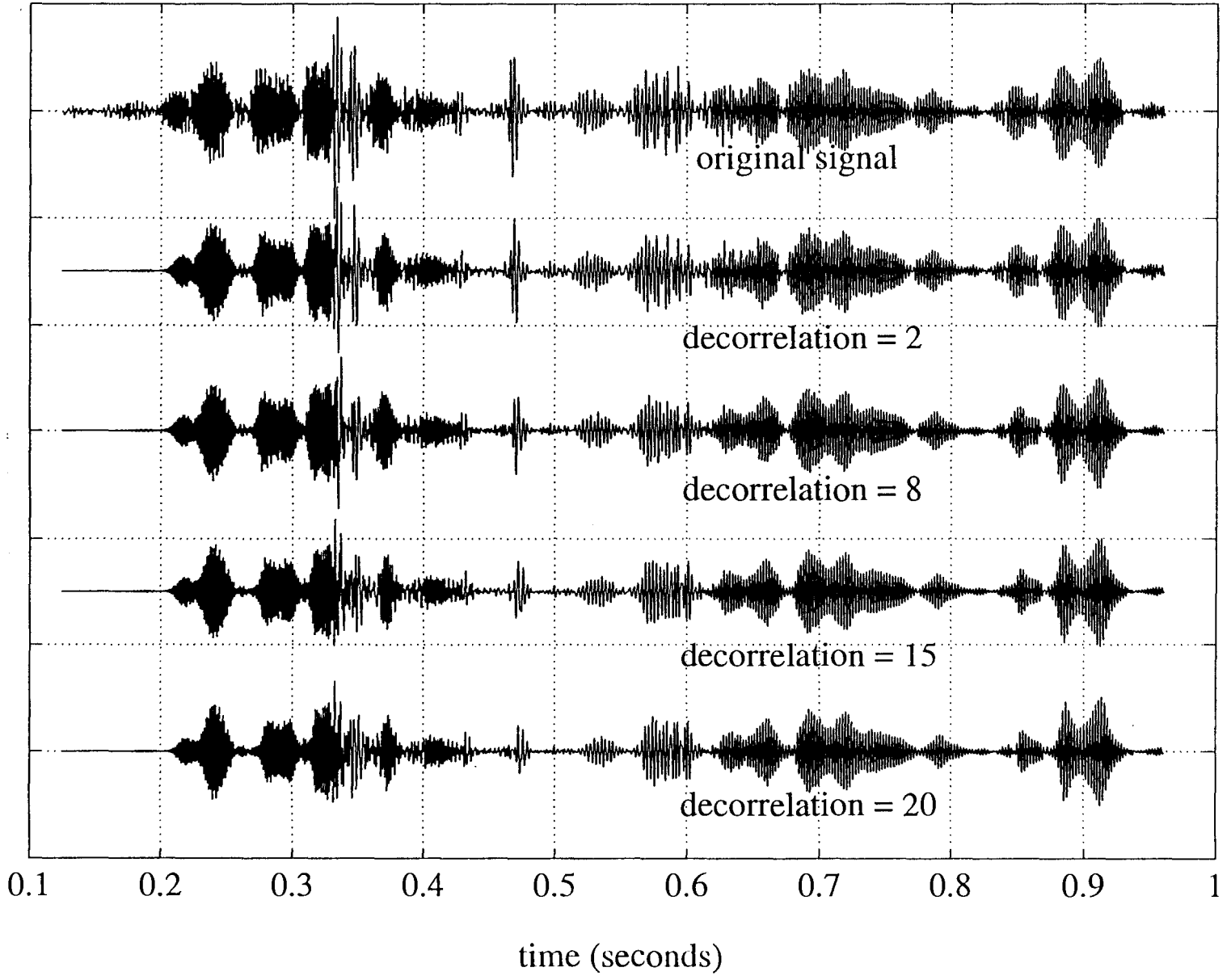


Figure 10.6 *Non-Stationary Output Variable μ (LMS)*

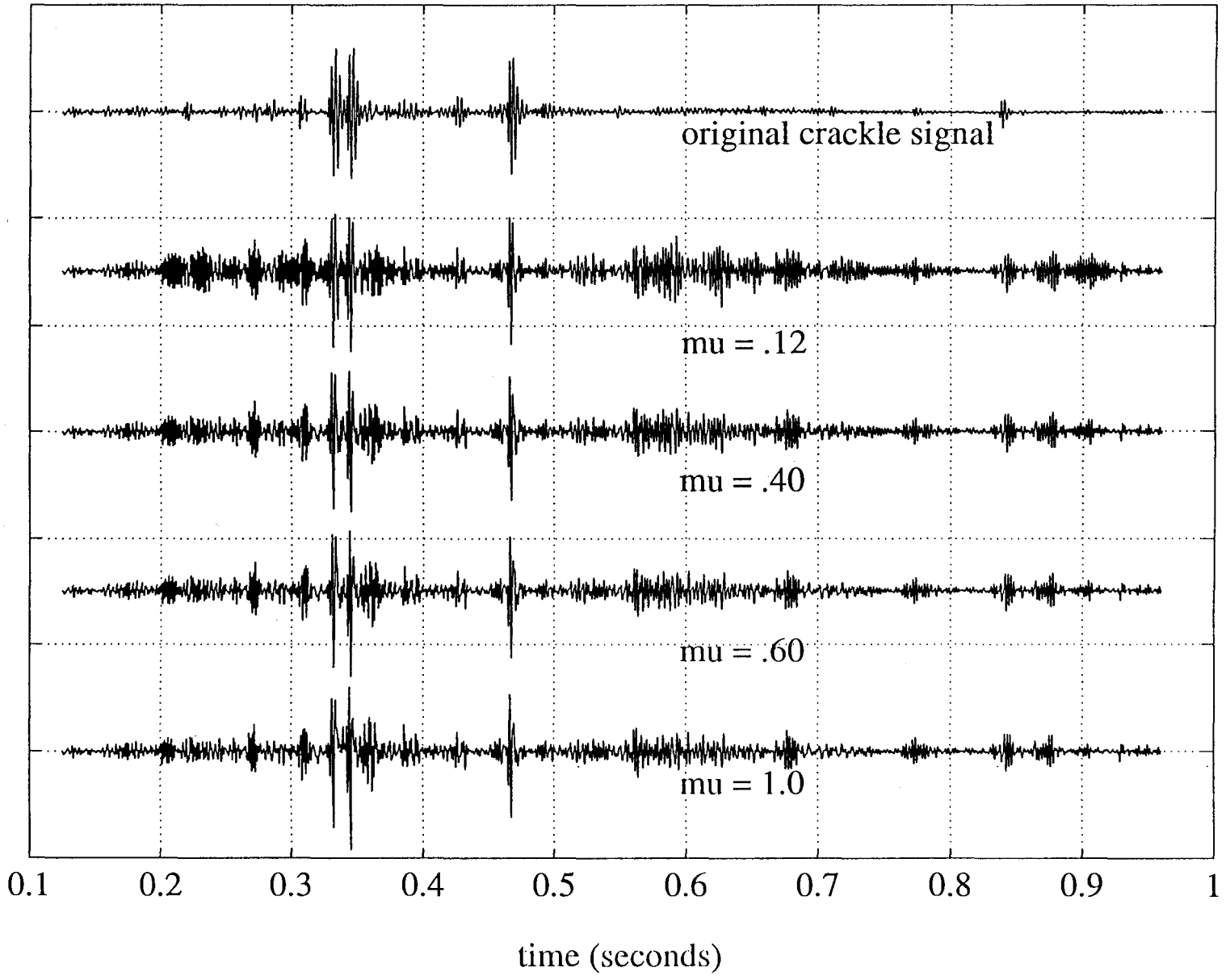


Figure 10.7 *Stationary Output Variable μ (LMS)*

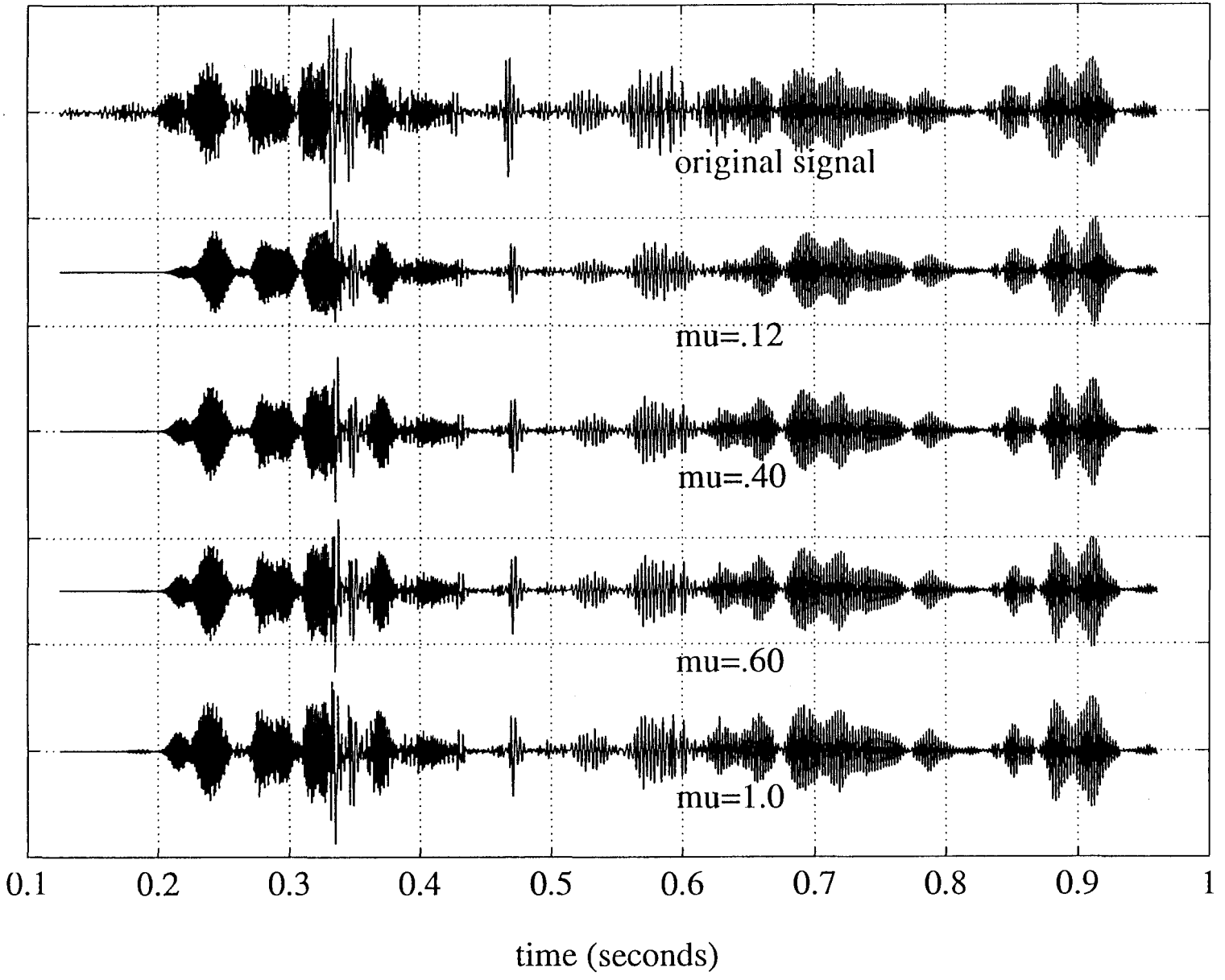


Figure 10.8 *Non-Stationary Output Variable # of Taps (LMS)*

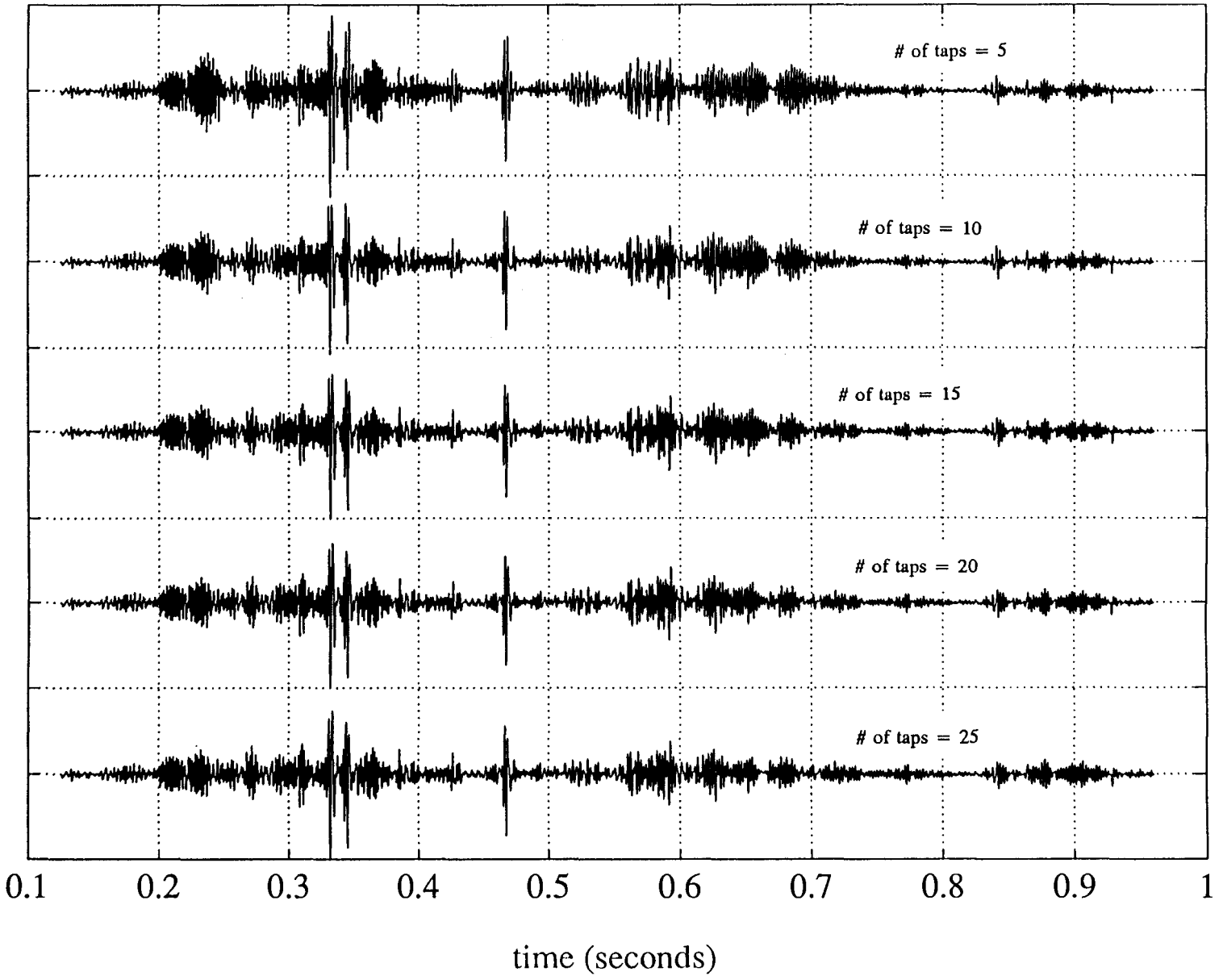


Figure 10.9 *Stationary Output Variable # of Taps (LMS)*

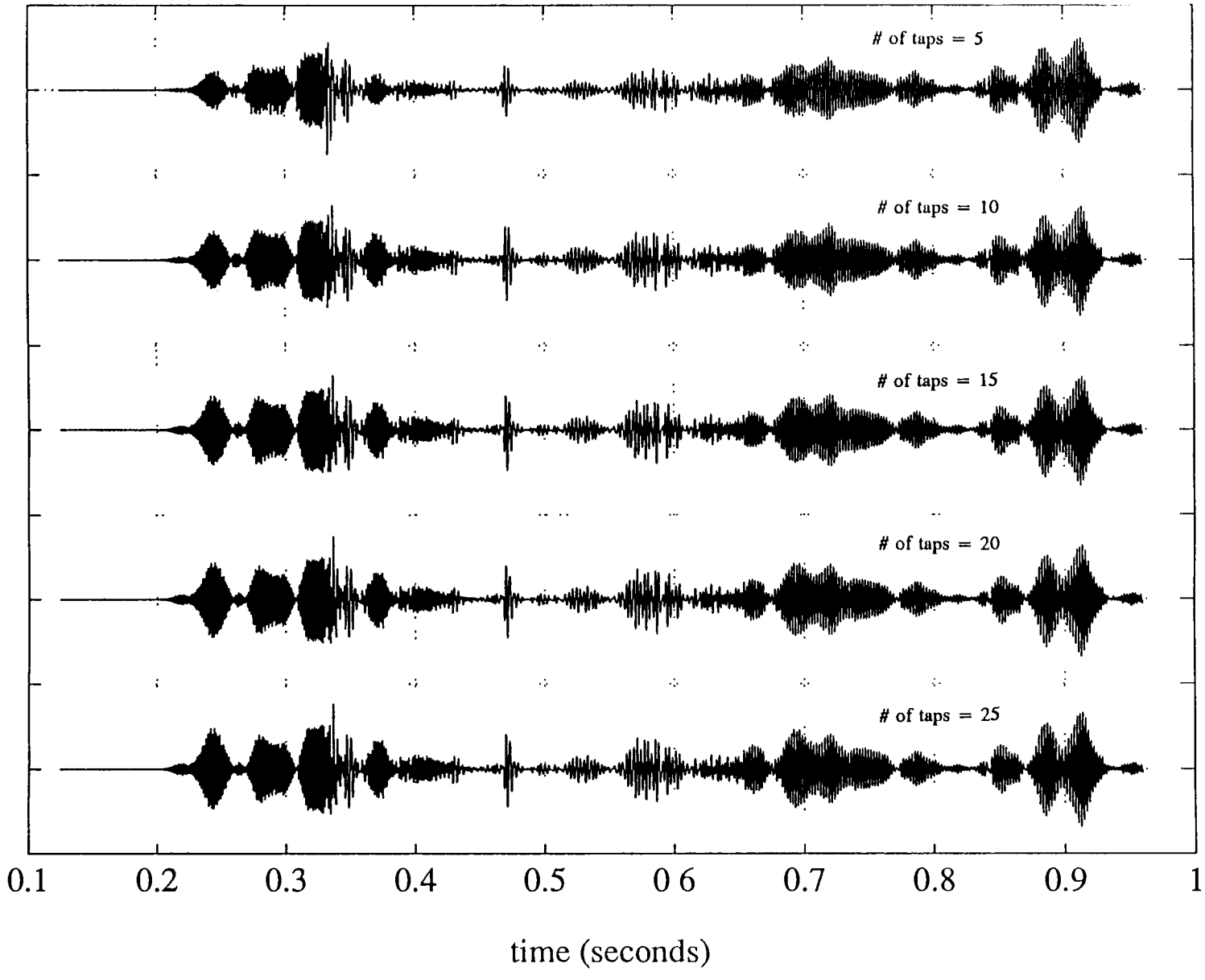
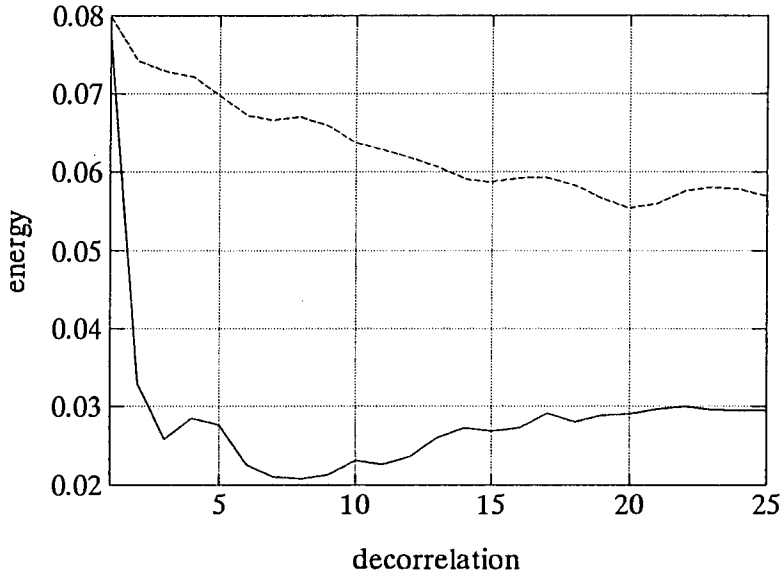
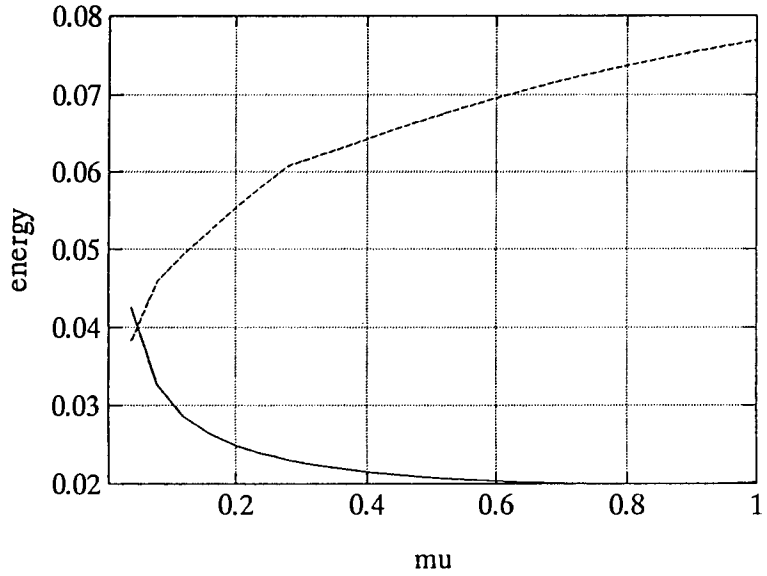


Figure 10.10 *Optimal Parameter Selection for ALE (LMS)*

10.10a Effect of decorellation
10.10b Effect of mu
10.10c Effect of number of tap weights

solid line - non-stationary (error) output
broken line - stationary (filter) output





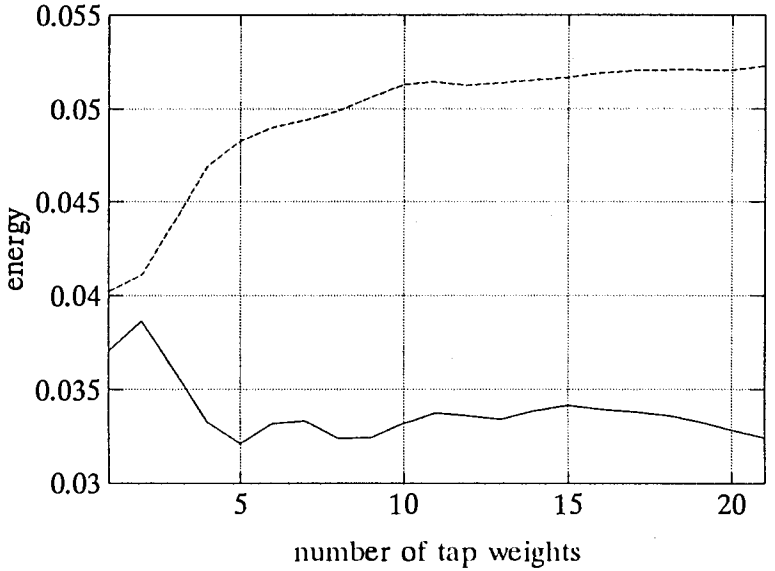


Figure 10.11 *Non-Stationary Output Variable Decorrelation (RLS)*

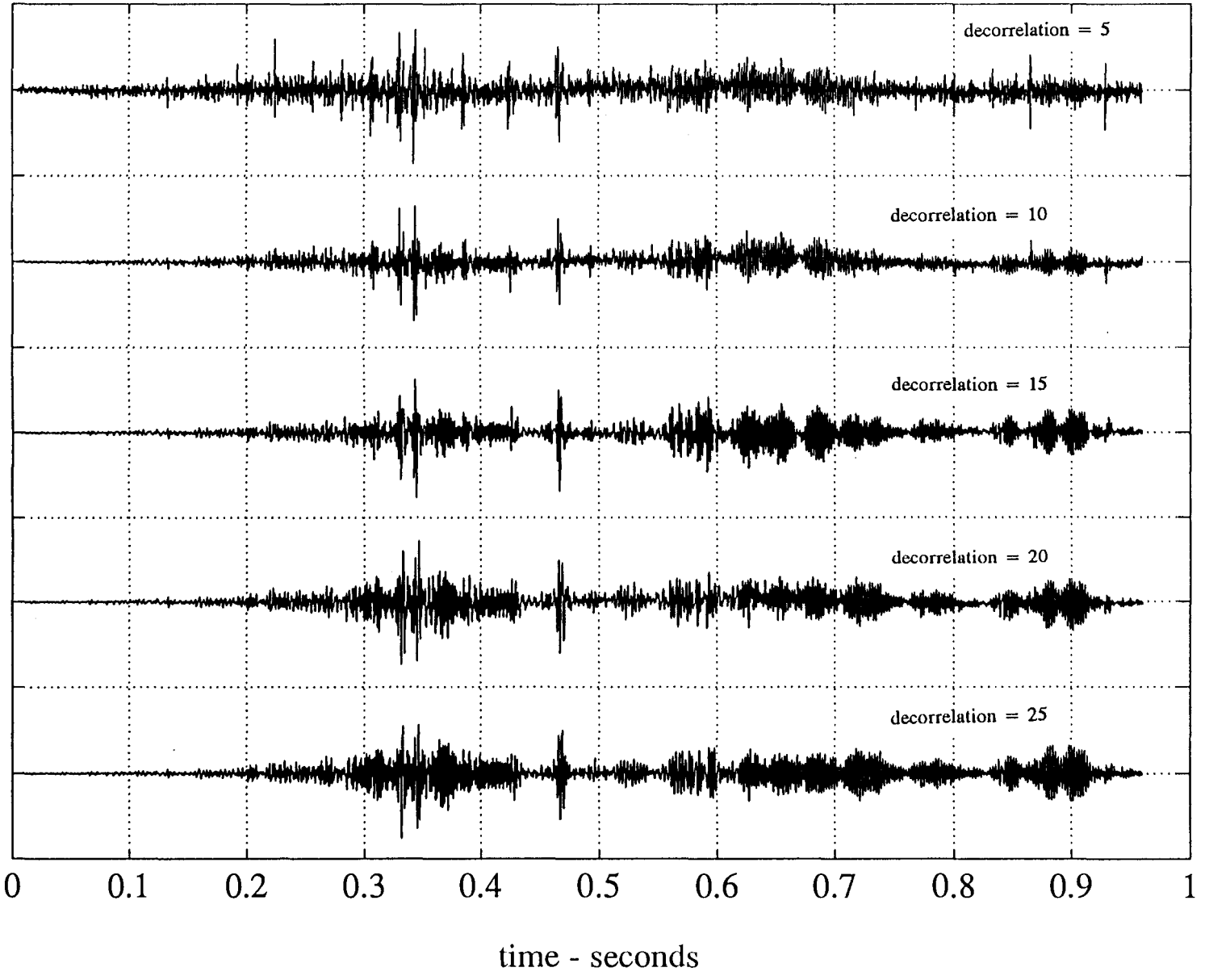


Figure 10.12 *Stationary Output Variable Decorrelation (RLS)*

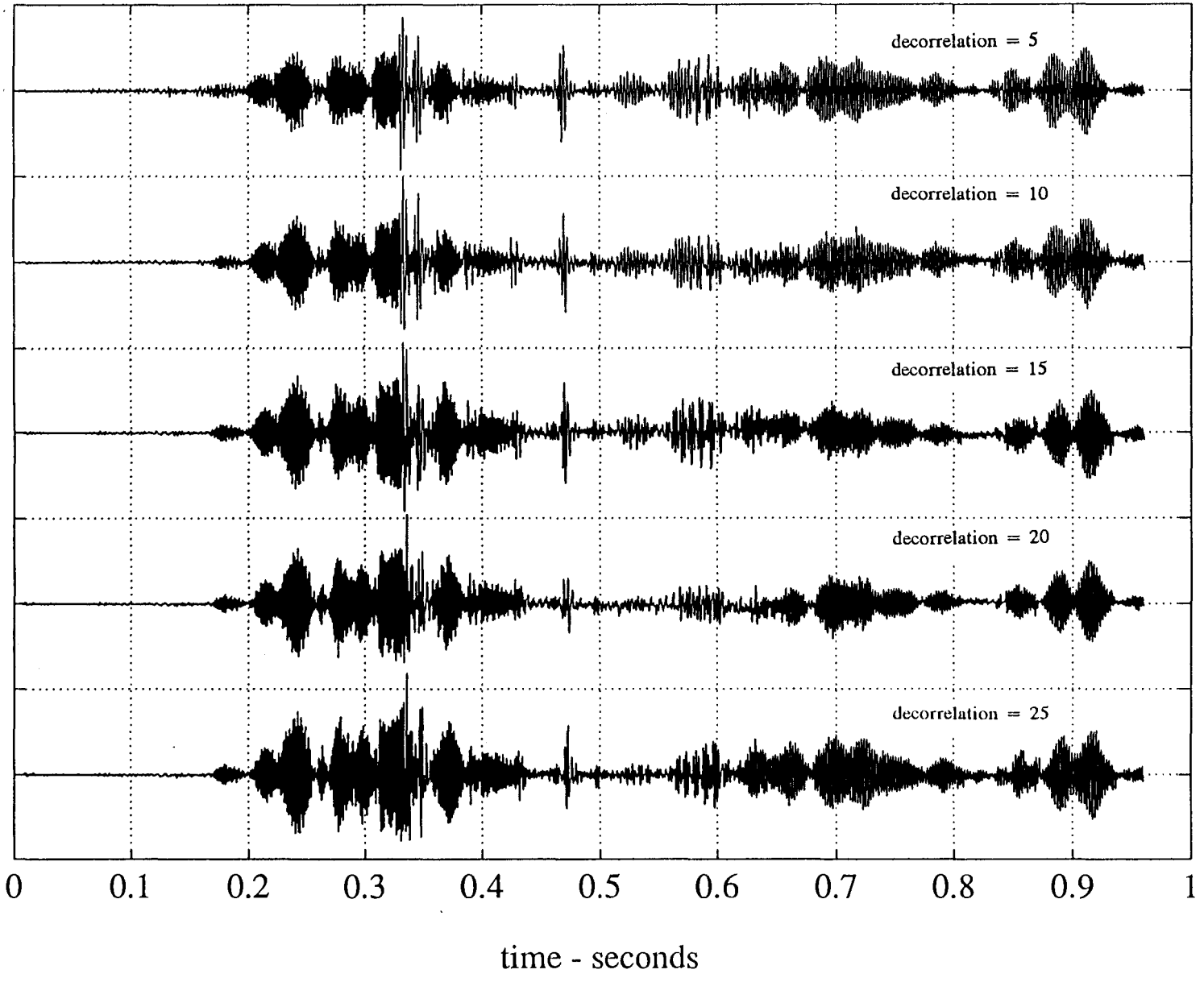


Figure 10.13 *Non-Stationary Output Variable # of Taps (LMS)*

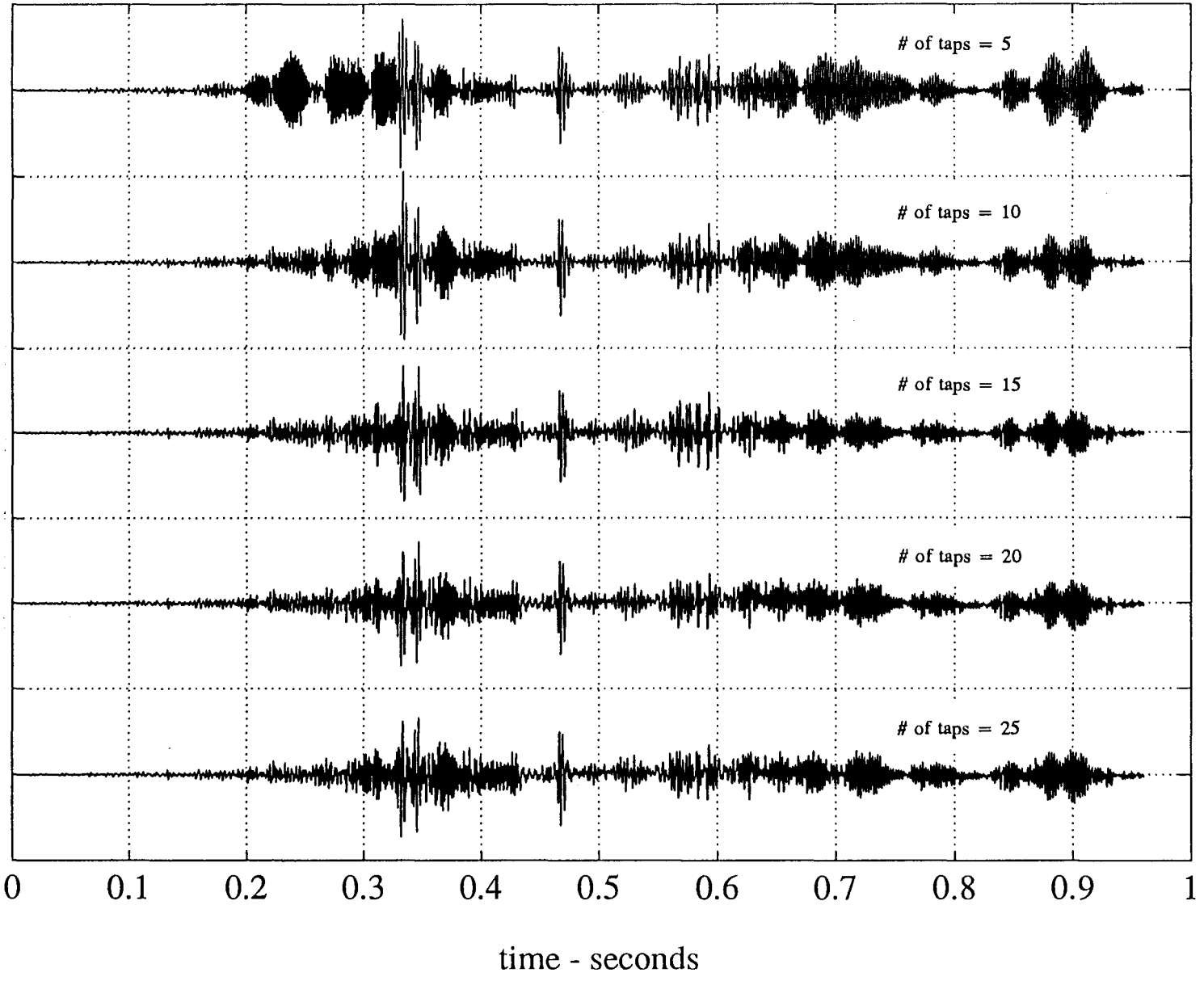


Figure 10.14 *Stationary Output Variable Decorrelation (RLS)*

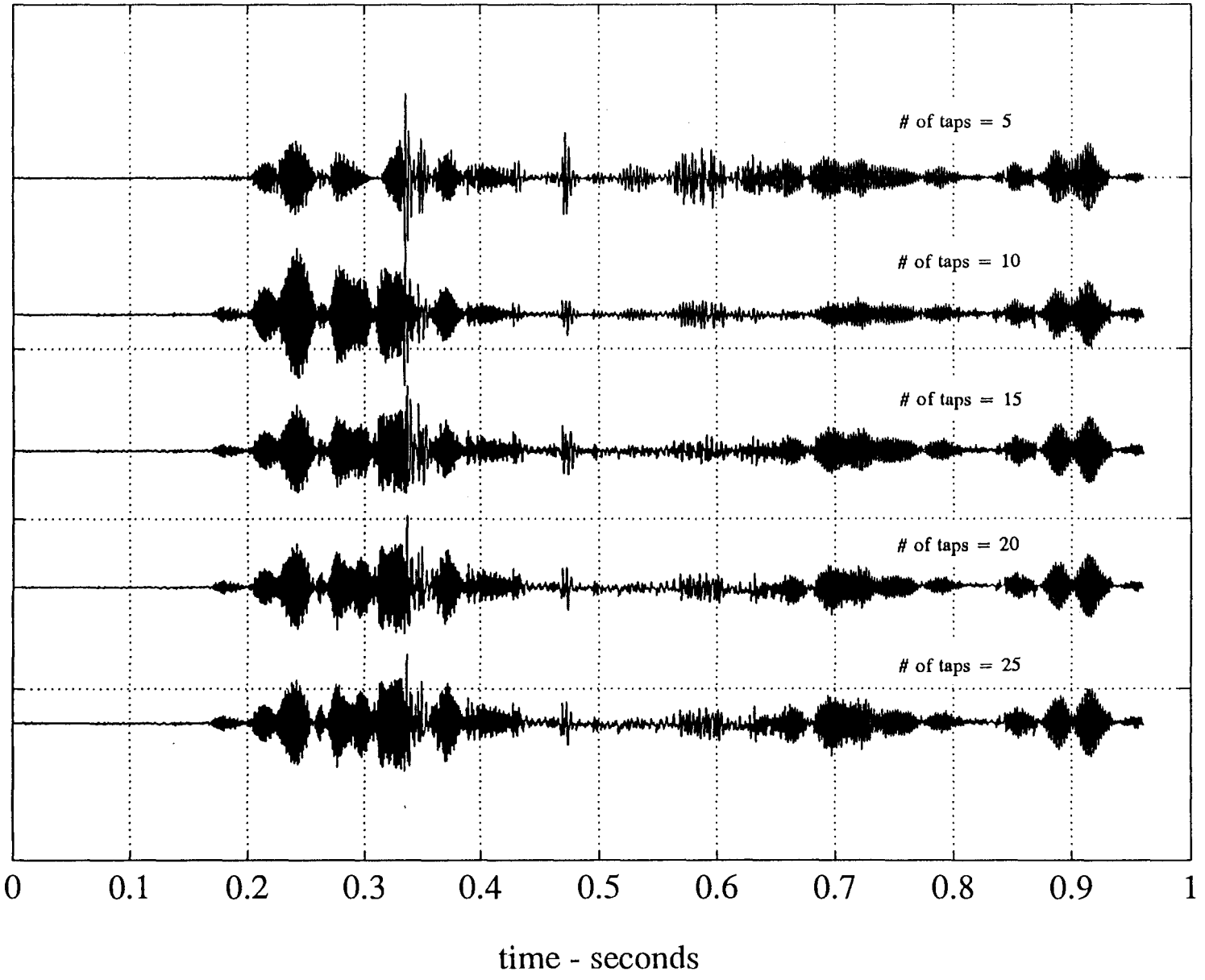


Figure 10.15 *Optimal Parameter Selection for ALE (RLS)*

10.15a Effect of decorrelation
10.15b Effect of the number of tap weights

solid line - non-stationary output

broken line - stationary output

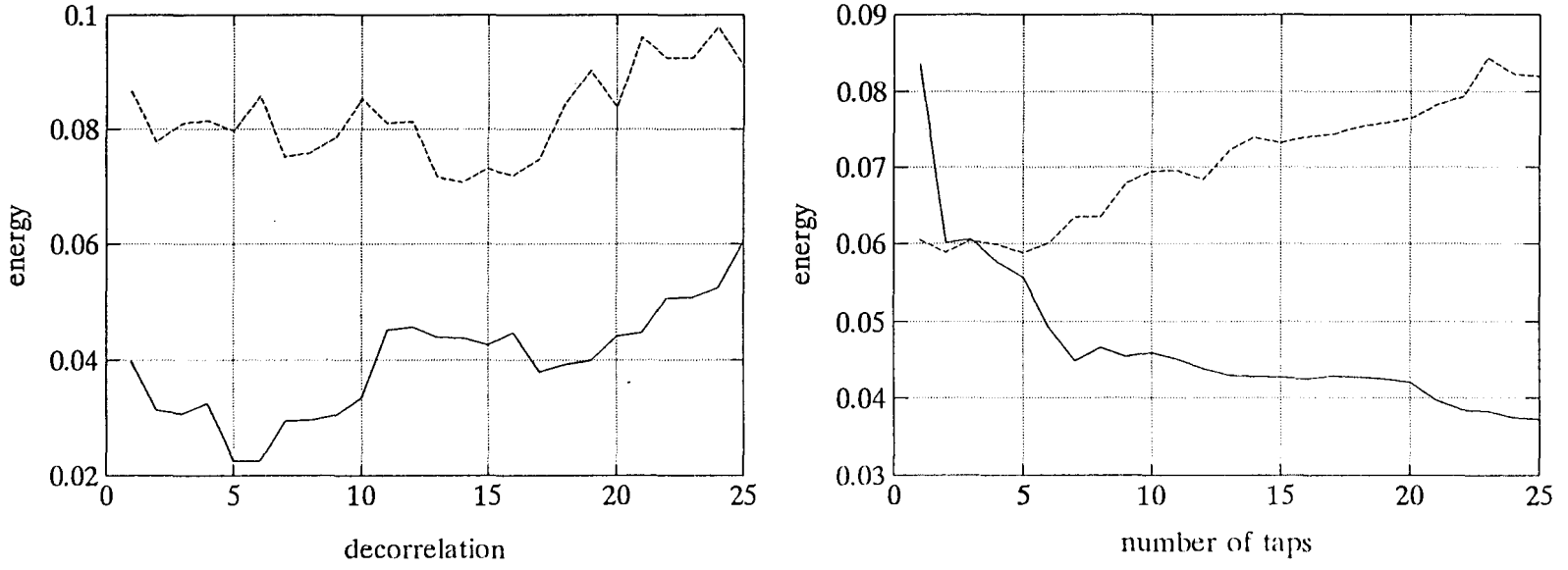


Figure 10.16a
original signal

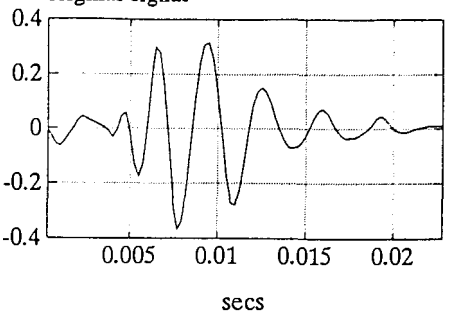


Figure 10.16b
dcp = 8 (RLS)

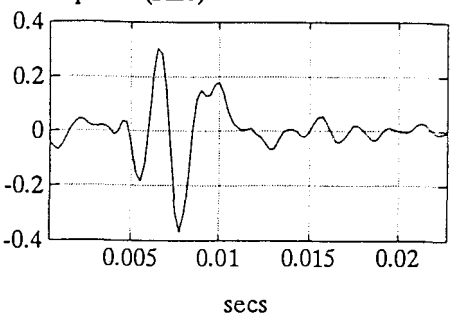


Figure 10.16c
dcp = 8, mu =.375

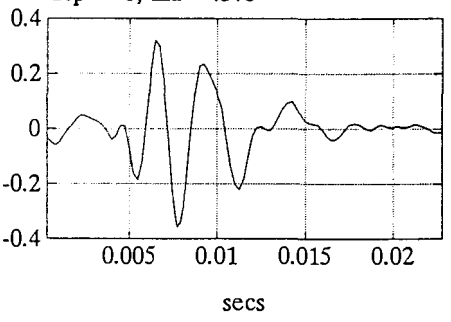


Figure 10.16d
dcp = 8, mu =.65

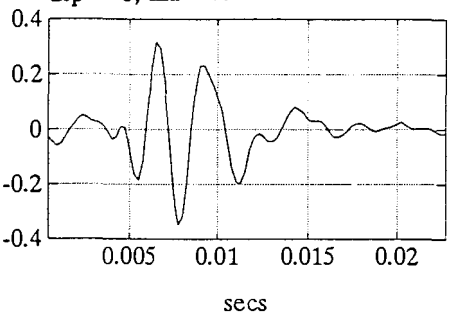


Figure 10.16e
dcp = 15 (RLS)

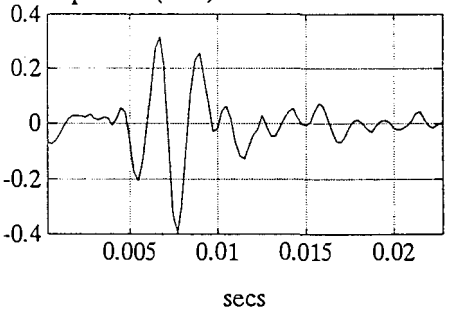


Figure 10.16f
dcp = 15, mu =.65

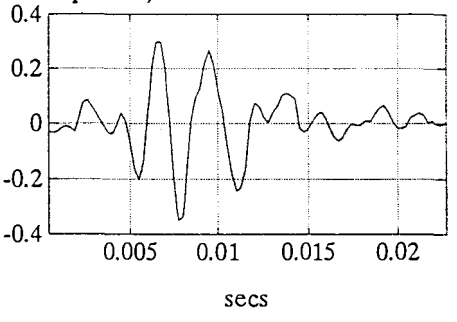


Figure 10.16g
dcp = 30, (RLS)

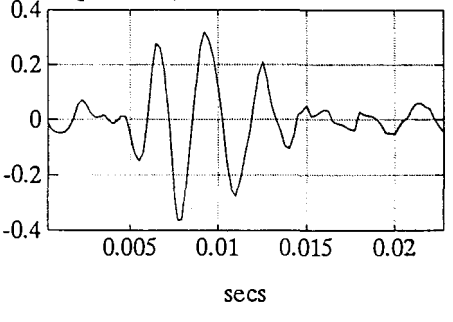


Figure 10.16h
dcp = 30, mu =.65

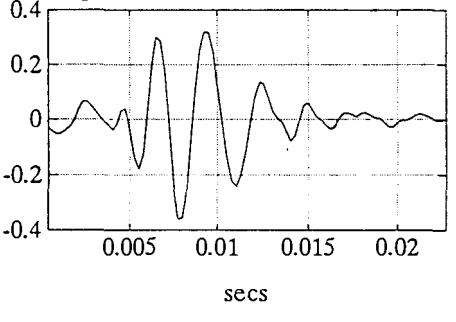


Figure 10.17a
original signal

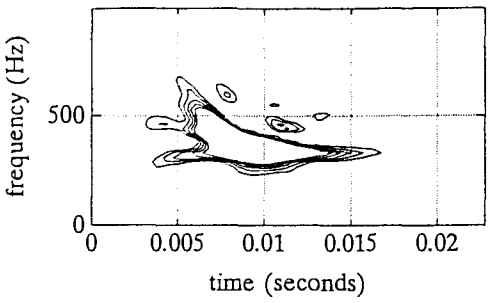


Figure 10.17b
dcp = 8 (RLS)

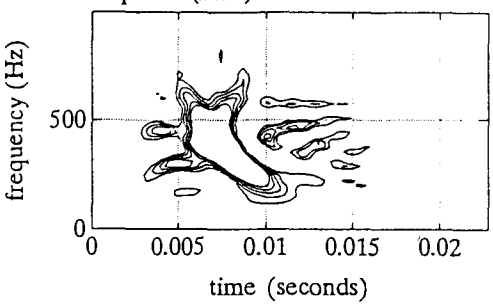


Figure 10.17c
dcp = 8, mu =.375

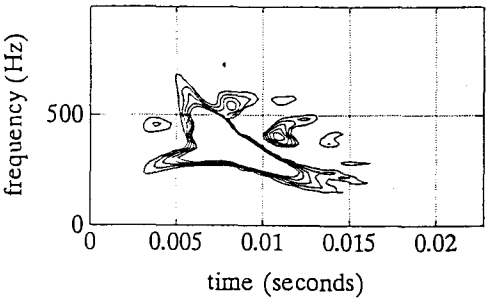


Figure 10.17d
dcp = 8, mu =.65

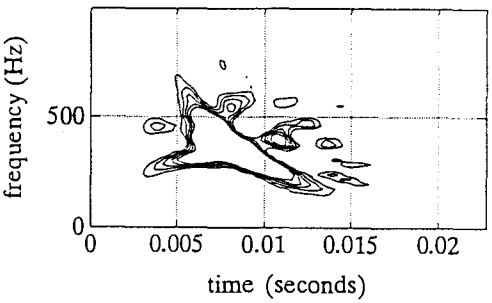


Figure 10.17e
dcp = 15 (RLS)

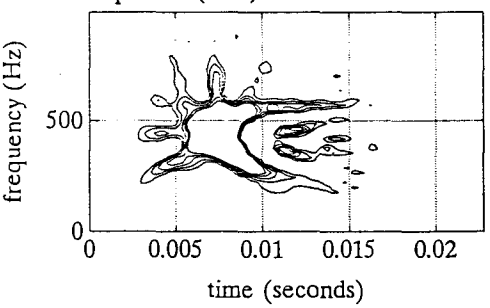


Figure 10.17f
dcp = 15, mu =.65

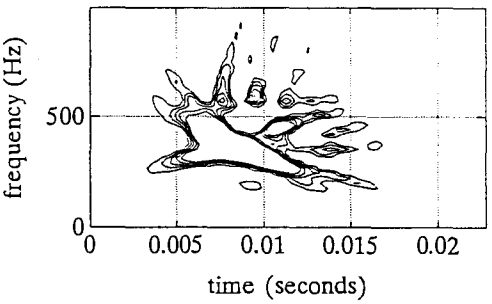


Figure 10.17g
dcp = 30, (RLS)

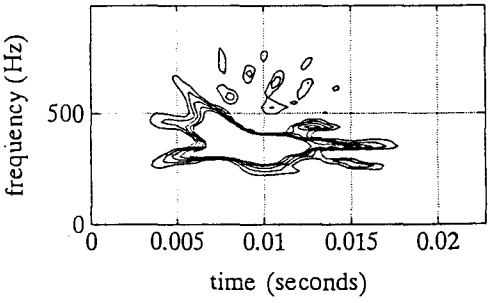


Figure 10.17h
dcp = 30, mu =.65

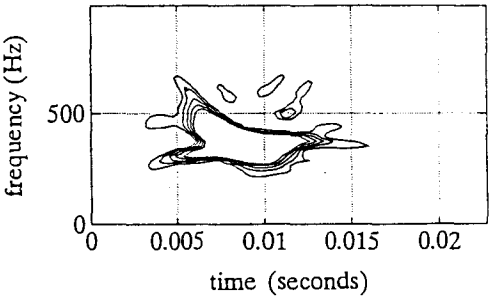


Figure 10.18 *Comparison of LMS and RLS Performance at a Decorrelation of 30*

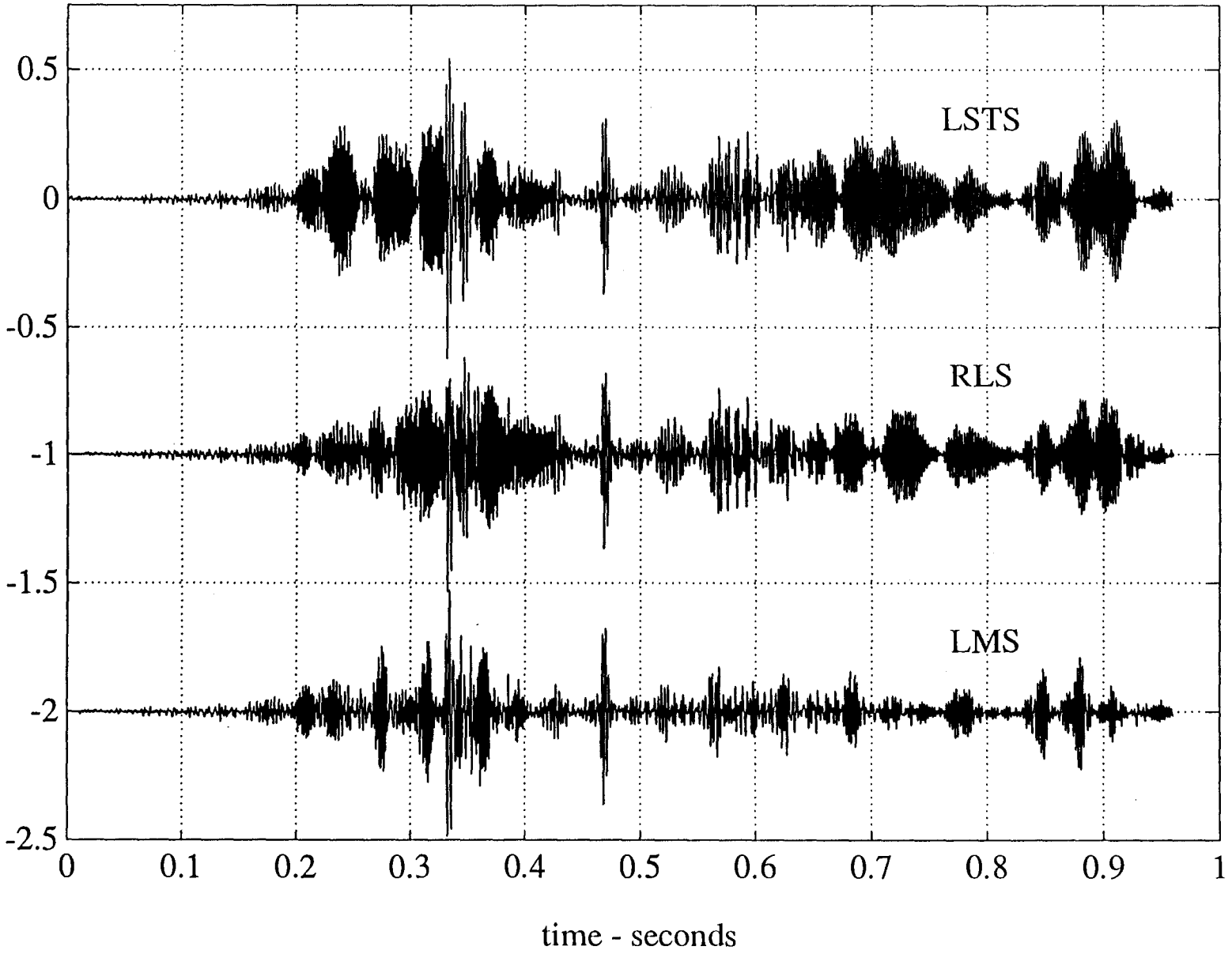
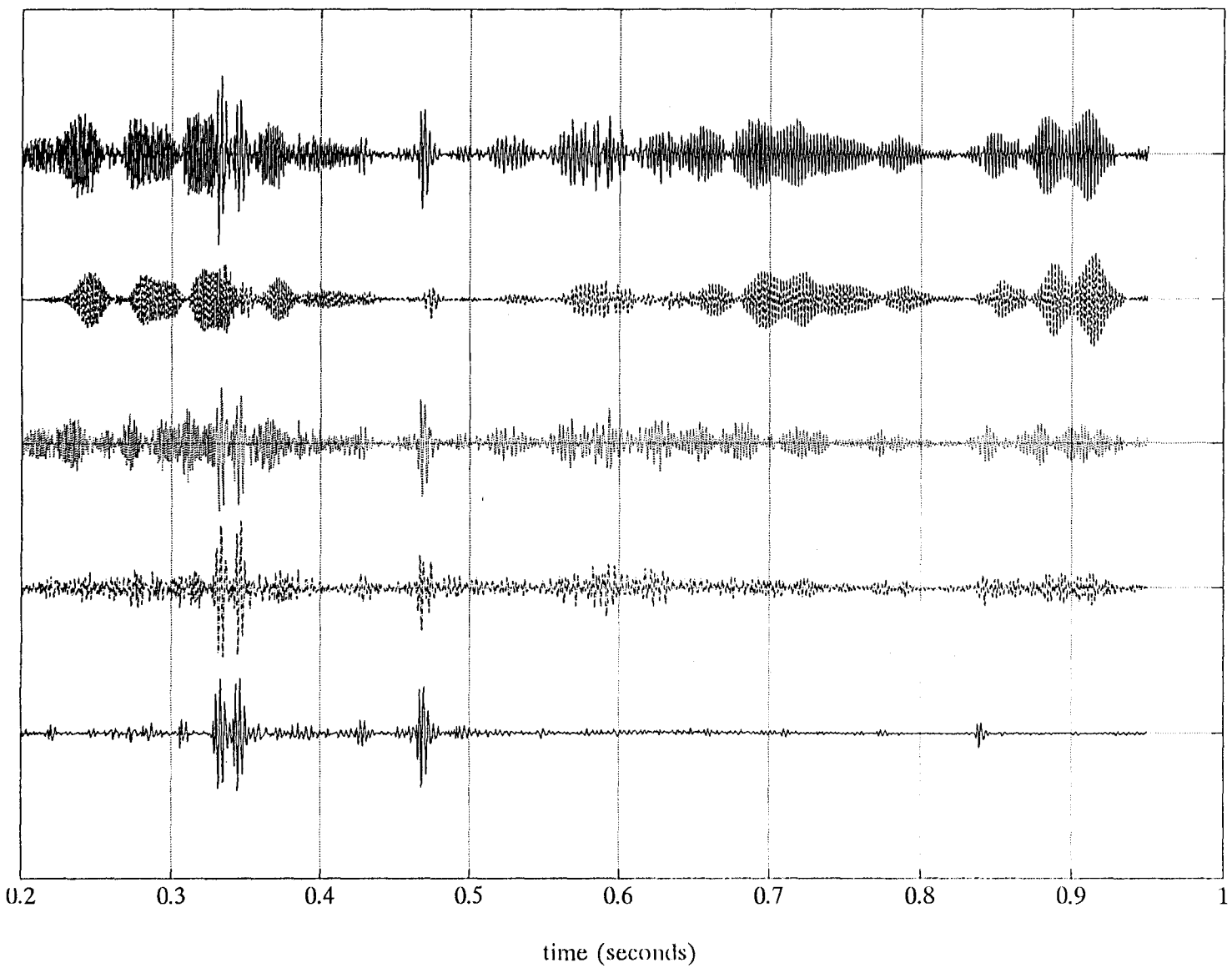


Figure 10.19 *Steps in Crackle Isolation using the Self-Tuning Block Filter*



11.0 Display

Previously demonstrated in figure 7.2, was the short time Fourier transform's limitation in displaying transient signals, namely an inability for good resolution simultaneously in time and frequency. In this section, GTFR methods are utilized to better display the time-frequency information contained in the crackles through an increase in resolution. Also, the criteria for the choice of the most appropriate kernel is developed.

11.1 Mono-component Signals and Inner Interference Terms

The first case to be considered is that of the single isolated crackle. The time series representation of this signal is shown in figure 11.1a. Its corresponding Wigner Ville distribution is shown in figure 11.1b. Since this is a mono-component signal, by previous argument, interference terms should not exist (sec 7.2). However, they are still present and can be seen as the extra structures surrounding the main term that describes the crackle.

Application of the ZAM and Choi-Williams (CW) kernels to the signal results in significantly smoothed time frequency distributions. The CW kernel used in this example attenuated some of the ambiguity function's signal terms (figure 11.2b) resulting in a slightly distorted and enlarged TFR (fig. 11.3b). The ZAM kernel provided a better 'fit'

to the signal terms and a very crisp time-frequency representation (figs 11.2a & 11.3a). The differences between the WVD and kernelled WVD indicates that mono-component signals are also capable of generating cross terms.

These cross terms are referred to as 'inner' interference terms since their genesis stems from the interference of signal components 'inside' a mono-component signal (Hla92). These inner interference terms follow the same rules of interference geometry that apply to the 'outer' interference terms which are due to different signal components interacting with each other (section 7.2.1). Although the terms in the ambiguity plane that are responsible for the signal terms in the TFR have previously been shown to cluster around the origin, it appears that there are in fact sub-classes of terms in the auto-term group. In this group, there exist terms that are strictly responsible for the structure describing the signal. These occur closer to the origin than those that describe the inner interference terms.

To further reveal the nature of the inner ITs, a mono-component signal of a chirp that first increased in frequency then decreased in frequency was constructed (figure 11.4a). When viewed in the time-frequency domain, this signal has the appearance of the letter C rotated so that it stands on its open face. The 'inner' interference terms appear as parallel 'C's of decreasing size inside the signal term. This phenomena occurs in the TFR of the single crackle and may be seen in figure 11.1b at the top of the signal term. The use of the kernelling process is of benefit to mono-component signals since it smooths the inner interference terms.

11.2 Display of Multiple Crackles

Utilizing the kernel approach is an absolute must for the display of multi-component signals. Figures 11.5a,b show the time series representation and the WVD for a signal consisting of two closely spaced crackles. The WVD exhibits a great number of interference terms. The CW kernel is applied to the data as shown in figure 11.6a. Whatever the exact scaling parameter for the kernel is, it can be seen the retention of some of the interference term information occurs. This will always be the case for displaying a signal containing multiple crackles since their similarity will produce interference terms that lie on the tau=0 axis in the ambiguity plane. A distorted time-frequency representation is the result (fig. 11.6b). This problem is overcome with the use of the ZAM kernel as shown in figure 11.7a,b. Here, the kernel surgically cuts around the signal terms located at the origin resulting in a clean TFR with little distortion (fig. 11.7b). Clearly, the ZAM kernel is better for displaying the crackles.

Figure 11.1a *Single Crackle*

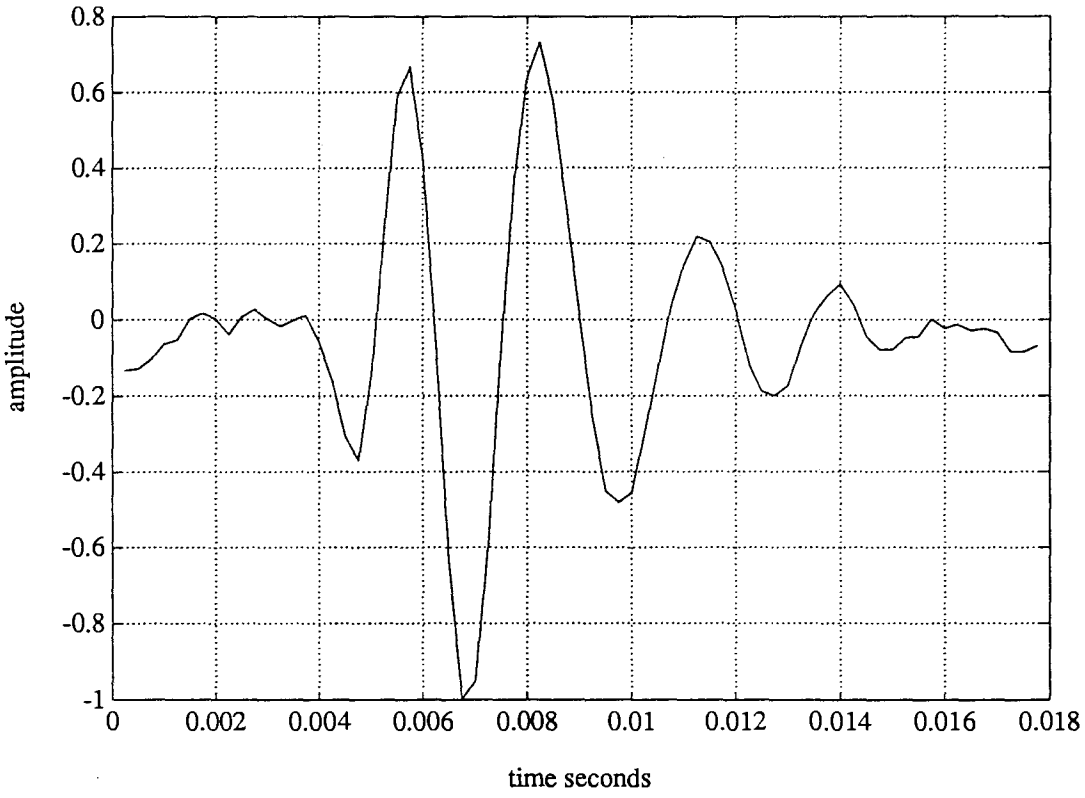


Figure 11.1b *Wigner Distribution of 'Single Crackle'*

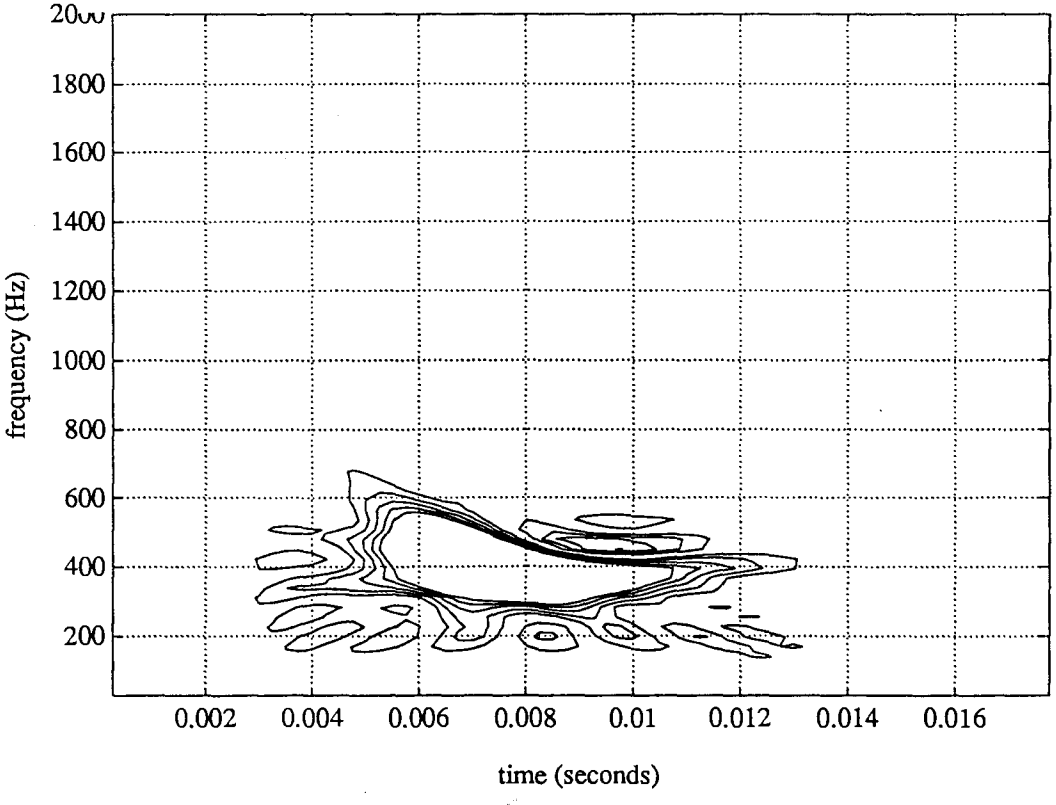


Figure 11.2a *Ambiguity Function of 'Single Crackle' with ZAM Kernel Superimposed*

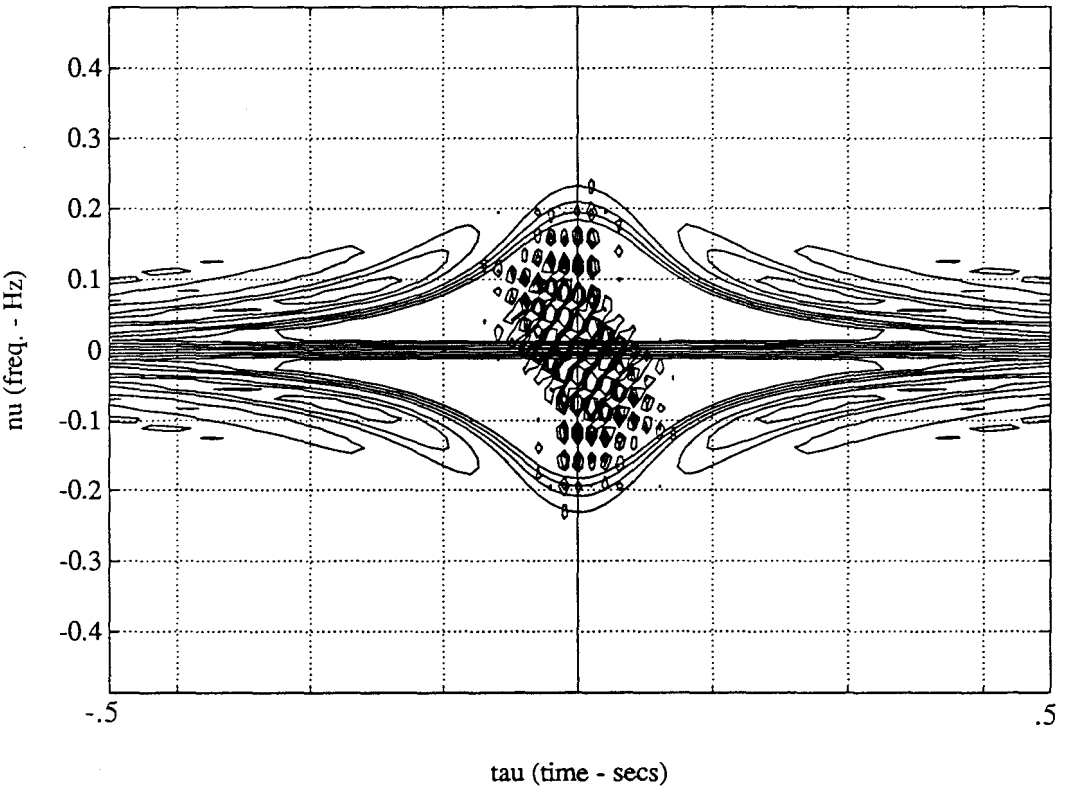


Figure 11.2b *Ambiguity Function of 'Single Crackle' with Choi- Williams Kernel Superimposed*

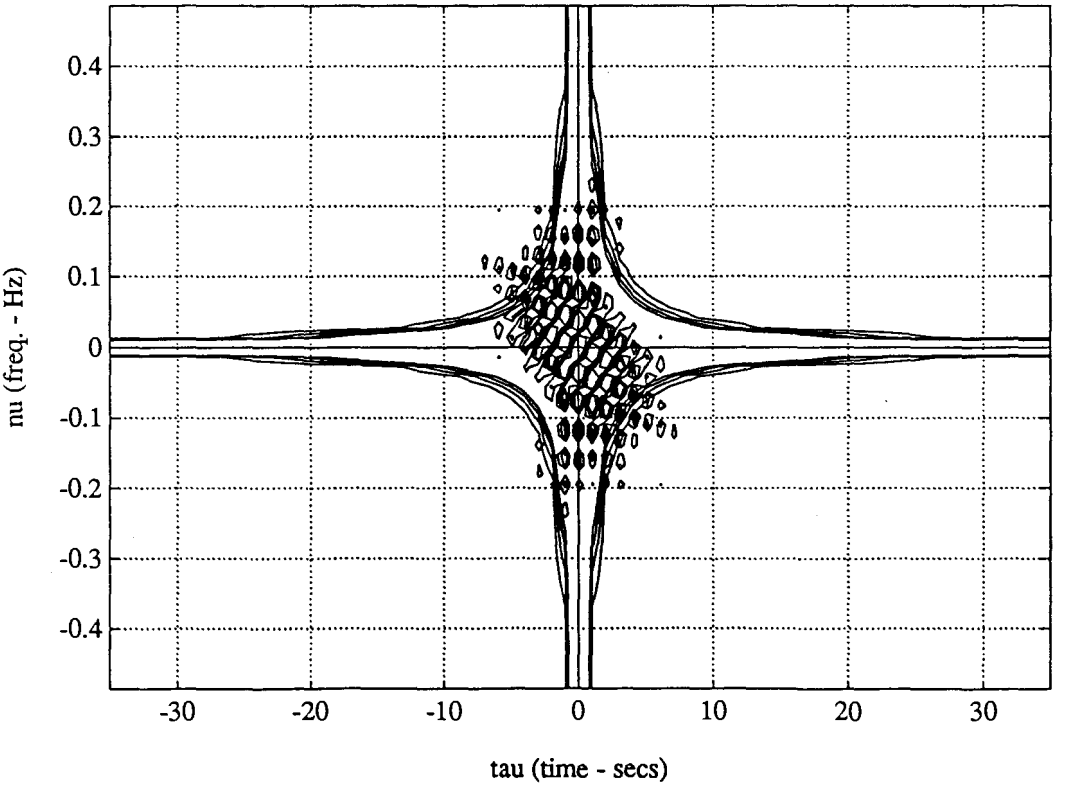


Figure 11.3a *TFR Resulting from Figure 2a*

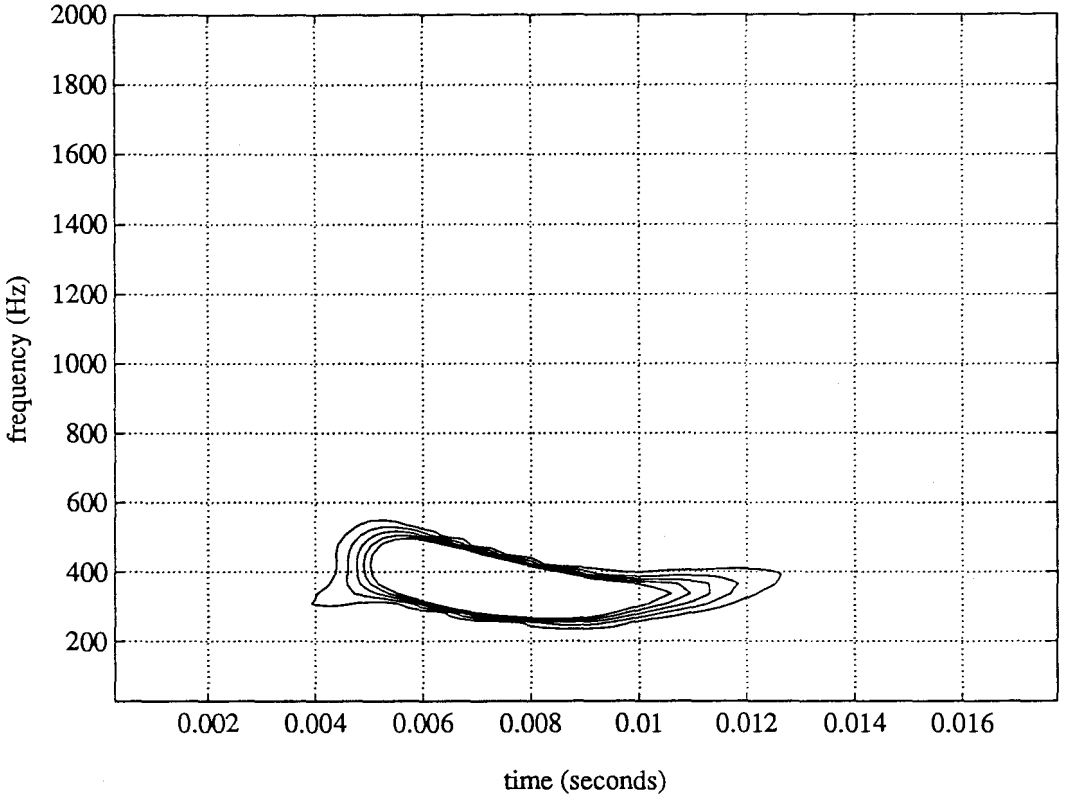


Figure 11.3b *TFR Resulting from Figure 2a*

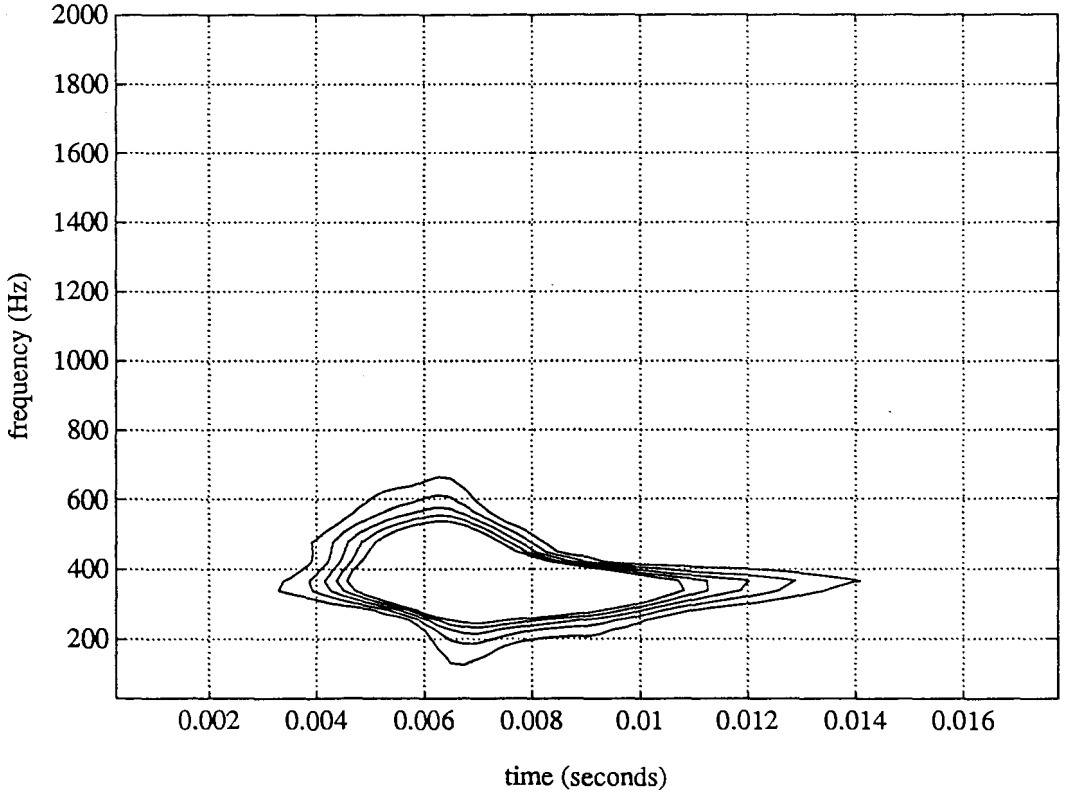


Figure 11.4a Time Series of a Chirp

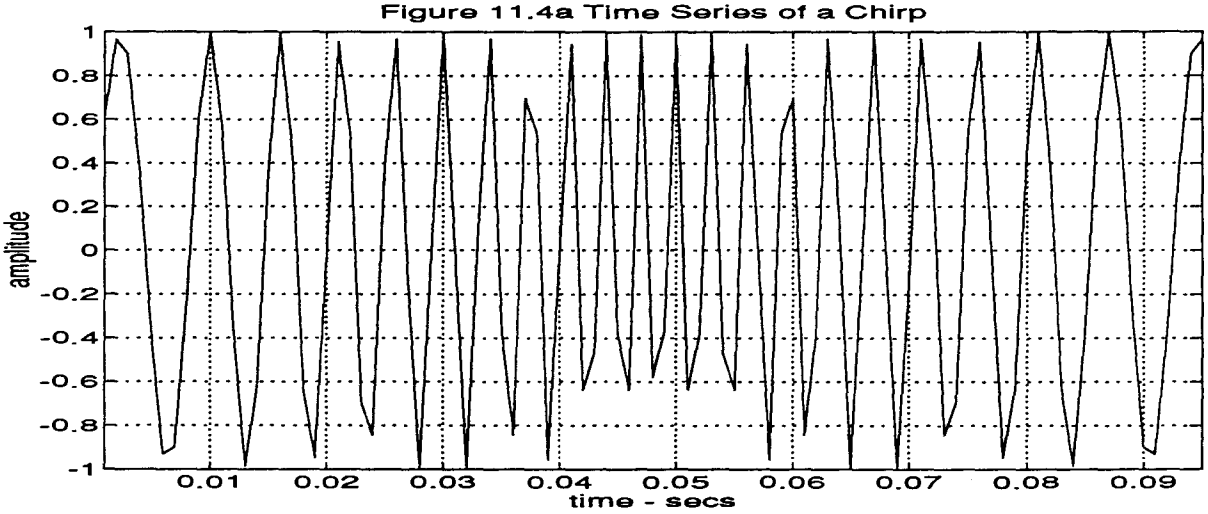


Figure 11.4b Ambiguity Function of the Chirp Showing the Inner ITs

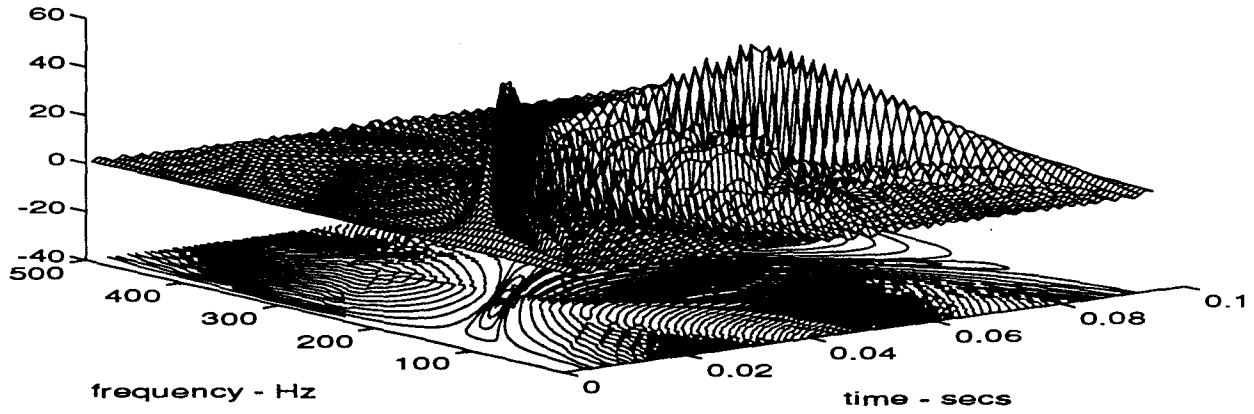


Figure 11.5a *Double Crackle*

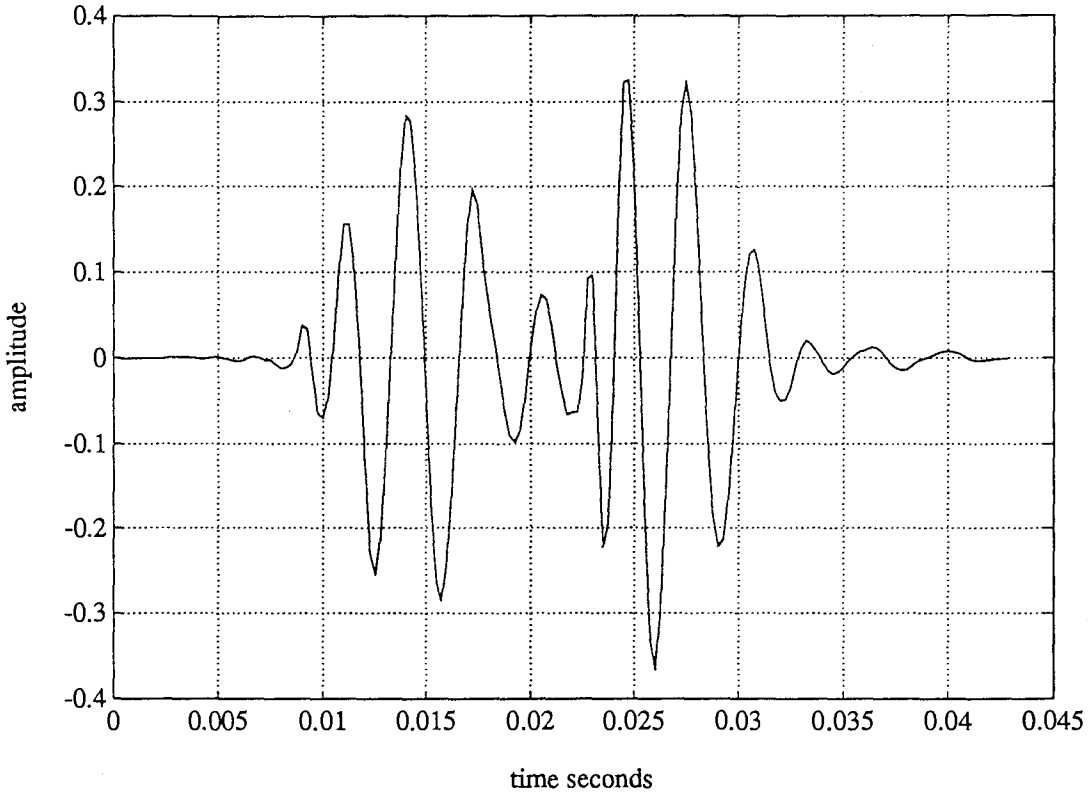


Figure 11.5b *Wigner Distribution of 'Double Crackle'*

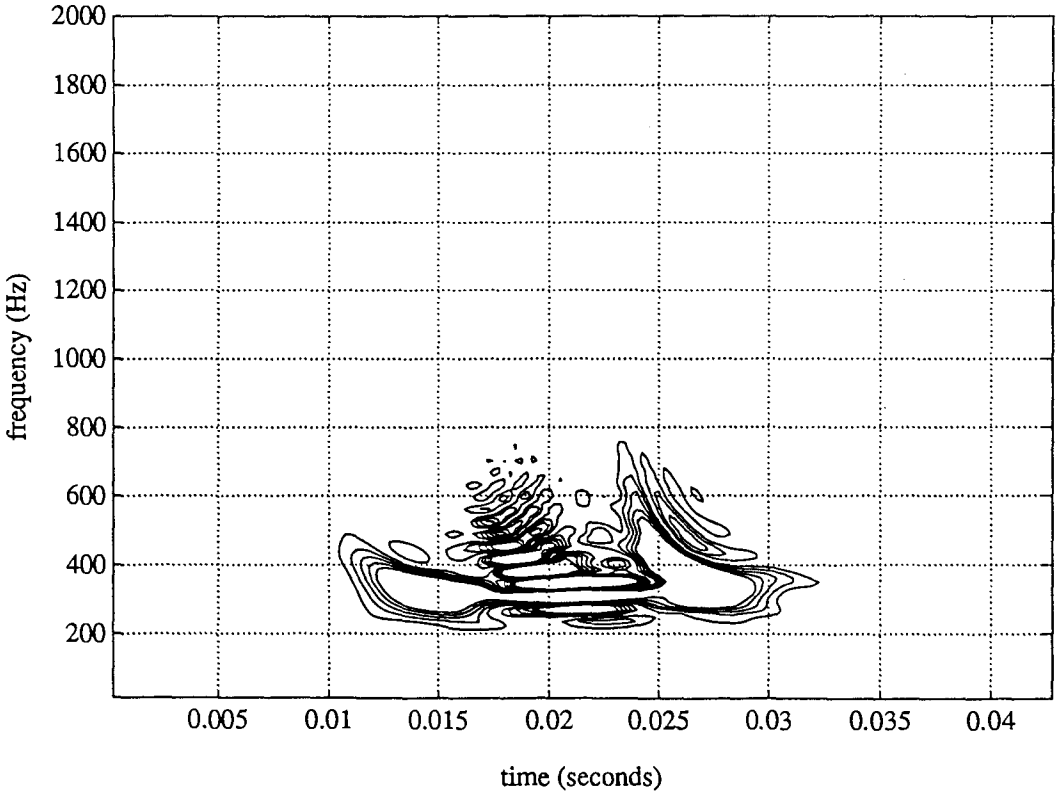


Figure 11.6a *Ambiguity Function of 'Double Crackle' with Choi- Williams Kernel Superimposed*

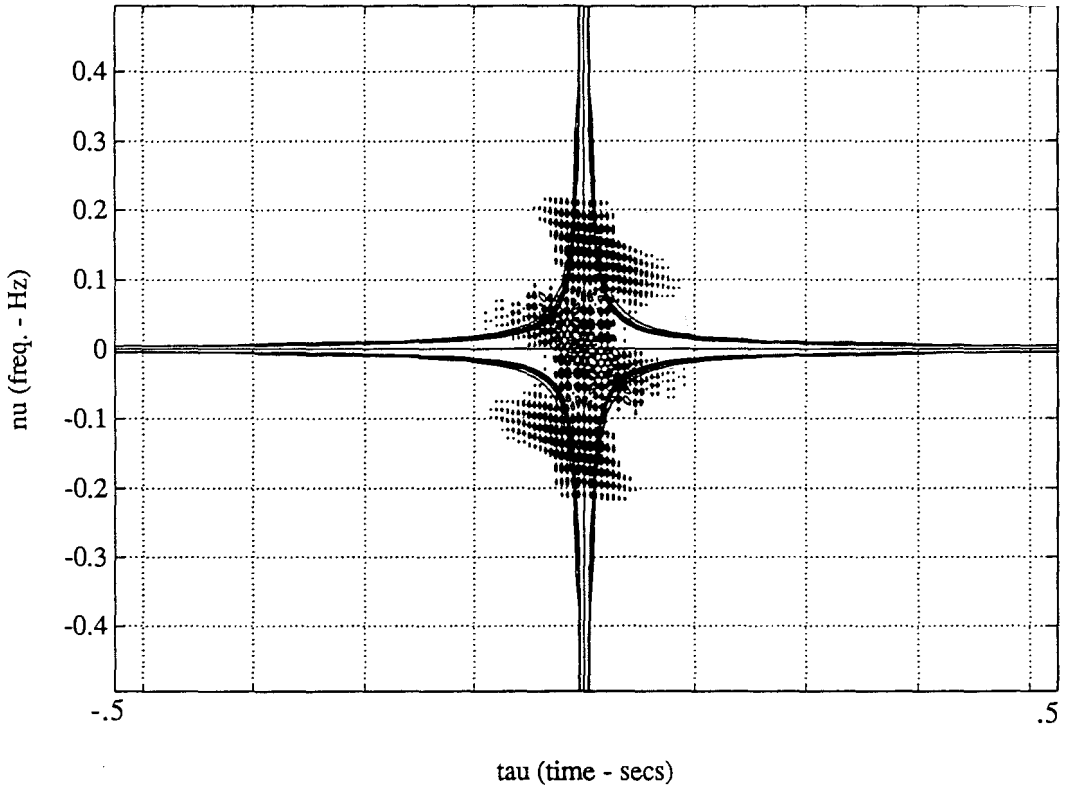


Figure 11.6b *TFR Resulting from Figure 6a*

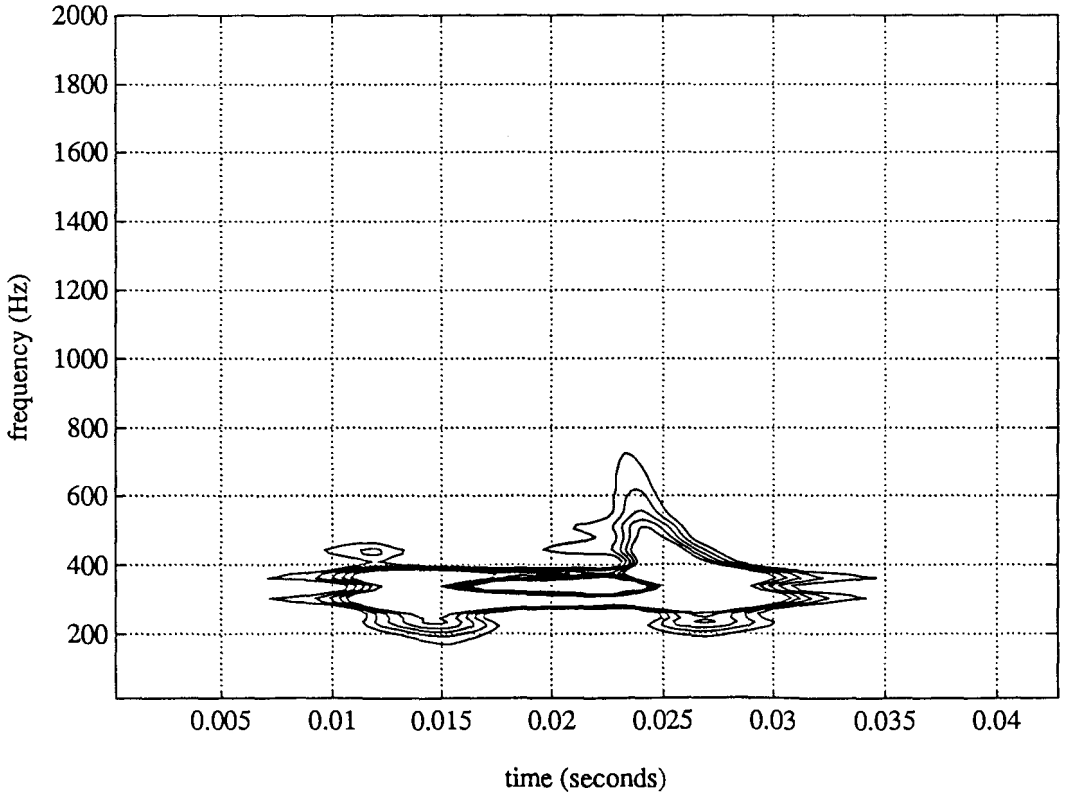


Figure 11.7a *Ambiguity Function of 'Double Crackle' with ZAM Kernel Superimposed*

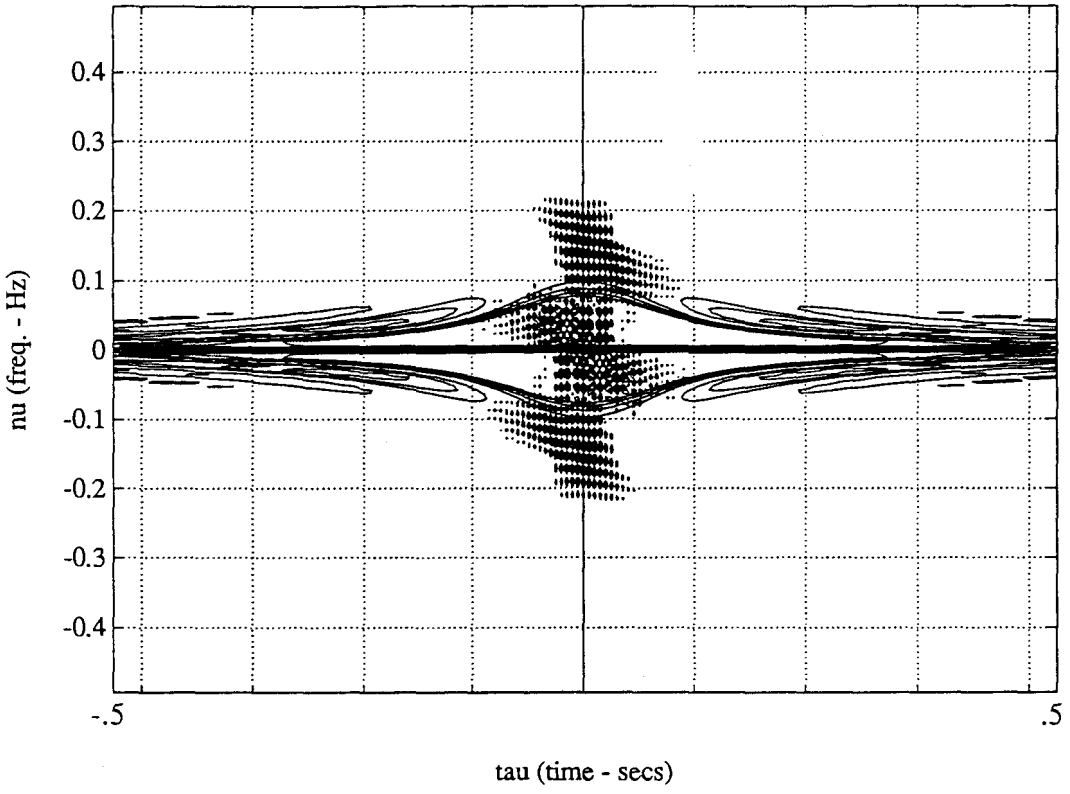
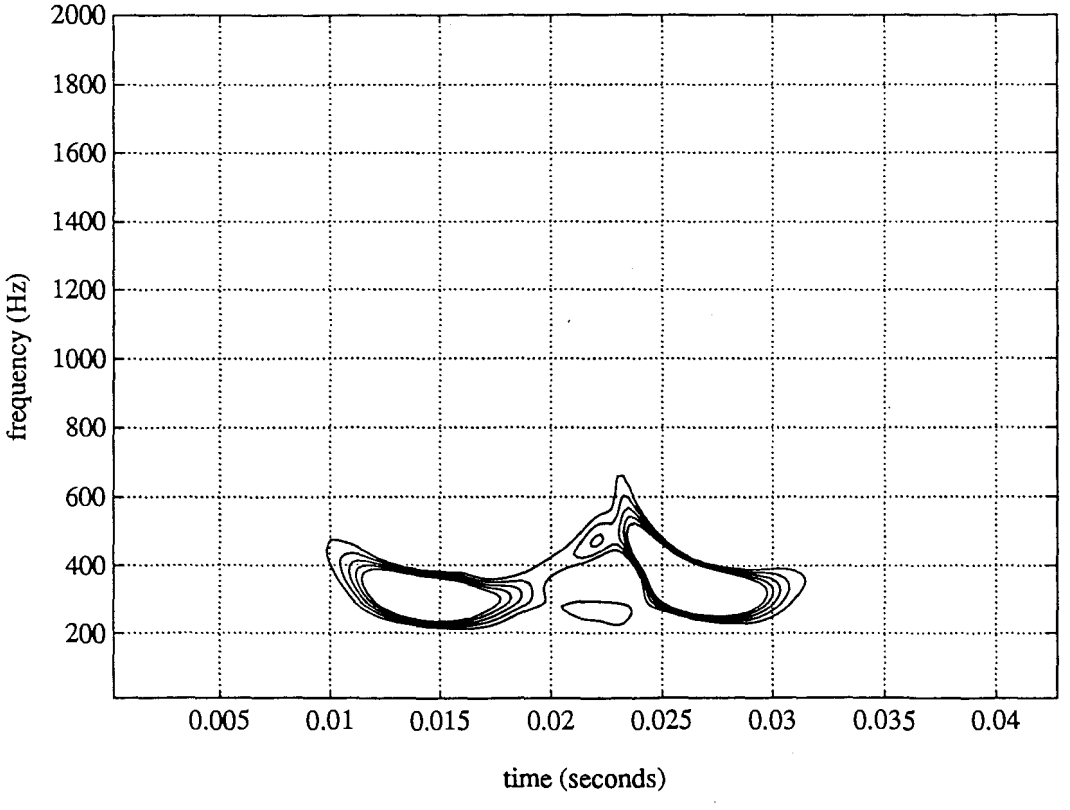


Figure 11.7b *TFR Resulting from Figure 7a*



Section 4 - Conclusion

12.0 Conclusion

13.0 Future Direction

12.0 Conclusion

12.1 Separation of Lung Sounds

Prediction Error Filtering

Prediction error filtering did not prove to be useful for separation. In the cases of fixed weight and RLS adaptive prediction error filter, the predictor co-efficients are derived from the signal's statistics. The adaptive LMS routine was better able to distinguish between crackles and wheezes due to its greater freedom to track statistical variations. Although the wheeze component is stationary it is not stationary 'enough' to provide accurate predictions of the wheeze and thereby allow isolation of the crackles only as prediction error. However, neither one of the three methods worked with great accuracy. Although the literature shows success in using PE filtering for crackle extraction (Ara91, Ono89), these applications have been used in signal conditions where only crackles and noise exist. In this environment, the PE filter needs only to discriminate between noise and signal rather than to discriminate between signals.

Adaptive Line Enhancement

Of the methods for separating the wheezes and crackles, adaptive line enhancement proved the most successful. As mentioned, some of the problems that were encountered are attributable to the data used. Testing with data taken in a controlled

manner where flow rate is monitored would probably provide better results.

The LMS routine performed best in driving the ALE. The reason for this is same as the explanation for the ALE PE filter out performing the RLS version. Stated again, the LMS is better able to track statistically noisy data since memory of previous statistics of the signal is adapted out quickly.

The degree of signal distortion introduced by the ALE in crackles is a concern. It has been determined that for a crackle free of distortion to exist after the ALE process, a large decorrelation parameter is required. Unfortunately, this permits wheeze artifact to find its way in to the non-stationary output. There are most likely only four factors of importance in examining a crackle. These are, timing, pitch, duration and energy content. Timing and pitch may be determined from a filtered crackle before distortion occurs. Analysis of the early part of the crackle could possibly give enough of an indication of duration and energy. If this is the case, distortion of the later part of the crackle is not an issue. Since very little is known about the physical conditions that produce or control a crackle and its properties, the problems introduced by crackle distortion is uncertain.

Self Tuning Block Filter

The self-tuning block filter performed well in reducing the amount of wheeze artifact in the non-stationary output of the adaptive line enhancer. Although some wheeze artifact exists widening of the filter's bandwidth would help eliminate this. The one problem that arises with this approach is the non-linear phase characteristics introduced

by the use of local filters with differing characteristics, which may also introduce distortion in the waveforms.

12.2 Display

Short Time Fourier Transform

While suitable for the display of the stationary wheezes, the short Time Fourier transform performs poorly when used on crackles. This is attributed to the reciprocal relationship between high frequency resolution and high temporal resolution which is an inherent property of the STFT.

Generalized Time Frequency Representations

Crackle display using generalized time frequency methods, provides detail that was previously unavailable. Earlier attempts utilized the zero crossings of a crackle to determine a crackle's frequency. This necessitates isolation first and presents the information in a manner that is somewhat alien. Using time frequency methods crackles may be more easily identified in a high resolution time frequency representation and may be replaced with accurate visual representations of a crackle's frequency content.

13.0 Future Direction

Continuation of this research can be pursued on three fronts. The first is implementing the algorithms on a portable system that is able to perform the processing in real time. With current technology, a data acquisition system can be coupled with the ALE, STBF algorithms in order to separate the lung sounds. Presently, the crackle display using the GTFR methods would need to be performed after signal collection due to the vast number of calculations required. In the quickly growing field of time-frequency analysis, research on real-time implementation of these processes is transpiring and integrated circuits to perform this have been developed. Soon the problem of the computational overhead will be solved.

The second area of research is to expand the system to include detection of the lung sounds as well as classification through pattern recognition techniques. Analysis of crackles could greatly benefit from this technique since the only currently accepted manner is manual examination. The type of analysis leads in to the third area of research.

The final area is to use the techniques as tools in research to establish links between the lung sounds and their genesis. As previously stated, auscultation is a technique that is waning due to other more rigorous and quantifiable procedures. With the tools presented here and further development, it is hoped that stronger links between the acoustic phenomena and the physical processes generating them may be forged creating a simple and reliable technique to aid in diagnosis.

References

(Ara91) Arakawa, K. et al. Non-Linear Digital Filters for Extracting Crackles from Lung Sounds. Frontiers of Medical and Biological Engineering 3:4 245-257 1991.

(Atl91) Atlas, L. et al. Signal Analysis with Cone Kernel Time-Frequency Distributions and their Applications to Speech. 375-388 Time-frequency signal analysis methods and applications. edited by Boashash, B.. Longman Cheshire, Melbourne. 1991.

(Cha83) Charbonneau, G. et al. An accurate recording system and its use in breath sounds spectral analysis. 1120-1127 1983.

(Cho89) Choi, H. and Williams, W. Improved Time-Frequency Representation of Multicomponent Signals Using Exponential Kernels. IEEE Transactions on Acoustics, Speech, and Signal Processing 37:6 862-871 1989.

(Coh85) Cohen, L. and Posch, T. Positive Time-Frequency Distribution Functions. IEEE Transactions on Acoustics, Speech, and Signal Processing 33:1 31-38 1985.

(Coh89) Cohen, Leon. Time-frequency distributions - A review. Proceedings of the IEEE 77:7 941-981 1989.

(Coh91) Cohen, L.. Introduction: A Primer on Time-Frequency Analysis. 3-42 Time-frequency signal analysis methods and applications. edited by Boashash, B.. Longman Cheshire, Melbourne. 1991.

(Dud73) Duda, R and Hart. Pattern Classification and Scene Analysis. Wiley, New York. 1973.

(Epl78) Epler, G. et al. Crackles (rales) in the interstitial pulmonary diseases.

Chest 73 333-339 1978.

(Fei71) Feigen, L. Physical Characteristics of Sound and Hearing. American Journal of Cardiology 28:Aug 130-133 1971.

(Fle90) Flenley, P.. Respiratory Medicine. Baillière Tindal, London. 1990.

(For78) Forgacs, Paul. Lung Sounds. Baillière Tindal, London. 1978.

(For78b) Forgacs, P. The Functional Basis of Pulmonary Sounds. Chest 73:3 399-405 1978.

(Gav81) Gavriely, N. et al. Spectral characterisitics of normal breath sounds. 307-314 1981.

(Har91) Harris, F.. Techniques and Limitations of Spectrum Analysis with the DFT. 184-207 Time-frequency signal analysis methods and applications. edited by Boashash, B.. Longman Cheshire, Melbourne. 1991.

(Hay78) Haykin, Simon. Communication Systems. John Wiley and Sons, New York. 1978.

(Hay91) Haykin, Simon. Adaptive Filter Theory. Prentice-Hall, Toronto. 1991.

(Hla92) Hlawatsch, F. and Boudreaux-Bartels, G.F. Linear and quadratic time-frequency signal representations. IEEE Signal Processing Magazine April 21-67 1992.

(Hoe90) Hoevers, J., Loudon, R.G. Measuring Crackles. Chest 98 1240-1243 1990.

(Iye86) Iyer, V. et al. Reduction of heart sounds from lung sounds by adaptive filtering. IEEE Trans.Biomed.Eng. 33 1141-1148 1986.

(Iye89) Iyer, V. et al. Autoregressive modelling of lung sounds: Characterization

of source and transmission. IEEE Trans.Biomed.Eng. 36 1133-1137 1989.

(Iye90) Iyer, V., Ploysongsang, Y., Ramamoorthy, P. Adaptive Filtering in Biological Signal Processing. Critical Reviews in Biomedical Engineering 17:6 531-584 1990.

(Joo84) Joo, T., et al. Pole-Zero Modeling and Classification of Phonocardiograms. IEEE Transactions on Biomedical Engineering 30:Feb 110-118 1983.

(Kai91) Kaisla, T. et al. Validated method for automatic detection of lung sound crackles. Med. & Biol. Eng. & Comp. 29 517-521 1991.

(Kat91) Katila, T., Piirila, P., Kallio, K., Paajanen, E., Rosqvist, T. and Sovijarvi, A.R.A. Original waveform of lung sound crackles: A case study of the effect of high-pass filtration. J. Appl. Physiol. 71 2173-2177 1991.

(Kos90) Koster, M. et al. Continuous adventitious lung sounds. Journal of Asthma 27:4 237-249 1990.

(Lee90) Lee, A., and Chou, J. Computer Classification of the PCG Waveform. International Journal of Systems Science 21:3 593-609 1990.

(Les86) Lessard, C. and Wong, W. Correlation of Constant Flow Rate with Frequency Spectrum of Respiratory Sounds when Measured at the Trachea. IEEE Transactions on Biomedical Engineering 33:4 461-463 1986.

(Mar87) Martin, D. Respiratory Anatomy and Physiology. CV Mosby, St. Louis. 1987.

(Mun91) Munakata, M., Ukita, H., Doi, I., Ohtsuka, Y., Masaki, Y., Homma, Y. and Kawakami, Y. Spectral and waveform characteristics of fine and coarse crackles. Thorax 46 651-657 1991.

(Mus90) Mussel, M. et al. Distinguishing normal and abnormal tracheal breathing sounds by principal component analysis. Jap. J. of Physiol. 40 713-721 1990.

(Nan84) Nandagopal, D. et al. Spectral Analysis of Second Heart Sound in Normal Children by Selective Linear Prediction Coding. Medical & Biological Engineering & Computing 22 229-239 1984.

(Nat74) Nath, A., Capel, L. Inspiratory crackles-early and late. Thorax 29 223-227 1974.

(Nat81) Nath, A.. Lung Sounds. 9-32 Clinical Investigation of Respiratory Disease. edited by Clark, T.. Chapman and Hall, London. 1981.

(Ogi90) Ogilvie, C.. Symptoms and Signs in Respiratory Disease. 207-221 Respiratory Medicine. edited by Brewis, R. et al. Bailliere Tindall, London. 1990.

(Ols78) Olsen, D. and Mitchell, R.. History and Physical Examination in Patients with Pulmonary Problems. 9-16 Synopsis of Clinical Pulmonary Disease. edited by Mitchell, R.. CV Mosby, St. Louis. 1978.

(Ono89) Ono, M., Arakawa, K., Mori, M., Sugimoto, T. and Harashima, H. Separation of fine crackles from vesicular sounds by a nonlinear digital filter. IEEE Trans. Biomed. Eng. 36 286-291 1989.

(Opp75) Oppenheim, A.. Digital Signal Processing. Prentice Hall, Englewood Cliffs, New Jersey. 1975.

(Pas89) Pasterkamp, H., Carson, C., Daien, D. and Oh, Y. Digital respirosonography. New images of lung sounds. Chest 96 1405-1412 1989.

(Plo89) Ploysonsang, Y., et al. Characteristics of normal lung sounds after adaptive filtering. Am. Rev. Respir. Dis. 139 951-956 1989.

(Ran88) Rangayyan, R., Lehner, R. Phonocardiogram Signal Analysis: A Review. CRC Critical Reviews in Biomedical Engineering 15:3 211-235 1988.

(Ray89) Raymond, L.H. et al. Validation of an automatic crackle (rale) counter. Am. Rev. Respir. Dis. 140 1017-1020 1989.

(Rio90) Rioul, O. and Vetterli, M. Wavelets and signal processing. IEEE Signal Processing Magazine October 14-38 1990.

(Urq81) Urquhart, R. et al. The diagnostic value of pulmonary sounds: A preliminary study by computer-aided analysis. Compu. Biol. Med. 11:3 129-139 1981.

(Wid75) Widrow, B. et al. Adaptive Noise Cancelling: Principles and Applications. Proceedings of the IEEE 63:12 1692-1718 1975.

(Wid85) Widrow B., and Stearns S.. Adaptive Signal Processing. Prentice-Hall, Englewood Cliffs, New Jersey. 1985.

(Wil88) Wilkins, R. Lung Sounds: A Practical Guide. CV Mosby, St. Louis. 1987.

(Wil91) Williams, W. and Jeong, J.. Reduced Interference Time-Frequency Distributions. 75-97 Time-frequency signal analysis methods and applications. edited by Boashash, B.. Longman Cheshire, Melbourne. 1991.

(Zha91) Zhao, Y. et al. The use of cone-shaped kernels for generalized time-frequency representations of non-stationary signals. IEEE Transactions on Acoustics, Speech, and Signal Processing 38:7 1084-1091 1991.