Higher-Order Statistics (HOS)-Based Deconvolution for Ultrasonic Nondestructive Evaluation (NDE) of Materials

by

Lahouari Ghouti

A Thesis Presented to the

FACULTY OF THE COLLEGE OF GRADUATE STUDIES

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS

DHAHRAN, SAUDI ARABIA

In Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

In

ELECTRICAL ENGINEERING

June, 1997

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.
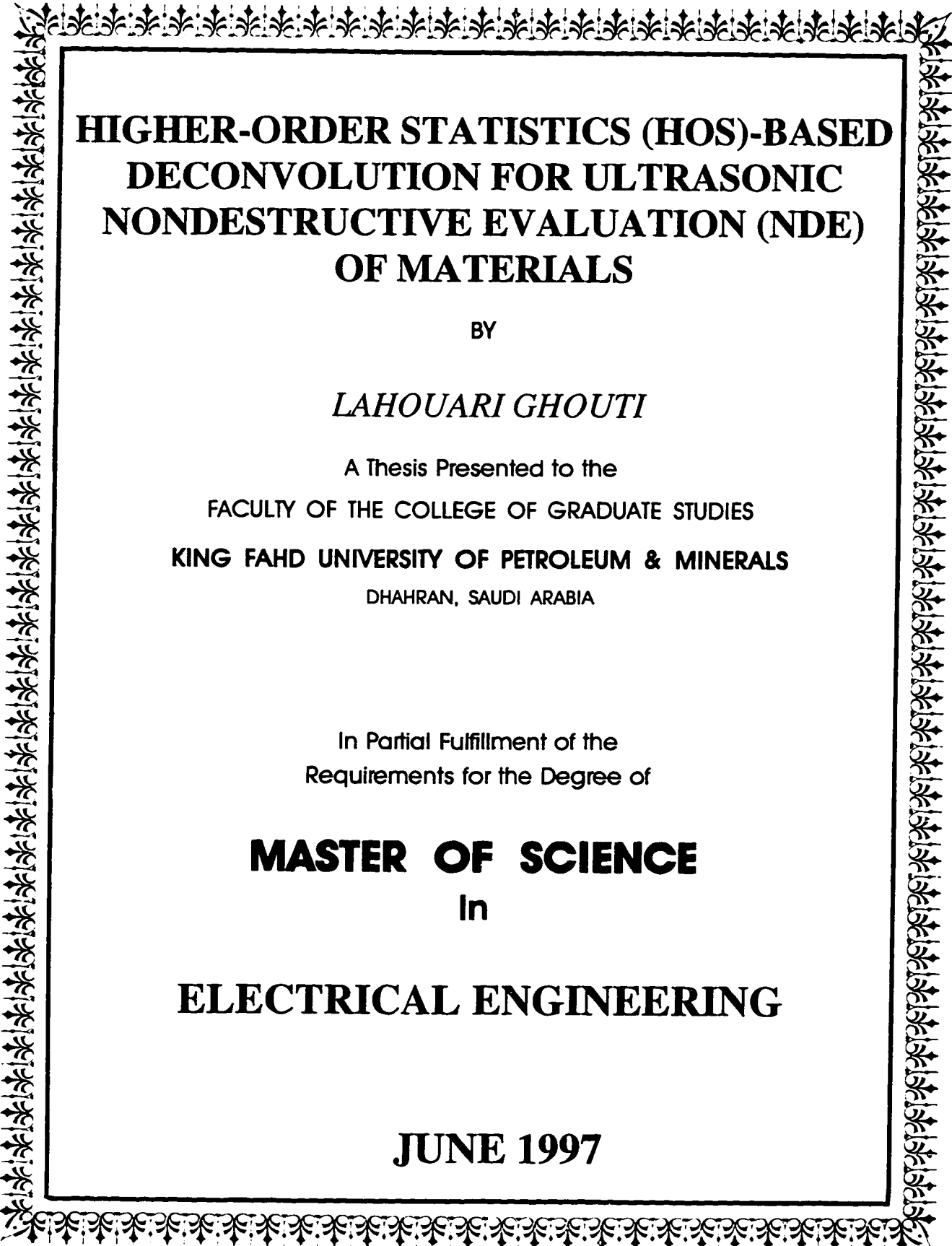
In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

HIGHER-ORDER STATISTICS (HOS)-BASED DECONVOLUTION FOR ULTRASONIC NONDESTRUCTIVE EVALUATION (NDE) OF MATERIALS

BY

LAHOUARI GHOUTI

A Thesis Presented to the

FACULTY OF THE COLLEGE OF GRADUATE STUDIES

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS

DHAHRAN, SAUDI ARABIA

In Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

In

ELECTRICAL ENGINEERING

JUNE 1997

UMI Number: 1386577

UMI
300 North Zeeb Road
Ann Arbor, MI 48103

KING FAHD UNIVERSITY OF PETROLEUM AND MINERALS

DHAHRAN, SAUDI ARABIA

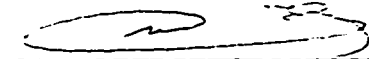## COLLEGE OF GRADUATE STUDIES

This thesis. written by

LAHOUARI GHOUTI

under the direction of his Thesis Advisor and approved by his Thesis Committee.

has been presented to and accepted by the Dean of the College of Graduate Studies.

in partial fulfillment of the requirements for the degree of

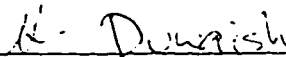## MASTER OF SCIENCE IN ELECTRICAL ENGINEERING

Thesis Committee

_____

Dr. Ahmed Yamani (Chairman)

_____
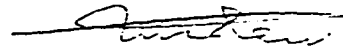
Dr. Maamar Bettayeb (Co – Chairman)

_____

Dr. Adil Belghonaim (Member)

_____

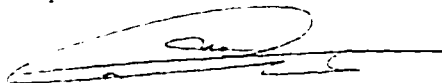Dr. Hussain N. Al – Duwaish (Member)

_____

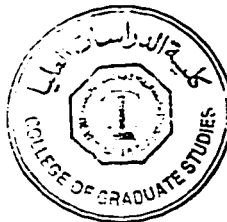Dr. Saud A. Al – Semari (Member)

_____

Department Chairman

_____

Dean. College of Graduate Studies

16 - 2 - 92
_____

Date

Dedicated to

THE MEMORY OF MY FATHER, LOVELY MOTHER,

BROTHERS, SISTERS & MY OLD FRIEND AND WIFE

whose patience, perseverance and strong prayers

led to this accomplishment.

Contents

List of Figures

List of Tables

Abstract

Name: Lahouari Ghouti

Title: Higher-Order Statistics (HOS)-Based Deconvolution For Ultrasonic NonDestructive Evaluation Of Materials

Major Field: Electrical Engineering

Date of Degree: June 1997

High resolution signal processing techniques involving higher-order statistics (HOS) and artificial neural networks (ANN) which are useful in ultrasonic nondestructive evaluation (NDE) of materials systems subject to additive white Gaussian noise (AWGN) and masking effects of measurement systems used and propagation paths, are investigated in this Thesis. The proposed techniques are: i) a batch-type deconvolution method using the complex bicepstrum algorithm, and ii) automatic ultrasonic defect classification system using a modular learning strategy. Performance evaluation of the proposed methods and comparisons with existing methods are made by means of Monte-Carlo simulations, experimental data and analysis.

The first scheme makes use of the complex cepstrum of the third-order cumulants (complex bicepstrum) of the ultrasonic signals. The second scheme based on a modular learning strategy consisting of three functional blocks, takes into account the nonstationary character of the ultrasonic NDE system and makes use of the "information preserving rule" which allows accurate and reliable classification procedure.

It is demonstrated that the proposed techniques perform very efficiently in both white or colored Gaussian and symmetrically-distributed noise classes at moderate and low signal-to-noise ratios (SNR). Comaprisons with existing methods demonstrate improved performance characterized by high resolution properties, robustness and low sensitivity to additive Gaussian noise. However, the improved performance is achieved at the expense of higher computational complexity and data requirements.

Master of Science Degree

King Fahd University of Petroleum and Minerals

Dhahran, Saudi Arabia

June 1997

خلاصة الرسالة

إسـم الطـالـب : الهواري معمر غوثي .

عنوان الرسالة : إستخراج الإستجابة النبضية للخلل في الفحص الإتلافي للمواد بالأمواج الفوق صوتية بإستعمال إحصاءات عالية الرتبة .

التخصص : الهندسة الكهربائية .

تاريخ الرسالة : صفر ١٤١٨ هـ (يونيو ١٩٩٧ م) .

تناقش هذه الرسالة تقنيات معالجة الإشارة ذات الدقة العالية ، والتي تستعمل إحصاءات عالية الرتبة والشبكات العصبية الإصطناعية ، المفيدة في نظم الفحص الإتلافي بالأمواج الفوق صوتية ، المتعرضة للضجيج الأبيض الإضافي ذي التوزيع الطبيعي والتأثيرات السلبية لأجهزة القياس المستخدمة وطرق تنقل الموجة الفوق صوتية في المواد تحت الفحص .

١ - تقنيـة إستخراج الإستجابة النبضية للخلل المتواجد في المواد تحت الفحص بإستعمال الطيف اللوغاريتمي لإحصاءات الرتبة الثالثة للإشارات الفوق صوتية .

٢ - نظام آلي لتصنيف الخلل بإستعمال إستراتيجية تعلم وحداوي للشبكات العصبية الإصطناعية . يتكون النظام من ثلاث وحدات عملية تأخذ بعين الإعتبار الطبيعية المتغيرة للإشارات الفوق صوتية مستندة في ذلك إلى ، قاعدة حفظ المعلومات ، التي تسمح بعملية تصنيف مضمونة وعالية الدقة .

قمنا بتحليل آداء التقنيات المقترحة ومقارنتها مع الحلول المستعملة عن طريق المحاكاة بإستخدام بيانات إصطناعية وأخرى تجريبية . ولقد اثبتت هذه الرسالة تميز التقنيات المقترحة بآداء عالي مع وجود ضجيج أبيض أو متداخل ذي توزيع طبيعي أو متناظر في حالة نسبة إشارة إلى ضجيج متوسطة أو ضعيفة .

تجدر الإشارة إلى أنه يترتب على الدقة المكتسبة زيادة باهضة في العمليات الحسابية وحجم البيانات المطلوبة .

درجة الماجستير في العلوم
جامعة الملك فهد للبترول والمعادن
الظهران - المملكة العربية السعودية

Chapter 1

Introduction

1.1   Definitions

Detecting defects in materials is an unconditional stage in the quality control process of engineering systems. Such a procedure is variously known as *NonDestructive Testing* (NDT), *NonDestructive Evaluation* (NDE), *NonDestructive Inspection* (NDI), but also as *Quality Control* of materials. The American society for NDT adopted the following definition for such a procedure.

Definition 1.1.1 *[1] NonDestructive evaluation (NDE) is an examination of an object or material in a manner which will not impair its future usefulness.*

Besides the ultrasonic NDE method, other techniques are being utilized as well. These are :

1

- Radiographic Inspection

- Magnetic-Particle Inspection

- Liquid-Penetrant Inspection

Ultrasonic NDE is the main concern of this thesis, a brief definition of the technique is given below.

Definition 1.1.2 *[2] Ultrasonic nondestructive inspection is a method used for detection of hidden flaws in the material being tested. It employs an ultrasonic signal that propagates through the material, then the reflected echoes from defects and discontinuities in the material are received by an ultarasonic transducer and then recorded for further processing.*

A general ultrasonic inspection setup is shown in Fig. 1.1. Complete review and definitions of the other techniques can be found in [3].

1.2 Problem Formulation

Mathematically, the model usually assumed for the ultrasonic NDE setup consists of a system whose response $y(k)$ to an input $x(k)$ are related by :

$$y(k) = x(k) * h_d(k) + n(k) \qquad (1.1)$$
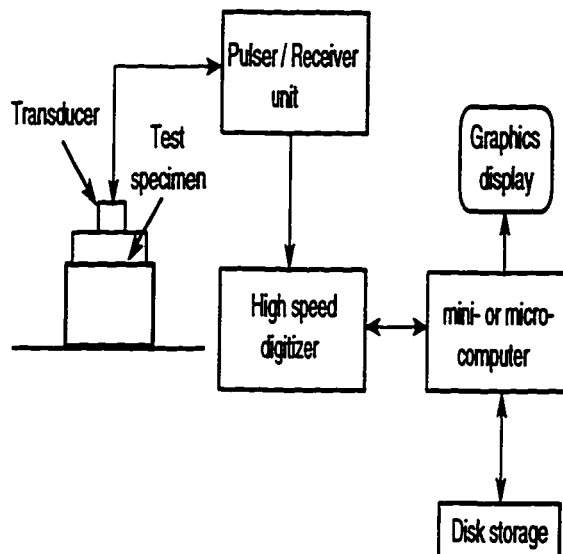
Figure 1.1: A general ultrasonic inspection system.

where $n(k)$ and $*$ represent the system and/or measurement noise and the linear convolution operator, respectively. The reflector and/or the propagation medium is characterized by its impulse response $h_d(k)$. The main assumptions leading to a mathematically tractable solution of Eq. 1.1 are summarized below :

- The ultrasonic pulse propagates through a linear medium.

- $n(k)$ and $x(k)$ are *statistically independent.*

The first assumption allows us to isolate the impulse response of the defect from the effects of the transducer and the propagation medium as much as possible. While, the second assumption is at the heart of a variety of well known identification methods [4, 5]. If one models the overall system as a cascade of many linear time-invariant (LTI) systems, such that (see Fig. 1.2b) :

- $P_1(k)$: Forward path impulse response (IR).

- $P_2(k)$: Backward path IR.

- $T_1(k)$: Forward transducer IR.

- $T_2(k)$: Backward transducer IR.

then, Eq. 1.1 becomes :

$$y(k) = u(k) * T_1(k) * P_1(k) * h_d(k) * P_2(k) * T_2(k) + n(k) \qquad (1.2)$$

where $u(k)$ is the driving impulse to the transducer which is assumed ideal. It is assumed that the IR of a flawless material approximates closely to an impulse, i.e., a Dirac impulse (see Fig. 1.2a) then, the corresponding echo from the material, which is $x(k)$, can be expressed as :

$$x(k) = u(k) * T_1(k) * P_1(k) * \delta(k) * P_2(k) * T_2(k) \qquad (1.3)$$

since $\delta(k)$ is the neutral element of convolution operation, Eq. 1.3 becomes :

$$x(k) = u(k) * T_1(k) * P_1(k) * P_2(k) * T_2(k) \qquad (1.4)$$

As suggested by Eq. 1.2, the defect signature, $h_d(k)$, is subject to :

1) Transducer effects.

2) Propagation paths effects.

3) Additive noise effects.

The problem is therefore that of *deconvolution* or identification of $h_d(k)$.

By deconvolution, it is desired to unravel the masking effects mentioned above on

the defect impulse respone. For instance, the ultrasonic signals for a particular

reflector recorded under the same conditions, but using different transducers can

be quite different. Also, the same reflector existing in different materials can be

characterized by different signatures. Deconvolution task therefore, aims to unravel

the masking effects and extract the defect signature which is an essential step for the

identification and classification of defects. A major cause that strongly influences the

role and success of ultrasonic NDE in structural engineering is the inspector using

the instrumentation [6]. A number of factors contribute to make the job tedious and

increase the likelihood of error through boredom. Complete automation of defect

classification is viewed as a potential solution for the classification problem. Taking

these problems into account, high resolution signal processing techniques have been

proposed for ultrasonic NDE processing to achieve two main objectives :

1) Deconvolution of the defect IR to unravel the masking effects of the propagation

paths and measurement system and reduce the noise effects on the defect

signature.
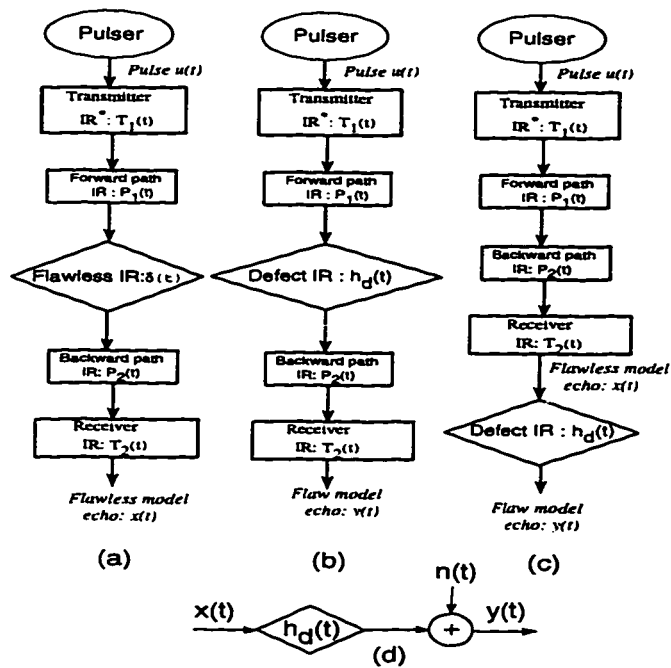
Figure 1.2: The linearity property of the model of Eq. 1.1. (a) Flawless specimen. (b) Exact model of a defective specimen. (c) rearranged LTI systems by linearity principle. (d) Equivalent model of Eq. 1.1.

2) Defect classification in an automatic fashion to icrease the decision making reliability.

1.3 Summary of the Thesis Work

Aiming to contribute with some methods for use in ultrasonic NDE processing, a deconvolution scheme based on higher-order statistics (HOS) is developed using polycepstra principles. HOS are known to be robust to additive Gaussian noise and preserve the true phase character of the underlying system. For the task of ultrasonic defect classification, a modular system is introduced.

1.4 Thesis Objectives and Organization

1.4.1 Thesis Objectives

The overall objective of this thesis is twofold: (1) apply a robust approach for ultrasonic NDE deconvolution, subject to additive white or colored Gaussian and white symmetrically-distibuted noise classes, by exploiting the properties of HOS, and (2) propose a modular learning strategy for ultrasonic defect classification. These objectives are achieved through the following tasks :

- Review the application of the existing deconvolution techniques to the UNDE deconvolution problem and identify their limitations and drawbacks.

- Investigate a non-adaptive technique for ultrasonic NDE deconvolution (subject to additive noise), by looking beyond the second-order statistics (SOS) which are sensitive to additive Gaussian noise and blind to the phase information.

- Evaluate the performance of the proposed method by using synthetic and experimental ultrasonic NDE data.

- Provide a comprehensive comparison between the proposed method and the existing ones.

- Develop an automatic modular defect classification paradigm.

In what follows, a description of the thesis organization is given.

1.4.2 Thesis Organization

Following this chapter where definitions and problem formulation are given, Chapter 2 concentrates on the review and analysis of the existing ultrasonic NDE processing techniques that have been applied for deconvolution and defect classification. Some of the problems encountered when applying these techniques on experimental ultrasonic NDE data are discussed, and possible solutions for performance enhancement are provided whenever possible. Chapter 3 deals with the mathematical preliminaries important to the development of the proposed scheme for ultrasonic NDE deconvolution. Exploiting the attractive properties of the higher-order cumulants,

important results, useful to the proposed scheme are summarized. In the first part of Chapter 4, the proposed HOS-based deconvolution method is introduced. The method estimates the defect or flaw IR by using the complex cepstrum of the third-order cumulants. As such, the proposed method is capable of reconstructing the impulse response of the defect without resorting to any prior knowledge about the noise statistics and perform in high signal-to-noise ratio (SNR) domain. In the second part of Chapter 4, details about the proposed automatic defect classification are given. The adopted network architecture is presented along with the different training paradigms used in the system. Comparison, by means of computer simulation using synthetic and experimental data, between the proposed deconvolution scheme and the existing methods is carried out in Chapter 5. Data length effects and performance at different SNR levels are also investigated. It is demonstrated in the simulation that the HOS-based scheme outperforms the existing methods when operating in severe contexts (such as short data length and low SNR levels), at the expense of higher computational complexity. The performance of the new defect classification system is demonstrated using artificial defects modeled by LTI systems. Finally, summary of the main results and conclusions are given in Chapter 6. Some problems associated with the proposed techniques are mentioned and potential solutions are suggested as future research work. Some results and properties of HOS and other topics which are important but not essential for the continuity of the text, are given in appendices.

Chapter 2

Review of Existing Ultrasonic NDE Processing Techniques

2.1 Ultrasonic NDE Deconvolution

2.1.1 Preliminaries and Definitions

High resolution methods of digital signal processing have been applied to the ultrasonic NDE deconvolution problem. An excellent review of these applications is given by Chen and Sin [7, 8]. Hayward and Lewis [9] compared the application of some non-adaptive deconvolution techniques to the same problem. Because of the inherent ill-conditionning nature of the deconvolution, the proposed methods suffer from this severe limitation, although they are known to be of high resolution. The-

oretically, an optimal solution can always be found if the noise statistics are known a priori. Neal [10] used the prior knowledge about the noise statistics to derive an optimal Wiener filter that overcomes the ill-conditionning problem. Neal and Thompson [10, 11] recognized that the proposed method is far from being realistic since the assumptions made are not easily verified experimentally. This puts a serious limitation on the applicability of Neal's optimal Wiener filter to experimental ultrasonic NDE deconvolution problems and calls into question its practical usefulness. An interesting adaptation of the available deconvolution methods commonly used in seismic deconvolution problems has been proposed by Chen and Sin [7, 8, 12]. Chen and Sin noticed the strong similarity between the two problems (seismic and ultrasonic deconvolution) then, tested most of the available techniques on the ultrasonic problem. Based on the obtained results, they retained and recommended for use those methods which gave satisfactory performance. However, it should be pointed out that although the seismic and ultrasonic deconvolution problems are in many ways similar, they are also substantially different from each other. Caution must be used when attempts are made to generalize the findings from the seismic problem to the ultrasonic one [7, 8] [1]. The deconvolution methods, in general, can be categorized in three different broad classes [4, 5, 13] :

1. *Time domain* deconvolution techniques.

---

[1]Seismic deconvolution looks for the spikes in the deconvolved IR, while ultrasonic NDE deconvolution is interested in the shape of the deconvolved IR.

2. *Frequency domain* deconvolution techniques.

3. *Cepstral or homomorphic* domain deconvolution techniques.

The first two classes can be implemented either in a *parametric* or *nonparametric* fashion [4, 5]. The cepstral method is nonparametric since no model-based formulation is involved. Ljung and Glover [14] compared time and frequency domain techniques and stated that they complete each other. The presented techniques are said to be SOS-based, as opposed to HOS-based methods, requiring only the knowledge of the first two statistics of the involved processes, since the Gaussianity of the noise is inherently assumed [4, 15]. The time domain class contains among others [4, 5, 7] :

1. Correlation Method.

2. Least Squares (L_2 norm) Method.

3. Least Absolute Value (L_1 norm) Method.

4. Minimum Variance Deconvolution MVD Method.

While the frequency domain class contains [4, 5, 7] :

1. Spectral Analysis Method.

2. Wiener Filter A Method.

3. Wiener Filter B Method.

4. Spectral Extrapolation Method.

The cepstral or homomorphic class consists of two methods :

1. Complex Cepstrum method.

2. Power Cepstrum method.

Chen and Sin [7, 8, 12] found that the Wiener filter method is the most suitable SOS-based method for ultrasonic NDE deconvolution. In what follows details pertaining to this technique are given since the obtained results in this Thesis are compared to those obtained using Wiener filter method. Details about other techniques can be found in [7, 8, 12].

2.1.2   Wiener Filter A

A schematic representation of the ultrasonic NDE model given by Eq. 1.1 is shown in Fig. 2.1(a). Exploiting the *commutativity* of the linear deconvolution operation of Eq. 1.1, we can use $x(k)$ or $h_d(k)$ as input to the system of Fig. 2.1(a). Hence, the system shown in Fig. 2.1(a) can be put in the form shown in Fig. 2.1(b). The classical Wiener filter used as an inverse filter to retrieve the "*pseudo*" input to system $h_d(t)$ is shown in Fig. 2.1(c). This inverse filter is obtained by minimizing the following cost function [16] :

$$E\left[\left(h_d(k) - \hat{h}_d(k)\right)^2\right]$$
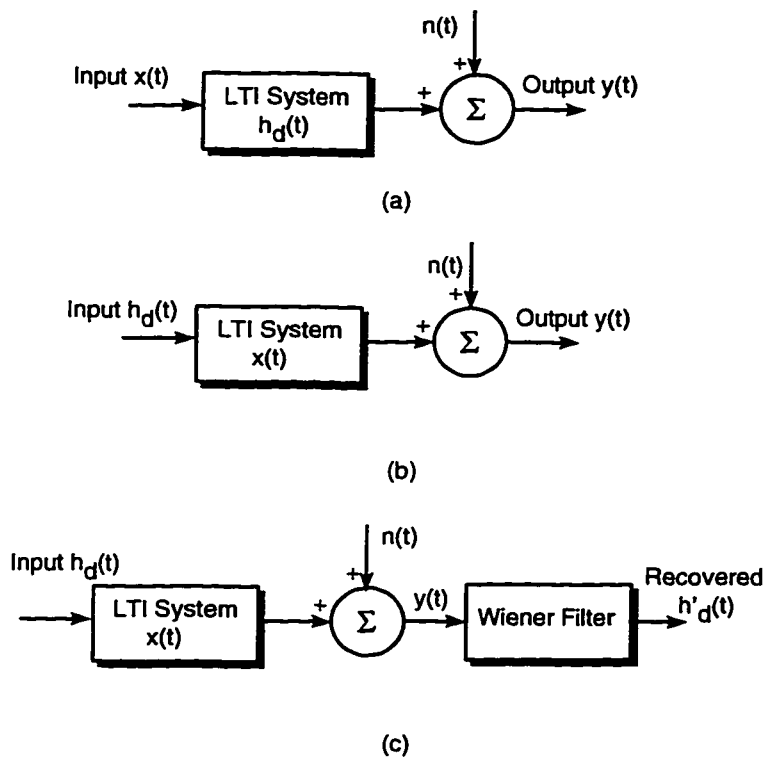
Figure 2.1: Wiener Filter Structure for Ultrasonic NDE Deconvolution.

and which is given by [16] :

$$\hat{H}_d(f) = \frac{Y(f)X^*(f)}{|X(f)|^2 + q} \tag{2.1}$$

where f is the frequency variable and $X^*(f)$ is the complex conjugate of the Fourier

transform of $x(k)$. q is a positive real number called the *noise-desensitizing factor*.

Walden [16] gave its optimal form for a robust Wiener deconvolution filter as (see [8,

12]) :

$$q(f) = p\frac{S_n(f)}{S_{h_d}(f)} \tag{2.2}$$

where p is a positive real number, $S_n(f)$ is the power spectral density (PSD) of the

additive noise $n(k)$ and $S_{h_d}(f)$ is the PSD of the $h_d(k)$. It should be pointed out

the major difficulties associated with Eq. 2.2 are :

- $S_{h_d}(f)$ is an unknown parameter.

- $S_n(f)$ is the noise PSD that is difficult to estimate a priori. A method for

  estimating the noise PSD is given in [4]. Such an estimate gives information

  on the actual SNR.

Neal [10] in his PhD Thesis developed a different optimal form of the Wiener filter,

and a corresponding value of the q parameter, with a priori statistical knowledge.

But, this seems quite far from the ultrasonic NDE problems reality. Neal [11], in

a private communication, recognized that this does not hold in practice. In the formulation of Wiener filter A of Eq. 2.1, $X(f)$ and $Y(f)$ are estimated directly as Fourier transforms of $x(k)$ and $y(k)$, respectively. This constitutes a violation to the stochastic nature of the signals $x(k)$ and $y(k)$ since they are corrupted by measurement noise. However, the asymptotic equivalence between the various classical *Fourier-Type* methods of PSD estimation is a known result in the literature [13, 17]. For small sample size, this result does not hold anymore as in the ultrasonic NDE case at hand. So caution must be used when using these estimation techniques for small sample size. Also, Wiener filter A suffers from its limited resolution and high variance. Besides these shortcomings, the computational efficiency and the implementation simplicity of this filter, must be highlighted as attractive and desirable properties. A summary of the Wiener filter A is given in Table 2.1.

2.1.3 Wiener Filter B

In an attempt to take advantage of the attractive properties of Wiener filter A and remedy to its shortcomings mentioned above, Wiener filter B is formulated using the following assumptions :

1. $x(k)$ in Eq. 1.1 is considered as the output of an AutoRegressive (AR) system whose input is a white driving noise.

2. $y(k)$ in Eq. 1.1 is considered as the output of another AR system whose input is a different white driving noise.

This formulation has removed the violation associated with the computation of the PSD's of $x(k)$ and $y(k)$. Hence, the high resolution characteristic of the parametric method of AR spectral estimation is gained at the expense of increased computational complexity (see Table 2.3 for details). Then, Eq. 2.1 becomes :

$$\hat{H}_d(f) = \frac{Y_{AR}(f)X_{AR}^*(f)}{|X_{AR}(f)|^2 + q} \tag{2.3}$$

where $X_{AR}(f)$ and $Y_{AR}(f)$ are the AR spectra of $x(k)$ and $y(k)$, respectively. A major difficulty is associated with the parametric AR modeling and consists of the selection of the orders of the models to be used in Eq. 2.3 of Wiener Filter B [17, 18, 19] . If, the Maximum Entropy Method (MEM) is used to estimate the involved PSD's in Eq. 1.2, the result will have higher resolution and the required order selection procedure wil be bypassed [20], since MEM is known to be *auto-order selector* [20]. However, the increased resolution may be affected by an artifact known as *"Spectrum Line Splitting"* (See Chen [21] for solutions to this problem). A summary of the Wiener filter B is given in Table 2.2.

2.1.4 Comparison and Discussion

Chen and Sin [2, 8, 12] discussed the usefulness and the limitations of some of the existing SOS-based deconvolution methods, and compared their computational complexity. First, they noted that the Wiener filter A suffers from the oscillations effects which are due to the band-limited nature of the involved signals. Wiener filter B has better resolution. The resolution enhancement is mainly due to the fact that this technique models better the signals $x(t)$ and $y(t)$ by AR models of Eq. 2.3. It is known that this representation (AR) is capable of perfectly modeling single spike signals, which have a transient nature [4, 15, 22]. This improvement is achieved at the expense of increased computational complexity. The spectral extrapolation can also well handle the modeling of the involved UNDE signals that have single spike nature [12]. However, this technique presents a limitation that consists of the determination of the spectral region where the SNR is known to be high. In the spectral analysis, where spectral windows are introduced to reduce the variance of the PSD estimates, the resolution of the estimation is increased without increasing much the computations. Among the time domain methods, the correlation analysis seems to require less computations and can even be efficiently implemented if FFT routines are used [13]. The LS approach is a time domain technique that requires perfect alignment of the signals before being run. Any misadjustment may cause severe degradation in the quality of the estimated IR (see

Chapter 5 for details). The L_1 method involves intensive computation load and is rarely implemented on Personal Computers (PC's) [8]. Besides this, this method is not really appropriate for the ultrasonic NDE deconvolution [7, 8, 12]. On the other hand, the use of a phase unwrapping procedure [13, 23], in the homomorphic deconvolution will increase its computational load. For comparison purposes, the following simplifications are needed [12] :

- The record length of $x(k)$ and $y(k)$ are both equal to N.

- The number of real multiplications required by an N-point FFT is the same as that required by an inverse FFT of the same order, which is approximately equal to $(3N/2)Log_2N$.

- A real division has the same complexity as a real multiplication.

- One complex multiplication is equivalent to four real multiplications.

An additional assumption is needed for the homomorphic method :

- *Phase Unwrapping* algorithm requires at minimum N complex multiplications.

The comparison is summarized in Table 2.3.

2.2 Automatic Defect Classification

Aiming to help in making correct decisions about possible defects present in the material under test, attempts have been made to classify these defects according to

their :

a) Geometry.

b) Size.

c) Orientation in the case of angular flaws.

To achieve such an objective, two different approaches have been adopted in the literature, namely :

- Pattern recognition approach using cluster analysis [2].

- Intelligent approach using artificial neural networks (ANN) [24].

The first approach is based on statistical pattern recognition techniques such as [2, 25] :

1) Bayes or maximum likelihood decision rules.

b) Nearest neighboor decision rule.

c) Tree classification method.

The second approach is motivated by the universal approximation capabilities of ANN [26]. Both approaches make direct use of $x(k)$ and $y(k)$ [2]. Hence, the classification is affected by the following factors :

---

[2]In the proposed classification system (see Chapters 4 and 5), the deconvolution using the proposed HOS-based scheme is considered as a preprocessing step.

- The additive noise $n(k)$ effects.

- The propagation paths effects.

- The transducers effects.

2.2.1 Pattern Recognition Approach

In his pioneering work, Chen [2, 25] developed the main ideas of the first approach, borrowing ideas from the statistical pattern recognition area. He proposed different features that convey information about the defects in the material under test. The proposed features can be cast into :

• Time domain features.

• Frequency domain features.

• Correlation domain features.

• Impulse response features.

Then, identified certain effective features to classify the defects according to their geometry and size. However, it should be noted that [2] :

1) Feature extraction is key to any pattern recognition problem. Features can be extracted mathematically, or based on the physical understanding of the problem at hand.

2) Features must provide significant discriminating information among the pattern classes considered, while they can properly describe the patterns.

3) It is unrealistic to exhaustly search for the best features. A good subset of features can always be selected from given ones.

4) Feature extraction is inherently problem dependent. Thus the most suitable features are not the same for different ultrasonic defects classification problems.

5) Most feature extraction/selection work has assumed that the signals are stationary which is not always valid in practice. A better classification performance can be achieved using joint time-frequency analysis (JTFA). The role of JTFA can be quite important in defect classification problems (see Chapters 4 and 5).

6) It is always difficult to evaluate theoretically the feature performance in pattern recognition.

After a plethora of tests, Chen [2] established several effective mathematical features for most UNDE classification problems. However, it should be noted that it is not possible to find a one-to-one mapping between the selected features and the defect class although there is good correlation between them. Details about effective features proposed by Chen can be found in [25] where they were used with the tree classification method.
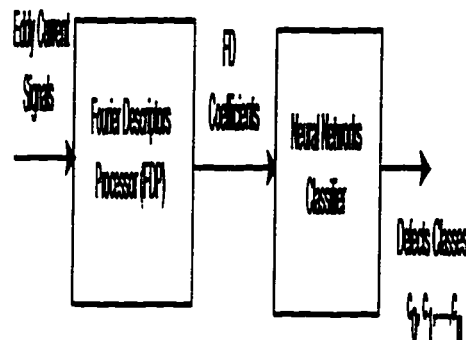
Figure 2.2: Block Diagram of the Classification Scheme.

2.2.2 Neural Networks Approach

Chen [2] suggested the use of ANN without giving details about the architecture and the implementation. Udpa and Udpa [24] proposed an ANN architecture for defect classification using Fourier descriptors (FD) [1, 24] . They motivated the use of Fourier descriptors for their invariance under [24] :

- Rotation.

- Scaling.

- Translation.

Their system is based on Eddy current [1] signals for which they derive FD coefficients. Then, these coefficients constitute the input pattern to the ANN for classification using the backpropagation (BP) algorithm. Fig. 2.2 shows the block diagram of the classification scheme. The network architecture is :

1) An Input layer of eight (08) neurons. The input vector consists of eight (08) FD coefficients.

2) A hidden layer of five (05) neurons.

3) An output layer of two (02) neurons. The output is coded in binary to represent one of the four possible defect classes.

2.3 Summary

In this chapter, a review of the existing ultrasoni NDE processing techniques is given. Two major signal processing tasks are reviewed. The first task is the SOS-based deconvolution methods that are used to unravel the masking effects of the propagation paths and measurement system on the defect IR. The second one consists of automatic ultrasonic defect classification using artificial intelligence approach. Most of the merits of the SOS-based deconvolution methods are highlighted, and their limitations revealed. For ease of analysis, tables summarizing their implementation are given. A comparative study of their relative computational complexity is presented in a summary table, to allow better analysis and comparison between these methods. Fig. 2.3 shows a tree structure of the algorithms being applied to ultrasonic NDE deconvolution. It is shown that *ill-conditionning* is a difficulty common to all the methods since deconvolution is involved. Also, their vulnerability to additive Gaussian noise is evident. A need for a remedy to these problems urged
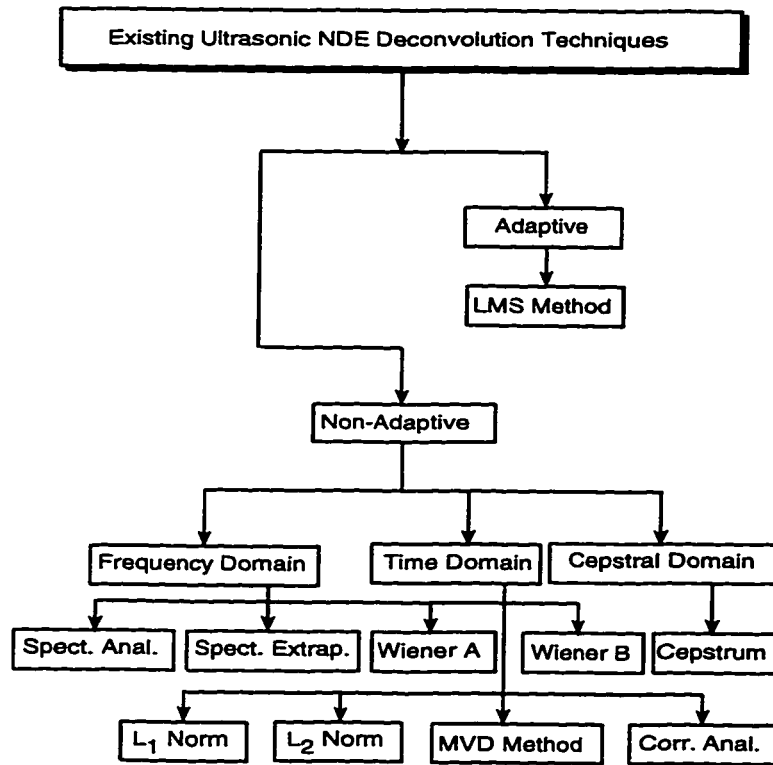
Figure 2.3: Tree structure of Classical Ultrasonic NDE Deconvolution Techniques.

the search for methods that are more robust to the noise present in the ultrasonic NDE model. Solutions using emerging HOS that proved robustness and immunity to additive Gaussian noise and have been successfully applied in communications, system identification and other applications [27, 28, 29] are applied to ultrasonic NDE deconvolution in the following chapters. Also, a new modular learning paradigm is proposed for ultrasonic defect classification is proposed.

The implementation of Wiener filter A requires:

$x(k)$: The reference signal

$y(k)$: The echo signal

(1) Take the Fourier Transform of $x(k)$ via FFT

(2) Take the Fourier Transform of $y(k)$ via FFT

$$X(f) = FFT[x(k)]$$
$$Y(f) = FFT[y(k)]$$

(3) Use Wiener A Equation to estimate the defect IR :

$$\hat{H}_d(f) = \frac{Y(f)X^*(f)}{|X(f)|^2+q}$$

where the factor q is given by :

$$q = 0.1 * max(|X(f)|)$$

(4) Take the inverse FFT of $\hat{H}_d(f)$ as the estimated IR :

$$\hat{h}_d(k) = IFFT(\hat{H}_d(f))$$

Note: The factor q may be changed to improve the quality of the estimated IR.

Table 2.1: Summary Table of Wiener Filter A Deconvolution Algorithm.

The implementation of Wiener filter B requires:

$x(k)$: The reference signal

$y(k)$: The echo signal

(1) Model $x(k)$ as the output of pth-order AR model.

(2) Model $y(k)$ as the output of qth-order AR model.

$$x(k) = \sum_{j=1}^{p} a_j x(k - j) + w(n)$$

$$y(k) = \sum_{j=1}^{q} b_j y(k - j) + v(n)$$

where the a_k's and b_k's coefficients can be found via LP or MEM methods.
where LP: Linear Prediction and MEM: Maximum Entropy (or Burg) Method.
Note 1: Both signals $x(k)$ and $y(k)$ can assume the same order p.

(3) Estimate the Power spectrum of both signals $x(k)$ and $y(k)$ via:

$$X_{AR}(f) = \frac{1}{\sum_{i=1}^{p} a_i e^{-j2\pi f i}}$$

$$Y_{AR}(f) = \frac{1}{\sum_{i=1}^{q} b_i e^{-j2\pi f i}}$$

(4) Use Wiener A Equation to estimate the defect IR :

$$\hat{H}_d(f) = \frac{Y_{AR}(f) X^*_{AR}(f)}{|X_{AR}(f)|^2 + q}$$

where the factor q is given by :

$$q = 0.1 * max(|X(f)|)$$

(5) Take the inverse FFT of $\hat{H}_d(f)$ as the estimated defect IR:

$$\hat{h}_d(k) = IFFT(\hat{H}_d(f))$$

Note 2: The factor q may be changed to improve the quality of the estimated IR.

Table 2.2: Summary Table of Wiener Filter B.

Table 2.3: Compuational Complexity of the deconvolution algorithms.

Method	No. of Real Multiplications Required	No. of Real Additions Required
Wiener Form A	$6N.Log_2(N) + 8N$	$9N.Log_2(N) + 2N$
Wiener Form B	$6N.Log_2(N) + 13N + 2NP$ (P is the order of the AR filter)	$9N.Log_2(N) + 3N + 2NP$
Spectral Extrapolation	$6N.Log_2(N) + 8N + 2NP$	$9N.Log_2(N) + 2N + 2NP$
Spectral Analysis	$6N.Log_2(N) + 9N$	$6N.Log_2(N) + 3N$
Cepstrum Method	$6N.Log_2(N) + 8N$	$6N.Log_2(N) + 9N$
Correlation Analysis	$\frac{9N}{2}.Log_2(N) + 6N$	$\frac{9N}{2}.Log_2(N) + 3N$
Least Squares (L_2 norm)	$8N.Log_2(N) + 6N + 2N^2$	$12N.Log_2(N) + 3N + 2N^2$

Chapter 3

Review Of Higher-Order

Statistics

3.1 Introduction

HOS is a rapidly evolving signal analysis area which finds broad applications in

engineering and science. Last years, have witnessed a great interest in HOS-based

techniques. Applications of HOS methods span a broad range that include telecom-

munications, sonar, radar, geophysics, image processing, speech processing and other

applications [27, 28, 29, 30]. Most of the efforts made in the last three decades in the

area of power spectrum estimation, were based on the properties of the autocorre-

lation function. However, the information contained in the autocorrelation function

provides a complete statistical description of Gaussian processes only, since then the

SOS-based methods almost exclusively rely on the Gaussianity assumption. Such an assumption has been justified in the past for leading to mathematically tractable solutions in the Least-Squares sense, and by the central limit theorem (CLT). By virtue of CLT theorem, the sum of many random processes tends to be normally distributed [31]. Furthermore, SOS are blind to the phase information and hence are appropriate only for the description of Gaussian processes governed by linear mechanisms. Recently, Johnson and Rao [32] have questioned the mere existence of Gaussian data in real problems. Also, in many applications where non-Gaussian processes or nonlinearities are present, SOS-based methods fail to provide correct description of the involved models. In such situations, one must look beyond the power spectrum of the process, into higher-order spectrum; i.e., polyspectra, to extract phase information, as well as information about deviation from Gaussianity and the presence of nonlinearities, such as quadratic and cubic phase coupling [30]. In the sequel, the motivations behind the use of HOS are given :

- Extract information due to deviation from Gaussianity.

- Recover the correct phase character of signals and systems.

- Detect and quantify some nonlinearities in time series.

- Suppress additive Gaussian noise of unknown spectrum characteristics (white or colored), also the third-order statistics suppress non-Gaussian white noise with symmetric probability density function (pdf).

The first motivation is based on mathematical relations between Gaussian and Non-Gaussian processes having same SOS description (statistical mean and autocorrelation), given in the following section. The second one relies on the attractive feature of HOS (cumulants and polyspectra) to preserve the true phase character of signals and systems, while the third motivation is a natural result of the introduction of HOS in the analysis of nonlinear systems operating under a random input. It should be noted also, that HOS are nonlinear in their formulation. The property that for Gaussian processes only, all cumulants spectra of order greater than two are identically zero, represents the basis for the fourth motivation. As it can be deduced from the motivations given above, HOS-based techniques become high signal-to-noise ratio (SNR) domains where detection, parameter estimation and signal reconstruction can be performed, provided that a sufficiently large number of signal samples are available [30]. Henceforth, any statistical solution using SOS can be recast into the HOS formulation automatically boosting the effective SNR. A polyspectra classification map is shown in Fig. 3.1. Reasons behind the use of moment or cumulant polyspectra will be given in the following sections.

3.2 Definitions

The material of this section is taken, for the most part from [27], [29], [30] and [33]. Given a set of n real random variables $\{x_1, x_2, \ldots, x_n\}$, their joint moments of order
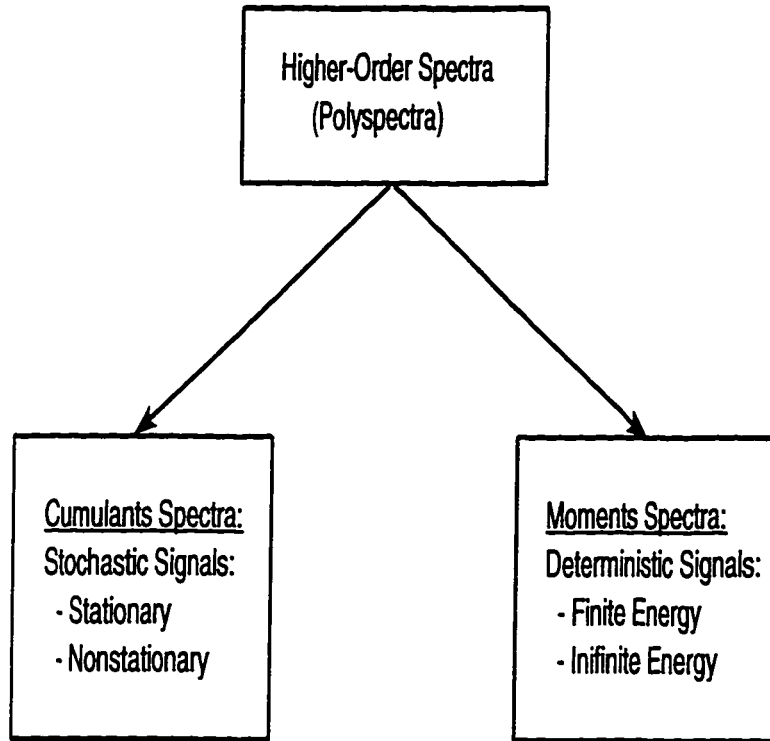
Figure 3.1: Polyspectra Classification Map.

$r = k_1 + k_2 + \cdots + k_n$ are given by [31] :

$$Mom[x_1^{k_1}, x_2^{k_2}, \ldots, x_n^{k_n}] \triangleq E\{x_1^{k_1} x_2^{k_2} \ldots x_n^{k_n}\}$$

$$= (-j)^r \frac{\partial^r \Phi(w_1, w_2, \ldots, w_n)}{\partial w_1^{k_1} \partial w_2^{k_2} \ldots \partial w_n^{k_n}} \bigg|_{w_1 = w_2 = \cdots = w_n = 0} \tag{3.1}$$

where

$$\Phi(w_1, w_2, \ldots, w_n) \triangleq E\{\exp(j(w_1 x_1 + w_2 x_2 + \cdots + w_n x_n))\} \tag{3.2}$$

is their joint characteristic function. $E\{.\}$ denotes the expectation operator. Another form of the joint characteristic function is defined as the natural logarithm of $\Phi(w_1, w_2, \ldots, w_n)$; i.e.,

$$\Psi(w_1, w_2, \ldots, w_n) \triangleq \log[\Phi(w_1, w_2, \ldots, w_n)] \tag{3.3}$$

This form of joint characteristic function is called *Cumulant-generating function* The joint cumulants of order r, denoted by $Cum[x_1^{k_1}, x_1^{k_1}, \ldots, x_n^{k_n}]$, of the same set of random variables, are defined as the coefficients associated with the Taylor series expansion of the cumulant-generating function about zero [30, 31]; i.e.,

$$Cum[x_1^{k_1}, x_2^{k_2}, \ldots, x_n^{k_n}] \triangleq (-j)^r \frac{\partial^r \Psi(w_1, w_2, \ldots, w_n)}{\partial w_1^{k_1} \partial w_2^{k_2} \ldots \partial w_n^{k_n}} \bigg|_{w_1 = w_2 = \cdots = w_n = 0} \tag{3.4}$$

Thus, the joint cumulants can be expressed in terms of the joint moments of a set of random variables (see Kendall and Stuart [34] for such relations up to order $r = 12$). For example the moments up to order 4 of one random variable $\{x_1\}$ are given by :

$$m_1 = Mom[x_1] = E[x_1]; \qquad m_2 = Mom[x_1, x_2] = E[x_1^2]$$

$$m_3 = Mom[x_1, x_2, x_3] = E[x_1^3]; \quad m_4 = Mom[x_1, x_2, x_3, x_4] = E[x_1^4] \tag{3.5}$$

The moments are related to the cumulants by :

$$c_1 = Cum[x_1] = m_1; \quad c_2 = Cum[x_1, x_2] = m_2 - m_1^2$$

$$c_3 = Cum[x_1, x_2, x_3] = m_3 - 3m_2m_1 + 2m_1^3$$

$$c_4 = Cum[x_1, x_2, x_3, x_4] = m_4 - 4m_3m_1 - 3m_2^2 + 12m_2m_1^2 - 6m_1^4 \qquad (3.6)$$

If the mean $E[x_1](i.e., m_1)$ is zero, then it follows that :

$$c_2 = m_2; \quad c_3 = m_3; \quad c_4 = m_4 - 3m_2^2$$

A similar representation of cumulants of vector processes is given in the work of Manzano [35]. From the definitions given above, it should be noted that the computation of nth-order cumulants requires the knowledge of the nth- and lower-order moments.

3.3 Moments and cumulants of stationnary processes

Let $\{x(t)\}$ be a real <u>nth-order stationary</u> random process and its moments up to order n exist, then

$$Mom[x(k), x(k + \tau_1), \ldots, x(k + \tau_{n-1})] \triangleq$$

$$E[x(k), x(k + \tau_1), \ldots, x(k + \tau_{n-1})] \tag{3.7}$$

Because of the stationarity assumption (weak of order n), the nth-order moments are only a function of the time difference $(\tau_1, \tau_2, \ldots, \tau_{n-1})$. Now, the moments of nth-order stationary process can be written as:

$$m_n^x(\tau_1, \tau_2, \ldots, \tau_{n-1}) \triangleq E[x(k), x(k + \tau_1), \ldots, x(k + \tau_{n-1})] \tag{3.8}$$

Similary, the nth-order cumulants of $\{x(t)\}$ are also (n-1)-dimensional functions of (n-1) variables, written in the form :

$$C_n^x(\tau_1, \tau_2, \ldots, \tau_{n-1}) \triangleq Cum[x(k), x(k + \tau_1), \ldots, x(k + \tau_{n-1})] \tag{3.9}$$

It is interesting to note that the nth-order cumulant function of a non-Gaussian stationary random process $\{x(t)\}$ can be written as (for $n = 3, 4$ only) [36] :

$$C_n^x(\tau_1, \tau_2, \ldots, \tau_{n-1}) = m_n^x(\tau_1, \tau_2, \ldots, \tau_{n-1}) - m_n^G(\tau_1, \tau_2, \ldots, \tau_{n-1}) \qquad (3.10)$$

where $m_n^x(\tau_1, \tau_2, \ldots, \tau_{n-1})$ is the nth-order moment function of $\{x(t)\}$ and $m_n^G(\tau_1, \tau_2, \ldots, \tau_{n-1})$ is the nth-order moment of an equivalent Gaussian process that has the same mean value and autocorrelation sequence as $\{x(t)\}$. Clearly, if $\{x(t)\}$ is Gaussian, $m_n^x(\tau_1, \tau_2, \ldots, \tau_{n-1}) = m_n^G(\tau_1, \tau_2, \ldots, \tau_{n-1})$ and thus $C_n^x(\tau_1, \tau_2, \ldots, \tau_{n-1})$ are identically zero. Note, however, that this is only true for orders $n = 3, 4$. Eq.3.10 gives a measure of the deviation of the random process from Gaussianity. Variance, skewness and kurtosis measures are given in terms of cumulant lags.

3.3.1   Moments Versus Cumulants

When developing algorithms based on HOS, cumulants are preferred to moments for the following reasons :

- As the covariance/correlation function of a white process is an impulse function and its spectrum is flat, the cumulants of higher-order white processes (definition will be given later) are multidimensional impulse functions, and their polyspectra (definition will be given later) are flat as well. Hence, the whiteness and flatness properties are conserved with cumulants, unlike with

moments. For illustration, Figure 3.2 gives 3-d plots of third-order cumulants and moments of third-order white one-sided exponential sequence.

- The cumulants of the sum of two statistically independent random processes equals the sum of the cumulants of each individual process, whereas the same is not true for moments. This property allows an easy manipulation of cumulants as a linear operator when processes are statistically independent.

- Higher-order moments of Gaussian processes are not zero, unlike cumulants. However, all the statistical information about Gaussian processes is contained in the moments of order $(n < 2)$ (i.e., mean and autocorrelation). Therefore, all moments of order equal or greater than two $(n \geq 2)$ have no new information to provide. This property is illustrated with the third-order moments and cumulants of a Gaussian process shown in Fig. 3.3.

3.4 Properties of Cumulants

Some useful properties of cumulants which are of key importance to this thesis, are summarized below for ease of later use, and reference.

P1. If λ_i, $i = 1, \ldots, n$ are constants and x_i, $i = 1, \ldots, n$ are random variables, then

$$Cum[\lambda_1 x_1, \lambda_2 x_2, \ldots, \lambda_n x_n] = \prod_{i=1}^{n} \lambda_i \, Cum[x_1, x_2, \ldots, x_n] \qquad (3.11)$$
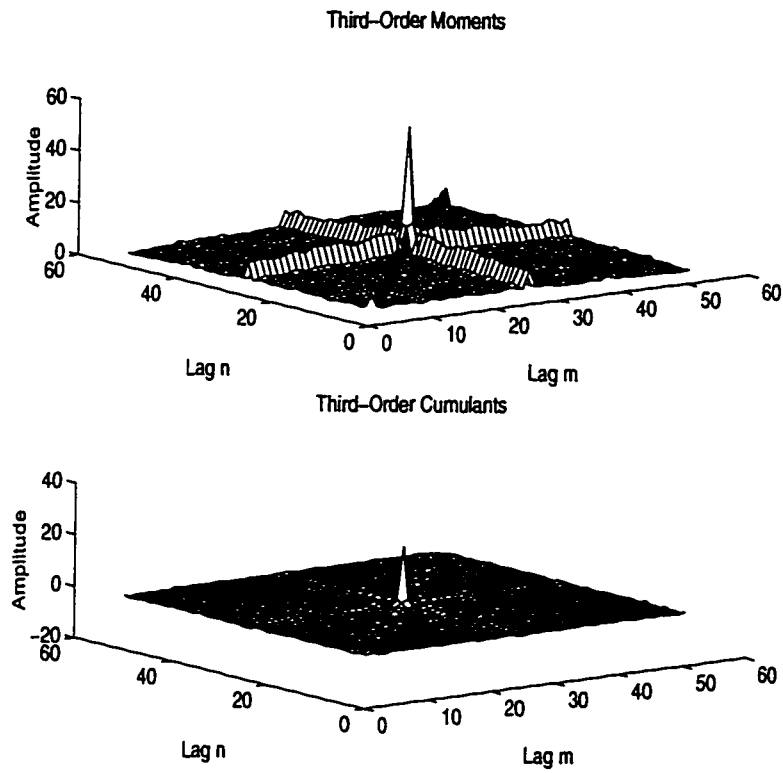
Figure 3.2: Third-Order Cumulants and Moments of One-Sided Exponential Sequence.
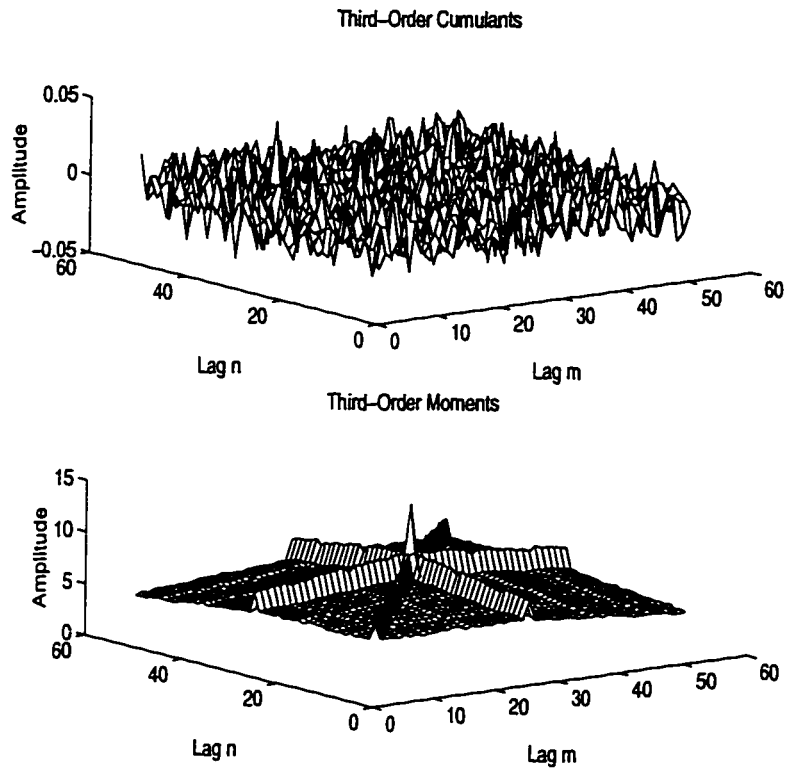
Figure 3.3: Third-Order Cumulants and Moments of non zero-mean Gaussian Sequence.

P2. Cumulants are symmetric in their arguments, i.e.,

$$Cum[x_1, x_2, \ldots, x_n] = Cum[x_{i1}, x_{i2}, \ldots, x_{in}] \qquad (3.12)$$

where $(i_1, i_2, \ldots, i_k)$ is a permutation of $(1, 2, \ldots, n)$. The symmetry property of cumulants will be further detailed in the sequel.

P3. If α is a constant, then

$$Cum[\alpha + x_1, x_2, \ldots, x_n] = Cum[x_1, x_2, \ldots, x_n] \qquad (3.13)$$

P4. If the random variables $\{x(t)\}$ are independent of the random variables $\{y(t)\}$, then for $t = 1, 2, \ldots, n$

$$Cum[x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n] =$$

$$Cum[x_1, x_2, \ldots, x_n] + Cum[y_1, y_2, \ldots, y_n] \qquad (3.14)$$

P5. If the random variables $\{x(t)\}$ with $t = 1, 2, \ldots, n$ can be divided into any two or more groups which are statistically independent, their nth-order cumulants are identical to zero, i.e.,

$$Cum[x_1, x_2, \ldots, x_n] = 0$$

P6. Cumulants are additive in their arguments :

$$Cum[x_1 + y_1, x_2, \ldots, x_n] =$$

$$Cum[x_1, x_2, \ldots, x_n] + Cum[y_1, x_2, \ldots, x_n] \tag{3.15}$$

this property is shared also by moments

$$Mom[x_1 + y_1, x_2, \ldots, x_n] =$$

$$Mom[x_1, x_2, \ldots, x_n] + Mom[y_1, x_2, \ldots, x_n] \tag{3.16}$$

P7. Cumulants are blind to any linear phase shift. Given a zero-mean stationary random process $\{x(t)\}$ with finite moments up to order n, and its delayed version $\{y(t)\}$ defined as: $y(t) = x(t - d)$, where d is a constant integer, then it follows that :

$$Cum[y(t), y(t + \tau_1), \ldots, y(t + \tau_{n-1})] =$$

$$Cum[x(t - d), x(t - d + \tau_1), \ldots, x(t - d + \tau_{n-1})] \tag{3.17}$$

because of the stationarity assumption, $Cum[x(k - d), x(k - d + \tau_1), \ldots, x(k - d + \tau_{n-1})]$ will depend only on the time differences $(\tau_1, \tau_2, \ldots, \tau_{n-1})$, hence,

Eq. 3.17 becomes :

$$Cum[y(t), y(t + \tau_1), \ldots, y(t + \tau_{n-1})] = C_n^x(\tau_1, \tau_2, \ldots, \tau_{n-1})$$

$$= Cum[x(t), x(t + \tau_1), \ldots, x(t + \tau_{n-1})] \qquad (3.18)$$

which implies that process $\{x(t)\}$ and $\{y(t)\}$ have identical cumulants. Stated in a different way, *cumulants suppress any linear phase shift*. However, cross-cumulants do preserve linear phase shifts, which makes them suitable for time delay estimation [30].

3.5 Cumulants Spectra (Polyspectra)

Given a real random process $\{x(t)\}$, stricly (or weakly) stationary, up to order n, with nth-order cumulants sequence defined as :

$$C_n^x(\tau_1, \tau_2, \ldots, \tau_{n-1})$$

its nth-order spectra (or *polyspectra*) are defined below :

Definition 3.5.1 *[30, 37, 38] The nth-order cumulants spectra of the process $\{x(t)\}$, denoted by $C_n^x(w_1, w_2, \ldots, w_{n-1})$ exist and are defined as the (n-1)-dimensional Fourier transforms of the nth-order cumulants, if $C_n^x(\tau_1, \tau_2, \ldots, \tau_{n-1})$ satisfy the following*

condition :

$$\sum_{\tau_1=-\infty}^{\infty} \sum_{\tau_2=-\infty}^{\infty} \cdots \sum_{\tau_{n-1}=-\infty}^{\infty} |C_n^x(\tau_1, \tau_2, \ldots, \tau_{n-1})| < \infty \tag{3.19}$$

in other words, the nth-order cumulants sequence must be stable. Other expressions for the condition of polyspectra existence can be found in the seminal work of Brillinger and Rosenblatt [37, 38] [1]. Eq. 3.19 represents the higher extension of the usual conditions for a Fourier transform to be well defined. The nth-order cumulant spectra is thus defined :

$$C_n^x(w_1, w_2, \ldots, w_{n-1}) \triangleq \sum_{\tau_1=-\infty}^{\infty} \sum_{\tau_2=-\infty}^{\infty} \cdots \sum_{\tau_{n-1}=-\infty}^{\infty} C_n^x(\tau_1, \tau_2, \ldots, \tau_{n-1})$$
$$\exp\{-j(w_1\tau_1 + w_2\tau_2 + \cdots + w_{n-1}\tau_{n-1})\} \tag{3.20}$$

In general, $C_n^x(w_1, w_2, \ldots, w_{n-1})$ are complex, i.e., they can be written in terms of magnitude and phase :

$$C_n^x(w_1, \ldots, w_{n-1}) \triangleq |C_n^x(w_1, \ldots, w_{n-1})| \exp\{j\Psi_n^x(w_1, \ldots, w_{n-1})\} \tag{3.21}$$

---

[1]Brillinger and Roseblatt [37, 38] make use of the following condition :

$$\sum_{\tau_1=-\infty}^{\infty} \sum_{\tau_2=-\infty}^{\infty} \cdots \sum_{\tau_{n-1}=-\infty}^{\infty} (1 + |\tau_j|) |C_n^x(\tau_1, \tau_2, \ldots, \tau_{n-1})| < \infty$$

for $j = 1, 2, \ldots, n - 1$, which is stronger than Eq. 3.19. It ensures faster decay to zero of the cumulants than Eq. 3.19.

where $|C_n^x(w_1, \ldots, w_{n-1})|$ and $\Psi_n^x(w_1, \ldots, w_{n-1})$ are the polyspectra magnitude and phase, respectively. The cumulant spectra is also periodic with period 2π with respect to all its arguments, i.e., :

$$C_n^x(w_1, w_2, \ldots, w_{n-1}) = C_n^x(w_1 + 2\pi, w_2 + 2\pi, \ldots, w_{n-1} + 2\pi)$$

Historical notes regarding the definition and terminology of polyspectra can be found in [30, 37, 38]. Other definition than that given in Eq. 3.20 is possible using the Cramer spectral representation [30], in this thesis the definition of Eq. 3.20 is preferred. However, the Cramer spectral representation allows the introduction of the nth frequency variable w_n, that when added to the remaining frequency variables will yield zero. When using the latter definition, the notion of the hidden frequency $w_n = -w_1 - w_2 - \cdots - w_{n-1}$ should be kept in mind [30]. In practice, only few special cases of polyspectra gained interest. This is mainly due to the fact that higher-dimensional Fourier transforms do not have easy interpretation and their manipulation is quite difficult.

3.6 Polyspectra of Linear Systems

Let $\{x(t)\}$ be a non-Gaussian random process with all its moments finite up to order n (nth-order stationarity) and with cumulants spectra $C_n^x(w_1, w_2, \ldots, w_{n-1})$, and $h(t)$ be the IR of an LTI system. The Barlett-Brillinger-Rosenblatt [35, 37, 38]

summation formula relates the input $\{x(t)\}$, and output $\{y(t)\}$ of the LTI system through $h(t)$ as follows :

$$C_n^y(w_1, \ldots, w_{n-1}) \triangleq$$

$$H(w_1) \ldots \ldots H(w_{n-1}).H^*(w_1 + \cdots + w_{n-1}).C_n^x(w_1, \ldots, w_{n-1}) \qquad (3.22)$$

where

$$H(w) \triangleq \int_{-\infty}^{\infty} h(t) \exp\{-jwt\}dt$$

is the stable LTI system TF. Since $H(w)$ and $C_n^x(w_1, w_2, \ldots, w_{n-1})$ are complex, in general, they can be written as :

$$H(w) = |H(w)| \exp\{j\phi_h(w)\} \qquad (3.23)$$

where $|H(w)|$ and $\phi_h(w)$ are the magnitude and phase of the transfer function of the LTI system, respectively, and

$$C_n^x(w_1, w_2, \ldots, w_{n-1}) = |C_n^x(w_1, w_2, \ldots, w_{n-1})|$$

$$\exp\{j\Psi_n^x(w_1, w_2, \ldots, w_{n-1})\} \qquad (3.24)$$

where $|C_n^x(w_1, w_2, \ldots, w_{n-1})|$ and $\Psi_n^x(w_1, w_2, \ldots, w_{n-1})$ are the magnitude and phase of the nth-order cumulants spectra of the input $\{x(t)\}$, respectively. Then, Eq. 3.22 is expressed as follows:

$$|C_n^y(w_1, w_2, \ldots, w_{n-1})| = |H(w_1)|.|H(w_2)| \ldots |H(w_{n-1})|.$$

$$|H^*(w_1 + w_2 + \cdots + w_{n-1})|.|C_n^x(w_1, w_2, \ldots, w_{n-1})| \qquad (3.25)$$

and

$$\Psi_n^y(w_1, w_2, \ldots, w_{n-1})\} = \phi_h(w_1) + \phi_h(w_2) + \cdots + \phi_h(w_{n-1})$$

$$-\phi_h(-w_1 - w_2 - \cdots - w_{n-1}) + \Psi_n^x(w_1, w_2, \ldots, w_{n-1}) \qquad (3.26)$$

Eq. 3.26 shows that the output cumulants spectra of order $n > 2$ carry phase information about the LTI system. By setting n to specific values in Eq. 3.22, useful special cases of polyspectra are obtained.

1. n= 2 Power Spectrum.

2. n= 3 Bispectrum.

Details about the case $(n = 3)$ are given in [39].

3.6.1 LTI systems driven by white noise

Theorem 3.6.1 *[40] Let $H(w)$ be the transfer function of a finite-dimensional, LTI, causal, exponentially stable system whose impulse response is $h(t)$. System input $\{x(t)\}$ is an independent and identically distributed (i.i.d.) sequence with nth-order cumulants given by :*

$$C_n^x(\tau_1, \tau_2 \ldots, \tau_{n-1}) \triangleq \begin{cases} \gamma_n^x & \tau_1 = \tau_2 = \cdots = \tau_{n-1} = 0 \\ 0 & otherwise , \end{cases} \qquad (3.27)$$

where γ_n^x denotes the nth-order cumulant of the random variable $\{x(t)\}$, which exists if the nth-and lower-order central moments of $\{x(t)\}$ exist. Then, the nth-order output cumulants are give by :

$$C_n^y(\tau_1, \tau_2 \ldots, \tau_{n-1}) = \gamma_n^x \sum_{i=0}^{\infty} h(i)h(i + \tau_1) \ldots h(i + \tau_{n-1}) \qquad (3.28)$$

Using Theorem 3.6.1, Eq. 3.22, in the special case where $\{x(t)\}$ is non-Gaussian nth-order white, will have the following form [30, 36, 37, 38] :

$$C_n^y(w_1, \ldots, w_{n-1}) \triangleq$$

$$\gamma_n^x H(w_1) \ldots . H(w_{n-1}).H^*(w_1 + \cdots + w_{n-1}). \qquad (3.29)$$

where γ_n^x is the nth-order cumulant spectra.

3.7 Cumulants Cepstra (Polycepstra)

3.7.1 Preliminary

Extensions of the power spectrum to higher orders have been introduced in the last section. Similarly, in this section, the complex cepstrum of higher-order cumulants (polycepstra) is discussed. This section provides an exposition of the properties and computation of polycepstra, where their applications to nonminimum phase signal reconstruction and identification are highlighted. Concepts useful to the derivation of the polycepstra-based deconvolution technique presented in the next chapter, are introduced.

3.7.2 Definitions

The impulse response $h_d(t)$ of Eq. 1.1, is in general modeled as an AutoRegressive Moving Average (ARMA) energy sequence. Its Z transform is generally nonminimum phase (has zeros outside the unit circle) and can be written as [30] :

$$H(z) = Az^{-r}I(z^{-1})O(z) \qquad (3.30)$$

where A is a constant defined as "*system gain*", and r is an integer representing the "*system delay*", and

$$I(z^{-1}) = \frac{\prod_{i=1}^{L_1}(1 - a_i z^{-1})}{\prod_{i=1}^{L_3}(1 - c_i z^{-1})} \qquad (3.31)$$

$$O(z) = \prod_{i=1}^{L_2}(1 - b_i z^{-1}) \qquad (3.32)$$

are the minimum and maximum phase components, respectively with $|a_i| < 1$, $|b_i| < 1$ and $|c_i| < 1$. Poles $\{c_i\}$ and zeros $\{a_i\}$ are inside the unit circle whereas zeros $\{\frac{1}{b_i}\}$ are outside. There is no restriction to have also poles outside the unit circle. The minimum phase component of $h_d(k)$ is given by :

$$i(k) = \frac{1}{2\pi} \sum_{0}^{N} I(w) \cdot e^{jwk}, \quad k \geq 0, \quad \text{(causal part)}, \qquad (3.33)$$

whereas the maximum phase component is given by

$$o(k) = \frac{1}{2\pi} \sum_{0}^{N} O(w) \cdot e^{jwk}, \quad k < 0, \quad \text{(anti-causal part)}, \qquad (3.34)$$

where $I(w) = I(e^{-jw})$ and $O(w) = O(e^{jw})$. A linear convolution operation relates $i(k)$ and $o(k)$ with $h_d(k)$ as follows :

$$h_d(k) = \sum_{-N}^{N} i(\tau).o(k - \tau)$$

(3.35)

Now, to compute the complex cepstrum, the natural logarithm of Eq. 3.30 is used :

$$\log[H(z)] = \log(|A|) + \log(z^{-r}) + \sum_{i=1}^{L_1} \log(1 - a_i z^{-1})$$
$$+ \sum_{i=1}^{L_2} \log(1 - b_i z) - \sum_{i=1}^{L_3} \log(1 - c_i z^{-1})$$

(3.36)

The term z^{-r} corresponds to a linear phase shift which is suppressed in HOS domain (see properties of cumulants in section 3.4). The gain parameter A does not affect the shape of the signal, but causes a scale ambiguity that may be serious in some cases [13]. The algebraic sign of A can be determined via a procedure due to Hayes et al. [41] given specific conditions. Hatzinakos and Nikias [42] estimate A using second- and fourth-order cumulants at zero lags (i.e., $c_2^x(0)$ and $c_4^x(0,0,0)$) in the case of communications signals. Now, taking the inverse Z transform of Eq. 3.36, yields the complex cepstrum [13] :

$$c_h(m) = \frac{1}{2\pi} \oint_C \log[H(z)]z^{m-1}dz = Z_1^{-1}[\log[H(z)]]$$

(3.37)

where the contour of integration C is within the region of convergence which includes the unit circle. Using FT, Eq. 3.37 is written as :

$$c_h(m) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log[H(w)] \exp\{jwm\}dw = F_1^{-1}[\log|H(w)| + j\phi_h(w)] \quad (3.38)$$

where $|H(w)|$, and $\phi_h(w)$ are the magnitude and phase of the transfer function $H(w)$, respectively. $Z_1[.]$, $Z_1^{-1}[.]$ denote the 1-d forward and inverse Z transforms, respectively. Also, $F_1[.]$, $F_1^{-1}[.]$ denote the 1-d forward and inverse Fourier transforms, respectively. Problems related to the continuity of the complex logarithm have been adressed in Chapter 2 where phase unwrapping technique due to Tribolet [23] was presented, to overcome the phase discontinuity problem (see Chapter 2 for details). The complex cepstrum as defined, can be related to the underlying ARMA model parameters (poles and zeros) given in Eqs. 3.30, 3.32 as follows [13, 30, 43] :

$$c_h(m) = \begin{cases} \log|A|, & m = 0 \\ -\frac{A^{(m)}}{m}, & m > 0 \\ \frac{B^{(-m)}}{m}, & m < 0 \end{cases} \quad (3.39)$$

where

$$A^{(m)} = \sum_{i=1}^{L_1} a_i^m - \sum_{i=1}^{L_3} c_i^m \quad (3.40)$$

$$B^{(m)} = \sum_{i=1}^{L_2} b_i^m \tag{3.41}$$

are *cepstral parameters* which contain the minimum and maximum phase information, respectively. Note that for MA signals (systems), the second term of $A^{(m)}$ in Eq. 3.40 is identically zero.

3.7.3   Bicepstrum of Deterministic and Random Signals

Since, the complex cepstrum is a nonparametric signal processing technique, where no statistical model is assumed a priori, a common framework for both deterministic and random signals is possible. Using Theorem 3.6.1, any stochastic process $\{x(k)\}$ can be modeled as the output of an LTI system driven by a zero-mean non-Gaussian white process with skewness γ_3^ξ such that :

$$C_3^x(z_1, z_2) = \gamma_3^\xi H(z_1) H(z_2) H(z_1^{-1}.z_2^{-1}) \tag{3.42}$$

with

$$C_3^x(z_1, z_2) = \sum_{\tau_1=-\infty}^{\infty} \sum_{\tau_2=-\infty}^{\infty} c_3^x(\tau_1, \tau_2) z_1^{-\tau_1} z_2^{-\tau_2} \tag{3.43}$$

and $c_3^x(\tau_1, \tau_2)$ are the third-order cumulants of $\{x(t)\}$, defined previously. On the other hand, if $h(k)$ is a deterministic energy signal, the 2-d Z transform of its third-

order moments is given by [30] :

$$M_3^h(z_1, z_2) = H(z_1)H(z_2)H(z_1^{-1}.z_2^{-1}) \tag{3.44}$$

Note that Eq. 3.42 and Eq. 3.44 are similar apart a scaling factor γ_3^ξ, although they represent signals of different classes (random and deterministic). This similarity allows to have a common framework for both types of signals, assuming γ_3^ξ equal to unity, without loss of generality. No distinction will be made further between the two types of signals, the term $C_3^x(z_1, z_2)$ is used henceforward. Pan and Nikias [43] were the first to define the complex cepstrum of higher-order cumulants.

Definition 3.7.1 *[43] The bicepstrum is defined as the inverse 2-d Z transform of the natural logarithm of the bispectrum, i.e.,*

$$b_h(m, n) = Z_2^{-1}\{\log[C_3^h(z_1, z_2)]\} \tag{3.45}$$

where $C_3^h(z_1, z_2)$ is the bispectrum of $H(z)$ defined in Eq. 3.30.

Pan and Nikias [43] showed that :

$$
b_h(m, n) = \begin{cases}
\log|A^3| & , \quad m = 0, \ n = 0 \\[1.5ex]
-\frac{1}{n}A^{(n)} & , \quad m = 0, \ n > 0 \\[1.5ex]
-\frac{1}{m}A^{(m)} & , \quad m > 0, \ n = 0 \\[1.5ex]
\frac{1}{m}B^{(-m)} & , \quad m < 0, \ n = 0 \\[1.5ex]
-\frac{1}{n}B^{(-n)} & , \quad m = 0, \ n < 0 \\[1.5ex]
-\frac{1}{n}B^{(n)} & , \quad m = n > 0 \\[1.5ex]
\frac{1}{n}B^{(-n)} & , \quad m = n < 0 \\[1.5ex]
0 & , \quad \text{otherwise}
\end{cases}
\tag{3.46}
$$

where $A^{(m)}$ and $B^{(m)}$ are the cepstral parameters defined in Eq. 3.40 and Eq. 3.41, respectively. When comparing Eq. 3.46 with that giving the complex cepstrum, i.e., Eq. 3.39, it clearly appears that :

$$
c_h(m) = b_h(m, 0) = b_h(0, m) = b_h(-m, -m), \quad m \neq 0
\tag{3.47}
$$

The definition of the bicepstrum in Eq. 3.45 suffers from the same difficulty as that of the complex cepstrum, i.e., the phase unwrapping problem [13, 23, 43]. In an attempt to overcome this problem, Pan and Nikias [43] derived the differential bicepstrum corresponding to the differential cepstrum proposed by Polydoros and

Fam [44]. The differential bicepstrum is defined as :

$$
db_h(m, n) = \begin{cases}
A^{(m-1)} & , \ m \geq 2, \ n = 0 \\[2mm]
B^{(1-m)} & , \ m \leq 2, \ n = 0 \\[2mm]
-A^{(-n)} & , \ m = n + 1, \ n \leq -1 \\[2mm]
B^{(n)} & , \ m = n + 1, \ n \geq 1 \\[2mm]
0 & , \ \text{otherwise}
\end{cases}
\tag{3.48}
$$

Remarks:

Several remarks regarding the continuity of the complex logarithm used in Eq. 3.45 are made :

- $C_3^x(z_1, z_2)$ is a rational polynomial in z_1 and z_2.

- 2-d phase unwrapping algorithm [45] is needed.

- Analycity of $C_3^x(z_1, z_2)$ must be guaranteed.

Dudgeon [45] defined the conditions for the existence of cepstra for 2-d rational polynomials in the following lemmma.

Lemma 3.7.1 *[45] Any 2-d array having a rational Z transform will have a well defined 2-d complex cepstrum provided :*

(i) *Its Fourier transform is not equal to zero or infinity,*

(ii) *Any linear phase trend because of time shift is eliminated.*

With third-order cumulants, the first condition is satisfied by the use of 2-d exponential windows to move zeroes (or poles) away from the unit bicircle. The second condition of Lemma 3.7.1 is always satisfied when using HOS quantities, since any linear phase term is suppressed by higher-order cumulants. Pan and Nikias [43] developed an LS approach for the computation of cepstral parameters from third-order cumulants as follows :

$$\sum_{\tau=1}^{\infty} \{ A^{(\tau)} [m_3^h(n-\tau,l) - m_3^h(n+\tau,l+\tau)]$$
$$+ B^{(\tau)} [m_3^h(n-\tau,l-\tau) - m_3^h(n+\tau,l)] \} = -n.m_3^h(n,l) \qquad (3.49)$$

where $m_3^h(n,l)$ represents the third-order moments of $h(t)$. Eq. 3.49 provides a direct relationship between cepstral parameters $A^{(\tau)}$ and $B^{(\tau)}$, and the third-order statistics. Using the properties of the cepstral parameters given previously (exponential decay with τ), Eq. 3.49 can be truncated to obtain an approximate form as :

$$\sum_{\tau=1}^{p} A^{(\tau)} [m_3^h(n-\tau,l) - m_3^h(n+\tau,l+\tau)]$$
$$+ \sum_{j=1}^{q} B^{(j)} [m_3^h(n-j,l-j) - m_3^h(n+j,l)] \cong -n.m_3^h(n,l) \qquad (3.50)$$

The truncation points, p and q have been empirically chosen by Pan and Nikias [43]

as :

$$p = \frac{\log(c)}{\log(a)}$$
$$q = \frac{\log(b)}{\log(a)} \tag{3.51}$$

where

$$\max[|a_i|, |c_j|] < a < 1, \quad 1 \le i \le L_1, \text{ and } 1 \le j \le L_3$$

$$\max[|b_k|] < b < 1, \quad 1 \le k \le L_2 \tag{3.52}$$

where a_i, b_i and c_i are defined in Eq. 3.30. c is avery small number, such that $A^{(\tau)} = 0$ for $\tau > p$ and $B^{(\tau)} = 0$ for $\tau > q$. The truncation limits in Eq. 3.50 require some a priori knowledge about the pole-zero location of $H(z)$, i.e., the values a and b in Eq. 3.52. Pan and Nikias [43] used the following rule of thumb to determine limits of the arguments (n, l) of $m_3^h(n, l)$:

$$w = \max[p, q] \quad \text{and} \quad z = \lfloor \frac{w}{2} \rfloor$$

$$n = -w, \ldots, 0, \ldots, w, \quad \text{and} \quad l = -z, \ldots, 0, \ldots, z,$$

where $\lfloor a \rfloor$ denotes nearest integer to a towards $-\infty$.

z and w are integers to form the overdetermined system of equations:

$$Ya = y \qquad (3.53)$$

Y is a $[(2w+1) \times (2z+1)] \times (p+q)$ matrix whose entries are expressions of the form

$\{m_3^h(\mu, \lambda) - m_3^h(\sigma, \nu)\}$. a is the unknown $(p+q) \times 1$ vector of unknown parameters

in the form :

$$a = [A^{(1)}, A^{(2)}, \ldots, A^{(p)}, B^{(1)}, B^{(2)}, \ldots, B^{(q)}]^T$$

and y is a $[(2w+1) \times (2z+1)] \times 1$ vector whose entries are terms of the form

$\{n.m_3^h(n, l)\}$. The LS solution of Eq. 3.53 is given by [30, 43] :

$$\hat{a} = [Y^T Y]^{-1} Y^T y \qquad (3.54)$$

An alternative method for computing the cepstral parameters in Eq. 3.46 based on

FFT's was proposed by Pan and Nikias [30, 43] :

$$m.b_h(m, n) = F_2^{-1} \left\{ \frac{F_2[\tau_1.m_3^h(\tau_1, \tau_2)]}{F_2[m_3^h(\tau_1, \tau_2)]} \right\} \qquad (3.55)$$

where F_2 denotes the 2-d Fourier transform and F_2^{-1} its inverse.

Chapter 4

Proposed Solutions

4.1   Introduction

Solutions to two distinct signal processing tasks for UNDE are proposed in this chapter. These solutions are :

1) HOS-based deconvolution of ultrasonic defect IR to unravel the masking effects of the transducer and the propagation paths.

2) Ultrasonic defect classification using a new modular learning strategy based on different functional blocks.

4.2   HOS-Based Deconvolution

The general convolutional model of UNDE is given by Eq. 1.1, which is reproduced below for convenience :

$$y(k) = x(k) * h_d(k) + n(k) \qquad (4.1)$$

assumptions usually made about Eq. 4.1 are given in Chapter 2. The recovery of $x(k)$ or $h_d(k)$ from noisy environment, represents a difficult signal processing task with a broad range of applications. The techniques developed in Chapter 2 can be cast into HOS formulation, to produce more robust deconvolution schemes. The HOS-based deconvolution method, developed in this chapter, belong to the polycepstra approach. Fig. 4.1 gives a classification map of the different HOS approachs.

4.2.1   Model and Assumptions

In the HOS context, the reference and echo signals, $x(k)$ and $y(k)$, respectively, are assumed to satisfy the following assumptions :

1. Shanks or Prony [46] modeling of both signals does not produce poles or zeros on the unit circle to guarantee the existence of stable HOS quantities and yield a well-posed deconvolution problem.

```
┌─────────────────────────┐
│     HOS-Based UNDE       │
│   Deconvolution Methods  │
└─────────────────────────┘
```
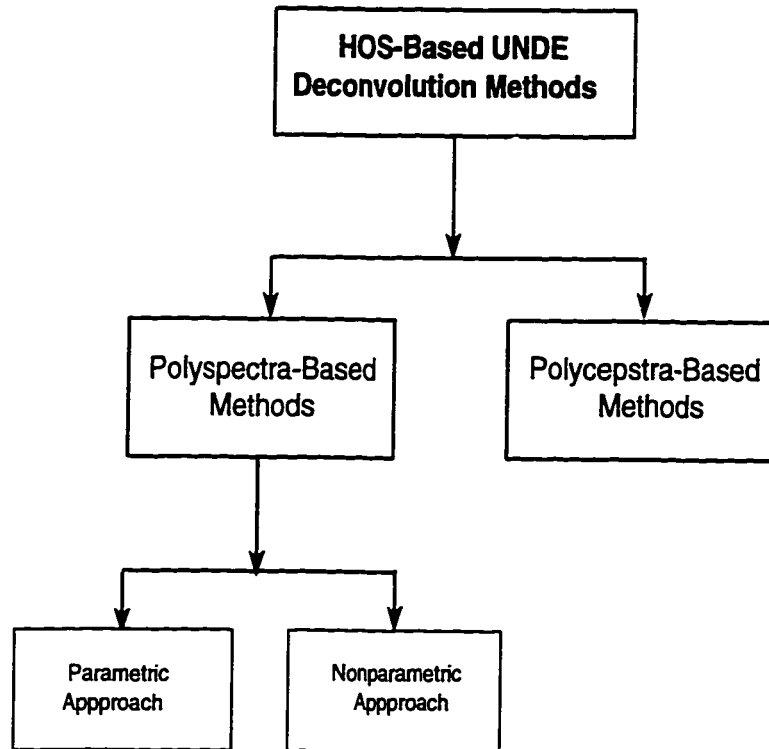
Figure 4.1: HOS-Based Ultrasonic NDE Deconvolution Classification Map.

2. No pronounced resonance and antiresonance exist in both signals i.e., poles and zeros are not located in the vicinity of the unit circle.

Some constraints made on the model of Eq. 4.1 in the SOS context are relaxed when using HOS formulation. Among the relaxed constraints, there are :

(i) The noise, $n(k)$, does not have to be white, it can be correlated (i.e., colored) in the Gaussian case.

(ii) When using third-order statistics, the noise $n(k)$ can assume any symmetric pdf (for example Laplacian pdf) if the whiteness condition is satisfied [47].

The proposed HOS-based deconvolution scheme utilizes the complex bicepstrum (see Chapter 3). Extensions to the fourth-order methods are quite straightforward. Other HOS-based schemes using the bispectrum can be proposed however they suffer from the large estimation errors engendered by the signal recovery method [48, 49].

4.2.2 The Complex Bicepstrum Method

Using results of Chapter 3 and the model assumptions discussed in the previous section, the third-order cumulants of Eq. 4.1 are [27, 28, 30, 37, 38] :

$$c_3^y(\tau_1, \tau_2) = c_3^x(\tau_1, \tau_2) . \sum_{k=0}^{\infty} h_d(k) h_d(k + \tau_1) h_d(k + \tau_2) \qquad (4.2)$$

where $c_3^y(\tau_1, \tau_2)$ and $c_3^x(\tau_1, \tau_2)$, are the third-order cumulants of $y(k)$ and $x(k)$, respectively. The noise $n(k)$ has zero third-order cumulants since it is assumed Gaussian (white or colored) [27, 30, 33]. Eq. 4.2 is the Barlett-Brillinger-Rosenblatt summation formula defined in Chapter 3. The form of Eq. 4.2 assumes causality of the system given by $h_d(k)$. Taking 2-d Fourier transform of both terms of Eq. 4.2, gives the following bispectral relation :

$$B_3^y(w_1, w_2) = B_3^x(w_1, w_2) . H(w_1) . H(w_2) . H^*(w_1 + w_2) \qquad (4.3)$$

$B_3^y(w_1, w_2)$, $B_3^x(w_1, w_2)$ and $H(w)$ are the bispectra of $y(k)$, $x(k)$ and the system TF. It should be noted that $H^*(w_1 + w_2) = H(-w_1 - w_2)$, since $h_d(k)$ is assumed a real sequence. In general, $H(w)$ is complex, hence :

$$H(w) = |H(w)| \cdot e^{j\phi(w)} \tag{4.4}$$

where $|H(w)|$ and $\phi(w)$ are the magnitude and phase response. The notation used in Eq. 4.4 applies also to polyspectra quantities, such that :

$$B_3^y(w_1, w_2) = |B_3^y(w_1, w_2)| \cdot e^{j\Psi_3^y(w_1, w_2)} \tag{4.5}$$

and

$$B_3^x(w_1, w_2) = |B_3^x(w_1, w_2)| \cdot e^{j\Psi_3^x(w_1, w_2)} \tag{4.6}$$

Similarly, Eq. 4.3 is written using Eqs. 4.4, 4.5 and 4.6 as :

$$|B_3^y(w_1, w_2)| = |B_3^x(w_1, w_2)| \cdot |H(w_1)| \cdot |H(w_2)| \cdot |H^*(w_1 + w_2)| \tag{4.7}$$

and

$$\Psi_3^y(w_1, w_2) = \Psi_3^x(w_1, w_2) + \phi(w_1) + \phi(w_2) - \phi(w_1 + w_2) \tag{4.8}$$

The last two relations form the basis for the deconvolution of $h_d(k)$. An alternative approach is to get directly the "*bispectrum* $H(w_1, w_2)$ ". The following lemma gives the definition of the IR bispectrum.

Lemma 4.2.1 *The discrete IR, $h(k)$, has a bispectrum given by :*

$$H(w_1, w_2) = H(w_1) \cdot H(w_2) \cdot H^*(w_1 + w_2) \tag{4.9}$$

where $H(w)$ is the Discrete Fourier Transform (DFT) of $h(k)$ given by :

$$H(w) = \sum_{k=0}^{N-1} h(k) \cdot e^{-jkw} \tag{4.10}$$
$$(w = -\frac{N}{2}, \dots, \frac{N-1}{2})$$

The proof of this lemma is given below.

Proof:

Since $h(k)$ is deterministic, then it has third-order moments given by [30, 37, 38] :

$$m_3^h(\tau_1, \tau_2) \triangleq \sum_{k=-\infty}^{\infty} h(k) \cdot h(k + \tau_1) \cdot h(k + \tau_2) \tag{4.11}$$

where τ_1 and τ_2 are specified by the domain of support of $m_3^h(\tau_1, \tau_2)$. Then, the bispectrum is simply the 2-d DFT of $m_3^h(\tau_1, \tau_2)$:

$$H(w_1, w_2) \triangleq \sum_{\tau_1 = -\infty}^{\infty} \sum_{\tau_2 = -\infty}^{\infty} m_3^h(\tau_1, \tau_2) \cdot e^{-j(w_1 \tau_1 + w_2 \tau_2)} \tag{4.12}$$

Eq. 4.12 is rewritten using Eq. 4.11 as :

$$
\begin{aligned}
H(w_1, w_2) &= \sum_{\tau_1 = -\infty}^{\infty} \sum_{\tau_2 = -\infty}^{\infty} \left(\sum_{k = -\infty}^{\infty} h(k) \cdot h(k + \tau_1) \cdot h(k + \tau_2) \right) \cdot e^{-j(w_1 \tau_1 + w_2 \tau_2)} \\
&= \sum_{k = -\infty}^{\infty} \sum_{\tau_2 = -\infty}^{\infty} h(k) \cdot h(k + \tau_2) \cdot e^{-j w_2 \tau_2} \left(\sum_{\tau_1 = -\infty}^{\infty} h(k + \tau_1) \cdot e^{-j w_1 \tau_1} \right) \\
&= \sum_{k = -\infty}^{\infty} \sum_{\tau_2 = -\infty}^{\infty} h(k) \cdot h(k + \tau_2) \cdot e^{-j w_2 \tau_2} \cdot H(w_1) \cdot e^{j w_1 k} \\
&= \sum_{k = -\infty}^{\infty} h(k) \cdot H(w_1) \cdot e^{j w_1 k} \cdot \left(\sum_{\tau_2 = -\infty}^{\infty} h(k + \tau_2) \cdot e^{-j(w_2 \tau_2)} \right) \\
&= \sum_{k = -\infty}^{\infty} h(k) \cdot H(w_1) \cdot e^{j w_1 k} \cdot H(w_2) \cdot e^{j w_2 k} \\
&= H(w_1) \cdot H(w_2) \cdot \left(\sum_{k = -\infty}^{\infty} h(k) \cdot e^{jk(w_1 + w_2)} \right)
\end{aligned}
\tag{4.13}
$$

since $h(k)$ is assumed real, then :

$$
\begin{aligned}
H(w_1, w_2) &= H(w_1) \cdot H(w_2) \cdot \left(\sum_{k = -\infty}^{\infty} h(k) \cdot e^{-jk(w_1 + w_2)} \right)^* \\
&= H(w_1) \cdot H(w_2) \cdot H^*(w_1 + w_2)
\end{aligned}
\tag{4.14}
$$

where * denotes complex conjuguation. □

Using Lemma 4.2.1, Eq. 4.3 is written as :

$$B_3^y(w_1, w_2) = B_3^x(w_1, w_2) \cdot H(w_1, w_2) \tag{4.15}$$

The complex cepstrum of Eq. 4.15 can be obtained, by considering $B_3^x(w_1, w_2)$ and $B_3^y(w_1, w_2)$ as "2-d deterministic signals, and using the technique developed by Dudgeon [45, 50]. However, this will necessarily call for the use of a phase unwrapping algorithm to remove the effects of the ambiguity associated with the use of the complex logarithm [13, 45, 50], under the conditions of Dudgeon's Lemma (see Chapter 3 for datails). It is shown in Chapter 3 that the bispectra $B_3^x(w_1, w_2)$, taken as the 2-d FT of the third-order cumulants [30], have a well-defined complex cepstrum since they always satisfy Dudgeon's Lemma. However, the elegant method of Pan and Nikias [30, 43] is used instead, since it cleverly uses the notion of differential cepstrum developed by Polydoros and Fam [13, 44] to bypass the computational difficulties associated with Dudgeon's technique. Applying Pan-Nikias method on Eq. 4.15 yields :

$$b_h(m, n) = b_y(m, n) - b_x(m, n) \tag{4.16}$$

where $b_h(m,n)$, $b_y(m,n)$ and $b_x(m,n)$ are the bicepstra of $h_d(k)$, $y(k)$ and $x(k)$, respectively. Hence, the knowledge of $b_y(m,n)$ and $b_x(m,n)$ directly yield the bicepstrum of the defect impulse response without any mathematical difficulty such as existence issue [50, 45] or ill-posedness. It is interesting to note that Eq. 4.16 can be efficiently used in the identification of linear systems subject to <u>multiplicative noise</u> or a certain class of nonlinear systems [30]. From the results of Chapter 3, it is clear that $b_h(m,n)$ leads to the terms $A^{(m)}$ and $B^{(m)}$ through Eq. 3.46 in the linear case [51]. $A^{(m)}$ and $B^{(m)}$ are used used to estimate the defect impulse response $h_d(k)$ using the recursive Oppenheim-Schafer method [13] which is reproduced below :

$$i_h(0) = 1 \quad and \quad o_h(0) = 1$$

$$i_h(k) = -\frac{1}{k} \sum_{m=2}^{k+1} A^{(m-1)} \, i_h(k-m+1) \quad k \geq 1$$

$$o_h(k) = \frac{1}{k} \sum_{m=k+1}^{0} B^{(1-m)} \, o_h(k-m+1) \quad k \leq -1 \tag{4.17}$$

where $i_h(k)$ and $o_h(k)$ are the minimum and maximum phase parts of $h_d(k)$, respectively [30, 43]. $h_d(k)$ can be considered as a cascade of two subsystems with IR's, $i_h(k)$ and $o_h(k)$ such as :

$$h_d(k) = i_h(k) * o_h(k) \tag{4.18}$$

Comments on the implementation of the bicepstrum method

The estimation of $b_y(m,n)$ and $b_x(m,n)$ aims to obtain the cepstral parameters $A^{(m)}$ and $B^{(m)}$. This step can be done using two different approaches that are :

- Linear system approach using Eq. 3.53 (LS).

- Fourier Transform approach using Eq. 3.55 (FFT).

The first approach is highly affected by the numerical unstability usually associated with the LS solution, however it allows to independently estimate $i_h(k)$ and $o_h(k)$ by using different lengths for the truncation points p and q parameters in Eq. 3.51. The FFT method is unaffected by the the numerical unstability problem. However, this is achieved at the expense of a limitation in the choice of the parameters p and q unlike the LS method. The minimum and maximum phase of $h_d(k)$ are treated equally. This certainly affects the quality of the estimated defect IR with a dominant part (maximum or minimum phase part). A thorough comparison of both methods can be found in Chapter 3.

<u>Summary of the Bicepstrum Deconvolution Method</u>

The different steps of the implementation of the direct bispectrum-based deconvolution are summarized below :

1) Reference and echo signals, $x(k)$ and $y(k)$ are available.

2) Check the nullity of the bispectra of both $x(k)$ and $y(k)$ using the available

   Gaussianity/linearity tests [52, 53].

3) Estimate the bicepstra $b_y(m,n)$ and $b_x(m,n)$ using :

   $\Longrightarrow$ LS method using Eq. 3.53.

   or

   $\Longrightarrow$ FFT method using Eq. 3.55.

4) Obtain the bicepstrum of the defect impulse response as :

$$b_h(m,n) = b_y(m,n) - b_x(m,n) \qquad (4.19)$$

5) Get the cepstral parameters $A^{(m)}$ and $B^{(m)}$ from $b_h(m,n)$ using Eq. 3.46.

6) Use Oppenheim-Schafer method to recover $h_d(k)$ from $A^{(m)}$ and $B^{(m)}$ using

   Eqs. 4.17, 4.18.

4.3   Automatic Defect Classification System

Once an accurate and reliable defect impulse response is obtained from the deconvolution step using any of the techniques described above, it is fed into an automatic defect classification system based on a sophisticated learning paradigm to faithfully

identify the type of defect present in the material under test. Automatic defect classification constitutes an important signal processing task on which relies the overall performance of the ultrasonic inspection system. It is a difficult signal processing task, particularly when the inspection system operates at moderate and low SNR levels. Existing automatic defect classification systems use features extracted from the time and frequency representations of the echo signals [54], which are usually corrupted by the material and noise effects. This approach combines both uncertain time and frequency information present in the echo signals and its performance may be severely affected by the inspection environment which can be in some practical situations very noisy, such as pipes welding inspection. Here, a novel modular learning strategy for defect classification and detection is described. The design of this technique is motivated by the *information preservation rule* [55]. The information preservation rules states that [55] :

" *The information content of a defect impulse response should be optimally preserved in a statistical sense and efficiently used in a computational sense, until the defect classifier is ready for final decision-making* ".

The classification strategy makes no assumption on the ultrasonic inspection environment. It incorporates three functional blocks :

(1) Joint Time-Frequency Analysis (JTFA).

(2) Features Extraction using Principal Components Analysis (PCA).

Input Patterns:
Defects IR's

↓

WVD
Computer

↓

Optimal Features Extraction
Using Sanger's Algorithm
with GHA Learning.

↓

Decision Making Using MLP Network with
BackPropagation Algorithm.

↓

Output Patterns:
Classification Decisions.
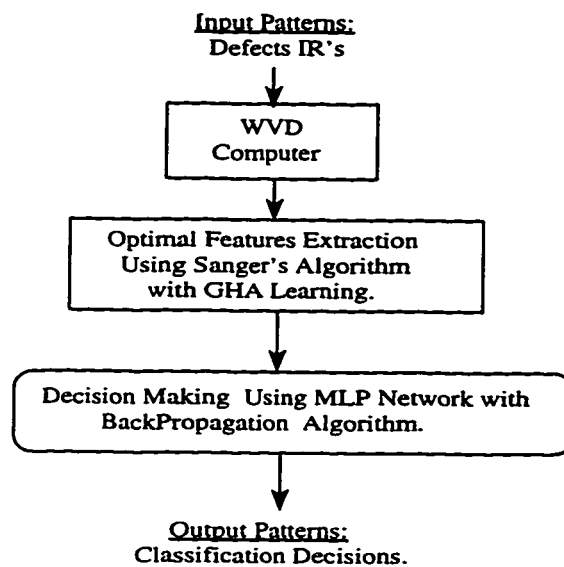
Figure 4.2: Block Diagram of Proposed Automatic Defect Classification System.

(3) Decision Making.

The joint time-frequency analysis, implemented using the Wigner-Ville Distribution (WVD) [56, 57], transforms the defect impulse response into a time-frequency "image" that accounts for any time-varying nature of the signal's spectral contents. This "image" provides the input to an unsupervised artificial neural network (ANN) that constitutes the PCA functional block used for feature extraction in an optimal fashion. The compressed features are then introduced into a supervised multilayer perceptron (ANN) performing the defect classification. The block diagram of this automatic and modular defect classification strategy is shown in Fig. 4.2. In what follows, details on each of the blocks shown in Fig. 4.2, are presented.

4.3.1 Joint Time-Frequency Analysis

Nonstationary signals have time-varying spectral properties, requiring the use of some form of joit time-frequency analysis. JTFA is a well-developed technique [56, 57]. In particular, Cohen's class [57] of time-frequency distributions (TFD), has been applied to a variety of signal processing problems [56]. This technique has two objectives [55, 58] :

(1) Bring out the nonstationary behavior of the defect IR in a visible fashion.

(2) Allow the separation of multiple components contained in the defect impulse response.

The aim of JTFA is to approximate the time-frequency energy distribution function $E_x(t, f)$, which is defined as the energy contained in a signal $x(t)$ within an infinitesimally small neighborhood around time t and frequency f [56, 57]. It is question of "approximate" function because of the disjoint nature of time-frequency concentration. That is, a signal $x(t)$ cannot be concentrated in both time and frequency simultaneously, by virtue of the *uncertainty principle* [56, 57, 58]. As mentioned earlier, in the proposed paradigm, JTFA is performed using WVD for the reasons that are given next.

Wigner-Ville Distribution

Cohen [57] has developed a general framework for time-frequency distributions. TFD's are usually classified into linear and linear methods. The latter category includes the subclass of *bilinear* TFD's (BTFD's). A particular BTFD that has gained interest in a broad range of science and engineering [58] is WVD. WVD of a signal $x(t)$ is given by [56, 58] :

$$W_x(t, f) = \int_{-\infty}^{\infty} x\left(t + \frac{\tau}{2}\right) x^*\left(t - \frac{\tau}{2}\right) e^{-j2\pi f\tau} d\tau \qquad (4.20)$$

where the lag variable τ plays the role of a dummy variable, $*$ denotes the complex conjuguate. It should be noted that any other TFD derived by smoothing in the time-frequency plane is singular [57], in which case there is loss of information due to the smoothing operation [55, 56, 57, 58]. This highlights the fact of the "*optimal information-preserving*" of the standard WVD. Three properties make the standard WVD highly desirable from a signal processing viewpoint [55] :

(1) $W_x(t, f)$ is real for any complex-valued signal $x(t)$.

(2) It exhibits the least amount of spread in the time-frequency plane.

(3) Optimal information-preserving, i.e., lossless transform.

For the reasons given above, WVD has been selected to be the tool for performing JTFA which is the first step in the proposed defect classification system. It should

be noted that the presence of *cross-terms* that are usually considered as a severe

limitation of the standard WVD, is positively used in the classification scheme,

since they bear information about the nature of defects present in the material

being tested [1]. Hence, cross-terms play a major role in the proposed system, and

based on their presence and nature, an accurate and reliable decison is made at the

level of the third stage of the classification system (see Fig. 4.2). Boashash et al. [58]

developed various algorithms for the digital computation of the standard WVD. Two

interesting algorithms are based of their counterpart for spectrum estimation [17],

namely :

(a) Low Resolution Classical Fourier-based WVD Estimation.

(b) High Resolution Parametric WVD Estimation.

The first algorithm is numerically efficient and is based on the use of FFT, but it

suffers from its low resolution [58]. However, the second algorithm uses a parametric

model (AR or ARMA) to produce a high resolution standard WVD at the expense

of an increased computational complexity [58]. In the first stage of the proposed sys-

tem, the standard WVD is estimated using the first algorithm for its computational

simplicity. The discrete version of the standard WVD is given by [58] :

$$W_x(n,k) = \sum_{m=-(N-1)/2}^{(N-1)/2} x(l+m)x^*(l-m)e^{-j4\pi mk/N}$$

---

[1]It is a common practice to use the analytic form of $x(t)$ by using its Hilbert transform in Eq. 4.20 to reduce the cross-terms [56].

$$0 \le n \le N \quad \text{and} \quad 0 \le k \le N \qquad (4.21)$$

where N is the length of the discrete version, $x(k)$, of the signal $x(t)$. It is that N is even for implementation reasons [58]. n, k and and m are the discrete variables corresponding to the continuous variables t, f and τ, respectively. It should be note that $W_x(n,k)$ represents an image of size N-by-N, where N represents the number of frequency samples in the discrete standard WVD for each time sample n.

4.3.2   Features Extraction

It is clear from the first processing step presented above that the use of WVD drastically increases the amount of data initially available, i.e., the deconvolved defect IR. This increase is actually due to to the combination of the information present in the time and frequency representations of the defect IR at hand. At the output of the first stage, the data is N times the initial one where N is the length of the defect IR. However, the obtained data is highly redundant [55]. A solution to this problem may be provided by a compression scheme such the Discrete Cosine Transform (DCT) or the Karhunen-Loeve Transform (KLT) [59]. The former method has been successfully used in still image compression, however it is used in a lossy fashion. The lossy nature of DCT makes it inappropriate for the proposed defect classification strategy. KLT is an optimal transform [59], variously known as PCA, Eigenvector Decomposition (EVD). PCA has been largely applied in chemometrics

and statistics [60]. However, KLT in its general form requires the knowledge of the exact autocovariance matrix (ACM) of the signal to be compressed [60]. This constitues a major limitation in its application in real world problems. However, PCA is a good candidate for optimal features extraction because it satisfies the *information-preserving rule* stated earlier.

Review of Principal Components Analysis :

An excellent review of PCA with its application to image compression can be found in [60].

Let X represent a vector that varies over the set of observations vectors and denote the cordinates of X by $x_1, x_2, \ldots, x_N$. The set of observations is denoted by $X_1, X_2, \ldots, X_N$. The sample mean, M, of the observations is given by [60] :

$$M_X = \frac{1}{N} (X_1 + X_2 + \cdots + X_N) \qquad (4.22)$$

The *mean-deviation* set of observations is given by [60] :

$$B = [X_1 - M_X, X_2 - M_X, \ldots X_N - M_X] \qquad (4.23)$$

From Eq. 4.23, the *sample covariance matrix* is given by [60] :

$$S = \frac{1}{N-1} BB^T \qquad (4.24)$$

where T denotes the matrix transposition. The goal of PCA is to find an orthogonal $p \times p$ matrix $P = [v_1, v_2, \ldots, v_p]$ that determines a change of variables, $B = PU$ such that :

$$
\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{bmatrix} = [v_1, v_2, \ldots, v_p] \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_p \end{bmatrix} \tag{4.25}
$$

with the property that the new variables $u_1, u_2, \ldots, u_p$ are _uncorrelated_ and are arranged in order of decreasing variance. The unit eigenvectors $v_1, v_2, \ldots, v_p$ of the ACM matrix S are called the principal components of the data (in the matrix of observations B). Eq. 4.25 shows that the new data U is uncorrelated and retains all the statistical information contained in the original data B. PCA cannot be "*optimal*" if the ACM matrix is not available or known a priori.

Principal Components Analysis Using Neural Networks :

Recently, attempts have been made to implement the PCA technique using ANN that _learn adaptively_ the ACM from the data directly. Oja [61] was the first to develop an unsupervised ANN that is capable of performing PCA without resorting to the computation of the ACM matrix. However, it has been shown [62] that Oja's paradigm can estimate only the first principal component v_1 of Eq. 4.25. Later,

Sanger [62] employed the generalized Hebbian algorithm (GHA) to estimate the p principal components of Eq. 4.25. The algorithm operates in a *recursive fashion* to implement Eq. 4.25. Hence, Sanger's algorithm is *capable of estimating the exact ACM*, making it a linear method for *dimensionality reduction* such as signal/image compression. Diamantaras and Kung [63] improved Sanger's algorithm such that the latter can estimate the (p+1)th principal component from the pth one recursively. The task of features extraction that constitues the second processing step of the proposed automatic defect classification is implemented using Sanger's algorithm for its implementation simplicity compared with that proposed by Diamantaras and Kung.

Sanger's Algorithm :

Fig. 4.3 shows the unsupervised ANN structure for the implementation of Sanger's algorithm. The goal of Sanger's algorithm is to produce an ANN that performs PCA of arbitrary size on the input vector [62]. Let the input vector be in *mean-deviation* form and denoted by :

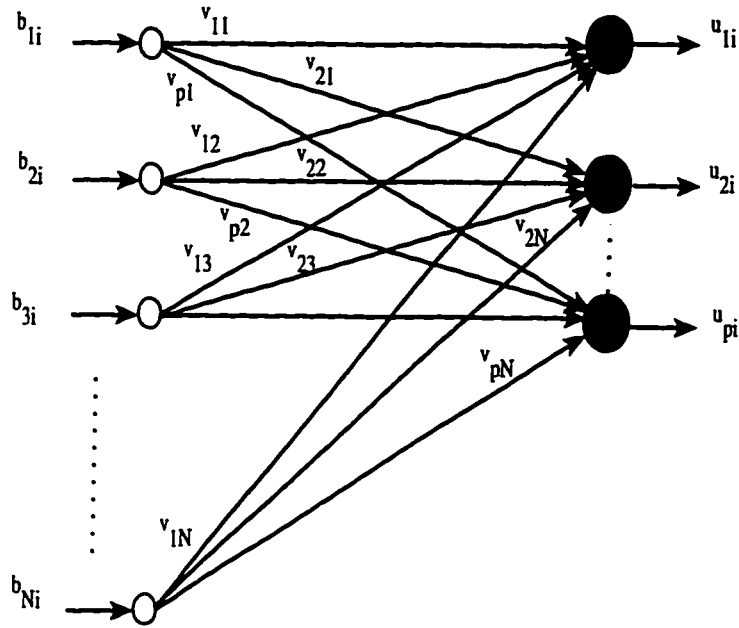$$B = [B(1), B(2), \ldots, B(L)] \qquad (4.26)$$

Figure 4.3: Implementation Structure for Sanger's Algorithm.

where N in the input vector dimension and L is the number of the available input vectors. Each input vectot $B(i)$ is given by :

$$B(i) = \begin{bmatrix} b_{1i} \\ b_{2i} \\ \vdots \\ b_{Ni} \end{bmatrix} \qquad (4.27)$$

The network input/output relation is given by :

$$U = VB \qquad (4.28)$$

Two assumptions about the feedforward network of Fig. 4.3, are made [64] :

(1) Each neuron in the output layer of the network is linear.

(2) The network has N inputs and p outputs, both of which are specified as the the input and output dimensions, respectively. Moreover, the network has fewer outputs than inputs (i.e., $p < N$).

$\{V\}$ constitute the synaptic weights of the network tha are subject to training [62]. $\{v_{ji}\}$ connects the source node i in the input layer to computation nodes j in the output layer, where $i = 1, 2, \ldots, N$ and $j = 1, 2, \ldots, p$. The output u_j of the neuron j, produced in response to the set of inputs $\{b_i \mid i = 1, 2, \ldots, N\}$, is given by (see Fig. 4.4) :

$$u_j = \sum_{i=1}^{p} v_{ji} b_j \qquad j = 1, 2, \ldots, p \tag{4.29}$$

The synaptic weight v_{ji} is adapted in accordance to the GHA technique as given by [62] :

$$\Delta v_{ji} = \eta \left[u_j b_i - u_j \sum_{k=1}^{j} v_{ki} u_k \right], \quad \begin{array}{l} i = 1, 2, \ldots, N \\ \\ j = 1, 2, \ldots, p \end{array} \tag{4.30}$$

where Δv_{ji} is the change applied to the synaptic weight v_{ji}, and η is the learning-rate parameter.

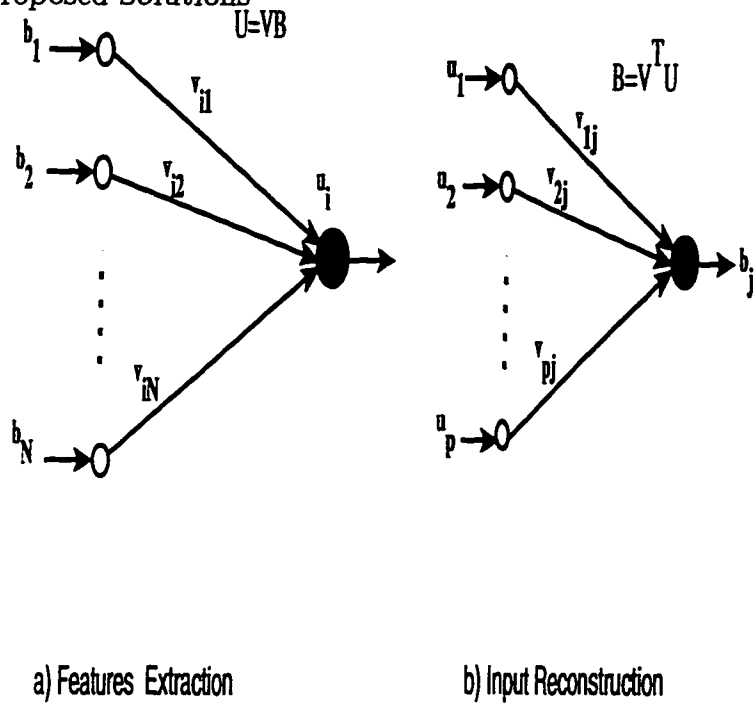a) Features Extraction   b) Input Reconstruction

Figure 4.4: Signal-Flow Diagram Representation of GHA Algorithm.

Algorithm Summary:

1. Initialize the synaptic weights of the network, v_{ji}, to small random values at time index $n = 1$. Assign a small positive value to the learning-rate parameter η.

2. For $n = 1$, $j = 1, 2, \ldots, p$, and $i = 1, 2, \ldots, N$, compute

$$u_j(n) = \sum_{i=1}^{p} v_{ji}(n)b_j(n)$$

$$\Delta v_{ji}(n) = \eta \left[u_j(n)b_i(n) - u_j(n) \sum_{k=1}^{j} v_{ki}(n)u_k(n) \right]$$

where $b_i(n)$ is the *ith* element of the $N \times 1$ input vector $B(n)$ and p is the desired number of principal components.

3. Increment n by 1, go to step 2, and continue until $n = L$ where L is the number of *input trainining vectors*.

4.3.3 Decision Making

The pattern classification is the last processing step in the proposed defect classification system, which is required to discriminate between different defects that may be present in the material under test. Features extracted from WVD "images" using Sanger's algorithm are processed by an ANN classifier for final decision-making about defect present in the inspected material. The classifier uses a supervised learning [64] based on a multilayer perceptron (MLP). The design of the latter is motivated by the fact that it is able to construct arbitrary decision boundaries between variours vectors in the defect feautures space [26]. The MLP consists of three different layers :

1. Input layer of dimension equal to that of the output of the second block in the classification scheme (see Fig. 4.2).

2. Hidden layer of variable dimension.

3. Output layer of dimension equal to $n = \log_2(D)$ where D represents the number of possible defects to be classified.

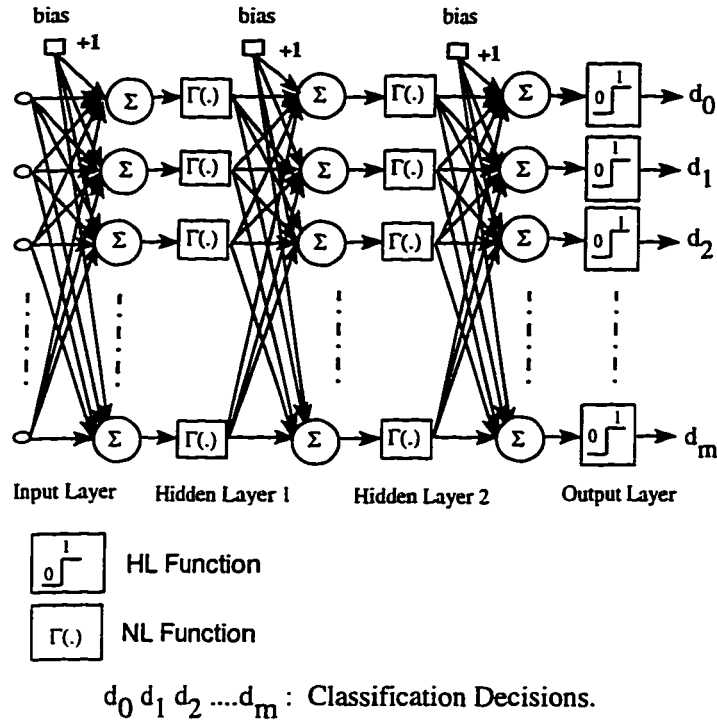$d_0\ d_1\ d_2\ ....d_m$: Classification Decisions.

Figure 4.5: MLP Classifier Network Architecture.

It should be noted that the output of the MLP is of binary type, where each binary combination indicates one defect type. To force the output of the MLP network to binary, a competitive learning rule is used [64]. In competitive learning, the output neuron with the largest analog value is set to one and all the other outputs are set to zero. Fig. 4.5 gives the architecture of the MLP classifier network.

4.3.4 Training Procedures

The proposed defect classification system, using a modular learning paradigm, incorporates two distinct learning strategies:

1) Unsupervised learning on WVD "images" using Sanger's algorithm.

2) Supervised learning on "*optimal*" features extracted from PCA network.

The overall training of the classification systems incorporates :

a) *A Preprocessing* step where the HOS-based deconvolution technique, described previously in this chapter, is used to estimate $h_d(k)$ from $x(k)$ and $y(k)$. The deconvolved IR is the input database for the next processing step.

b) WVD computation for each defect IR $h_d(k)$ using the discrete standard WVD formula given by Eq. 4.21. The obtained "images" constitute a new 2D database.

c) Training of the PCA network, using Sanger's algorithm, with each WVD database obtained from the WVD computer of the previous step. Details about the training are given below.

d) Training of the MLP network to learn the different possible defects with "*optimally*" reduced features obtained from the step (c) above, using the classical backpropagation (BP) algorithm [64].

PCA Network Training

The PCA network training is based on Eq. 4.30. To be specific, for a successful use of the proposed scheme, it is assumed that :

(i) The training database consists of D WVD "images" datasets corresponding to D-1 artificial defects and the flawless case.

(ii) The datasets are labeled $D^{(1)}$ for the flawless case and $D^{(i)}$ for the artificial defect i.

(iii) Each dataset contains A WVD "images".

The procedure for training the PCA network is given below :

PCA Network Training Algorithm:

1. For $i = 1$, start with dataset $D^{(1)}$ of the flawless case.

2. Initialize the synaptic weights of the network, $V^{(i)}$, (See Eq. 4.30) to small random values. Assign a small positive value to the learning-rate parameter η.

3. Pick a WVD "image" from $D^{(i)}$ at random.

4. Use Sanger's algorithm.

5. Put back the learned WVD "image" in the set $D^{(i)}$.

6. Go back to step 3 if Sanger's algorithm has not converged. Otherwise go to step 7.

7. Save the converged synaptic weights $V^{(i)}$ corresponding to dataset $D^{(i)}$.

8. Increment i by 1, to train with the next dataset $D^{(i+1)}$, go to step 2, and continue until $i = D$.

Decision Making Using MLP Network With BP Algorithm:

Only few differences exist between the training paradigms used in the PCA and MLP networks that can be summarized in :

(i) An unsupervised learning technique governs the PCA network using GHA algorithm.

(ii) WVD "images" constitute the training input to the PCA network.

(iii) No target outputs are needed for the PCA network.

(iv) The MLP network is trained with a supervised scheme using the BP algorithm.

(v) The MLP network is trained using the output of the PCA network as training vectors.

(vi) The MLP network requires the presence of target/reference outputs. The reference is the binary representation of each artificial defect .

Given these differences, it should be noted that the training procedure for the MLP network is similar to that used for PCA network that is presented previously, except that there is error feedback to the input of the MLP network. The procedure for

training the MLP network is as follows :

MLP Network Training Algorithm:

1. For $i = 1$, start with dataset $D_c^{(1)}$ of the flawless case, where $D_c^{(i)}$ represent the "*optimally*" reduced features obtained from the PCA network for defect i.

2. Initialize the synaptic weights of the network, $W_1^{(i)}$, and $W_2^{(i)}$ to small random values. Assign a small positive value to the learning-rate parameter μ.

3. Pick a features vector from $D_c^{(i)}$ at random, along with the corresponding target output.

4. Use Back Propagation BP algorithm [64].

5. Put back the learned features vector in the set $D_c^{(i)}$.

6. Go back to step 3 if the BP algorithm has not converged. Otherwise go to step 7.

7. Save the converged synaptic weights $W_1^{(i)}$ and $W_2^{(i)}$ corresponding to dataset $D_c^{(i)}$.

8. Increment i by 1, to train with the next dataset $D^{(i+1)}$, go to step 2, and continue until $i = D$.

Chapter 5

Computer Simulation and

Experimental Results

The applicability of the proposed HOS-based deconvolution and defect classification techniques presented in Chapter 4, to both synthetic and experimental ultrasonic signals, is examined in this chapter. For the sake of objectivity, the performance analysis of the proposed deconvolution technique is carried out to study the effects of different implementation parameters such as :

1) The data length of the involved ultrasonic signals.

2) Estimation procedure used in the computation of the HOS quantities such as cumulants, polyspectra and polycepstra.

3) Signal-to-Noise Ratio SNR level used.

4) Statistical properties of the additive noise $[n(k)]$.

It is worth mentioning that the performance analysis is carried out exclusively using synthetic signals assuming a valid ultrasonic model. In this case, comparisons are made possible given that the knowledge about the true defect model is available. In real ultrasonic data, the definition of a valid comparison criterion is a difficult task [65].

5.1 Computer Simulation

5.1.1 Introduction

The simulated ultrasonic signals, used throughout this chapter, follow the following simple model representing real ultrasonic inspection systems.

The Model: In this model, the transducer and the propagation paths are assumed to be LTI systems. The reference signal $x(k)$ is modeled as a shifted version of the transmitted ultrasonic pulse. In this case, the reference signal $x(k)$ is written as [1] :

$$x(k) = A(k - \tau_p) \cos\left[2\pi f_0(k - \tau_p)\right] \tag{5.1}$$

where τ_p is the time delay introduced by the propagation paths, f_0 is the transducer center frequency, and $A(k)$ is the envelope of the transmitted ultrasonic pulse that is a Gaussian pulse and is given by :

$$A(k) = \exp\left[-\left(\frac{2\pi f_0 k}{\sigma}\right)^2\right] \qquad (5.2)$$

where $\sigma = 1.5\ \pi$. This simulation model presents certain characteristics that can be summarized in :

1) It is convenient for use in nonparametric deconvolution schemes.

2) $x(k)$ is the output of an analog modulation system, i.e., amplitude modulation (AM).

3) Modulated signals are usually characterized by pronounced resonance modes [30].

4) Parametric modeling of pronounced resonance and antiresonance models yields poles at the vicinity of the unit circle [30].

A simulated reference signal following this model is illustrated in Fig. 5.1. A parametric model for this signal is obtained using the prediction error method (PEM) [4, 5]. The pole-zero map of the obtained parametric ARMA model is shown Fig. 5.2. The closeness of the poles to the unit circle has the undesirable effect of making the cepstral parameters $A^{(m)}$ and $B^{(m)}$ of longer duration. This leads to the use of larger values of p and q in Eq. 3.51 (See Chapter 3 for details),
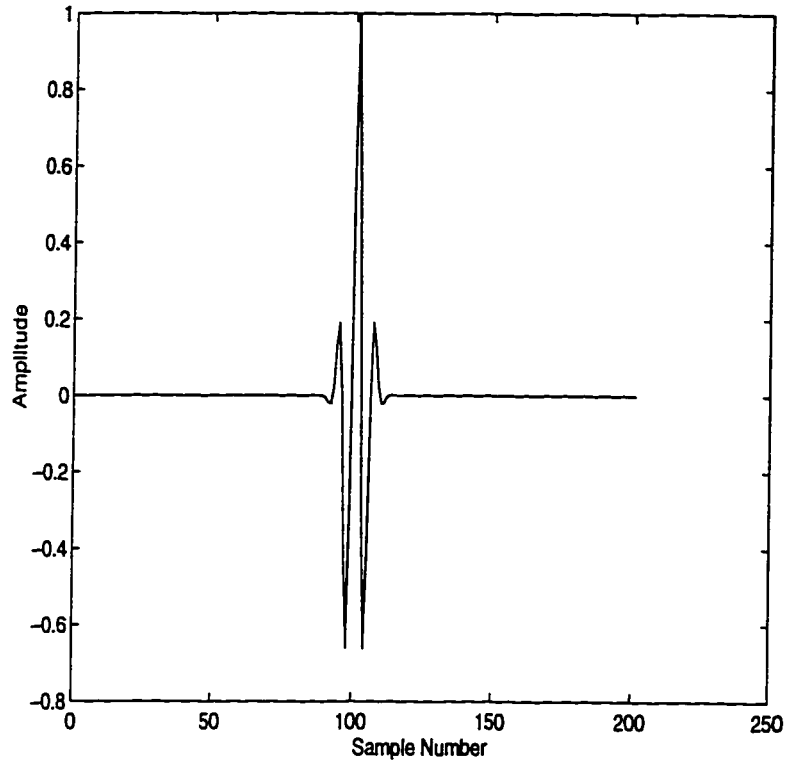
Figure 5.1: Synthetic Ultrasonic Reference Signal x(k).

requiring a larger system of linear equations in Eq. 3.53, or the use of larger FFT

lengths in the FFT-based solution of Eq. 3.55. The cepstral parameters $A^{(m)}$ and

$B^{(m)}$ corresponding to the ARMA model of the reference signal $x(k)$ are shown in

Fig. 5.3. For a fast decay of $A^{(m)}$ and $B^{(m)}$, exponential windowing may be used

in the HOS domain to keep the poles far away from the vicinity of the unit circle.

Such a solution preserves the stationary character of the involoved signals unlike 1-d

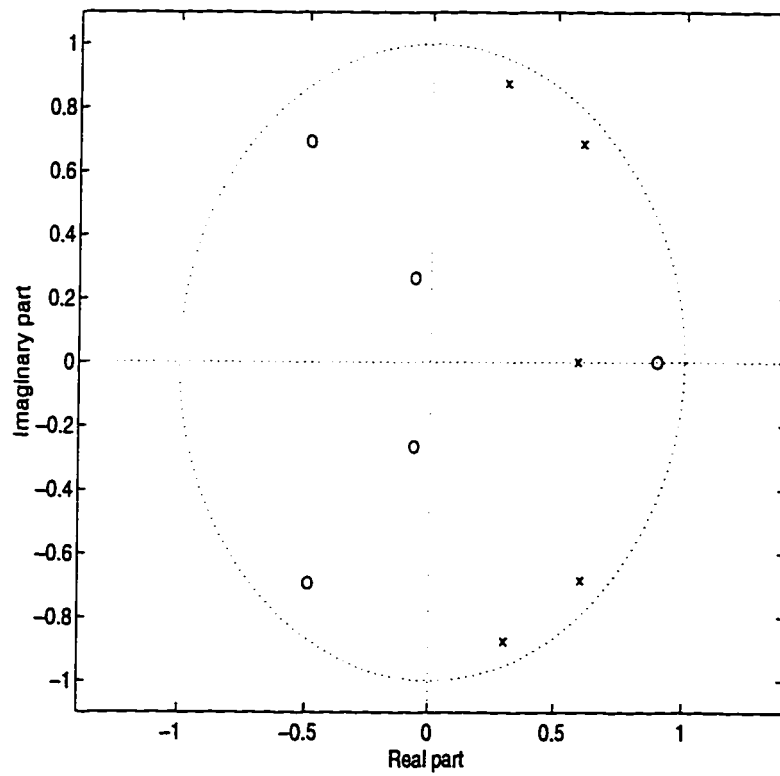exponential windowing on the signals themselves.

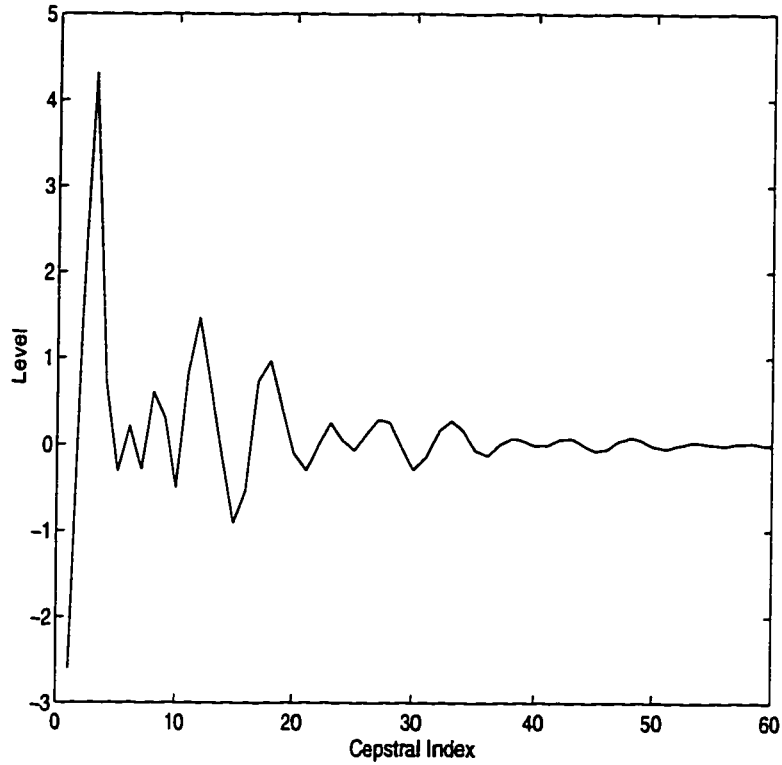Figure 5.2: Pole-Zero Map of (5,5) ARMA Model of Reference Signal x(k).

Figure 5.3: Cepstral Parameters $A^{(m)}$ and $B^{(m)}$ of Reference Signal x(k).

5.1.2 Artificial Defect Models

The adopted ultrasonic model follows the linear formulation of Eq. 1.1, i.e., :

$$y(k) = x(k) * h_d(k) + n(k)$$

For a qualitative analysis of the proposed HOS-based deconvolution method, different models for $h_d(k)$ are proposed. These models are selected based on :

⟹ The model must quantify the behavior of the real defects existing in material.

⟹ The model must represent different possible experimental situations.

⟹ The model can be either <u>parametric</u> expressed in terms of parameters coefficients or <u>nonparametric</u> formulated in the time or frequency domain (i.e., IR or TF).

⟹ The real defects IR's can be either minimum or mixed phase systems, so must be the selected models.

Bearing in mind these facts, five (05) models are selected to represent different defects situations, namely;

1) Minimum phase (MP) moving average (MA) system, i.e., nonparametric model to compare results obtained using SOS-based deconvolution methods. This model is called System I.

2) Nonminimum phase (NMP) MA system, to illustrate the contribution of the proposed HOS-based schemes to preserve the true phase character of the defect IR. It will be refered to this model as System II.

3) MP AR model, i.e., a parametric model which better models the oscillatory character of the defects IR's. The third model is called System III.

4) MP ARMA model that represents the general case in the deconvolution step. Minimum phase is selected to be able to compare with results obtained using SOS-based algorithms such as PEM. This model is named System IV.
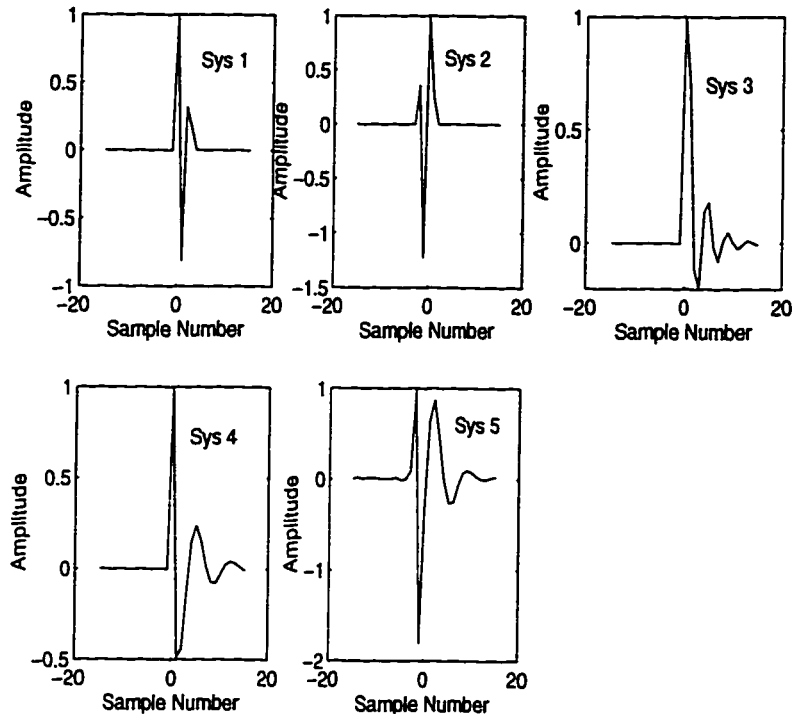
Figure 5.4: IR's of the Different Selected Models.

5) NMP ARMA model to be in agreement with experimental results using real signals that are known to be of NMP character. In this case, it is shown later that the proposed methods are able to discriminate between spectrally equivalent ARMA systems (SE). It will be refered to this model by System V in the remaining of this chapter.

The selected models are illustrated, by their IR's, in Fig. 5.4. Their corresponding TF's are shown in Fig. 5.5. Fig. 5.6 illustrates the cepstral parameters $A^{(m)}$ and $B^{(m)}$ associated with all the selected models. Only the first 20 parameters are shown (i.e., $m = 1, 2, \ldots, 20$). Note from Fig. 5.6 that the $B^{(m)}$'s parameters associated with the MP systems are vanishing since they convey information about any maximum
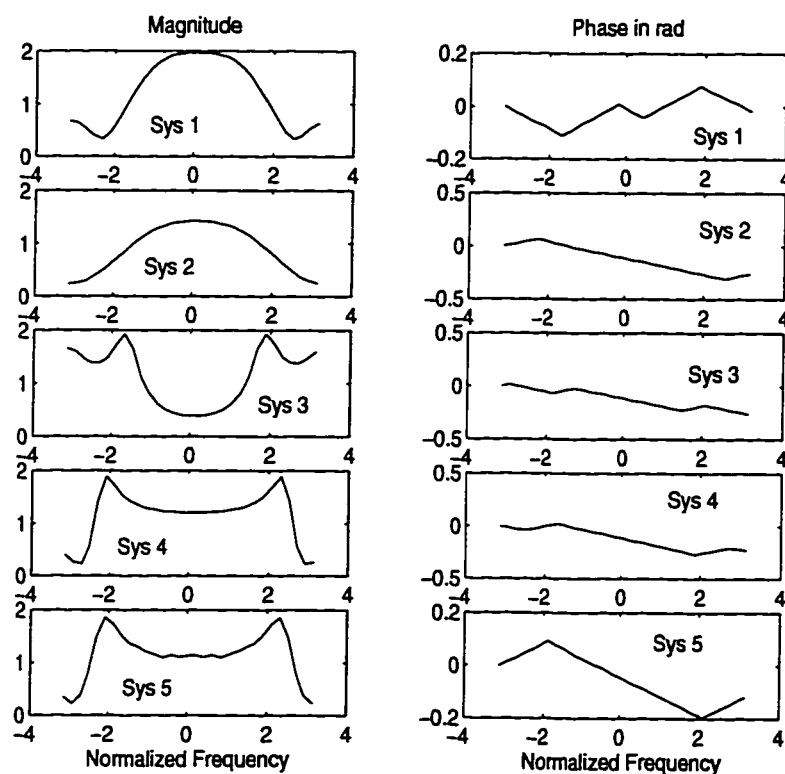
Figure 5.5: TF's of the Different Selected Models (Magnitude and Phase).
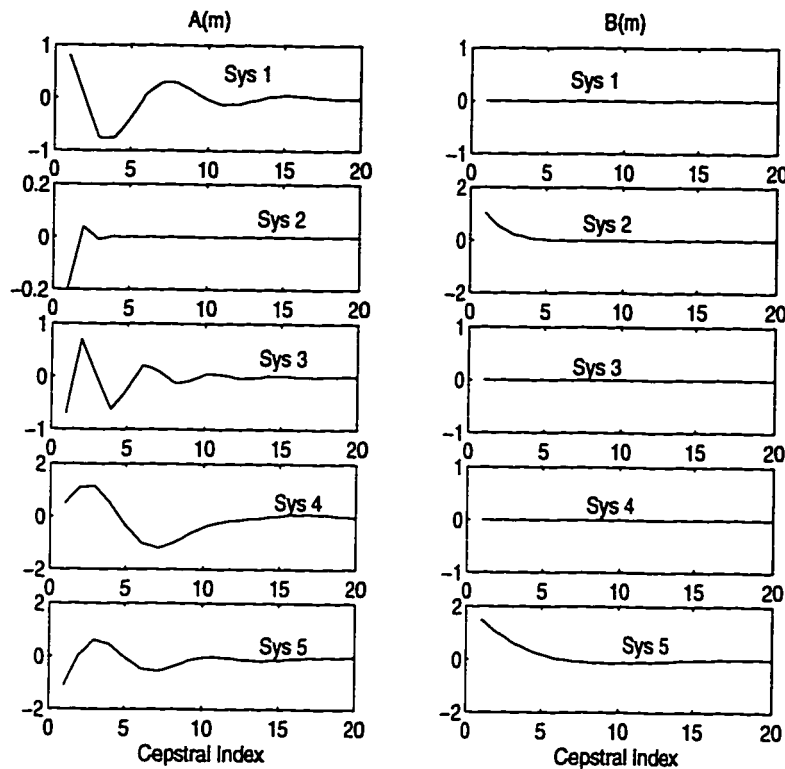
Figure 5.6: Cepstral Parameters $A^{(m)}$, $B^{(m)}$ of the Different Selected Models.

phase term existing in the model.

5.1.3 The Deconvolution Method

To illustrate the efficiency of the HOS-based method, the deconvolution of System II is investigated. It should be noted that System I can be deconvolved using SOS-based methods since it is a MP system.

Deconvolution of System II

System II represents a nonparametric model that is characterized by its IR or TF only. The deconvolution of this model is performed at different SNR levels where
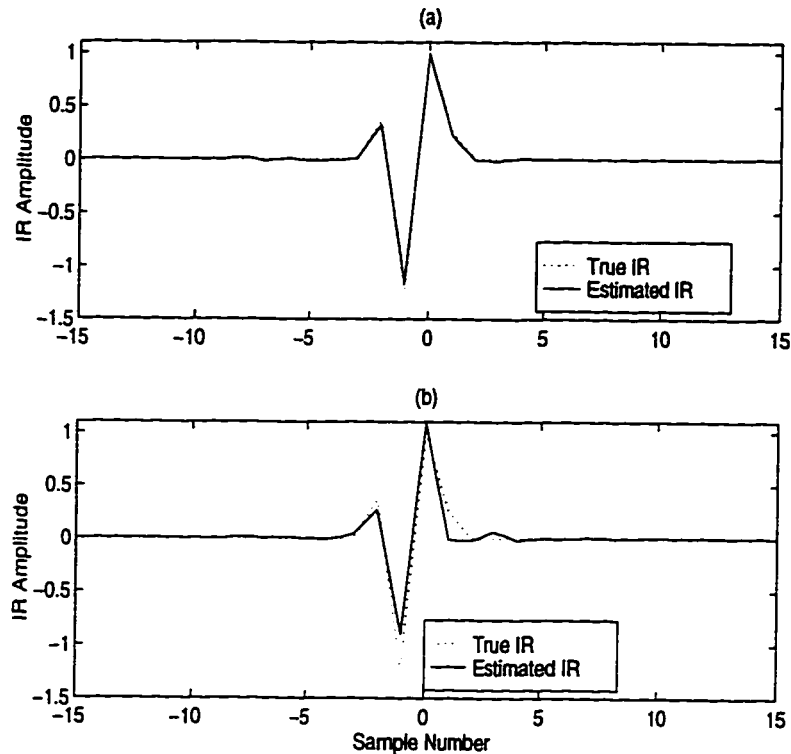
Figure 5.7: True and Estimated IR's of System II Using the Bicepstrum Method at 30 dB (a) and 0 dB(b).

the additive noise is white Gaussian. Fig. 5.7a shows the true and estimated IR's at

SNR = 40 dB. The case of SNR = 10 dB is illustrated in Fig. 5.7b. The truncation

points for the cepstral parameters are both set to 7 (see Eq. 3.53 for details). The

selection of the truncation point is based on Fig. 5.3 which gives information about

the extent of the cepstral parameters $A^{(m)}$ and $B^{(m)}$ in the quefrency domain [13].

However, it should be noted that the selection of p and q follows an empirical rule

suggested by Pan and Nikias [43]. Some remarks regarding the performance of the

bicepstrum method in the deconvolution of System II are made :

⟹ Due to its <u>nonparametric formulation</u>, the bicepstrum method does not require any a priori knowledge about the order of the underlying model.

⟹ The empirical rule for the selection of the truncation points p and q can be based on a knowledge about the positions of the system poles and zeros in the Z-plane. This can drastically reduce the dimension of the system of Eq. 3.53.

⟹ The scale ambiguity can be removed by using Hatzinakos-Nikias gain recovery method [42].

Performance in Noise Analysis

As stated previously, the bicepstrum method is characterized by its robustness to the addditive noise. This proposed deconvolution scheme produces accurate estimates of the true IR at extremely low SNR levels. Also, the estimated IR has excellent resolution of the peaks and smoothness. Additive white Gaussian noise is generated corresponding to different SNR levels ranging from 40 dB to -5 dB. The performance measure is given by the norm of the error defined as :

$$E = \sum_{i=1}^{L} \left[h_d(i) - \hat{h}(i) \right]^2 \qquad (5.3)$$

where $\hat{h}(k)$ represents the estimated defect IR using the deconvolution method. The performance of the bicepstrum method at different SNR levels is shown in Fig. 5.8. The error norm is normalized with respect to that of the Wiener filter method.

For comparison, the performance of Wiener filter A method is also included. From Fig. 5.8, it is clearly seen that the proposed bicepstrum method outperforms all its SOS-based counterpart schemes. For the sake of analysis completeness, it should be kept in mind that :

- Wiener filter method (both A and B schemes) has satisfactory performance at relatively high SNR levels.

- At moderate and low SNR levels, Wiener filter method performance is severely affected by the noise. This is also valid for Neal's optimal Wiener filter. This calls into question the practical usefulness of the optimality of Neal's Wiener filter.

- The IR estimates obtained using the bicepstrum method are usually high resolution and accurate.

- Bicepstrum method is robust to the additive Gaussian noise, and produces accurate estimates at very low SNR levels. However, this achievement is made possible only at the expense of a drastic increase in the computational complexity.

Analysis of the Noise Statistics Effects

The analysis of the noise statistics effects on the performance of the bicepstrum method is carried out using the following noise statistics :
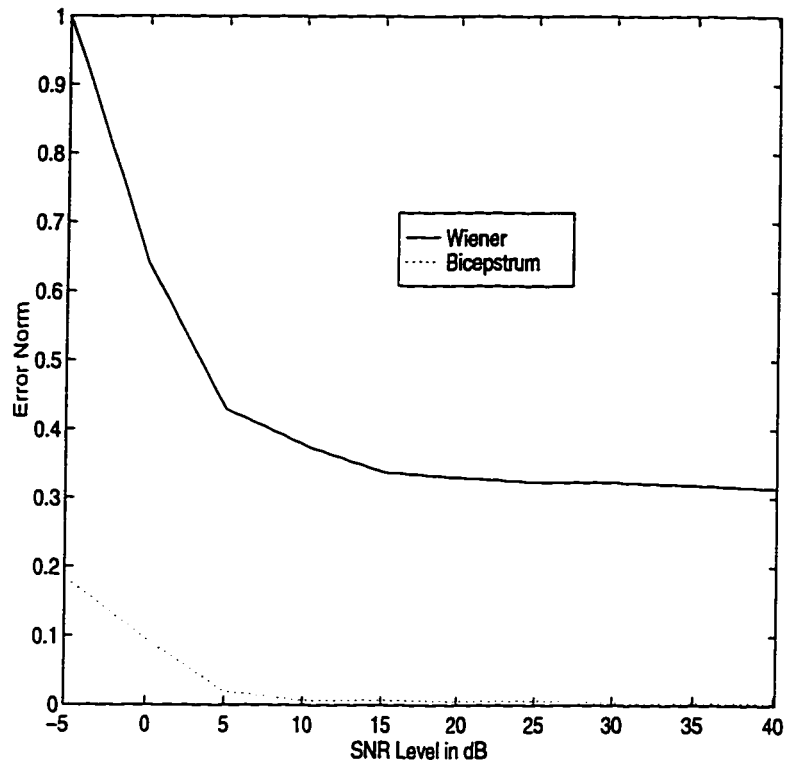
Figure 5.8: Bicepstrum and Wiener Filter A Methods versus SNR.

1. White Gaussian noise.

2. Correlated Gaussian noise using an AR correlating filter.

3. White Uniform noise.

4. White Laplacian noise.

5. White two-sided exponential noise.

It is worth mentioning that all the noise types listed above belong to the symmetrically-distributed noise class, which is known to have zero third-order statistics (cumulants and bispectra) [30] if the whiteness assumption is satisfied [47]. The colored Gaussian noise is obtained by passing a white Gaussian noise through a correlating AR
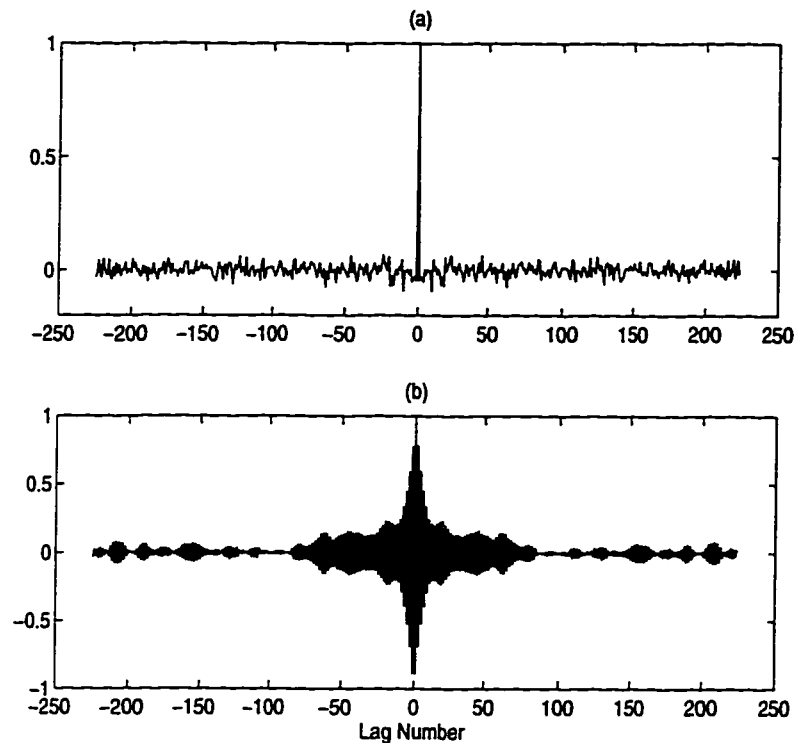
Figure 5.9: ACF of : (a) typical white Gaussian noise. (b) Correlated gaussian Noise (Factor = 0.9).

filter that has the following transfer function :

$$H_{cor}(z) = \frac{1}{1 + kz^{-1}} \tag{5.4}$$

where the factor k is a parameter that controls the correlation of the colored Gaussian noise at the AR filter output. The ACF of a white Gaussian noise is shown in Fig. 5.9(a). The ACF of the output of the correlating AR filter for $k = 0.9$ is illustrated in Fig. 5.9(b). The performance of the bicepstrum method in the presence of the different noise distributions previously introduced is shown in Fig. 5.10 in the case of System II. From this figure, it is seen that the bicepstrum method has a
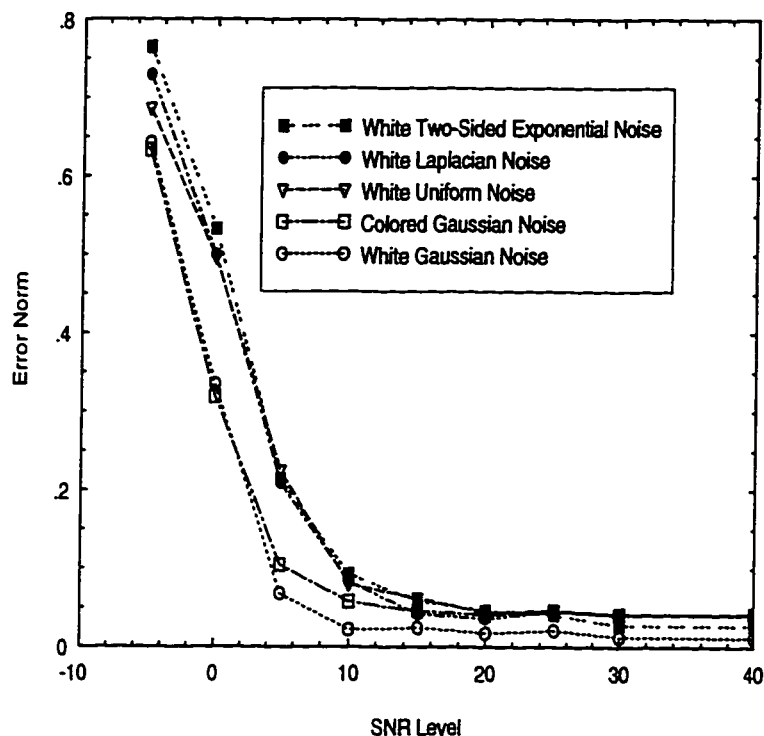
Figure 5.10: Bicepstrum Method Performance in Different Noise Statistics (System II).

similar performance in the different cases, with a slightly better performance in the case of the white Gaussian noise. From a theoretical viewpoint, HOS are assumed to be insensitive to all these noise distributions, the slight variations in performance in our simulation are possibly due to :

- Gaussian and uniform noises are generated using professional routines in the $MATLAB©$ environment. The other distributions are generated based on personal routines using basic random variables transformations [31]. These routines are not tested in a statistical sense to verify their exactitude.

Analysis of the Truncation Points Effects

The performance of the bicepstrum method in the deconvolution of the five differ-

ent selected systems is shown in Fig. 5.11. The additive noise is white Gaussian. It

clearly appears that the bicepstrum method performs better in the case of the de-

convolution of System II. This is due to the fact that the selected truncation points

in this case; $p = 7$ and $q = 7$; are more suitable for the case of System II. This fact

highlights the importance of the choice of the truncation points and hence suggests

the study of the effects of this choice on the performance of the bicepstrum method.

To better see these effects, consider Figs. 5.12- 5.14 where the true and estimated

IR's of System II, System III and System V, respectively, are shown for SNR level

of 5 dB. It is worth mentioning that the truncation points (p and q) in the previ-

ous figures are set to 7. This gives us more insight about the performance of the

bicepstrum method when these points are not appropriately selected. The cepstral

parameters shown in Fig. 5.3 represent a good tool for the selection of these points.

It is clear from Fig.5.3 that a choice of $p = 7$ and $q = 7$ in the case of System III

and System V has the undesirable effect of truncating many non-zero cepstral terms

that contribute in the system IR. This truncation can be thought of as <u>Gibbs effects</u>

in the quefrency domain. Hence, Fig. 5.3 suggests that for the last two systems, a

choice of a larger order is more appropriate say 10 or 12 for the parameters p and

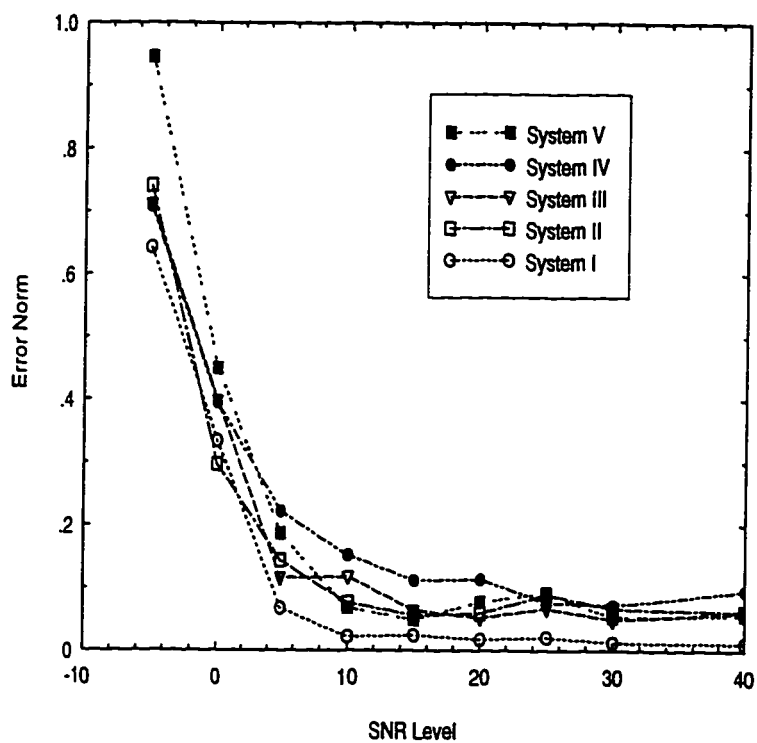q. The performance of the bicepstrum method for diffent choices of the parameters

Figure 5.11: Bicepstrum Method Performance for the Five Selected Systems in White Gaussian Noise.
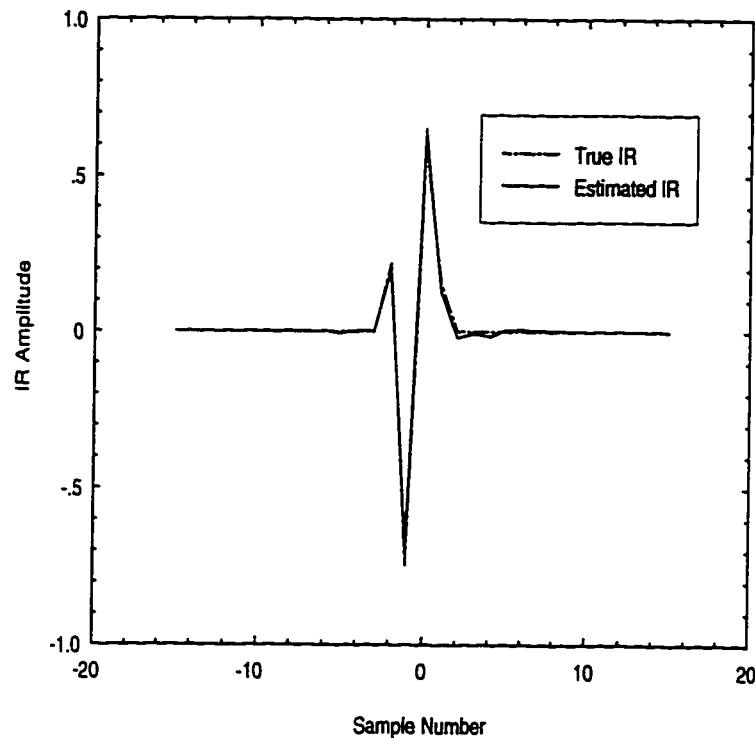
Figure 5.12: True and Estimated IR's of System II at SNR = 5 dB.

p and q is given in Fig. 5.15 for SNR = 30 dB in the case of System III and System

V. It should be noted that the combinations of the values of p and q range from (5,

5) to (9, 9). In Fig. 5.15, the obtained performance does not necessarily correspond

to the higher values of p and q. The performance index (i.e., the error norm) is not

directly proportional to the values of p and q. Table. 5.1 shows the relation between

the error norm and the values of p and q. Table 5.1 reveals interesting insights about

the influence of the choice of (p, q) values on the performance of the bicepstrum.

Some remarks about Table 5.1 are given below :

• System III is of MP type. However, it requires the use of q which corresponds

    to the maximum phase terms that do not exist theoretically in its IR. This is
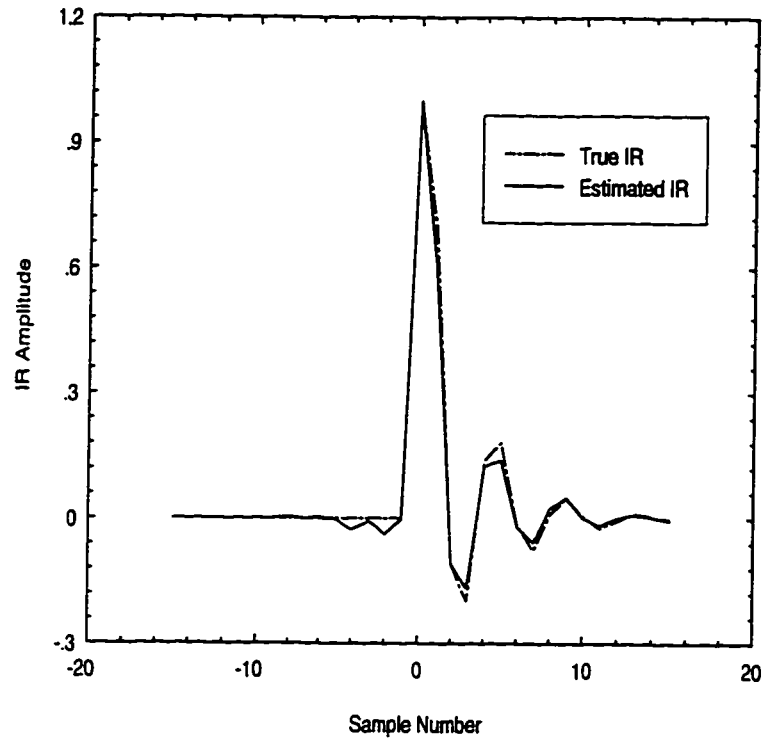
Figure 5.13: True and Estimated IR's of System III at SNR = 5 dB.

probably due to the inherent phase shift intrduced by HOS.

• The bicepstrum with System V achieves better performance when using greater values for p and smaller ones for q (see Table 5.1).

Fig. 5.15 indicates that the choice of the truncation points highly affects the quality of the estimated IR. This is due to the loss of information when some contributing cepstral points ($A^{(m)}$ and $B^{(m)}$ for large values of m) are not included in the solution of Eq. 3.53. This problem clearly appears when dealing with the deconvolution of real ultrasonic signals as shown in the results obtained using experimental data. Also from the knowledge gained in this simulation study, the following must be
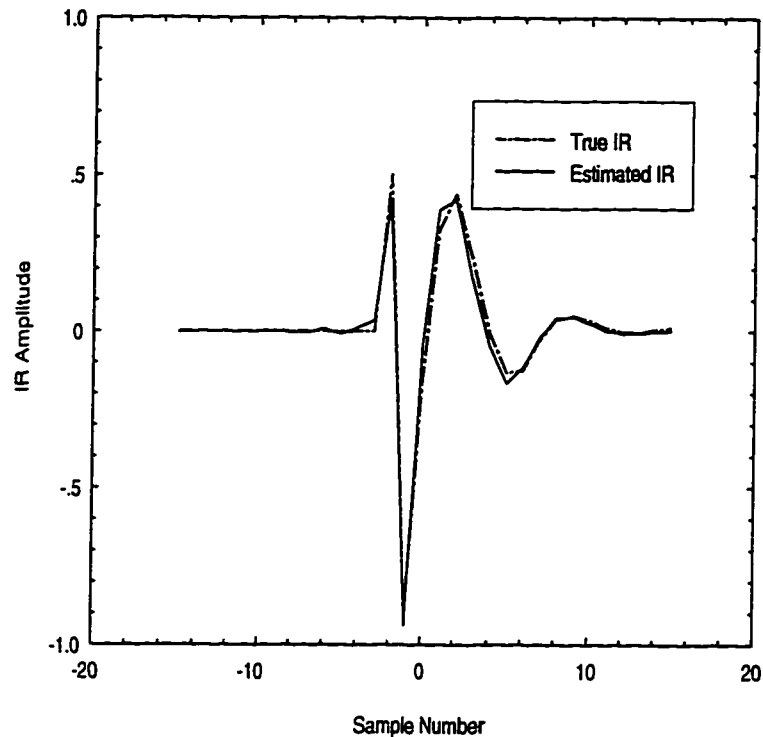
Figure 5.14: True and Estimated IR's of System V at SNR = 5 dB.

observed :

a) Any oscillation present in the estimated defect IR must be contributed to the selected truncation points.

b) HOS-based parametric modeling methods represent a good tool for extracting a-priori information about the positions of the poles and zeros of the system under analysis. This procedure may constitute a <u>preprocessing step</u> for a successfull deconvolution procedure using the bicepstrum technique. Note also that the bicepstrum itself may give an idea about the extent of the cepstral parameters $A^{(m)}$ and $B^{(m)}$ in the quefrency domain for the linear case [51].
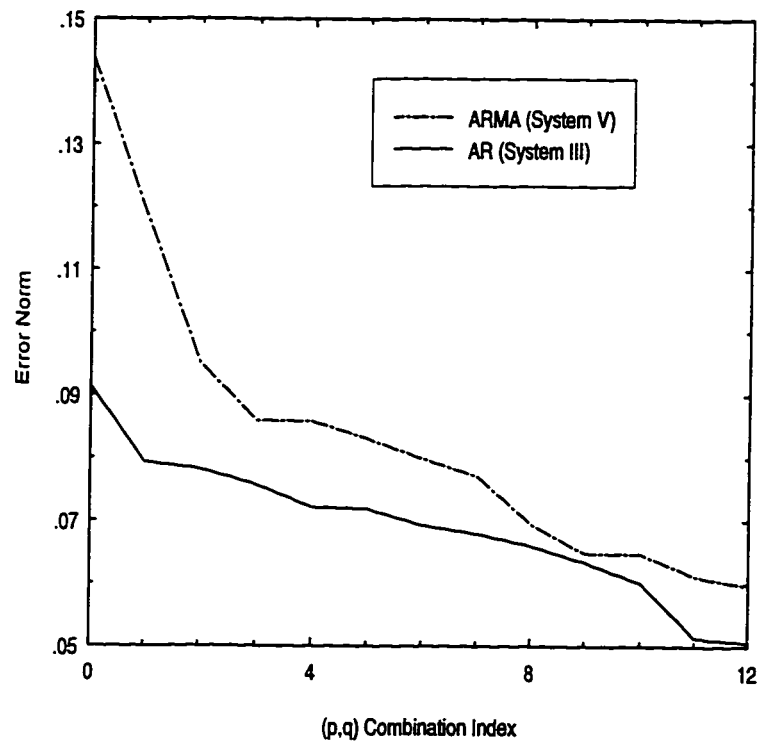
Figure 5.15: Performance of the Bicepstrum Method for Different Choices of p and q.

To conclude this analysis part, it is worth mentioning that one of its merits is that the analysis of the effects of p and q has not been documented in any published work related to polycepstra.

5.1.4   Automatic Defect Classification System

To highlight the practical merit of the new defect classifier system, a case study is performed using synthetic ultrasonic signals involving the classification of defects using their IR's obtained from the proposed HOS-based deconvolution scheme. The collection of the defects data is performed as follows :

1) Use the five synthetic defects systems (System I, ..., System V).

2) Set the SNR to 25 dB.

3) Apply the bicepstrum-based deconvolution technique to estimate the systems IR's.

4) Obtain for each system 50 different IR estimates.

5) The overall 250 estimates constitute the training (exemplar) set for the proposed intelligent classification system.

6) Set the SNR to 20 dB.

7) Repeat Steps 3-4.

System I

System II

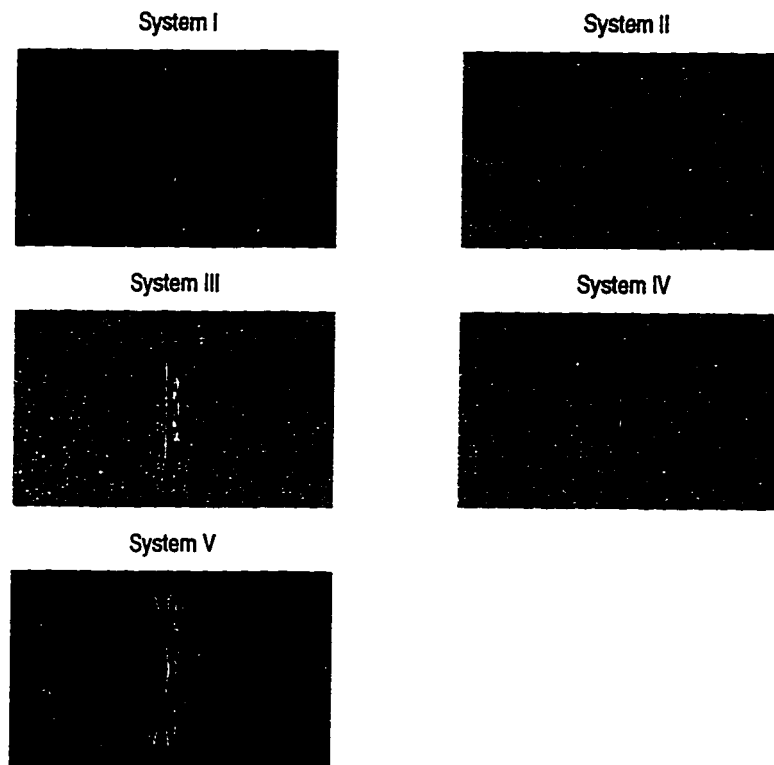System III

System IV

System V

Figure 5.16: WVD "Images" of the Five Systems IR's.

8) The new 250 IR estimates form the testing set for the proposed defect classifier

system.

WVD Computer

To appreciate the importance of the WVD block for the defect classification problem

at hand, WVD "images" of the artificial defects IR's and real ultrasonic signals are

presented. Fig. 5.16 shows the WVD "images" of the IR's of the five defects systems.

From these figures, it is clear that WVD "images" present a combined information

about the behavior of the defects in both time and frequency domains. Also note

the existence of cross-terms in the different WVD "images" that make the latter

an attractive tool in the problem at hand. To clearly see the cross-terms effects, consider the following signal :

$$s(t) = s_1(t) + s_2(t)$$

where

$$s_1(t) = \begin{cases} \exp\left(j2\pi f_1 t\right) & 0 \le t \le t_1. \\ 0 & \text{otherwise.} \end{cases}$$

and

$$s_2(t) = \begin{cases} \exp\left(j2\pi f_2 t\right) & t_2 \le t \le t_3. \\ 0 & \text{otherwise.} \end{cases}$$

with $f_1 < f_2$ and $t_1 < t_2 < t_3$. Fig. 5.17 shows the signal s(t). Theoretically, any JTFA distribution must have only two peaks corresponding to f_1 and f_2 in the case of the signal $s(t)$. However, WVD will have in addition to these two peaks, a third one corresponding to the interaction between f_1 and f_2 located at $f_c = \frac{f_1 + f_2}{2}$ [56]. These cross-terms allow WVD to preserve the information pertaining to the interaction between the different frequency components existing in the defects IR's [56]. The cross-terms effects are illustrated in Fig. 5.18. The cross-terms appear in Fig. 5.18 as the center line located at the frequency $f_c = \frac{f_1 + f_2}{2}$. In Fig. 5.18, the frequency
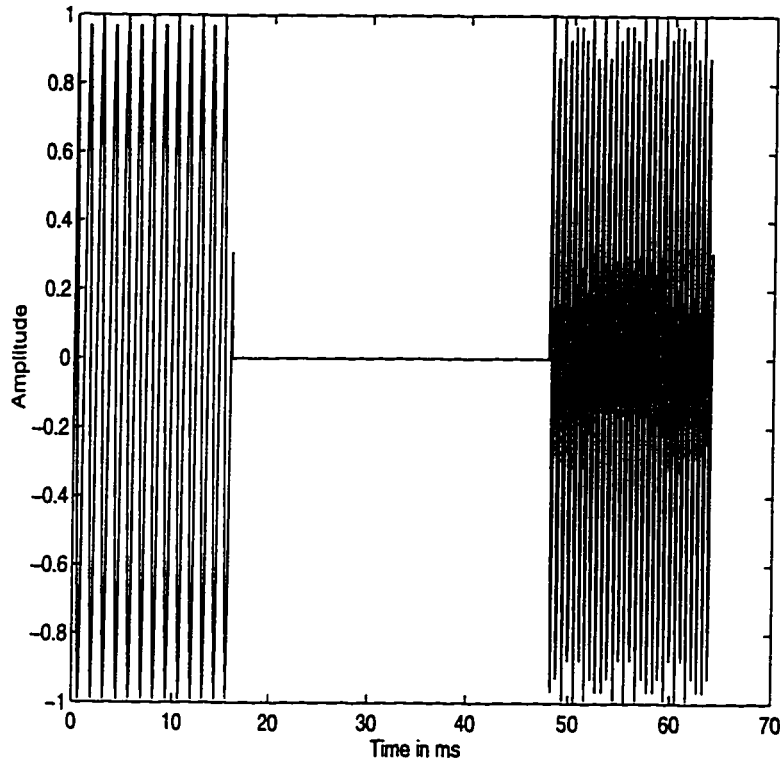
Figure 5.17: Time Representation of Signal s(t).

increases in the column direction from left to right and the time in the rows direction from up to down. Both positive and negative frequencies are shown. The cross-terms clearly appear and are located at the midfrequency $f_c = \frac{f_1 + f_2}{2}$. The analytic form of s(t) using the Hilbert transform reduces the effects of the cross-terms. Fig. 5.19 shows the WVD of s(t) using its analytic form. Only the positive frequencies are present in Fig. 5.19 due to the analytic form of the signal used. Note from Fig. 5.19 the clear representations in the WVD "images" when the analytic form of the signals is used.
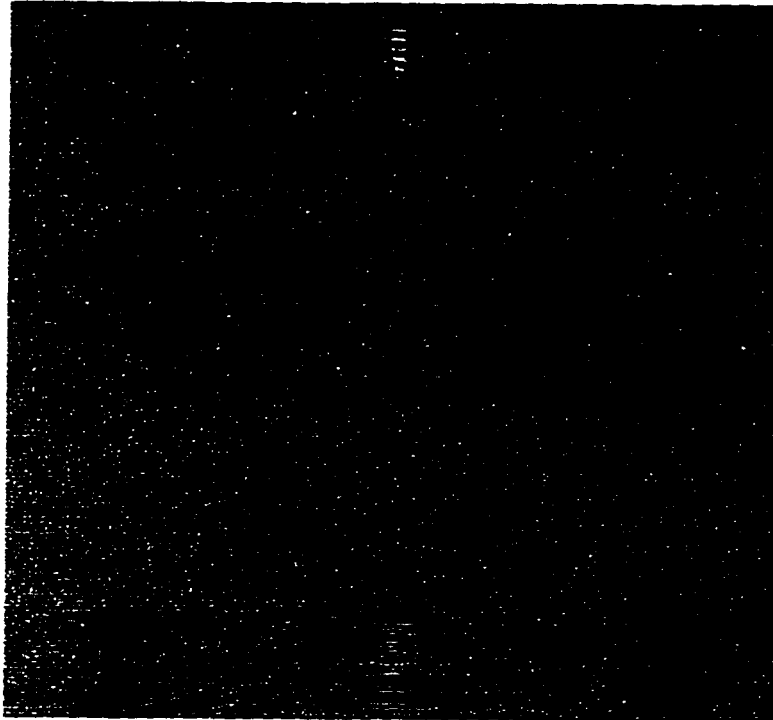
Exact WVD

Figure 5.18: Cross-Terms Effects in the Exact WVD Image of s(t).

Feature Extraction

Feature extraction is a difficult task and problem-dependent as reported by Chen [2]. In the proposed scheme, WVD "images" corresponding to each defect are compressed using GHA algorithm in an unsupervised fashion. This kind of training requires many passes through the images of interest to optimally and globally learn their statistical characteristics, unlike the KLT method which is locally optimal [62] and relies on the stationarity assumption of the images being coded or compressed [64]. Fig. 5.20 shows the learning curves of the unsupervised algorithm when applied on the WVD representing the five different defects. The step size is set to 5.10^{-3} for the first four systems. However the algorithm failed to converge in the case of System

Analytic WVD

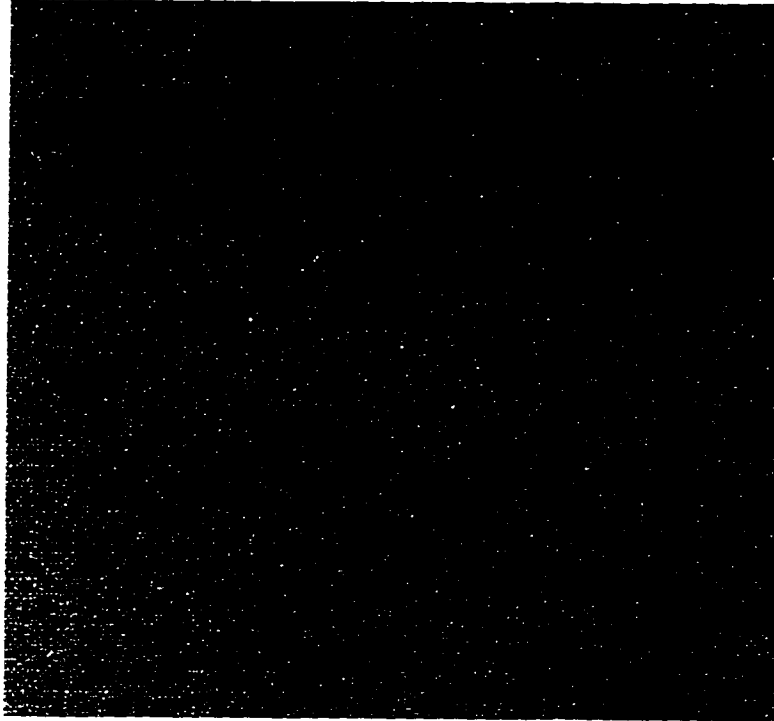Figure 5.19: Analytic WVD Image of s(t) with Reduced Cross-Terms.

V when using this step size. In last case, the step size is set to with 5.10^{-4}. The learning curves are shown only for the first 200 epochs [1] of the convergence process. To see the effect of the step size on the convergence rate of the GHA algorithm, the latter is set to 10^{-2} for the first four systems and to 10^{-3} for System V. The corresponding learning curves are illustrated in Fig. 5.21. Note that the learning curves in both cases are not smooth because no ensemble averaging is performed, i.e., the curves are obtained for only initialization of the PCA random weights. From the carried simulation, it is concluded that the step size cannot set to a value greater than 10^{-2} for the first four systems and 10^{-3} for the last system. From the

---

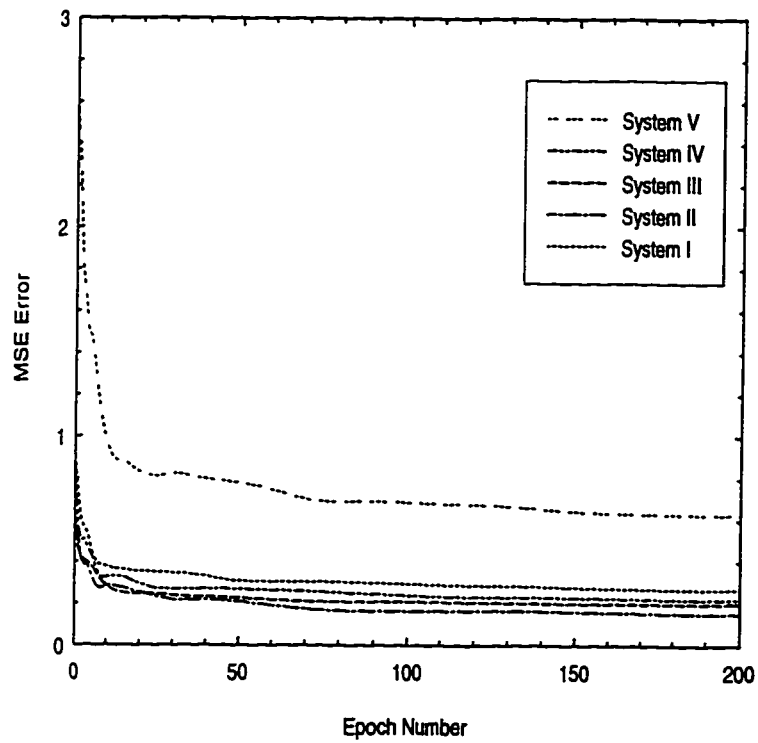[1] One epoch represents a complete pass through one WVD "image".

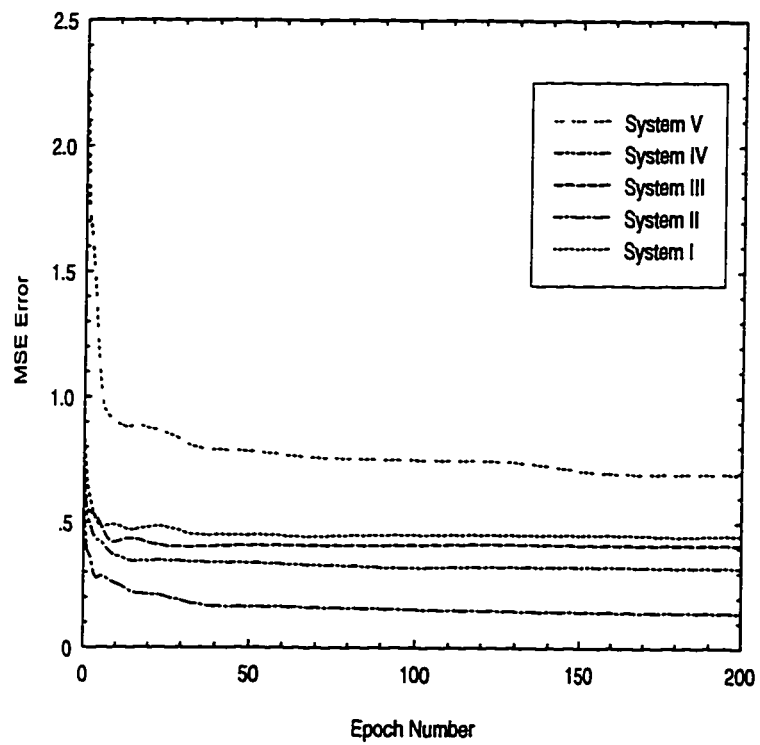Figure 5.20: Learning Curve of the GHA Algorithm for the WVD "Images" of the Five Systems (Case I).

Figure 5.21: Learning Curve of the GHA Algorithm for the WVD "Images" of the Five Systems (Case II).
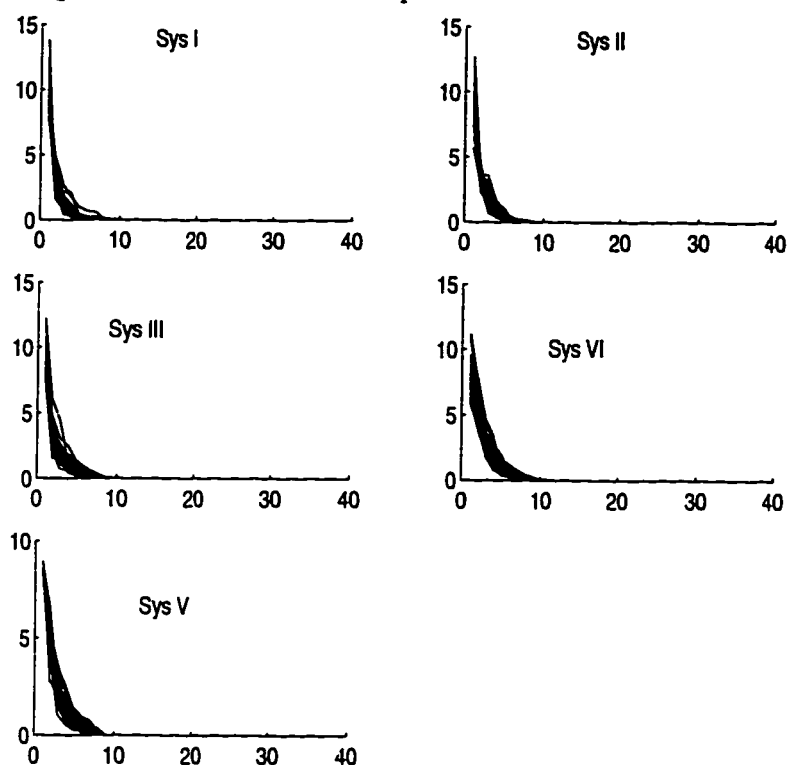
Figure 5.22: Singular Values of the Compressed WVD of the Defects IR's.

comparison of the last two figures, it appears that the GHA algorithm converges after 50 epochs. Also, it should be noted that the steady state cannot be reached after only two passes as reported by Sanger [62]. In both cases, the step size is set fixed unlike Sanger [66] who used a variable step size according to the convergence of the PC vectors to their final values. Using the SVD of the compressed WVD "images", the most dominant singular values are retained as effective features and fed into the ANN defect classifier block. All singular values of the compressed WVD "images" are illustrated in Fig. 5.22. From this figure, it clearly appears that the first 10 singular values, in all cases, are sufficient to describe the different feature spaces.

Defect Classification

Two different sets of features are fed to the ANN classifier block as follows :

1) Features set of estimated IR's at 25 dB. These features constitute the training set.

2) Features set of estimated IR's at 20 dB. These features constitute the testing set.

The performance of the classifier block is represented by its learning curve in Fig. 5.23. Five independent training realizations are shown in Fig. 5.23(a). Their ensemble average is shown in Fig. 5.23(b). It should be noted that one epoch represents a complete pass through all the vectors of the training set. The ANN classifier block has the following architecture :

1) Input layer consisting of 10 neurons. The input vector is formed by the 10 most dominant SV's of the compressed WVD "images".

2) First hidden layer with 8 neurons.

3) Second hidden layer with 5 neurons.

4) Output layer having 3 neurons. The output is of binary form, where each binary combination represents one defect class as shown in Table 5.2.

It is clear from Table 5.2 that not all the binary combinations correspond to defects classes since among 16 combinations only 5 ones are being used. The remaining
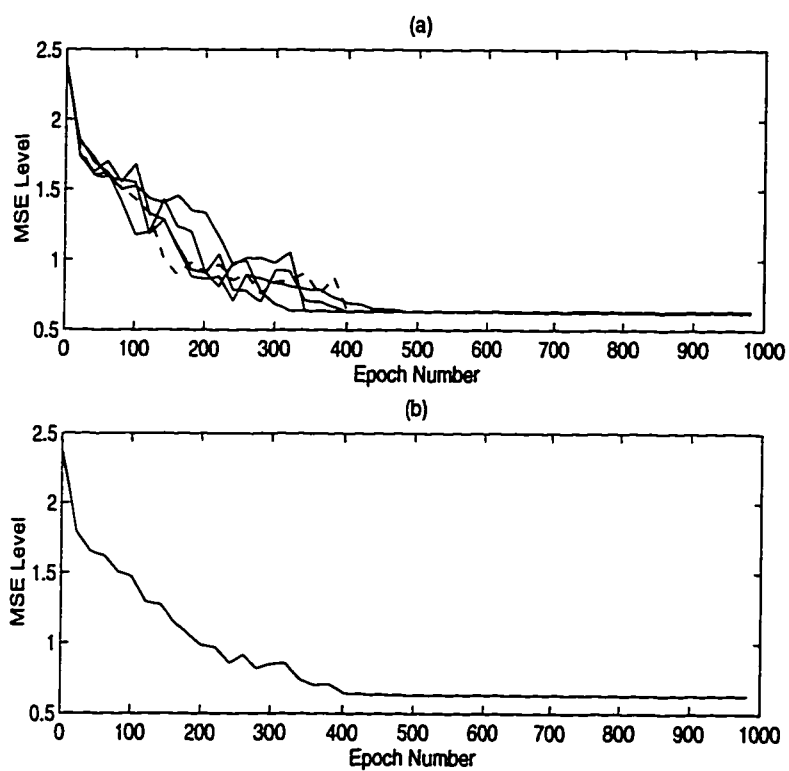
Figure 5.23: Learning Curve of the ANN Classifier Block. (a) Five Independent Realizations. (b) Ensemble Average.

binary combinations are considered as <u>unknown</u> or <u>unclassified</u> defects classes. The results of the defects classification using both sets of features are summarized in Table 5.3. A second network architecture is proposed using :

1) Input layer consisting of 10 neurons. The input vector is formed by the 10 most dominant SV's of the compressed WVD "images".

2) One hidden layer with variable size.

3) Output layer having 5 neurons. The output is of binary form, where only one output is activated to indicate the classified defect class.

However, with an increasing number of hidden neurons, the second network did not achieve the desired classification. Comparing both networks architectures highlights the following :

• The first network has two hidden layers which enables it to <u>efficiently</u> and <u>rapidly</u> learn the defects feature space.

• The defects space is too complicated to be resolved by only one hidden layer. Hence, the defect classification problem at hand constitutes a multidimensional space approximation problem [26, 64] that cannot be solved using only one hidden layer.

• The computational complexity is higher when using the first network architecture.

From the carried simulation, the ANN classifier block needs at least two hidden layers to converge.

5.2 Experimental Results

To highlight the practical merits and limitations of the proposed HOS-based deconvolution scheme, different artificial defects are created in the test material. The experimental data is obtained from the inspection of the test material. The different defects geometries and sizes considered are illustrated in Fig. 5.24. Results using SOS-based deconvolution techniques reviewed in Chapter 2 are also shown for comparison purposes. The main experimental results shown here are those corresponding to the specimen with defect of angular cut at $10°$. The test is performed using an ultrasonic transducer whose center frequency is 15 Mhz. The sampling frequency is 100 MHz. It is refered to this defect by T15A2. This flaw specimen is illustrated in Fig. 5.24. The reference signal obtained from the flawless specimen A0 using a 15 MHz transducer is shown in Fig. 5.25a. The signal is called T15A0. The echo from the defect A2 at 15 MHz is shown Fig. 5.25b. It is refered to this signal by T15A2.
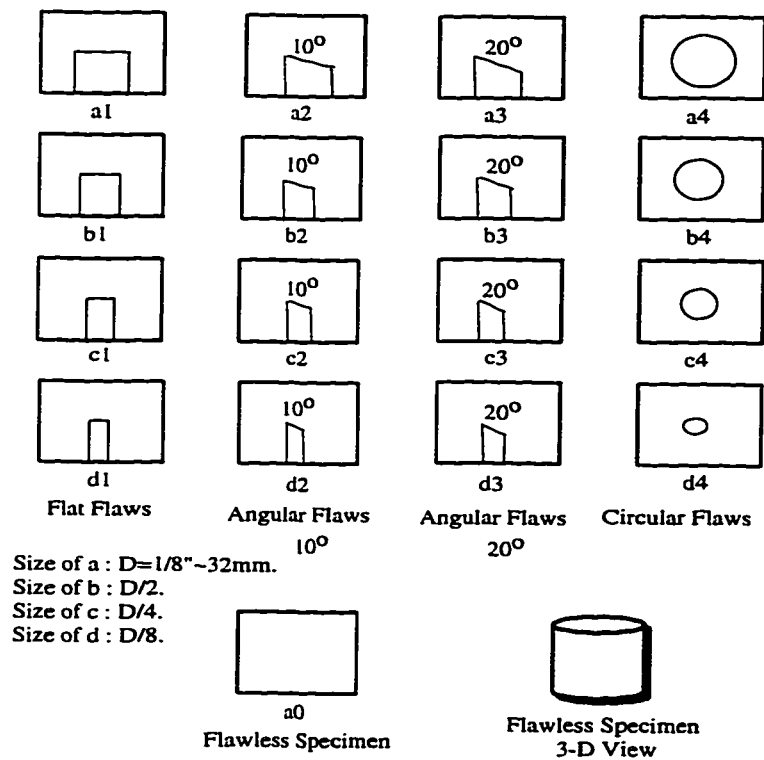
Figure 5.24: Artificial Defects Considered.
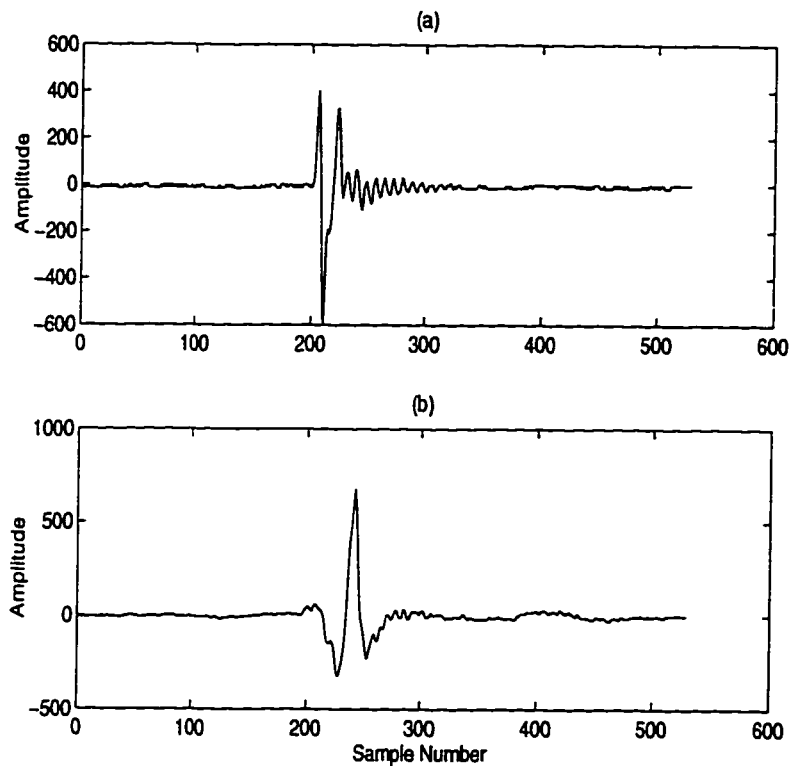
Figure 5.25: Experimental Reference (a) and Echo (b) Signals of Defect A2 at 15 MHz.
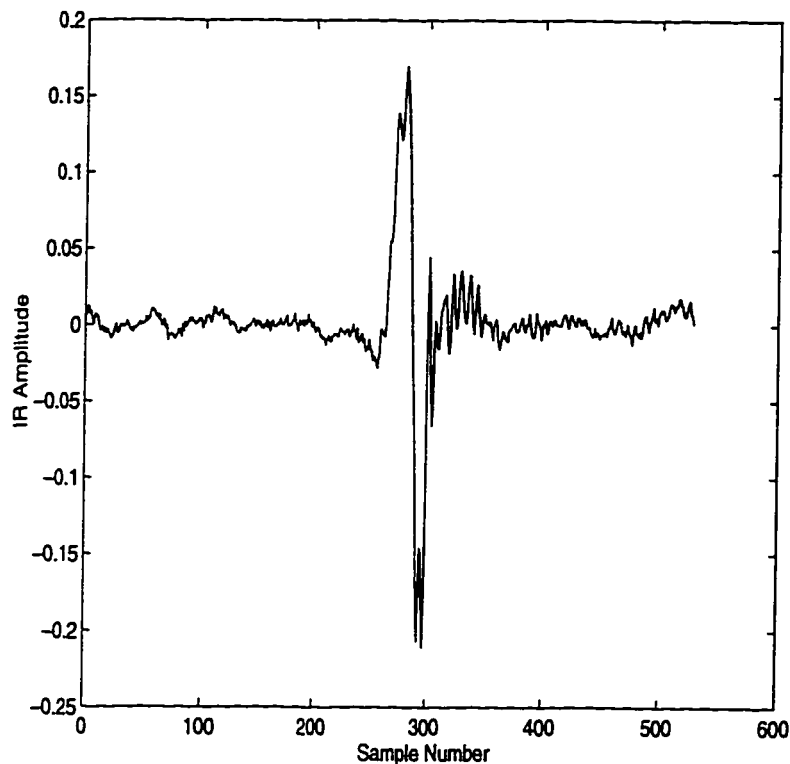
Figure 5.26: IR Using Wiener Filter A Method.

5.2.1 SOS-Based Deconvolution

Both Wiener A and Wiener B filters are used to estimate the IR of defect T15A2. The deconvolution results are shown in Fig. 5.26 and Fig. 5.27. The obtained results are in good agreement with those obtained by Chen and Sin [7, 8] for the same desensitizing factor q equal to $0.1 \max |X(f)|$. However, it should be noted that in the case of Wiener B, the prewhitening filter order is chosen according to the SVD of the correlation matrix of both reference and echo signals. The parametric formulation of Wiener B filter allows it to yield smoother and better resolution defect IR than that of Wiener A. Fig. 5.28 shows the singular values of the corrrelation
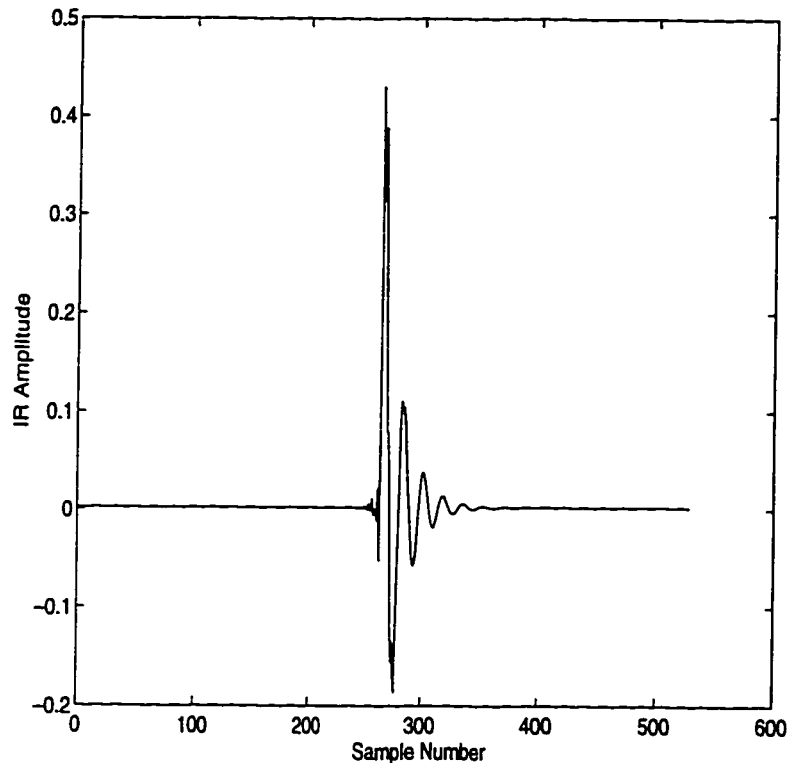
Figure 5.27: IR Using Wiener Filter B Method.

matrices of both reference and echo signals T15A0 and T15A2, respectively. The IR

obtained using the spectral extrapolation (SE) method is shown in Fig. 5.29. The

method extrapolates the spectral densities of the signals in both directions of the

frequency domain. This extrapolation is performed using an AR model of order 10.

The comparison of Figs. 5.26- 5.28 and Fig. 5.29 shows that the obtained IR's are

quite similar. Also, time domain SOS-based methods are tested. The estimated IR

using the LS method is shown in Fig. 5.30. Both signals were aligned before running

the algorithm [7, 8] using the cross-correlation lag at which occurs the maximum

correlation between the two signals. The lag of maximum correlation is found to be

equal to -17 with a maximum normalized correlation level of -0.601. Hence, the echo
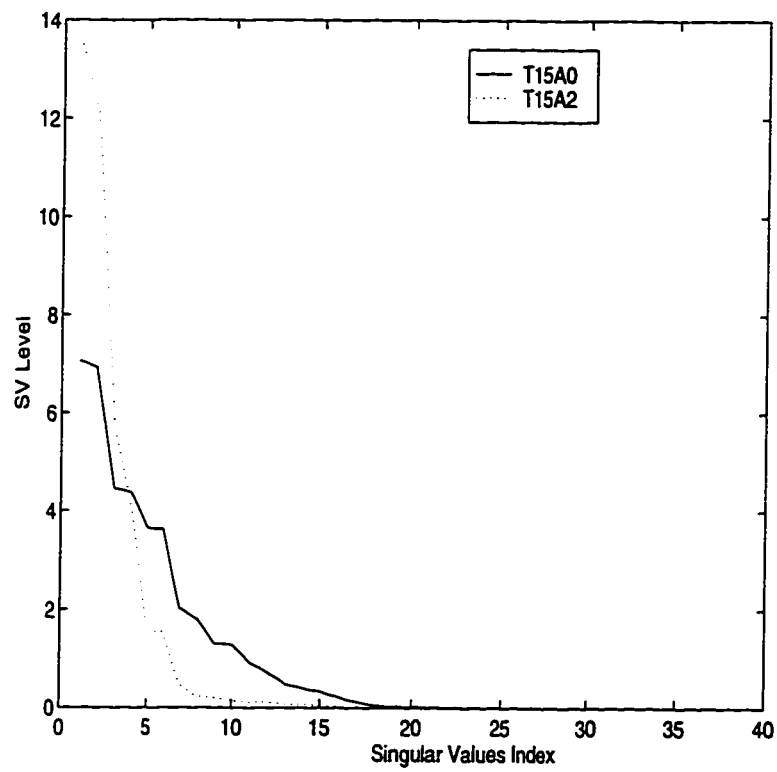
Figure 5.28: SVD of Correlation Matrices of T15A0 and T15A2 Signals.
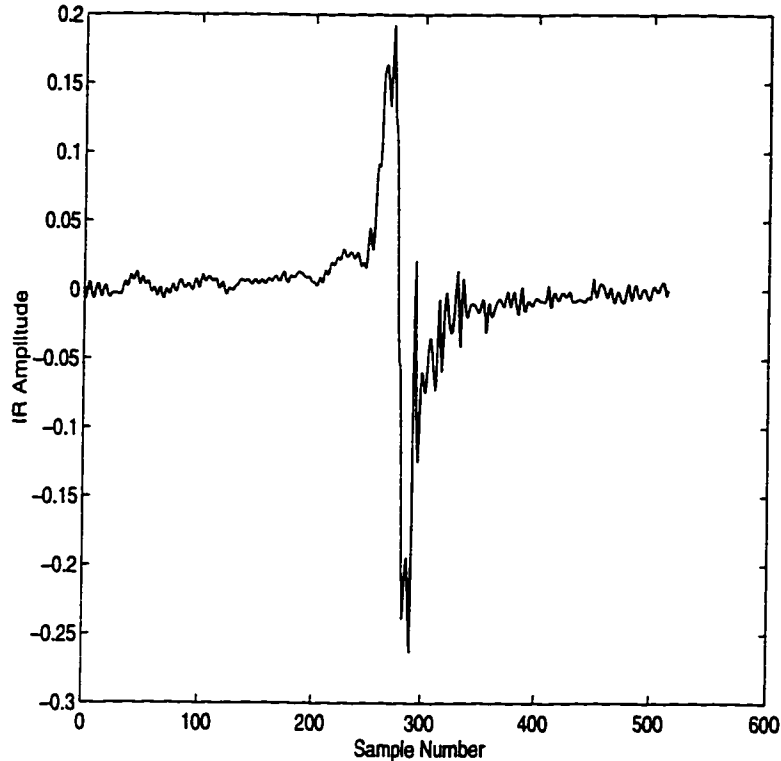
Figure 5.29: IR Using Spectral Extrapolation Method.

signal T15A2 is circularly shifted to the left side by an amount of 17 time samples. It clearly appears that the defect IR is highly affected by the noise. This is due to the ill-conditionning of the signal matrix using T15A0 which has many null entries making it almost singular [67] (see Chapter 2 for details). In an attempt to improve the resolution of the LS method, a solution based on SVD is proposed. Both signals are aligned and truncated in the time interval of [225 : 360]. The estimated IR is shown in Fig. 5.31. The truncation effects on the obtained IR clearly appear in this figure. However, it should be noted that this truncation interval are found to the best after many trials. The correlation method is implemented to yield the estimated defect IR shown in Fig. 5.32. This method has defined confidence limits for the
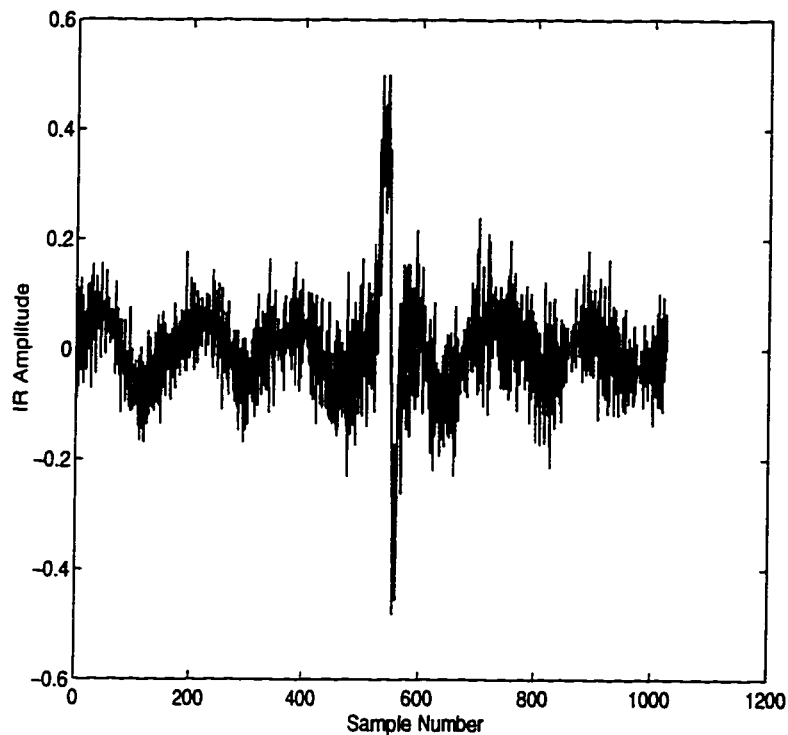
Figure 5.30: IR Using Time Domain LS Method.

estimated IR. The length of the correlation used is 128. The prewhitening filter is selected to be an AR model with order 10. The computed confidence limits at 95 % clearly show that the estimated IR is not consistent. This is mainly due to the nature of the ultrasonic signals where the poles closeness to the unit circle make their true IR's larger which will necessitate the solution of a very large system of equations which is usually ill-conditioned [67]. Finally, the homomorphic deconvolution is applied using the complex cepstrum method. In theory, this methods works only if both signals are well separated in the quefrency domain, which does not hold in the case of experimental ultrasonic NDE signals. The result shown in Fig. 5.33 confirms this assertion. Clearly, the performance of homomorphic deconvolution reported
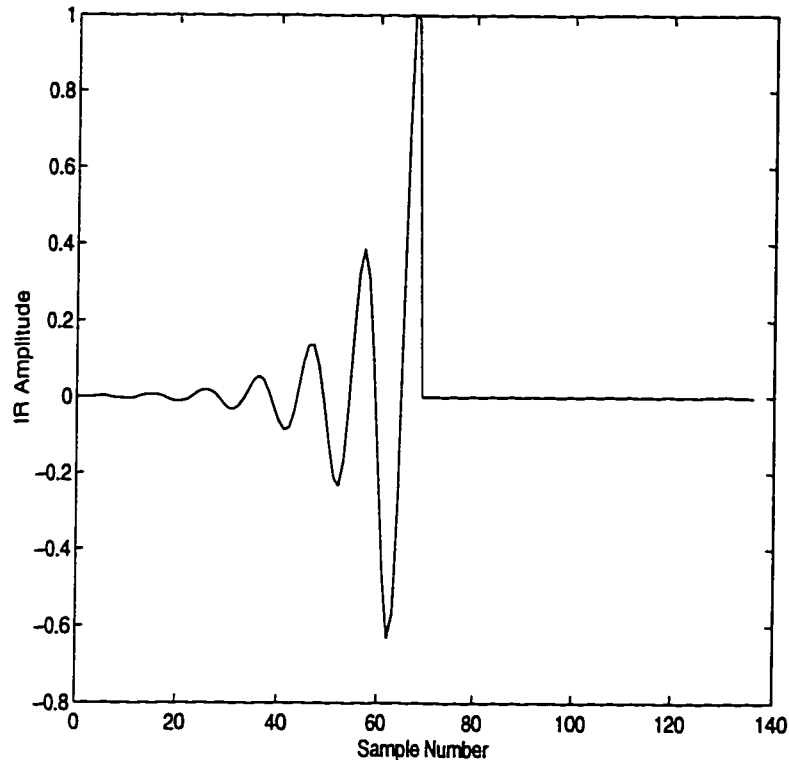
Figure 5.31: IR Using SVD-Based LS Method.

by Bhagat and Shimmin [68] cannot be achieved when dealing with real ultrasonic signals. This is due mainly to the fact these signals do not necessarily follow the models assumed by Bhagat and Shimmin [68].

5.2.2 HOS-Based Deconvolution

The bicepstrum method yields the IR shown in Fig. 5.34 for cepstral orders set both to 6. The obtained IR has a high resolution and satisfactory smoothness. The proposed method requires the estimation of a cumulants matrix of size (91 x 36) in the case of $p = q = 6$. It is worth mentioning that the data used contains only 512 time samples. To demonstrate the performance of the bicepstrum, the
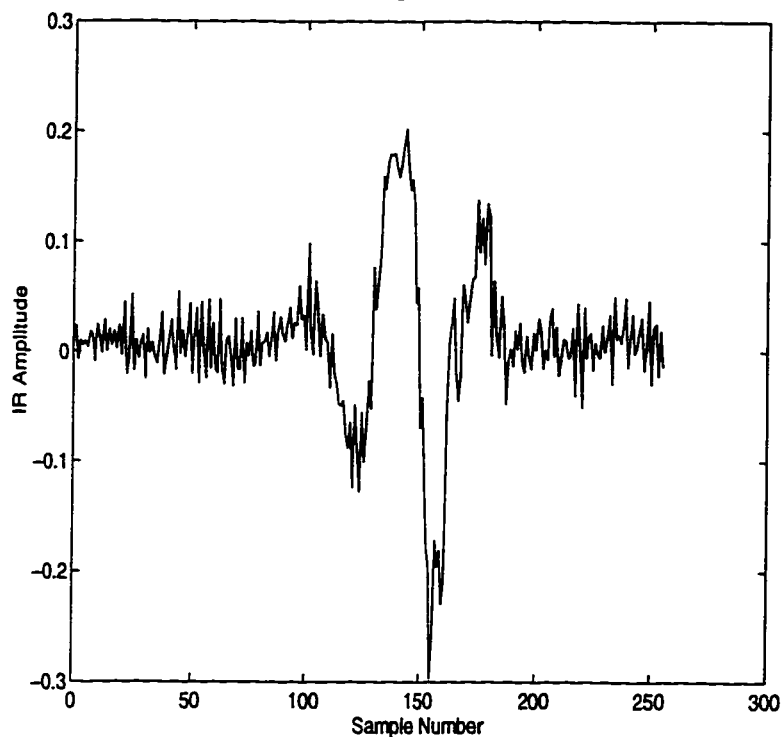
Figure 5.32: IR Using Correlation Method.

bispectrum has been used to estimate the defect IR. The deconvolution result is shown in Fig. 5.35. As shown in the figure, the method yields a noisy IR estimate. This is due to the short data used in the estimation which is of 512 length. Elgar et al. [69, 70] have investigated the statistics of the bispectrum and the trispectrum and noticed that the required data length to yield low variance polyspectral estimates is of the order of 10000 data samples. The oscillations seen in Fig. 5.35 are also due to the signal recovery method which is based on the Matsuoka-Ulrych algorithm [48]. The effects of short data length on the bispectrum estimation are given in [30].
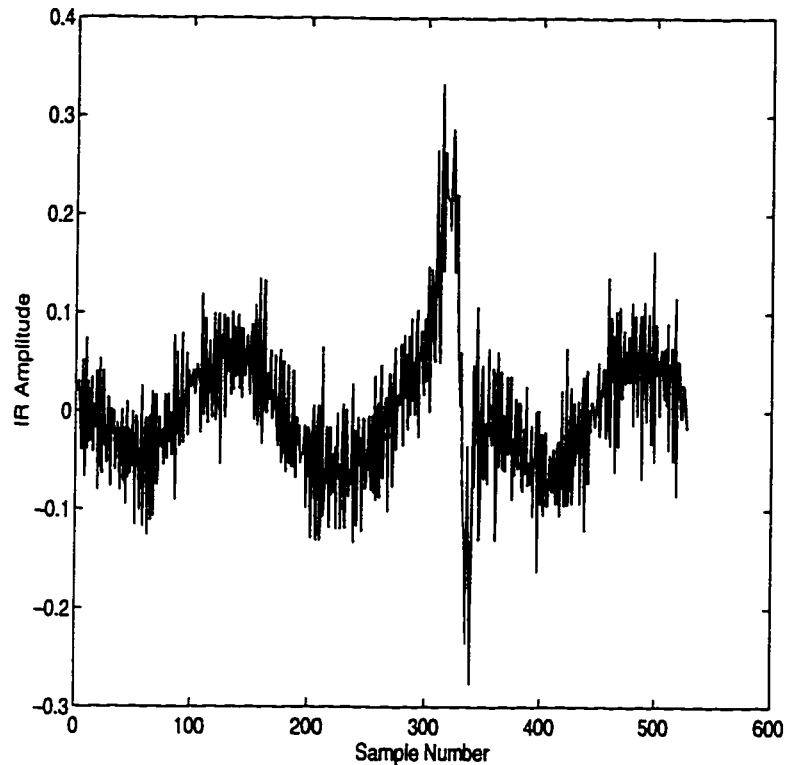
Figure 5.33: IR Using Complex Cepstrum Method.

5.3 Discussion and Conclusion

5.3.1 Discussion

In this chapter, the application of both SOS- and HOS-based techniques to ultrasonic NDE deconvolution is treated. A new HOS-based approach is proposed. In the computer simulation part, to make the deconvolution results comparable and accomodate both deconvolution classes (SOS and HOS), appropriate data lengths and simulation conditions are selected. For an objective comparison between both approaches, it must be kept in mind that :

- SOS-based methods produce longer defects IR's.
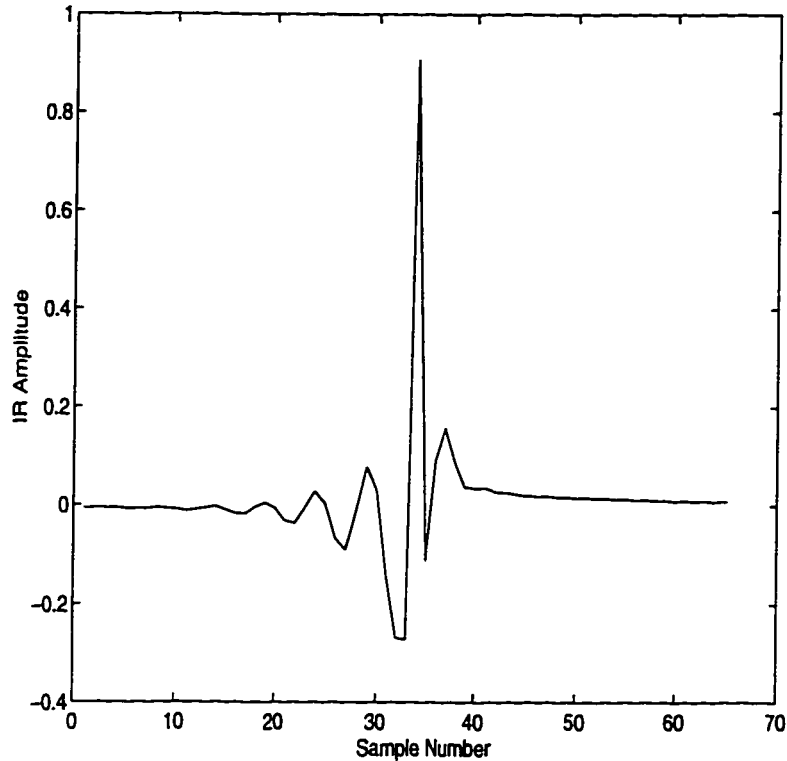
Figure 5.34: IR Using the Bicepstrum Method.

- Wiener filter method is the best SOS-based deconvolution method as reported by Chen and Sin [7, 8] and verified experimentally in this chapter. Hence, most comparisons are made using the latter technique.

- An improved version of the LS method is proposed in this part. It is based on the SVD solution of the linear system.

- Results using the complex cepstrum-based deconvolution are presented for comparison with those obtained using the bicepstrum and verify the failure of the assertion made by Bhagat and Shimmin [68] in practice.
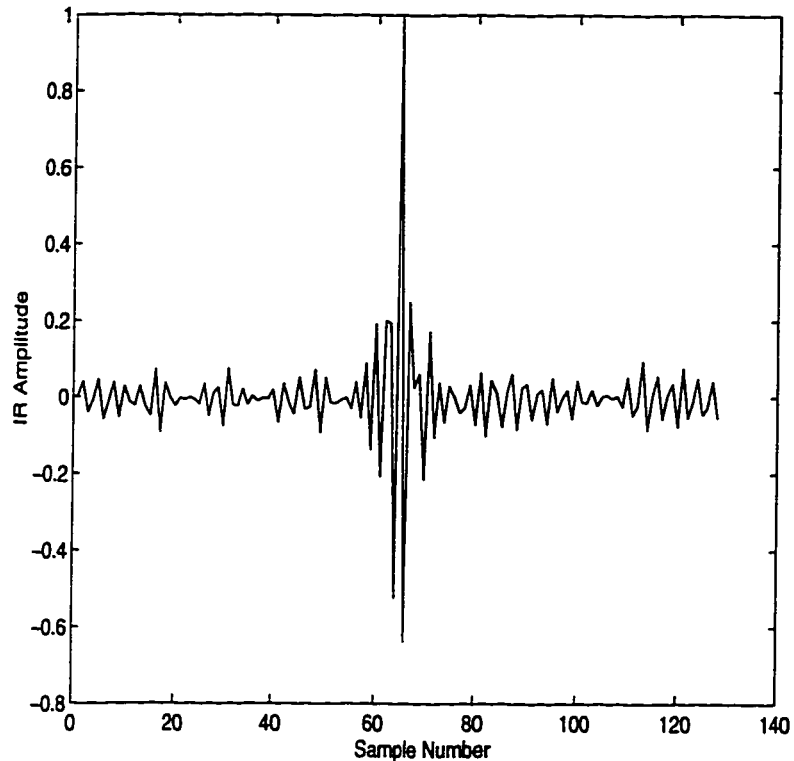
Figure 5.35: IR Using The Bispectrum Method.

• The proposed HOS-based method produces shorter defects IR's. This is due to the fact that the IR length is governed by the choice of the cepstral orders p and q.

The obtained results are generally in good agreement with the theory which states that HOS are insensitive to Gaussian and any white symmetrically-distributed noises (in the case of TOS). In the case where existing defect in the material is characterized by a system which has singularities are located in the vicinity of the unit circle, the underlying system IR will be larger enough to make its output more likely Gaussian by virtue of the CLT theorem [71]. This problem does not rise with SOS-based methods which inherently assume the data Gaussian. In this case, HOS-based

schemes will fail to recover the defect IR [71]. However, for the proposed bicepstrum method, a solution is possible by using exponential windowing in the HOS domain. This windowing affects the statistical shape of the system output without altering its stationarity. From the implementation point of view, SOS-based methods are not as affected by the data length unlike their HOS counterparts. Also, they are known to be computationally efficient compared to the latter. Their major limitations are :

- Poor performance at moderate and low SNR levels.

- Phase insensitivity since they always yield minimum phase systems.

5.3.2 Conclusion

At the end of the simulation carried out in this chapter where the performance analysis of the proposed ultrasonic NDE processing techniques, some conclusions are given :

1) The bicepstrum method is a very efficient nonparametric HOS-based scheme with improved perfomance even in the case of small data length as shown in the case of experimental data. Hence, It is suggested for ultrasonic NDE deconvolution in severe contexts such as correlated noise and low SNR levels.

2) The nature of the experimental ultrasonic signals do not allow to have larger data lengths. The large data length will consist mainly of identically zero samples

which are located at both tails of the ultrasonic echo. Hence, the experimental ultrasonic NDE inspection system must have the possibilty to record different realizations of the echo signals in order to have a consistently large amount of data to produce low variance HOS estimates. This is made possible in the case of ultrasonic B-scans where different measurements are recorded at different transducer positions.

3) The proposed defect classification scheme due to its modular architecture can be adapted to learn new defects features where the training is performed independently. This enables the proposed scheme to "adaptively learn" the defects feature space where the size of the ANN network can be increased to handle more difficult contexts.

System III		System V	
(p,q) Comb. Index	(p,q) Values	(p,q) Comb. Index	(p,q) Values
0	(5,5)	0	(6,6)
1	(5,6)	1	(6,7)
2	(7,9)	2	(6,5)
3	(7,7)	3	(8,6)
4	(6,5)	4	(5,5)
5	(9,9)	5	(5,6)
6	(6,7)	6	(6,8)
7	(8,6)	7	(8,8)
8	(6,8)	8	(7,7)
9	(7,6)	9	(7,6)
10	(9,7)	10	(7,9)
11	(6,6)	11	(9,7)
12	(8,8)	12	(9,9)

Table 5.1: Effect of the (p,q) Values on the Bicepstrum Performance.

Defect / Output	Output b_0	Output b_1	Output b_2
System I	0	0	0
System II	1	0	0
System III	0	1	0
System IV	1	1	0
System V	0	0	1

Table 5.2: Output Binary Combination of Defects Classes.

Procedure	Training at 25 dB	Testing at 25 dB	Testing at 20 dB
Results	96 %	96 %	80 %

Table 5.3: Classification Performance Using Both Features Sets.

Chapter 6

Conclusions and Suggested

Future Work

6.1 Summary

Motivated by the recently increased interest in the use of HOS, deconvolution of ultrasonic NDE signals based on HOS and automatic defect classification using modular learning strategy are investigated in this thesis. A potential class of HOS-based deconvolution is proposed based on polycepstra principles, along with its performance evaluation through simulation, using computer generated and experimental data. The proposed method exploit the properties of cumulants and polycepstra. As such, its implementation depends on the knowledge of higher-order cumulants of the involved signals. Also, being blind to Gaussian and all white symmetrically-

distributed noises in the special case of third-order statistics (TOS), enables this scheme to :

- operate in high signal-to-noise ratio (SNR) domain.

- operate efficiently in both white or colored Gaussian noise environments.

- operate well in any white symmetrically-distributed noise environment.

- preserve the exact nonminimum phase character of the underlying defect IR's.

The bicepstrum method, of batch type, is described in Chapter 4. It is demonstrated by means of computer simulations and comparisons with existing methods (those described in Chapter 2), that the bicepstrum-based deconvolution method performs effeciently with low sensitivity to additive white or colored Gaussian and other white symmetrically-distributed noises. Main simulation results are summarized in Chapter 5. It is shown in Chapter 4 that the estimation methods associated with the proposed method can be complex and highly ill-conditionned due to the singularities location in the Z plane. The solution is based on the application of multidimensional exponential windowing in the HOS domain. The efficiency of the proposed method is demonstrated through computer simulation and performance analysis carried out in Chapter 5. For the sake of objectivity, it is highlighted that the improved performance of the HOS-based method is achieved at the expense of higher computational complexity and storage requirements, a limitation common

to most HOS-based techniques. Finally, an automatic and intelligent defect classification system is developed. The proposed classification system consists of three different functional blocks where each block is assigned a specific processing task. The system uses JTFA to correctly handle any nonstationary character of the defects IR's and simultaneously combine time and frequency information. WVD "images" of the simulated defects IR's are optimally compressed to yield efficient defects features. The extracted features are fed into an ANN classifier for decision making about the defect class using the backpropagation (BP) algorithm. The performance of the classifier is characterized by an efficiency index defined as the number of correctly classified defects versus the total number of defects. Last but not least, the existing and proposed deconvolution methods are collected in a software toolbox that has been developed at the growth of the computer simulation carried out during this thesis work. The software toolbox, called NDETOOL, operates under MATLAB© environment. It consists of :

1) Set of deconvolution routines based on SOS and HOS.

2) Automatic defect classification routines that implement the proposed defect classification system.

3) Set of utility routines for display, scaling, shifting and other capabilities.

The toolbox contains also a user-friendly demo routine called NDEDEMO. Further deatils about NDETOOL can be found in Appendix A.

6.2 Suggested Future Work

In this section, interesting investigations related to this thesis study are presented for further research. The investigation topics are briefly descriped in what follows.

6.2.1 Exploitation of HOS symmetries

As reported in Chapter 3, HOS have useful symmetry regions, that can drastically reduce the computational complexity associated with the method described in Chapter 4. The proposed method based on TOS can efficiently take account of these symmetries.

TOS Symmetry:

Instead of estimating the third-order cumulants in all (τ_1, τ_2) -plane, only cumulants for lags defined by the wedge $\tau_1 < \tau_2$, and $\tau_1 \geq 0$, $\tau_2 \geq 0$ are computed, then using the property of the symmetry of the third-order cumulants of real signals, the latter are estimated only in the principal domain. Then, to estimate the bispectrum, the method proposed by Bessios and Nikias [72] is used to compute the bispectrum on polar rasters giving only the principal domain (see Chapter 3), again the twelve symmetry regions of the bispectrum of real signals are exploited to determine the bispectrum over the whole (w_1, w_2)-plane. Some efforts are made to develop computer programs that estimate the HOS only in their respective principal domains.
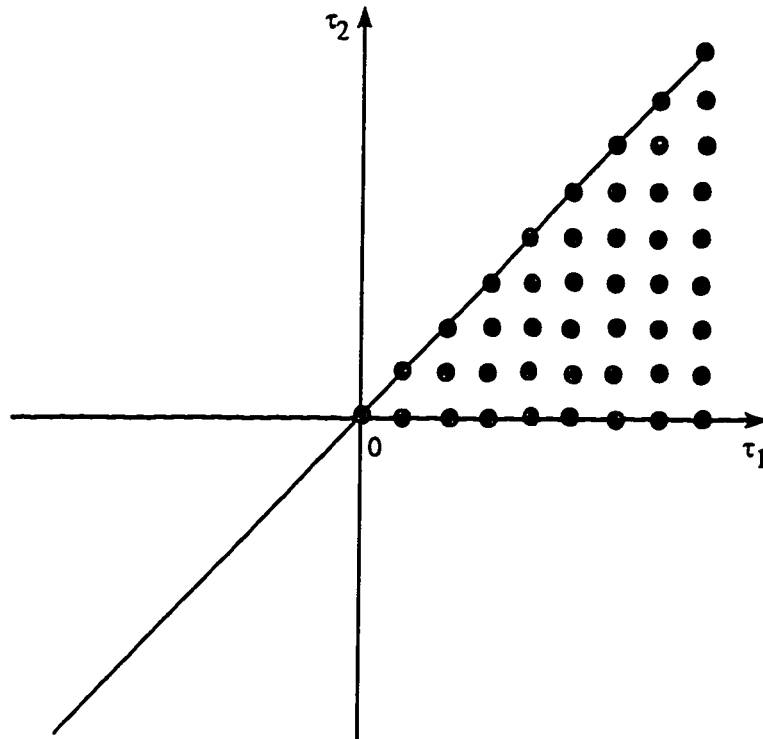
Figure 6.1: Principal Domain of Third-Order Cumulants $c_3(\tau_1, \tau_2)$.

For illustration the principal domain of the third-order cumulants consists of the wedge shown in Fig. 6.1.

6.2.2 HOS-Based Adaptive Deconvolution

It is assumed throughout this thesis that the used signals are stationary, hence, ultrasonic NDE deconvolution is performed off-line (i.e., batch). It is expected that the adaptive implementation will exhibit better performance when operating in slowly-varying environments, where the algorithm is equipped with nonstationary tracking capabilities.

6.2.3 Non-Linear Deconvolution

Throughout this thesis, the system is assumed to operate linearly. It can be shown that most of the existing nonlinearity statistical tests are not suitable for use with ultrasonic NDE signals [73]. An intuitive test based on bicepstrum can be used where the statistical analysis is possible, but appears to be more involved than that of Hinich and Rao-Gabr [52, 53]. HOS can be used for deconvolution of non-linear systems but the developed schemes are restricted to particular cases (i.i.d inputs and second-order Volterra systems). Results reported by Kim and Powers [74] can be investigated for a possible use in the ultrasonic NDE deconvolution problem.

6.2.4 Wavelets and wavelets Packets

Wavelets [75] is a rapidly evolving signal processing area with growing applications in science and engineering. The proposed research in this direction consists of exploiting the *denoising* property [75] of wavelets. Early work on noise reduction with wavelets was very much lossy compression using DCT [75]. A threshold is chosen and those transform coefficients with amplitude below the threshold are set to zero. The choice of the threshold is usually empirical and application-dependent. Most recently, Donoho [76] formalized this approach by an optimal soft thresholding technique. In soft thresholding, additional shrinkage by the threshold value is applied on those coefficients that are not set to zero [76]. Denoising using soft thresholding

achieves near optimal MSE that is obtained by using the theoretical KLT, subject

to the constraint that the denoised signal is at least as smooth as the pure one [75].

More details about soft thresholding can be found in [75, 76]. The proposed proce-

dure is as follows :

1) Take the wavelets decomposition of both reference and echo signals $x(k)$ and

   $y(k)$ using any type of the available wavelets family.

2) Apply denoising techniques using hard or soft thresholding [76]. Abbate et

   al. [77] proposed the use of a wavelet transform processor in ultrasonic flaw

   detection. Their work is based on the wavelet property of denoising.

3) Reconstruct both signals using reliable approximations that retain the most im-

   portant characteristics of the original signals. In wavelets literature, approx-

   imations correspond to the low frequency content of the signal while details

   correspond to their high frequency content which are usually affected by noise.

   The reconstructed signals are called $\hat{x}(k)$ and $\hat{y}(k)$, respectively.

4) Apply the proposed HOS-based deconvolution technique on $\hat{x}(k)$ and $\hat{y}(k)$.

A block diagram for the proposed scheme using wavelets and HOS is given in Fig. 6.2.

An interesting test signal is proposed by Donoho [75, 76]. The pure and noisy version

of this signal are shown in Fig. 6.3. The SNR level is set to -0.5 dB. The denoised

signal using the wavelet decomposition by symmetric mother wavelet is shown in
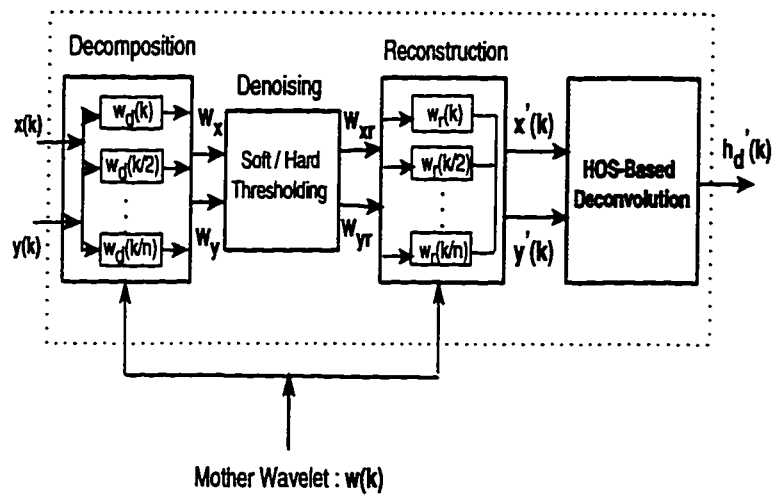
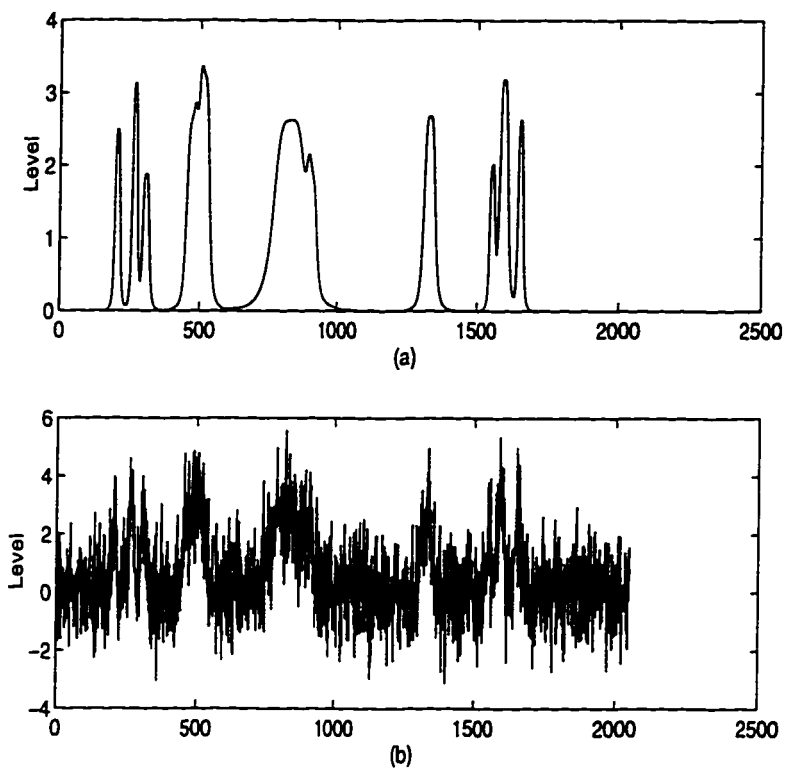Figure 6.2: Block Diagram of Proposed Scheme Using Wavelets and HOS.

Figure 6.3: Pure (a) and Noisy (b) Test Signal.

Fig. 6.4. The 5th decomposition level [75] is used in the reconstruction of the denoised signal.
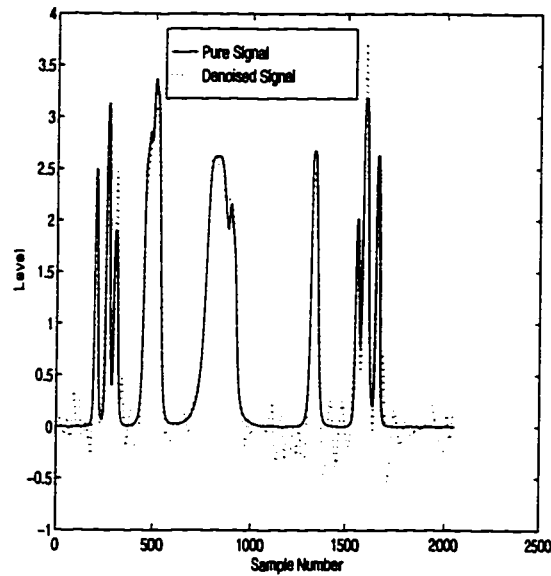
Figure 6.4: Denoised Signal Using Wavelet Decomposition With Symmetric Mother Wavelet.

Appendix A

The NDETOOL Toolbox

At the growth of the computer simulation carried out during the thesis research work, a set of different routines have been collected into a MATLAB-based toolbox, called *NDETOOL*. The toolbox contains four different categories :

1) Routines for some of the SOS-based deconvolution schemes reviewed in Chapter 2. Table A.1 gives summarizes the main SOS-based deconvolution techniques in the toolbox.

2) Routines for the HOS-based deconvolution methods presented in Chapter 4. It should be noted that the parametric methods are based on available routines in the Hi-spec toolbox [78]. Table A.2 summarizes the implemented HOS-based methods in the toolbox.

3) Routines for performing special transforms such as the Wigner-Ville distribution, and the parametric normalized bispectrum (bicoherence). The implemented transforms are shown in Table A.3.

4) Routines for performing automatic defects classification using a multilayer perceptron MLP, and WVD images compression using GHA algorithm. A summary is given in Table A.4.

The guided tour of the toolbox is implemented in the ndtdemo routine. The demo is based on a graphical user interface (GUI) which allows interaction between the toolbox and the user. Some of the interactive sessions are shown in Fig. A.1- A.5.
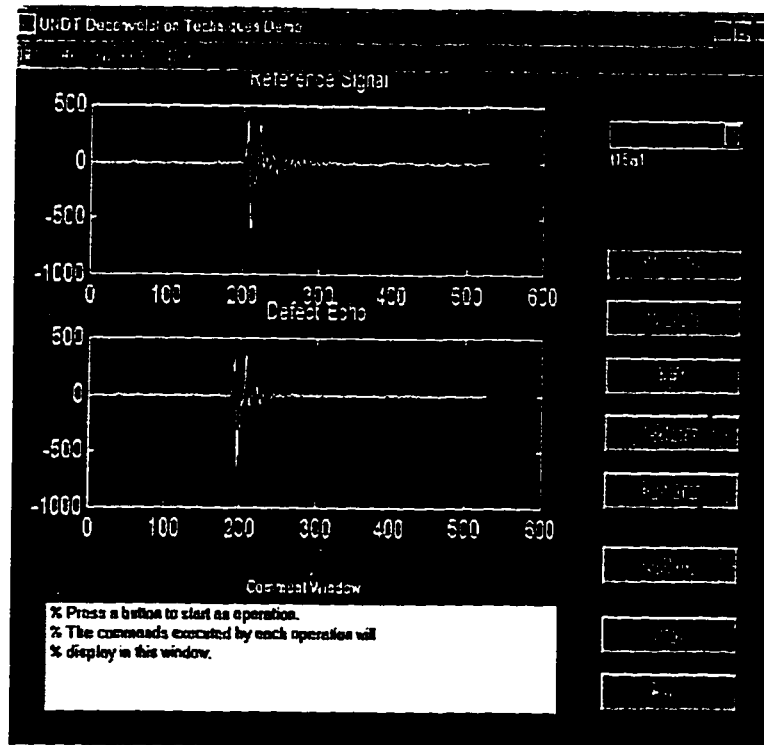
149

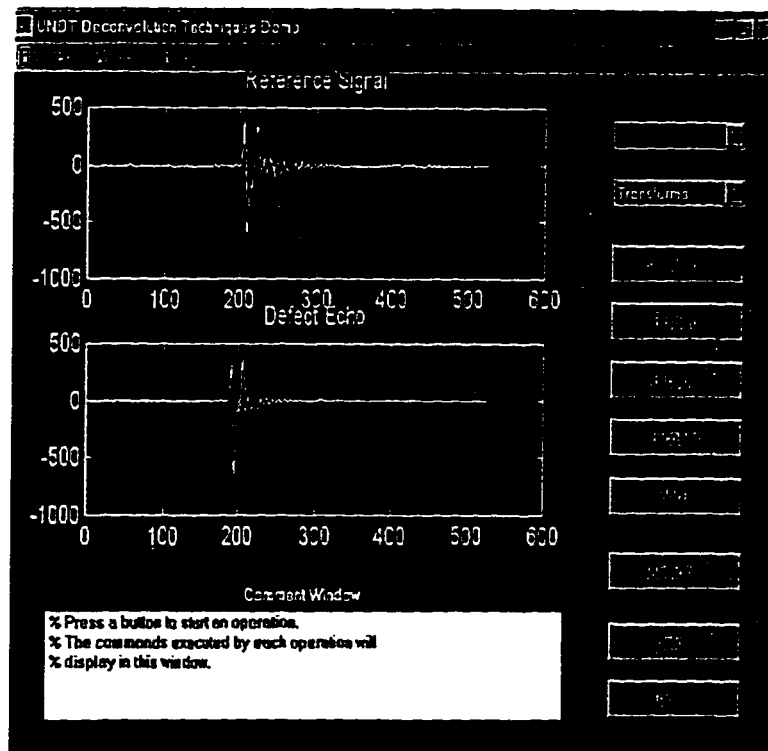Figure A.1: First Window of the Demo Routine.

Figure A.2: Second Window of the Demo Routine.
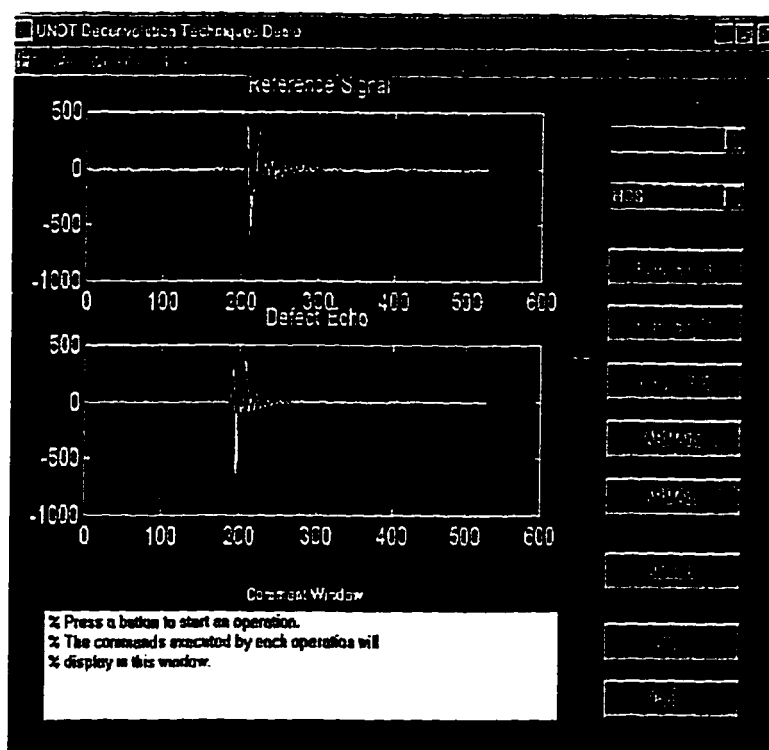
Figure A.3: Third Window of the Demo Routine.
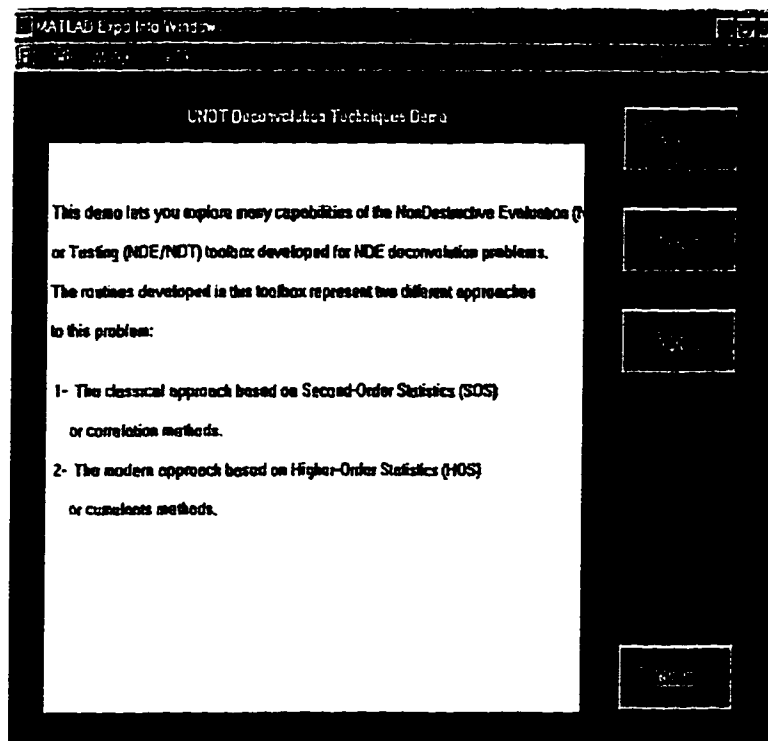
Figure A.4: Information Window of the Demo Routine.
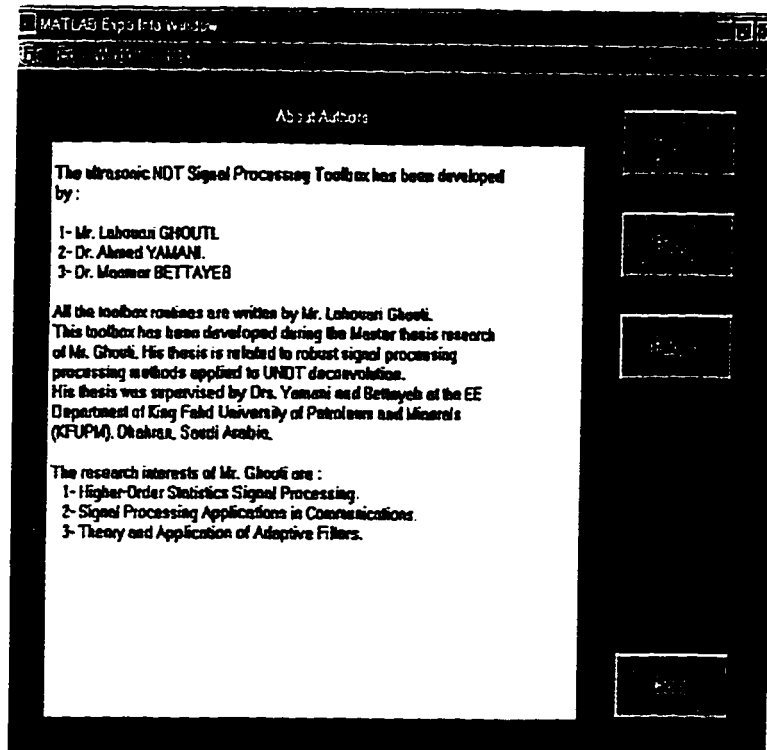
Figure A.5: Author Information Window of the Demo Routine.

SOS-Based Deconvolution Routines	
wiener	Both Wiener A and B Filters
xcorr	Correlation Method
ndtls	LS Method Using SVD Approach
cepsl	Complex Cepstrum Method

Table A.1: Implemented SOS-Based Deconvolution Routines.

HOS-Based Deconvolution Routines	
direbisp	Direct Bispectrum Method Using Both Approaches
indbisp	Indirect Bispectrum Method
bicep	Complex Bicepstrum Method
bispsim	Direct Bispectrum Simulation
inbsim	Indirect Bispectrum Simulation
bicsim	Complex Bicepstrum Simulation
noise	Generation of Different Noise Statistics

Table A.2: Implemented HOS-Based Deconvolution Methods.

Transforms	
wvd	Wigner-Ville Distribution
pwvd	Pseudo Wigner-Ville Distribution
bicoh	Normalized Bispectrum Using Parametric Approach

Table A.3: Implemented Special Transforms.

ANN	
PCA	WVD image compression using GHA Algorithm
matic	ANN for defect classification using a MLP

Table A.4: Implemented ANN Techniques.

Demo	
ndtdemo	Guided Tout to NDETOOL

Table A.5: GUI-Based Demo Routine

References

[1] A. Khan. Fourier descriptors approach for defect analysis in ultrasonic non-destructive evaluation. Master's thesis, EE Department KFUPM, June 1994.

[2] C. H. Chen. Geophysical and ultrasonic pattern recognition. In *11th IAPR: International Conference on Pattern Recognition*, Netherlands Congress Center, August 1992. IAPR, TU Delft and Erasmus University Rotterdam. Tutorial A2.

[3] S. K. Somuah. *Non-Destructive Testing of Materials by Ultrasound*. Electrical Engineering Dept., KFUPM. Dhahran 31261., 1990.

[4] L. Ljung. *System Identification: Theory for the User*. Prentice Hall, Englewood Cliffs, 1987.

[5] T. Soderstrom and P. Stoica. *System Identification*. Prentice Hall International, 1988.

[6] R. B. Thompson and D. O. Thompson. Ultrasonics in nondestructive evaluation. *IEEE Proceedings*, 73(12):1716–1755, December 1985.

[7] S. K. Sin and C. H. Chen. A comparison of deconvolution techniques for the ultrasonic nondestructive evaluation of materials. *IEEE Trans. On Image Processing*, 1(1):3–10, January 1992.

[8] C. H. Chen and S. K. Sin. High-resolution deconvolution techniques and their applications in ultrasonic nde. *International Journal of Imaging Systems and Technology*, 1:223–242, 1989.

[9] G. Hayward and J. E. Lewis. Comparison of some non-adaptive deconvolution techniques for resolution enhancement of ultrasonic data. *Ultrasonics*, 27:155–164, 1989.

[10] S. P. Neal. *A Prioir Knowledge Based Optimal Wiener Filtring Approach to Ultrasonic Scattering Amplitude Estimation*. PhD thesis, Iowa State University, Ames, Iowa, 1988.

[11] S. P. Neal. Private Communication, 1996.

[12] C. H. Chen and S. K. Sin. Applications of high-resolution deconvolution techniques to ultrasonic nde. In H. Lee and G. Wade, editors, *Acoustical Imaging*, pages 163–177. Plenum Press, New York, 1991. Vol. 18.

[13] A. V. Oppenheim and R. W. Schafer. *Discrete-Time Signal Processing*. Prentice Hall, Englewood Cliffs, 1989.

[14] L. Ljung and K. Glover. Frequency domain versus time domain methods in system identification. *Automatica*, 17(1):71–86, January 1981.

[15] S. Haykin. *Adaptive Filter Theory*. Prentice-Hall, Englewood Cliffs, 1991. Second Edition.

[16] A. T. Walden. Robust deconvolution by modified wiener filtering. *Geophysics*, 53(2):186–191, February 1988.

[17] S. M. Kay and S. L. Marple Jr. Spectrum analysis–a modern perspective. *IEEE Proceedings*, 69(11):1380–1419, November 1981.

[18] H. Akaike. A new look at the statistical model identification. *IEEE Trans. Autom. Control*, 19:716–723, March 1974.

[19] Y. T. Chan and J. C. Woods. A new order determination technique for arma processes. *IEEE Trans. On Acoustics, Speech, and Signal Processing*, 32(3):517–521, June 1984.

[20] J. P. Burg. Maximum entropy spectral analysis. In *37th Ann. Int. Meet. Soc. Explor. Geophys.*, Oklahoma City, Okla., 1967. Soc. Explor. Geophys.

[21] C. H. Chen. *Nonlinear Maximum Entropy Spectral Analysis Methods For Signal Recognition.* John Wiley and Sons, 1982.

[22] L. Ljung. *Matlab System Identification Toolbox.* The Mathworks Inc., 3 edition, 1993.

[23] J. M. Tribolet. A new phase unwrapping algorithm. *IEEE Trans. Acoustics, Speech, and Signal Processing,* 25(2):170–177, April 1977.

[24] L. Udpa and S. S. Udpa. Neural networks for the classification of nondestructive evaluation signals. *IEE Proceedings, Pt. F,* 138(1):41–45, February 1991.

[25] C. H. Chen. Pattern analysis for the ultrasonic nondestructive evaluation of materials. *International Journal of Pattern Recognition and Artificial Intelligence,* 1(2):251–260, 1987.

[26] R. P. Lippmann. An introduction to computing with neural nets. *IEEE Acoustics, Speech and Signal Processing Magazine,* 4(1):4–22, January 1987.

[27] C. L. Nikias and M. R. Raghuveer. Bispectrum estimation: A digital signal processing framework. *IEEE Proceedings,* 75(7):869–891, July 1987.

[28] M. R. Raghuveer and C. L. Nikias. Bispectrum estimation: A parametric approach. *IEEE Trans. On Acoustics, Speech, and Signal Processing,* 33(4):1213–1230, October 1985.

[29] J. M. Mendel. Tutorial on higher-order statistics (spectra) in signal processing and system theory: Theoretical results and some applications. *IEEE Proceedings*, 79(3):278–305, March 1991.

[30] C. L. Nikias and A. P. Petropulu. *Higher-Order Spectra Analysis : A Nonlinear Signal Processing Framework.* Prentice Hall, Englewood Cliffs, 1993.

[31] A. Papoulis. *Probability, Random Variables, and Stochastic Processes.* McGraw-Hill International Editions, 1991. 3rd Edition.

[32] D. H. Johnson and P. S. Rao. On the existence of gaussian noise. In *IEEE Int. Conference on Acoustics, Speech and Signal Processing*, pages 1673–1676, Toronto, 1991.

[33] C. L. Nikias and J. M. Mendel. Signal processing with higher-order spectra. *IEEE Signal Processing Magazine*, 10(7):10–37, July 1993.

[34] M. G. Kendall and A. Stuart. *The Advanced Theory of Statistics, Vol.1.* Charles Griffin and Co. Ltd., 1977. 4th Edition.

[35] J. M. Vidal. *Non-Gaussian ARMA Estimation Based on Higher-Order Statistics.* PhD thesis, Universitat Politecnica De Catalunya, Barcelona, Spain, 1992.

[36] A. Swami. *System Identification Using Cumulants.* PhD thesis, University of Southern California, Los Angeles, CA, 1988.

[37] D. R. Brillinger and M. Rosenblatt. Asymtotic theory of kth-order spectra. In B. Harris, editor, *Spectral Analysis of Time Series*, chapter 3, pages 153–188. John Wiley and Sons, 1967.

[38] D. R. Brillinger and M. Rosenblatt. Computation and interpretation of kth-order spectra. In B. Harris, editor, *Spectral Analysis of Time Series*, chapter 4, pages 189–232. John Wiley and Sons, 1967.

[39] L. Ghouti. M. Sc. Thesis Proposal, 1996.

[40] G. Giannakis and J. M. Mendel. Stochastic realization of non-minimum phase systems. In *Proceedings of American Control Conference*, pages 1254–1259, Seattle, WA., June 1986.

[41] M. H. Hayes, J. S. Lim, and A. V. Oppenheim. Signal reconstruction from phase or magnitude. *IEEE Trans. On Acoustics, Speech, and Signal Processing*, 28(6):672–680, December 1980.

[42] D. Hatzinakos and C. L. Nikias. Blind equalization using a tricepstrum based algorithm. *IEEE Trans. On Communications*, 39(5):669–682, May 1991.

[43] R. Pan and C. L. Nikias. The complex cepstrum of higher order cumulants and nonminimum phase system identification. *IEEE Trans. On Acoustics, Speech, and Signal Processing*, 36(2):186–205, February 1988.

[44] A. Polydoros and A. T. Fam. The differential cepstrum: Definition and properties. In *IEEE Int. Symposium on Circuits and Systems*, April 1981.

[45] D. E. Dudgeon. The computation of two-dimensional cepstra. *IEEE Trans. On Acoustics, Speech, and Signal Processing*, 25(6):476–484, December 1977.

[46] G. B. Giannakis. Cumulants : A powerful tool in signal processing. *IEEE Proceedings*, 75:1333–1334, 1987.

[47] M. R. Raghuveer. Third-order statistics : Issue of pdf symmetry. *IEEE Trans. On Signal Processing*, 43(7):1736–1738, July 1995.

[48] T. Matsuoka and T. J. Ulrych. Phase estimation using the bispectrum. *IEEE Proceedings*, 72(10):1403–1411, October 1984.

[49] M. Rangoussi and G. B. Giannakis. Fir modeling using log-bispectra: Weighted least-squares algorithms and performance analysis. *IEEE Trans. On Circuits and Systems*, 38(3):281–296, March 1991.

[50] S. J. Lim. *Two-Dimensional Signal And Image Processing*. Prentice Hall, Englewood Cliffs, 1990.

[51] A. M. Tekalp and A. T. Erdem. Higher-order spectrum factorization in one and two dimensions with applications in signal modeling and nonminimum phase system identification. *IEEE Trans. On Acoustics, Speech, and Signal Processing*, 37(10):1537–1549, October 1989.

[52] M. J. Hinich. Testing for gaussianity and linearity of a stationary time series. *Journal of Time Series Analysis*, 3(3):169–176, July 1982.

[53] T. S. Rao and M. M. Gabr. An introduction to bispectral analysis and bilinear time series models. In *Lecture Notes in Statistics*. Springer Verlag, 1984.

[54] C. H. Chen. Pattern analysis for the ultrasonic nondestructive evaluation of materials. *International Journal Of Pattern Recognition and Artificial Intelligence*, 1(2):251–260, 1987.

[55] S. Haykin and T. K. Bhattacharya. Modular learning strategy for signal detection in a nonstationary environment. *IEEE Trans. On Signal Processing*, 1997. Accepted for publication.

[56] S. Qian and D. Chen. *Joint Time-Frequency Analysis : Methods and Applications*. Prentice Hall, Englewood Cliffs, 1996.

[57] L. Cohen. Time-frequency distribution : A review. *IEEE Proceedings*, 77(7):941–981, July 1989.

[58] B. Boashash. Time-frequency signal analysis. In S. Haykin, editor, *Advances in Spectrum Analysis and Array Processing*, chapter 9, pages 418–517. Prentice-Hall, Englewood Cliffs, 1991.

[59] R. C. Gonzalez and P. Wintz. *Digital Image Processing*. Addison Wesley Pubublishing Company, 1987. 2nd Edition.

[60] D. C. Clay. *Linear Algebra and its Applications.* Addison Wesley Pububblishing Company, 1993.

[61] E. Oja. A simplified neuron model as a principal component analyzer. *Journal of Mathematical Biology,* 15:267–273, 1982.

[62] T. D. Sanger. Optimal unsupervised learning in a single-layer feedforward neural network. *Neural Networks,* 12:459–473, 1989.

[63] S. Y. Kung and K. I. Diamantaras. A neural network learning algorithm for adaptive principal component extraction (apex). In *IEEE Int. Conference on Acoustics, Speech and Signal Processing,* pages 861–864, 1990.

[64] S. Haykin. *Neural Networks: A Comprehensive Foundation.* Mcmillan College Publishing Company, 1994.

[65] A. K. Nandi, D. Mampel, and B. Rosher. Comparative study of deconvolution algorithms with applications in non-destructive testing. In *IEE Colloquim on Blind Deconvolution - Algorithms and Applications,* September 1995.

[66] T. D. Sanger. Private Communication, 1997.

[67] M. P. Ekstrom. A spectral characterization of the ill-conditionning in numerical deconvolution. *IEEE Trans. On Audio Electroacoustics,* 21(4):344–348, 1973.

[68] P. K. Bhagat and K. D. Shimmin. Homomorphic processing in ultrasonic nde. In D. O. Thompson and D. E. Chimenti, editors, *Review Of Progress in Quantitative NDE*, pages 351-358. Plenum Press, New York, 1984. Vol. 3.

[69] S. Elgar and R. T. Guza. Statistics of bicoherence. *IEEE Trans. On Acoustics, Speech, and Signal Processing*, 36(10):1667-1668, October 1988.

[70] V. Chandran, S. Elgar, and B. Van Hof. Statistics of tricoherence. *IEEE Trans. On Signal Processing*, 42(12):3430-3440, December 1994.

[71] P. S. Naidu. *Modern Spectrum Analysis of Time Series*. CRC Press, Boca Raton, 1996.

[72] A. G. Bessios and C. L. Nikias. Fft-based bispectrum computation on polar rasters. *IEEE Trans. On Signal Processing*, 39(11):2535-2539, November 1991.

[73] J. K. Tugnait. Testing for linearity of noisy stationary signals. *IEEE Trans. On Signal Processing*, 42(10):2742-2748, October 1994.

[74] K. I. Kim and E. J. Powers. A digital method for modelling quadratically non-linear systems with a general random input. *IEEE Trans. On Acoustics, Speech, and Signal Processing*, 36(11):1758-1769, November 1988.

[75] G. Strang and T. Nguyen. *Wavelets and Filter Banks*. Wellesley-Cambridge Press, 1996.

[76] D. Donoho. De-noising by soft thresholding. *IEEE Trans. On Info. Theory*, 41(3):613–627, May 1995.

[77] A. Abbate, J. Koay, J. Frankel, S. C. Shroeder, and P. Das. Signal detection and noise suppression using a wavelet transform signal procesor: Application to ultrasonuic flaw detection. *IEEE Trans. On Ultra., Ferro., and Freq. Control*, 44(1):14–25, January 1997.

[78] A. Swami, J. L. Mendel, and C. L. Nikias. *Matlab Hispec Toolbox*. The Mathworks Inc., 1 edition, 1993.

[79] C. L. Nikias and M. Shao. *Alpha-Stable Signal Processing*. John Wiley and Sons, 1995.

[80] L. Fausett. *Fundamentals of Neural Networks: Architecture, Algorithms and Applications*. Prentice-Hall International Inc., 1994.

[81] G. B. Giannakis. Cumulants : A powerful tool in signal processing. *IEEE Proceedings*, 75:1333–1334, 1987.

[82] F. Hlawatsch and G. F. Boudreaux-Bartels. Linear and quadratic time-frequency signal representations. *IEEE Signal Processing Magazine*, 9(4):21–67, April 1992.

[83] J. K. Tugnait. Counterexamples to : On estimatimating non-causal non-minimum phase arma models of non-gaussian processes. *IEEE Trans. On Signal Processing*, 40(4):1011–1013, April 1992.

[84] T. Olofsson and T. Stepinski. Deconvolution of ultrasound images in nondestructive material evaluation. Paper Under Review, 1997.

[85] I. Jouny and R. L. Moses. The bispectrum of complex signals: Definitions and properties. *IEEE Trans. On Signal Processing*, 40(11):2833–2836, November 1992.

[86] J. K. Tugnait. Testing for linearity of noisy stationary signals. *IEEE Trans. On Signal Processing*, 42(10):2742–2748, October 1994.

[87] A. Benveniste, M. Goursat, and G. Ruget. Robust identification of a nonminimum phase system : Blind adjustment of a linear equalizer in data communications. *IEEE Trans. On Automatic Control*, 25(6):385–399, June 1980.

[88] G. B. Giannakis and J. M. Mendel. Cumulant-based order determination of non-gaussian arma models. *IEEE Trans. On Acoustics, Speech, and Signal Processing*, 38(8):1411–1423, August 1990.

[89] M. R. Raghuveer annd S. A. Dianat. On the existence of autoregressive models for third-order cumulant matching. *IEEE Trans. On Acoustics, Speech, and Signal Processing*, 37(12):1931–1938, December 1989.

[90] F. J. Harris. On the use of windows for harmonic analysis with the discrete fourier transform. *IEEE Proceedings*, 66(1):51–83, January 1978.

[91] M. Shao and C. L. Nikias. Signal processing with alpha-stable processes. *IEEE Proceedings*, 75(7):869–891, July 1995.

[92] C. L. Nikias and M. R. Raghuveer. Discussion on: Higher order spectra by maximum entropy method. *Geophysics*, 50(1):165–166, January 1985.

[93] R. O. Plaisted and G. H. Pena. Short note: Higher order spectra by maximum entropy method. *Geophysics*, 48(10):1409–1410, October 1983.

[94] D. R. Hush and B. G. Horne. Progress in supervised neural networks: What's new since lippmann? *IEEE Signal Processing Magazine*, 10(1):8–39, January 1993.

[95] R. Yarlagada, J. B. Bednar, and T. L. Watt. Fast algorithms for lp deconvolution. *IEEE Trans. On Acoustics, Speech, and Signal Processing*, 33(1), January 1985.

[96] L. Ljung. Consistency of the least-squares identification method. *IEEE Trans. On Automatic Control*, 21:779–781, 1976.

[97] R. Pan and C. L. Nikias. Phase reconstruction in the trispectrum domain. *IEEE Trans. On Acoustics, Speech, and Signal Processing*, 35(6):895–897, June 1987.

[98] A. P. Petropulu and C. L. Nikias. The complex cepstrum and bicepstrum: Analytic performance evaluation in the presence of gaussian noise. *IEEE Trans. On Acoustics, Speech, and Signal Processing*, 38(7):1246–1256, July 1990.

[99] A. P. Petropulu and C. L. Nikias. Signal reconstruction from the phase of the bispectrum. *IEEE Trans. On Signal Processing*, 40(3):601–610, March 1992.

[100] J. F. Chen, J. A. Zagzebski, and E. L. Madsen. Non-gaussian versus non-rayleigh statistical properties of ultrasound echo signals. *IEEE Trans. On Ultra., Ferro., and Freq. Control*, 41(4):435–440, July 1994.

[101] Y. Zhu and J. P. Weight. Ultrasonic nondestructive evaluation of highly scattering materials using adaptive filtering and detection. *IEEE Trans. On Ultra., Ferro., and Freq. Control*, 41(1):26–33, January 1994.

[102] H. Bartelt, A. W. Lohmann, and B. Wirnitzer. Phase and amplitude recovery from bispectra. *Applied Optics*, 23(18):3121–3129, September 1984.

[103] P. J. Huber, B. Kleiner, T. Gasser, and G. Dumermuth. Statistical methods for investigating phase relations in stationary stochastic processes. *IEEE Tran. On Audio Electroacoustics*, 19:78–86, 1971.

[104] J. C. Hassab and R. Boucher. A probabilistic analysis of time delay extraction by the cepstrum in stationary gaussian noise. *IEEE Trans. On Info. Theory*, 22:444-454, 1976.

[105] L. A. Pflug, G. E. Ioup, J. W. Ioup, and R. F. Field. Properties of higher-order correlations and spectra for bandlimited, deterministic transients. *J. Acoust. Soc. Am.*, 91(2):975-988, February 1992.

[106] M. J. Hinich and H. Messer. On the principal domain of the discrete bispectrum of a stationary signal. *IEEE Trans. On Signal Processing*, 43(9):2130-2134, September 1995.

[107] D. R. Brillinger. The identification of a particular nonlinear time series system. *Biometrika*, 65:509-515, 1977.

[108] G. B. Giannakis. *Signal Processing Using Higher-Order Statistics*. PhD thesis, University of Southern California, Los Angeles, CA, 1987.

[109] T. Y. Naffouri. M. Sc. Thesis Proposal, 1996.

[110] B. Porat and B. Friedlander. Performance analysis of parameter estimation based on higher-order moments. *Int. J. Adaptive Control and Signal Processing*, 3:191-229, 1989.

[111] C. Nadeu and E. Lleida. On the cramer-rao bound of finite length cepstrum spectral estimators. *Electronic Letters*, 26(14):987-988, 1990.

[112] S. P Neal and D. O. Thompson. The measurement and analysis of acoustic noise as a random noise. In D. O. Thompson and D. E. Chimenti, editors, *Review Of Progress in Quantitative NDE*, pages 625–632. Plenum Press, New York, 1990. Vol. 9A.

[113] J. Le Roux. Reconstruction of a sampled signal fourier transform from its bispectrum. In J. L. Lacoume, editor, *Higher Order Statistics : Int. Sig. Proc. Workshop*, pages 323–326. Elsevier, Amsterdam, July 1991.

[114] M. R. Raghuveer and C. L. Nikias. Bispectrum estimation for short data length. In *IEEE Int. Conference on Acoustics, Speech and Signal Processing*, pages 1352–1355, 1985.

[115] C. K. An, S. B. Kim, and E. J. Powers. Optimized parametric bispectrum estimation. In *IEEE Int. Conference on Acoustics, Speech and Signal Processing*, pages 2392–2395, 1988.

[116] C. K. An, S. B. Kim, and E. J. Powers. Optimized parametric bispectrum estimation. In *IEEE Int. Conference on Acoustics, Speech and Signal Processing*, pages 2392–2395, 1988.

[117] Y Zhang, D. Hatzinakos, and A. N. Venetsanopoulos. Bootstrapping techniques in the estimation of higher-order cumulants from short data records.

In *IEEE Int. Conference on Acoustics, Speech and Signal Processing*, pages IV.200–IV.203, 1993. Vol. 4.

Vita

- LAHOUARI GHOUTI

- Born in Algeria.

- Permanent Address :

  27 Bd Mellah Ali. Oran 31026, ALGERIA.

- Received Bachelor of Engineering (**B.E.**) degree in Communications Engineering from Algerian National Telecommunications Institute, Oran, Algeria in June 1992.

- Joined KFUPM in April 1994 as a Research Assistant and enrolled into the Master of Science (**M.S.**) program in Electrical Engineering, with specialization in Communications and Digital Signal Processing Applications with interests in Higher-Order Statistics applied to Ultrasonic Nondestructive Evaluation (NDE) processing and software development for ultrasonic NDE processing.