7-21-2005

Wavelet-Based Monitoring and Analysis of Cardiorespiratory Response to Hypoxia

Vuslat Nazilli
University of South Florida

Follow this and additional works at: https://scholarcommons.usf.edu/etd
Part of the American Studies Commons

Scholar Commons Citation
Nazilli, Vuslat, "Wavelet-Based Monitoring and Analysis of Cardiorespiratory Response to Hypoxia" (2005). Graduate Theses and Dissertations.
https://scholarcommons.usf.edu/etd/789

This Thesis is brought to you for free and open access by the Graduate School at Scholar Commons. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Scholar Commons. For more information, please contact scholarcommons@usf.edu.
Wavelet-Based Monitoring and Analysis of Cardiorespiratory Response to Hypoxia

by

Vuslat Nazilli

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Industrial Engineering
Department of Industrial and Management Systems Engineering
College of Engineering
University of South Florida

Major Professor: José L. Zayas-Castro, Ph.D.
Tapas K. Das, PhD
Kendall F. Morris, Ph.D.
A.N.V. Rao, Ph.D

Date of Approval:
July 21, 2005

Keywords: principal curve, intermittent hypoxia, multichannel signal monitoring, factor analysis, wavelet transformation

© Copyright 2005, Vuslat Nazilli
DEDICATION

To My Family
ACKNOWLEDGEMENTS

I would like to thank my advisor Dr. José L. Zayas-Castro for the support and mentoring he has generously provided since the time I started my graduate studies at USF. Dr. Zayas-Castro has always kept my best interests in mind, and for that I sincerely thank him.

I also want to thank Dr. Qiang Huang for introducing me to this interesting research topic and allowing me to work on his method. I thank Dr. Tapas K. Das, Dr. A.N.V. Rao and Dr. Kendall F. Morris for serving in my committee and for their valuable time. I sincerely appreciate Dr. Morris’ support by providing me the data for my thesis and explaining the physical mechanisms related to my research.

Last but not the least, I would like to thank my family and friends for their never-ending love and support through all the struggles.
TABLE OF CONTENTS

LIST OF FIGURES ................................................................................................................ ii
LIST OF TABLES................................................................................................................ iv
ABSTRACT ........................................................................................................................ v

CHAPTER 1. INTRODUCTION .............................................................................................. 1
  1.1 A Physiological Application .................................................................................. 2
  1.1.1 Effects of Hypoxia ........................................................................................ 3
  1.1.2 Description of Data ....................................................................................... 5
  1.2 Thesis Organization ............................................................................................... 6

CHAPTER 2. LITERATURE REVIEW ......................................................................................... 8
  2.1 Research on the Analysis of Interaction between Biological Systems .............. 8
    2.1.1 Time-domain Methods .................................................................................. 9
    2.1.2 Frequency-domain Methods ...................................................................... 10
    2.1.3 Phase Locking ............................................................................................ 10
  2.2 Wavelet Analysis ........................................................................................................ 13
    2.2.1 Basics of Multiresolution Analysis (MRA) and Wavelet Transformation . 14
  2.3 Principal Curves .......................................................................................................... 19

CHAPTER 3. WAVELET TRANSFORMATIONS OF CARDIORESPIRATORY SIGNALS
            AND PRE-ANALYSIS ....................................................................................... 23
  3.1 Extraction of Detail and Approximation Coefficients .............................................. 23
  3.2 Analysis of Energy Redistributions in Wavelet Coefficients ..................................... 29

CHAPTER 4. DETECTION OF MULTISCALE CHANGES IN MULTICHANNEL DATA BY
            PRINCIPAL CURVE REGRESSION TEST (PCuRTest).................................... 34
  4.1 PCuRTest Algorithm ................................................................................................. 34
  4.2 Detection of Changes in Wavelet Coefficients by PCuRTest .................................... 37

CHAPTER 5. CONCLUSIONS AND FUTURE RESEARCH ..................................................... 42
  5.1 Conclusions ............................................................................................................ 42
  5.2 Future Research ....................................................................................................... 43
    5.2.1 Factor Analysis ........................................................................................... 43
    5.2.2 Multiple Battery Factor Analytic Model of Common Variation Patterns... 48

REFERENCES ..................................................................................................................... 52
LIST OF FIGURES

Figure 1.1 Schematic Representation of the Procedure Followed in this Thesis.........................2
Figure 1.2 Schematic Representation of Effects of OSA (*OSA: Obstructive Sleep Apnea, §LV: Left Ventricular, †BP: Blood Pressure, ‡HR: Heart Rate)...........................................5
Figure 1.3 1-min. Phrenic Nerve, Lumbar Nerve and Blood Pressure Signals Recorded Before Hypoxic Stimuli ............................................................................................................. 6
Figure 2.1 A Damped Sinusoidal Signal and Its Representation on Unit Circle…………………………...12
Figure 2.2 Some Wavelet Families ........................................................................................................ 13
Figure 2.3 Three-level Wavelet Decomposition Tree ........................................................................ 16
Figure 2.4 DWT of a Digital Signal .................................................................................................... 16
Figure 2.5 A Noisy Signal and the Reconstructed Signal After Denoising ........................................ 18
Figure 2.6 A Sample Signal and the Reconstructed Form After Hard- and Soft-thresholding ...... 19
Figure 2.7 A Sample Principal Curve by HS Definition ...................................................................... 20
Figure 3.1 Level-5 Decomposition of a Before-Hypoxia Phrenic Nerve Signal ............................ 24
Figure 3.2 Level-5 Decomposition of a Before-Hypoxia Lumbar Nerve Signal ............................. 25
Figure 3.3 Level-5 Decomposition of a Before-Hypoxia Blood Pressure Signal ............................ 26
Figure 3.4 Original and Thresholded Coefficients and Original and Reconstructed Signals for a Phrenic Nerve Signal .......................................................................................... 27
Figure 3.5 Original (red) and Reconstructed (blue) Forms of Lumbar Nerve and Blood Pressure Signal Samples ........................................................................................................ 27
Figure 3.6 Selection of Wavelet Coefficients ................................................................................... 28
Figure 3.7 Relative Energy Distributions at Different Scales of a Phrenic Nerve Signal ............ 30
Figure 3.8 Relative Energy Distributions at Different Scales of a Lumbar Nerve Signal ............ 31
Figure 3.9 Relative Energy Distributions at Different Scales of a Blood Pressure Signal ........... 32
Figure 3.10 Total Energy Distributions for Phrenic Nerve, Lumbar Nerve, and Blood Pressure Signals Respectively (Dashed Line Indicates the Hypoxic Period) ....................... 32
Figure 4.1 Principal Curve Regression .......................................................................................... 34
Figure 4.2 PCuRTtest Results for Some Samples from the Validation Dataset .......................... 38
Figure 4.3 PCuRTtest Results for Some Samples from the Validation Dataset ......................... 39
Figure 4.4 PCuRTtest Results for Some Samples from the Test Dataset .................................. 40
Figure 4.5 PCuRTtest Results for Some Samples from the Test Dataset........................................41
Figure 4.6 Relative Efficiency Graph for Selected Levels ............................................................ 41
Figure 5.1 Plots of Unrotated and Varimax-rotated Factor Loadings for a Sample Dataset ..........47
LIST OF TABLES

Table 3.1 Selected Wavelet Coefficients.................................................................28
Table 3.2 Mapping between Levels and Frequency/Period Values for Daubechies-4 Wavelet….33
WAVELET-BASED MONITORING AND ANALYSIS OF CARDIORESPIRATORY RESPONSE TO HYPOXIA

Vuslat Nazilli

ABSTRACT

Obstructive sleep apnea is a potentially life-threatening condition characterized by repetitive episodes of upper airway obstruction that occur during sleep, usually associated with a reduction in blood oxygen saturation. In US population, 9% of women, 24% of men, and 2% of children have been diagnosed with obstructive sleep apnea, suggesting that 18 million people may suffer from the consequences of nightly episodes of apnea. One of the most significant symptoms of obstructive sleep apnea is profound and repeated hypoxia. The analysis of the interaction between cardiovascular and respiratory signals has been a widely-explored area of research due to the significance of the results in describing a functional relationship between the underlying physiologic systems; however, statistical and analytical approaches to analyze the changes in these signals before and after hypoxia are still in their early stages of evolution. A major motivation for this research has been the lack of methodologies to detect mean and/or variance shifts and identify root sources of variation in time-frequency characteristics of multichannel data.

The contributions of this thesis are twofold. First, multiscale energy distributions based on wavelet transformations of the analyzed physiological signals are analyzed. This is followed by the development of an online multichannel monitoring approach based on principal curves that detects changes in the wavelet coefficients extracted from the analyzed signals.
CHAPTER 1. INTRODUCTION

Today’s advanced multielectrode array technologies allow researchers to obtain simultaneously large amounts of multichannel signal data. It is known that physiological signals have multiscale properties and generally represent processes with different localizations in time and frequency. External factors or faults may cause certain changes such as mean shifts and/or variance shifts in these signals simultaneously, but affect the signals at different scale levels. This type of data contains important but unobservable diagnostic information about the underlying mechanisms or sources of variation that contribute to process variability resulting in distinct variation patterns in the data. The effects of these variation sources can be categorized as spatial, describing the interaction of different signals and temporal, indicating the evolution of a variation source through time. The identification or diagnosis of these underlying mechanisms can only be accomplished through an interpretable separation of different variation patterns. Wavelet decomposition has proven to be an effective approach to obtain time-frequency localized information of both stationary and nonstationary signals. This thesis aims to extract this advantage of wavelet analysis for the detection of spatio-temporal changes by integrating the multiscale information from multiple signals and to identify the variation sources that cause these changes. The analysis procedure to be followed throughout this thesis is shown in Figure 1.1.

Section 1.1 includes a brief explanation of the analyzed problem and the motivation for undertaking this study. Section 1.2 explains the importance of exploring the potential effects of hypoxia on the malfunction of cardiovascular and respiratory systems and gives a short introduction on the characteristics of hypoxic effects. Section 1.3 gives a description of the
dataset to be used in the further analysis steps. Computations through the whole document have been performed in MATLAB and R.

Figure 1.1 Schematic Representation of the Procedure Followed in this Thesis

1.1 A Physiological Application

It is well known that cardiovascular and respiratory systems do not act independently although the nature of this interdependence is not completely understood. Most of the previous studies related to the analysis of hypoxic effects on these systems have been constrained to the analysis of these systems individually by using standard statistical methods. The studies related to the analysis of the interaction of these systems, on the other hand, do not include an online detection and exploration of changes in the relation between these systems subsequent to hypoxic conditions. The goal of this research is to detect and identify root causes of changes in the mutual activity of these two vital systems by integrating their responses to stimulation of hypoxia.
occurring at different scales. The next sub-section describes hypoxia and the different types of effects that it is known or hypothesized to incur on cardiorespiratory systems.

1.1.1 Effects of Hypoxia

Sleep-related breathing disorders represent a variety of abnormalities ranging from simple snoring to obstructive sleep apnea. Obstructive sleep apnea (OSA) is a potentially life-threatening condition characterized by repetitive episodes of upper airway obstruction that occur during sleep, usually associated with a reduction in blood oxygen saturation. The acute hemodynamic alterations of OSA include systemic and pulmonary hypertension, increased right and left ventricular afterload, and increased cardiac output. Figure 1.2 illustrates potential effects of OSA on cardiorespiratory mechanisms. The risks of undiagnosed obstructive sleep apnea include heart attacks, strokes, impotence, irregular heartbeat, pulmonary and systemic hypertension and heart diseases. One of the most significant symptoms of obstructive sleep apnea is profound and repeated hypoxia during sleep.

The ventilatory and cardiovascular responses during or following hypoxia are highly complex in mammals and are hypothesized to result from an interplay between several time-dependent facilitatory and inhibitory mechanisms. Some characteristics of mechanisms contributing to responses in these systems during or following hypoxia can be differentiated according to the following criteria:

i. Pattern, intensity and duration of the hypoxic exposure that triggers them

ii. Time domain of their effects (milliseconds to years)

iii. The direction of their effects (facilitatory vs. inhibitory)

iv. Type of stimulus (hypoxia vs. carotid sinus nerve (CSN) stimulation)

v. The subject model (species of the test animal, state of consciousness, etc.)
Figure 1.2 Schematic Representation of Effects of OSA (*OSA: Obstructive Sleep Apnea,

LV: Left Ventricular, BP: Blood Pressure, HR: Heart Rate)

To be specific, three types of ventilatory responses have been identified in mammals after brief (seconds to minutes) hypoxic exposures based on these criteria [2]: acute response (AR), short term potentiation (STP) and short term depression (STD). AR corresponds to the immediate increase and decrease in respiratory activity at the onset and completion of hypoxic stimulus respectively. STP is the further increase in respiratory activity characterized by an increase in amplitude and frequency of ventilation for a period of seconds to minutes and is a condition mostly observed in experiments with electrical stimulation of the CSN of cats, dogs and rats. STD, usually lasting from many seconds to a few minutes, describes a transient decrease in the frequency of phrenic nerve response at the termination of hypoxic stimulus and has been observed only in anaesthetized rats [3].

Responses observed in phrenic nerve activity during or after repeated episodes of hypoxic stimuli can be categorized into two types as progressive augmentation (PA) and long term
facilitation (LTF). PA represents an increase in the amplitude of hypoxic ventilatory response and has been observed in few cases including the inspiratory activity of anaesthetized cats. LTF refers to a progressive increase, lasting from many minutes to several hours, in respiratory output during non-stimulated intervals between successive hypoxic stimuli [5]. It is observed that LTF is induced by repeated hypoxia, chemical stimulation of carotid chemoreceptors and electrical stimulation of the carotid sinus nerve or brain stem midline but not by hypercapnia. The mechanisms behind LTF are not completely revealed yet; however it may have a significant role in maintaining stable respiration during sleep.

Some of the respiratory mechanisms identified in response to a hypoxic stimulus of several minutes to years are hypoxic ventilatory decline (HVD), ventilatory acclimatization to hypoxia (VAH), ventilatory deacclimatization from hypoxia (VDH) and hypoxic desensitization (HD).

The effects of intermittent hypoxia on the cardiovascular system are far less clear than those of continuous hypoxia. Recent studies support the hypothesis that repetitive airway obstruction can cause systemic hypertension [6]. A review on the beneficial and adverse effects of intermittent hypoxia on the blood pressure is available in [7].

A description of the physiological data is presented next.

### 1.1.2 Description of the Data

The data have been provided by the Department of Physiology at University of South Florida, College of Medicine. The dataset consists of two samples of phrenic nerve, lumbar nerve, and arterial blood pressure signals recorded before and after hypoxic stimulus generated by repeated carotid chemoreceptor stimulation in a vagotomized, artificially ventilated adult cat. The signals were amplified and digitized at 40 Hz. frequency. Phrenic and lumbar nerve discharges were integrated with a leaky resistance-capacitance circuit to obtain a moving time average of activity in the nerves. Detailed information related to the procedures and methods is available in [5].
Figures 1.2 and 1.3 display 1-min length integrated phrenic nerve signal, 1-min lumbar nerve signal and 1-min. blood pressure signal respectively.

![Graphs of Phrenic Nerve, Lumbar Nerve, and Blood Pressure Signals](image)

Figure 1.3 1-min. Phrenic Nerve, Lumbar Nerve and Blood Pressure Signals Recorded Before Hypoxic Stimuli

### 1.2 Thesis Organization

This thesis is organized as follows. Chapter 2 presents a review of the methodologies that has been proposed for the analyzed problem and introduces to the components of methodologies proposed in this thesis. Chapter 3 explains the procedure followed in extracting wavelet coefficients from phrenic nerve and blood pressure signals and presents a multiscale energy analysis of physiologic signals in response to hypoxia. Chapter 4 introduces the Principal Curve
Regression Test (PCuR Test) as a tool to detect changes in the wavelet coefficients obtained from phrenic nerve and blood pressure signals and analyzes the test results. Chapter 5 discussed future research directions and presents an approach based on Multiple Battery Factor Analysis for diagnosis of common variation patterns in multichannel data.
CHAPTER 2. LITERATURE REVIEW

Previous research on the analysis of interaction between respiratory and cardiovascular mechanisms has been performed under the domain of detection of synchronization between these two systems. This chapter includes a brief literature review on the synchronization studies for neurophysiologic systems based on the domain in which the systems have been analyzed. This is followed by a brief explanation of wavelet analysis with applications in process monitoring. The following sections describe principal curves and factor analysis methods which are to be used in the methodologies introduced in chapters 3 and 4 respectively.

2.1 Research on the Analysis of Interactions between Biological Systems

Most of the previous research on the analysis of the inter-relationships between biological systems has been performed under the field of synchronization analysis. Synchronization has been a well-studied concept since its discovery by C.H. Huygens [8] and has been used in the modeling of many biological systems. Detection of synchrony between pairs of oscillatory signals has been a problem of interest in many neuronal activities. Most of these studies are in the form of correlation detection between two neuronal signals in order to find a functional relationship between them.

One of the most interesting and important problems in this area has been the interaction of cardiovascular and respiratory systems. It is known for more than a century that heart rate is modulated by a respiratory related rhythm and this phenomenon is referred to as Respiratory Sinus Arrhythmia (RSA) [9]. Mechanisms resulting in RSA are found to be the modulation of baroreflex inputs to the autonomic neurones within the brainstem by ventilation, feedback from
arterial baroreceptors and a central rhythm generator in the brainstem [10]. Recent work suggests that RSA may improve energy efficiency by matching perfusion to ventilation within each respiratory cycle [11].

Another interaction between heart beats and respiratory activity is cardioventilatory coupling (CVC), which can be defined as mutual time conformity or synchronization of respiratory-cycle phase with cardiac-cycle phase observed in objects under low arousal (i.e. sleep, anaesthesia, sedation). Previous studies have suggested CVC as the triggering of inspiratory timing after a constant interval following a heart beat [13].

The existing approaches in the analysis of temporal correlations between neural signals can be categorized in three groups, based on the analysis of the time domain, the frequency domain or nonlinear dynamics of phase synchronizations. The traditional idea in all categories has been to obtain a correlation coefficient using either special marker events (e.g. times of R peaks in ECG signals) or the original raw data and evaluate its statistical significance. In the following sections, the synchronization studies on the interaction of cardiovascular and respiratory systems will be reviewed.

2.1.1 Time-domain Methods

A standard method used in time domain analysis when data consists of spike trains is cross-correlogram or its shuffle-corrected version [14]; however it has been observed that interpretation of the peaks in a cross-correlogram is quite ambiguous due to the possibility of various factors causing those peaks [15]. Lindsey et.al. [16] and Arata et.al [17] used gravity method in order to identify correlations between neuronal assemblies and visualize the fluctuations in the baroresponsive neuronal assemblies during the respiratory cycle and baroreceptor stimulation but were not able to make inferences about the strength of synchrony and duration of correlated activity in these studies. Orem and Dick [18] proposed the use of an index , \( \eta^2 \), the ratio of
variance occurring across fractions of the activity of a respiratory neuron to the total variance, to quantify the degree of respiratory-modulated activity in that neuron and also the consistency of this modulation from breath to breath. Dick and Morris [19] modified this index to assess the strength and consistency of arterial pulse modulation of respiratory-modulated neural activity.

2.1.2 Frequency-domain Methods

A common method used in frequency domain analysis is coherence spectrum. In [20, 21], heart rate fluctuations were analyzed by power spectrum in order to assess the functioning of short-term cardiovascular control systems. Galletly et.al. [22] explored the relationship between CVC and simple spectral and nonlinear measures (i.e. fractal dimension and approximate entropy) of HRV (Heart Rate Variability) motivated by the idea that CVC should also have an effect on the pattern of HRV as well as inspiratory timing and concluded that these methods were ineffective in revealing the presence of CVC in heart rate time series.

Another set of tools to analyze interactions between two (or more) signals in frequency domain has been spectral open-loop and closed-loop models [23]. Open-loop models simply define a linear input/output relationship between signals through a single transfer function in frequency domain. Closed-loop models have been developed in frequency domain to model the feed-forward mechanisms of cardiovascular control [24, 25].

2.1.3 Phase Locking

The third group of approaches is aimed at revealing the nonlinear dynamical interdependence between coupled neural signals, which cannot be described by methods based on linear correlations like cross-correlation. The studies under this domain brought about a new concept of phase locking, defined in general as

\[ |n\phi_2(t) - m\phi_1(t)| < \text{constant} \]  (2.1)
where \( \phi_{1,2} \) denote the phase series of the analyzed signals and integers \( n \) and \( m \) describe the orders of the synchronous states (i.e. the ratio of the frequencies of the signals is \( n:m \)). The values of \( n \) and \( m \) are usually determined by trial and error, using relative frequency plots.

A number of nonlinear statistical measures like mutual information [26, 27], maximal correlation, phase difference entropy [28] or phase coherence [29] have been suggested in phase synchronization analysis of neural signals. The motivation behind using these measures in synchronization analysis is the fact that, unlike conventional statistical approaches in time or frequency domains, they discard the effect of amplitude correlations in order to detect correlations in the timing only. In another neurophysiologic problem, Schiff et.al. [30] proposed the use of mutual prediction in order to verify the existence of nonlinear coupling in a neural ensemble.

A general problem encountered in synchronization analysis of neurophysiologic systems is that the temporal correlation between signals from such systems is usually disturbed by the nonstationarity in the interacting systems and/or coupling after some time, which conflicts with the underlying stationarity assumption in standard statistical methods used in this analysis. The possibility of such nonstationarities imposes a difficulty on the selection of an optimal analysis period length as well as the conclusions related to the synchronization between these systems. Although the analysis of more oscillation cycles would result in a more reliable analysis under stationarity conditions, in the case of nonstationarity, longer observations will avoid the detection of weaker or shorter interactions. One common method to deal with this situation is using sliding windows in the estimation of the correlation coefficient. Schafer et.al. [31] and Hurtado et.al [27] have addressed the nonstationarity problem of the analyzed oscillatory systems using Hilbert transforms of the signals in phase reconstruction of the time-series data since this method does not require the stationarity of the transformed signals. A tool used in defining the phase of an arbitrary signal is known in signal processing arena as the analytic signal concept [32].
process of phase extraction involves the projection of the time series onto the complex unit circle and measuring the angle of rotation over time (Figure 2.1). A review of other possible phase extraction methods is available in [33].

Another issue that should be taken into consideration in this analysis is whether or not to base the analysis on single-trial or multi-trial data. This factor has significant implications on the reliability of the estimated statistics because, particularly in the case of neurophysiologic data, the interactions may display great variability from trial to trial or may last for a very short duration.

Despite the extensive literature on the interaction between respiration and cardiovascular systems, the exact relationship between CVC and RSA still remains unknown. There is a general belief that the interactive relationship between CVC and RSA forms a complex feedback and feedforward mechanism; however they are distinct processes rather than just the reverse of each other [30, 31]. Galletly and Larsen [34] suggested that CVC aligns heart beats relative to inspiratory timing such that heart beats occur at positions in the ventilatory cycle where they were maximally affected by RSA, which might be due to the role of CVC in improving the variation in the heart beat interval with respiration during sleep. Galletly and Larsen [35] also noted that RSA appears to persist at all levels of arousal whereas CVC is observed best at low arousal and that RSA diminishes with age while there is no effect of age on CVC. In [31], the relationship between cardiorespiratory synchronization and RSA is analyzed using *synchrograms*, a graphic

Figure 2.1 A Damped Sinusoidal Signal and its Representation on Unit Circle
tool, and it is suggested that these phenomena might be stimulated by two competing physiological mechanisms. Dick and Morris [19] observed a correspondence between respiratory- and pulse-modulated activities and indicated that the hypothesized modulations may not be the only factors driving this association.

The next section explains the basic concepts of wavelet transformation and reviews its applications in statistical process monitoring.

2.2 Wavelet Analysis

Wavelets literally mean “small waves”. Wavelets are smooth continuous functions that form building blocks in function decompositions in a similar way to sine and cosine functions in Fourier transform. Unlike the sine and cosine bases used in Fourier transforms which are perfectly localized in frequency space, wavelets are localized in both time and frequency, decaying to zero as \( t \to \pm \infty \) (compact support). This property of wavelets enables the modeling of irregular or hidden patterns in signals and the extraction of process information at different frequency levels at different times. The ability of wavelets to “zoom-in” on details is also of significant value in visualizing complex data; on the other hand, details can also be suppressed easily and thus wavelets can be used for data smoothing. Figure 2.2 displays Daubechies-4, Daubechies-3, and Symmlet-8 wavelet functions.

![Figure 2.2 Some Wavelet Families](image)
The next section reviews the fundamental mathematical properties of wavelets and wavelet transformations. For a more detailed description, the reader is referred to [37, 38].

2.2.1 Basics of Multiresolution Analysis (MRA) and Wavelet Transform

The multiresolution analysis involves a decomposition of the function space into a sequence of closed subspaces \( V_j, j \in \mathbb{Z} \), of \( L^2(\mathbb{R}) \) such that

i. The subspace \( V_j \) is contained in all higher subspace, i.e.

\[
\ldots \subset V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset \ldots
\]

ii. All \( f(t) \in L^2(\mathbb{R}) \) be included at the finest resolution and only the zero function at the coarsest level, i.e.

\[
\bigcap j V_j = \{0\}, \bigcup j V_j = L^2(\mathbb{R})
\]

iii. \( f(t) \in V_j \iff f(2t) \in V_{j+1} \) (scale or dilation invariance),

iv. \( f(t) \in V_0 \Rightarrow f(t-k) \in V_0 \) (translation or shift invariance),

v. There exists a function \( \phi(t) \), called scaling function, such that \( \{\phi(t-k)\} \) is an orthonormal basis of \( V_0 \).

Following the properties outlined in (2.2), the scaling function \( \phi(t) \) generates all the basis functions \( \{\phi_j^k(t)\} \) through dilations and translations for each subspace \( V_j \), i.e.

\[
\phi_j^k(t) = 2^{j/2} \phi(2^j t - k); j, k \in \mathbb{Z}
\]

In addition to scaling functions, which represent the low frequency components or smoothness of a signal, a set of wavelet functions, \( \psi_j^k(t) \), are required to represent the high frequency components of the signal or fine scale deviations from the overall smoothness of the signal. \( \psi_j^k \)'s are generated through the translations and dilations of a mother wavelet, \( \psi \), i.e.

\[
\psi_j^k(t) = 2^{j/2} \psi(2^{j/2} t - k).
\]

Discrete Wavelet Transform (DWT) approximates any continuous function \( f(t) \in L^2(\mathbb{R}) \) as
where \( J \) denotes the number of resolution levels and \( K_j \) is the number of coefficients at the \( j \)th level. Obviously, unlike the case of Fourier transforms, there exists a large selection of wavelet families depending on the choice of the mother wavelet. The choice of a suitable set of basis functions depends on the level of smoothness, number of vanishing moments, and degree of time-frequency localization as required by the data or the application. Some of the wavelet bases available in literature are Haar, Daubechies, coiflets, Morlet wavelets, symlets, and biorthogonal wavelets. Haar basis is the simplest of wavelet bases; however, is not very effective in time-frequency localization and estimating smooth functions. The more complex functions like the Daubechies wavelets produce overlapping averages and differences that provide a better average than the Haar wavelet at lower resolutions.

In order to extract scaling coefficients \( \{ a_{j,k} \} \) and wavelet coefficients \( \{ d_{j,k} \} \), Mallat [39] developed an efficient algorithm referred to as Pyramidal Algorithm. This algorithm required the number of data points, \( n \), to be dyadic, i.e. \( n = 2^k, k \in \mathbb{Z} \). This algorithm is based on successive averaging and differencing referred to as low-pass and high-pass filtering respectively in signal processing terminology. Figure 2.3 displays a three-level wavelet decomposition three where \( G_0 \), \( H_0 \), and \( X[n] \) represent the low-pass filter, high-pass filter, and a signal of length \( n \) respectively. In this process, low-pass filter removes high frequencies of the data. Since details (e.g. sharp changes in the data) correspond to high frequencies, the averaging procedure tends to smooth the data. Therefore, at each level \( j \) of DWT, \( H_0 \) produces detail coefficients, \( d_j \), and \( G_0 \) produces the coarse approximations represented by \( a_j \). Figure 2.4 illustrates the concepts of averaging and differencing for a simple digital signal. The procedure of reconstructing a signal is the reverse process of decomposition.

\[
f(t) \approx \sum_{k=1}^{K} a_{j,k} \phi_{j,k}(t) + \sum_{j=1}^{J} \sum_{k=1}^{K_j} d_{j,k} \psi_{j,k}
\]
Note that during this process, each step produces a set of averages and coefficients that is half the size of the input data. For example, if the time series contains 256 data points, the first step will produce 128 averages and 128 detail coefficients. The averages then become the input for the next step (e.g., 128 averages resulting in a new set of 64 averages and 64 detail coefficients). This continues until one average and one coefficient (e.g., $2^6$) is calculated.

In practice, signals can be estimated using much fewer number of wavelet coefficients than the original number of data points in the signal. A major contribution of wavelet analysis has
been a reasonable estimation of the number of coefficients based on a certain criterion. This property is used in signal estimation algorithms also referred to as denoising that attempt to remove the parts of the signal that fall into a particular model of noise for data compression purposes. Figure 2.5 displays a sample noisy signal and the same signal after denoising.

A fundamental denoising method is shrinking the wavelet coefficients via hard thresholding and soft thresholding [40, 41, 42]. Both methods eliminates all the coefficients that are smaller than a specified threshold value with the only difference being that hard thresholding maintains the original values of the remaining coefficients whereas soft thresholding shrinks them toward zero by decreasing their values by the magnitude of the threshold value. Figure 2.6 illustrates how these thresholding methods work on a sample signal. The choice of an appropriate threshold value is important due to the fact that a very large threshold will possibly result in an over-smoothing of the data and a very small threshold will result in a rough reconstruction. One of the most common threshold measures in literature has been the universal threshold suggested by Donoho et al. [40] as

\[ t_j = \sigma_j \sqrt{\frac{2}{\log(n)}} \]  

(2.4)

where \( \sigma_j \) is the standard deviation of noise at level \( j \) and \( n \) is the signal length. An extensive review for threshold selection is available in [43]. Various other denoising schemes are available in literature. SUREShrink method [44] introduced an adaptive procedure that assigns a threshold value to each decomposition level based on the criterion of minimizing Stein’s Unbiased Risk Estimate. Saito [45] proposed the AMDL (Approximation Minimum Description Length) procedure which minimizes a cost objective which is a function of the number of eliminated coefficients. Other various thresholding schemes and their applications are available in references [46, 47, 48].
Figure 2.5 A Noisy Signal and the Reconstructed Signal After Denoising

Figure 2.6 A Sample Signal and the Reconstructed Form After Hard- and Soft-thresholding
(Note: all numbers are in generic values)
The next section describes principal curves. Principal curves form the basis for the methodology described in Chapter 4.

2.3 Principal Curves

Principal component analysis (PCA) is a standard analysis approach in multivariate statistics [49]. PCA finds the directions along which the data have the highest variability through an eigenvalue decomposition of the variance matrix of the data and has been used extensively for feature extraction and data compression purposes. First principal component of a random vector $\mathbf{x}$ is defined as the straight line which maximizes the variance of the projection of $\mathbf{x}$ to a line and minimizes the distance between the line and $\mathbf{x}$. It is known that, if the distribution of $\mathbf{x}$ is elliptical, the first principal component is “self-consistent”, meaning that any point of the line is the average of all points in $\mathbf{x}$ that project to this point.

Hastie [50] and Hastie and Stuetzle [51] (hereafter HS) generalized the self-consistency property of the first principal component to principal curves. HS defined principal curves as one-dimensional, self-consistent smooth curves that pass through the middle of a $p$-dimensional distribution or data cloud. Therefore, principal curves provide a good, one-dimensional summary of a high-dimensional dataset. Let $f(t) = (f_1(t), f_2(t), \ldots, f_p(t))$ be a smooth infinitely-differentiable curve in $\mathbb{R}^p$, where $f_1(t), f_2(t), \ldots, f_p(t)$ are coordinate functions in $p$-dimensional space and parameter $t$ is generally defined as the arc length along the curve. Projection index, $t_f : \mathbb{R}^p \to \mathbb{R}^1$, is defined as $t_f(x) = \sup_t \left\{ t : \|x - f(t)\| = \inf \|x - f(t)\| \right\}$, i.e. $t_f(x)$ is the he largest value of $t$ for which $f(t)$ is closest to $x$. According to the definition of HS, $f(t)$ is a principal curve if it satisfies the following properties:

i. $f$ is a unit-speed curve defined over a closed interval,

ii. $f$ does not intersect itself, i.e. $t_1 \neq t_2 \Rightarrow f(t_1) \neq f(t_2)$,
iii. \( f \) is self-consistent, i.e. \( E(x \mid t_f(x = t)) = f(t) \).

The algorithm proposed by HS for constructing principal curves starts with some prior summaries, such as the usual principal component line iterates between a projection step and an expectation step until convergence. The algorithm to obtain the principal curve of the distribution of \( x \) is summarized as follows:

Step 0. Set \( f^{(0)}(t) = \mu_x + ta \) where \( \mu_x \) is the mean of \( x \) and \( a \) is the first linear principal component of density function \( h \). Set \( t^{(0)}(x) = t_{f^{(0)}}(x) \) and \( j = 1 \).

Step 1. (Expectation) Define \( t^{(j)}(t) = E[ x \mid f^{(j-1)}(x) = t ] \).

Step 2. (Projection) Set \( t^{(j)}(x) = t_{f^{(j)}}(x) \) for \( \forall x \in h \) where

Step 3. Stop if \( \left\| D^2(h, f^{(j-1)}) - D^2(h, f^{(j)}) \right\| / D^2(h, f^{(j-1)}) < \epsilon \), where

\[
D^2(h, f) = E_h \left\| x - f(t_f(x)) \right\|^2.
\]

Otherwise, let \( j = j + 1 \) and go to Step 1.

In the algorithm for fitting principal curve from data sets, rather than from a density function \( h \), the expectations and distance function in Step 1 and 3 are replaced by their empirical counterparts respectively. In this analysis, cubic smoothing splines are used for local averaging in Step 1. Figure 2.7 is a display of a sample principal curve fit to a simulated data using the HS algorithm.
HS showed that no infinitesimally small perturbation to a principal curve will decrease
\[ E\|x - f(t_f(x))\|^2 \], which generalizes the minimum square error property of PCA to the
distribution of \( x \). HS did not answer in detail the questions on the existence and uniqueness of
principal curves; however proved the existence of principal curves for some special distributions,
such as elliptical and spherical distributions.

Tibshirani [52] stated that, with the algorithm proposed by HS, maximization of the
likelihood over all parameters may lead to inconsistent or inefficient estimates and developed an
alternative algorithm based on a mixture model. This approach makes use of the EM algorithm,
originally presented by Dempster [53], to compute maximum likelihood estimates of “missing
data” that can take the form of unrecorded observations, unobservable parameters, or
unobservable data.

Some studies exist on using principal curves in process monitoring in multi-sensor
manufacturing processes. Kim et.al [54] proposed a fault classification method for a forging
process based on the distances between the projections of multichannel tonnage signals and a principal curve estimated using these signals. Principal curve estimation has been used in a variety of nonlinear data analysis studies [55, 56, 57].

The next chapter explains the process followed to extract the wavelet coefficients to be used in Chapter 4. It also presents a preliminary multiscale analysis of the coefficients.
CHAPTER 3. WAVELET TRANSFORMATIONS OF CARDIORESPIRATORY SIGNALS

The methodologies to be presented in Chapters 4 and 5 consist of statistical tests and analysis tools where the inputs will be the wavelet coefficients obtained from each signal. In this chapter, the procedure for extracting the wavelet coefficients from phrenic nerve, lumbar nerve, and blood pressure signals is described and a preliminary multiscale analysis of these signals is performed.

3.1 Extraction of Scaling and Wavelet Coefficients

In order to extract the scaling and wavelet coefficients described in Section 2.2 using DWT, only the first 102.4 sec.s of each sample signal corresponding to $n=4096$ data points were taken. Daubechies-4 was chosen as the suitable wavelet family in DWT of the signals.

Figures 3.1, 3.2, and 3.3 display a level-5 wavelet decomposition of a sample phrenic nerve signal, lumbar nerve signal, and blood pressure signal respectively. The y-axes denote the signal value and x-axes indicate the observation index. The left column displays the approximation components and the right column shows the detail components of the signal respectively. Naturally, the highest frequencies appear at the lower levels of decomposition original signal while successive approximations appear less and less noisy, progressively losing more high-frequency information.

The coefficients were hard-thresholded using the scale-dependent universal threshold explained in Section 2.2. A visual check of the reconstructed signals displayed on the right columns of Figures 3.4, 3.5, and 3.6 shows that the thresholding strategy keeps the significant wavelet coefficients allowing almost perfect reconstruction.
Figure 3.1 Level-5 Decomposition of a Before-Hypoxia Phrenic Nerve Signal (X-axes Indicate Index Value and Y-axes Indicate Signal Value)
Figure 3.2 Level-5 Decomposition of a Before-Hypoxia Lumbar Nerve Signal (X-axes Indicate Index Value and Y-axes Indicate Signal Value)
Figure 3.3 Level-5 Decomposition of a Before-Hypoxia Blood Pressure Signal (X-axes Indicate Index Value and Y-axes Indicate Signal Value)

Figure 3.4 displays the original wavelet coefficients and those that are selected by hard thresholding for a sample phrenic nerve signal. Figures 3.5 and 3.6 show examples of reconstructed forms of lumbar and blood pressure signals respectively. The thresholding procedure naturally results in the extraction of different numbers of coefficients from each signal. The methodology to be introduced in Chapter 4 requires an equal number of wavelet coefficients as inputs; therefore, it is required to select a common set of coefficients from the signals. In order to eliminate the loss of information, the union of the selected coefficient positions is selected in
order to form the observation vectors as illustrated in Figure 3.6. This approach results in the extraction of 95 coefficients which are listed on Table 3.1.

![Figure 3.4 Original and Thresholded Coefficients and Original and Reconstructed Signals for a Phrenic Nerve Signal Sample](image)

![Figure 3.5 Original (red) and Reconstructed (blue) Forms of Lumbar Nerve and Blood Pressure Signal Samples](image)

Figure 3.4 Original and Thresholded Coefficients and Original and Reconstructed Signals for a Phrenic Nerve Signal Sample (Note: In Coefficient Tables Y-axis Indicates Decomposition Level and X-axis Indicates Coefficient Position)

Figure 3.5 Original (red) and Reconstructed (blue) Forms of Lumbar Nerve and Blood Pressure Signal Samples
Figure 3.6 Selection of Wavelet Coefficients (Note: In Coefficient Tables y-axis Indicates Decomposition Level and x-axis Indicates Coefficient Position)

Table 3.1 Selected Wavelet Coefficients

<table>
<thead>
<tr>
<th>Scale</th>
<th>A9</th>
<th>D9</th>
<th>D8</th>
<th>D7</th>
<th>D6</th>
<th>D5</th>
<th>D4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>a_{91}, a_{92}, a_{93}, a_{94}, a_{95}, a_{96}, a_{97}, a_{98}</td>
<td>d_{91}, d_{92}, d_{93}, d_{94}, d_{95}, d_{96}, d_{97}, d_{98}</td>
<td>d_{81}, d_{82}, d_{83}, d_{84}, d_{85}, d_{86}, d_{87}, d_{88}</td>
<td>d_{71}, d_{72}, d_{73}, d_{74}, d_{75}, d_{76}, d_{77}, d_{78}</td>
<td>d_{61}, d_{62}, d_{63}, d_{64}, d_{65}, d_{66}, d_{67}, d_{68}</td>
<td>d_{51}, d_{52}, d_{53}, d_{54}, d_{55}, d_{56}, d_{57}, d_{58}</td>
<td>d_{41}, d_{42}, d_{43}, d_{44}, d_{45}, d_{46}, d_{47}, d_{48}</td>
</tr>
</tbody>
</table>
3.2 Analysis of Energy Redistributions in Wavelet Coefficients

Wavelet transforms provide a set of powerful exploratory data analysis tools by giving insight about the dominant scales of variation in the analyzed time series. It is of potential value to quantitatively analyze the spectral energy redistribution in continuously monitored patients before and after hypoxia, since different effects of hypoxic stimuli may have different time-frequency localizations. It is also important to describe how the energy in different frequency bands is modified at different types of signals. The total energy contained in a signal $x(t)$ is defined as its integrated magnitude, i.e.

$$E_{\text{tot}} = \int_{-\infty}^{\infty} |x(t)|^2 \, dt = \|x(t)\|$$  \hfill (3.1)

The wavelet spectrum is defined as

$$E_j = \sum_{k=1}^{2^j} d_{j,k}^2$$  \hfill (3.2)

orthogonal transformations, $\forall j = 1,\ldots,J$ where $J$ is the cut-off frequency. Wavelet spectrum has proven to be useful for detecting of dominant scales of variation. Peaks in $E_j$ highlight the dominant energetic scales within the analyzed signal. It is beneficial to normalize $E_j$ by $E_{\text{tot}}$ to emphasize large peaks and sharp discontinuities and de-emphasize low-amplitude background noise.

Figures 3.7, 3.8, and 3.9 display the relative energy distributions $E_j / E_{\text{tot}}$ through time for phrenic nerve, lumbar nerve, and blood pressure signals at levels $j = 1,\ldots,9$ respectively. Figure 3.10 shows the total energy levels for these signals. For phrenic nerve signal, levels $j = 1,\ldots,5$ undergo an obvious relocation in energy at the termination of hypoxic stimulation. This period is characterized by a small through followed by a local peak in relative distribution. This phenomenon may indicate an increase in high-frequency activity. Lumbar nerve signal experiences an abrupt increase in relative energy at levels $j = 1,\ldots,7$ in response to hypoxia,
which indicates that relocation of energy in lumbar nerve signals cover both low-frequency and high-frequency levels. The energy relocation patterns in these decomposition levels are similar to each other and are mainly caused by the sudden decrease in total energy level in lumbar nerve activity as a consequence of hypoxic stimuli. In the case of blood pressure signal, the redistribution pattern is not unique among scales and is more complicated.

Figure 3.7 Relative Energy Distributions at Different Scales of a Phrenic Nerve Signal
Figure 3.8 Relative Energy Distributions at Different Scales of a Lumbar Nerve Signal
3.9 Relative Energy Distributions at Different Scales of a Blood Pressure Signal

3.10 Total energy distributions through time for phrenic nerve, lumbar nerve, and blood pressure signals respectively (dashed line indicates the hypoxic period)
It should be noted that for all wavelets there is a one-to-one relationship between the scale and period. This is important in the interpretation of the scale-dependent information. The scale refers to the width of the wavelet. As the scale increases and the wavelet gets wider, it includes more of the time series, and the finer details get smeared out. The scale can be defined as the distance between oscillations in the wavelet (e.g. for the Morlet), or it can be some average width of the entire wavelet (e.g. for the Marr or Mexican hat). The period (or inverse frequency) is the approximate Fourier period that corresponds to the oscillations within the wavelet. The relationship can be derived by finding the wavelet transform of a pure cosine wave with a known Fourier period, and then computing the scale at which the wavelet power spectrum reaches its maximum. Table 3.2 displays this relationship between decomposition levels and implied (or pseudo) frequencies and periods for Daubechies-4 wavelet.

Table 3.2 Mapping Between Levels and Frequency/Period Values for Daubechies 4 Wavelet

<table>
<thead>
<tr>
<th>Level</th>
<th>Pseudo-Frequency</th>
<th>Pseudo-Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>14.2857</td>
<td>0.07</td>
</tr>
<tr>
<td>2</td>
<td>7.1429</td>
<td>0.14</td>
</tr>
<tr>
<td>3</td>
<td>3.5714</td>
<td>0.28</td>
</tr>
<tr>
<td>4</td>
<td>1.7857</td>
<td>0.56</td>
</tr>
<tr>
<td>5</td>
<td>0.8929</td>
<td>1.12</td>
</tr>
<tr>
<td>6</td>
<td>0.4464</td>
<td>2.24</td>
</tr>
<tr>
<td>7</td>
<td>0.2232</td>
<td>4.48</td>
</tr>
<tr>
<td>8</td>
<td>0.1116</td>
<td>8.96</td>
</tr>
<tr>
<td>9</td>
<td>0.0558</td>
<td>17.92</td>
</tr>
</tbody>
</table>

The next chapter introduces PCuRTTest and its application on the wavelet coefficients selected from phrenic nerve, lumbar nerve, and blood pressure signals as described above.
CHAPTER 4. DETECTION OF MULTISCALE CHANGES IN MULTICHANNEL DATA USING PRINCIPAL CURVE REGRESSION TEST (PCuRTest)

Previous studies related to the exploration of the effects of hypoxia on biological systems are offline monitoring procedures where abnormalities or changes in the recorded signals are analyzed after the experiment is conducted. However, in cases where a quick response to a hypoxic condition has vital importance, an online monitoring system is necessary. The motivation in using PCuR Test [58], to monitor the wavelet coefficients of blood pressure and integrated phrenic nerve signals is to integrate their mutual multiscale behavior and to identify the coefficients or equivalently scale levels where a shift from the nominal conditions occur whenever new data is available. The data compression feature of wavelet coefficient is of significant value in this example, since it helps to decrease the number of tested data points by a great proportion.

In Section 4.1, a general description of PCuR Test will be given. Section 4.2 analyzes the results of applying PCuRTest to wavelet coefficients of phrenic nerve and blood pressure signals.

4.1 PCuR Test Algorithm

Let $\mathbf{x}_k$ be the observed value of $k$th feature of $p$-channel signals $(k = 1, ..., n)$, where $\mathbf{x}_k = (x_{k1}, \ldots, x_{kp})^T$ and the signal is of the form $\{x_k\}_{i=1}^n$. If signals with $n$ features are observed, denote by $\{x_{k,i}\}_{i=1}^N$ the $i$th observation of $\{x_k\}_{1}^n$, where $\mathbf{x}_{k,i} = (x_{k1,i}, \ldots, x_{kp,i})^T$ is the $i$th observation of $\mathbf{x}_k$, $i = 1, \ldots, N$ and $k = 1, \ldots, n$. The $N$ samples are of the form $\{X_k\}_{1}^n$, where the $N \times p$ matrix $\mathbf{X}_k = [\mathbf{x}_{k,1}, \mathbf{x}_{k,2}, \ldots, \mathbf{x}_{k,n}]^T$. 
The basic idea of PCuR is to extract a principal curve for each sample \{x_k\}_{1}^{n} and build a regression model between the principal curve and \{x_k\}_{1}^{n} (Figure 4.1). Using \(N\) in-control observations \{X_k\}_{1}^{n} under, \(N\) principal curves are extracted as the initial step. The principal curve \(f\) of \{x_k\}_{1}^{n} is a vector of \(p\) functions of single variable \(t\), i.e., \(f(t) = (f_1(t), f_2(t), \ldots, f_p(t))^T\), where \(f_1(t), f_2(t), \ldots, f_p(t)\) are coordinate functions in \(p\)-dimensional space. If \(y_k\) is the projection of \(x_k\) on \(f\), we can use \{y_k\}_{1}^{n} to describe \(f\). The \(i\)th principal curve fitted from \{x_k\}_{1}^{n} is denoted as \{y_{ki}\}_{1}^{n}. \{y_{ki}\}_{1}^{n} is then organized into an \(n \times p\) matrix as \([y_{1i}, y_{2i}, \ldots, y_{ni}]^T\). Define \(N \times p\) matrix \(Y_k\) as \(Y_k = [y_{k1}, y_{k2}, \ldots, y_{kn}]^T\). \(Y_k\) contains the projected points of \(X_k\) on the corresponding \(N\) principal curves.

With \(N\) principal curves in the form of \{\(Y_k\)\}_{1}^{n}, we establish regression models at each observation point. A multivariate linear regression model is assumed to adequately model “response” \(y_k\) and “predictors” \(x_k\) at the \(k\)th argument value,

\[
y_k^T = [1 \ x_k^T] B_k + \varepsilon_k \tag{4.1}
\]

\[E(\varepsilon_k) = 0, \text{ Cov}(\varepsilon_k) = \Sigma_k, \quad k = 1, \ldots, n. \tag{4.2}\]

Let \(B_k = [\beta_{k(1)}, \beta_{k(2)}, \ldots, \beta_{k(p)}]\). Define \(N \times 1\) vector \(1_N^T = (1, \ldots, 1)\). To simplify notation, the subscript \(n\) will be dropped when the dimension of the vector \(1_N\) is clear from the context.

For data sets \(X_k\) and \(Y_k\) of \(k\)th variable,

\[
Y_k = [1 \ X_k] B_k + E_k \quad k=1,2,\ldots,n, \tag{4.3}
\]

\[E_k = [\varepsilon_{k(1)}, \varepsilon_{k(2)}, \ldots, \varepsilon_{k(p)}], \ E(\varepsilon_{k(j)}) = 0, \text{ Cov}(\varepsilon_{k(j)}, \varepsilon_{k(g)}) = \sigma_{k(j,g)} I, \quad j, g = 1, 2, \ldots, p. \tag{4.4}\]

The \(N\) observations on \(k\)th trial have covariance matrix \(\Sigma_k = \{\sigma_{k(j,g)}\}\), but observations from different channels at time \(k\) are assumed to be uncorrelated. If \(\varepsilon_k\) follows normal distribution and \(\text{rank}(\begin{bmatrix} 1 & X_k \end{bmatrix}) = p+1, N \geq 2p+1\), the maximum likelihood estimator (MLE) of \(B_k\) is [49]
Estimated \((p+1)\times p\) projection matrix: 
\[
\hat{B}_k = \left( [1 \; X_k] \left[ [1 \; X_k] \right]^\intercal \right)^{-1} \left[ [1 \; X_k] \right]^\intercal Y_k
\]  
(4.5)

where \(\hat{B}_k\) has a normal distribution with 
\[E(\hat{B}_k) = B_k\]  
and 
\[\text{Cov}(\hat{B}_{k(j)}, \hat{B}_{k(g)}) = \sigma_{k(j,g)}\]

\(\left( [1 \; X_k] \left[ [1 \; X_k] \right]^\intercal \right)^{-1}\). \(\hat{B}_k\) is independent of MLE of \(\Sigma_k\) given by
\[
\hat{\Sigma}_k = \frac{1}{n} \hat{E}_k^\intercal \hat{E}_k
\]
(4.6)

, where \(N\hat{\Sigma}_k\) follows Wishart distribution with \((N-p-1)\) degree of freedom, i.e., \(W_{p,N-p-1}(\Sigma_k)\).

Accordingly, predicted values \(\hat{Y}_k\) and residuals \(\hat{E}_k\) at time \(k\) are computed as
\[
\hat{Y}_k = [1 \; X_k] \hat{B}_k = [1 \; X_k] \left( [1 \; X_k] \left[ [1 \; X_k] \right]^\intercal \right)^{-1} \left[ [1 \; X_k] \right]^\intercal Y_k
\]
(4.7)
\[
\hat{E}_k = Y_k - \hat{Y}_k = \left( I - [1 \; X_k] \left( [1 \; X_k] \left[ [1 \; X_k] \right]^\intercal \right)^{-1} [1 \; X_k]^\intercal \right) Y_k
\]
(4.8)

Once the PCuR model is established from in-control signals, it can be used to determine whether process change occurs in future observations. The principal curve of new observation \(\{z_k\}_1^n\) is treated as new response \(\{y_k\}_1^n\). By the results of multivariate regression [49], the predicted ellipsoids for new response \(y_k\) at time \(k\) are
Process change is believed to occur in a test sample if any point is outside control ellipsoid or indicating an out-of-control condition. In the next section, wavelet coefficients extracted from phrenic nerve signals and blood pressure signals as described in Chapter 3 are monitored by PCuRTTest.

### 4.2 Detection of Changes in Phrenic Nerve and Blood Pressure Signals using PCuRTTest

To apply PCuRTTest to the physiology data, the selected coefficients from both before- and after-hypoxia conditions were placed in $95 \times 58$ and $95 \times 15$ matrices respectively. Since there are three channels of signals in this case, $p=3$. The first 30 of the before-hypoxia samples are selected to form the training set and the remaining samples were used for validation of the model. The 15 samples from after-hypoxia conditions form the test dataset. Figures 4.1, 4.2, 4.3, and display the results of the test at a significance level of $\alpha = 0.01$. Each point on these figures corresponds to a different wavelet coefficient and the points above the threshold (in red) represent wavelet coefficients where a change in the process has occurred. Figures 4.1 and 4.2 indicate that there are some out-of-control coefficients in the before-hypoxia samples; however, these remain to be within a tolerable false alarm level. As Figures 4.3 and 4.4 display, PCuRTTest results in the detection of a significant amount of out-of-control coefficients.
Figure 4.2 PCuRTTest Results for Some Samples from the Validation Dataset
Figure 4.3 PCuRTTest Results for Some Samples from the Validation Dataset
Figure 4.4 PCuRTTest Results for Some Samples from the Test Dataset
Figure 4.5 PCuRTest Results for Some Samples from the Test Dataset

Figure 4.6 displays the average relative frequency of the selected levels computed by averaging the number of cases where the coefficients in that level were out-of-control over all trials.

Figure 4.6 Relative Efficiency Graph for Selected Levels

The next chapter discusses the results of this research and potential future directions.
CHAPTER 5. CONCLUSIONS AND FUTURE RESEARCH

5.1 Conclusions

This thesis presents an online process monitoring tool, PCurTest, to detect changes in wavelet coefficients of multichannel signals. The main contributions of this research can be summarized as follows:

i. Detecting changes in processes that involve the fusion of information from multiple channels

ii. Identifying the multiscale levels in signals where a change in relative wavelet energy occurs

iii. Identification of wavelet coefficients where a process change is observed using PCuRTest

iv. Application of PCuRTest for online monitoring of physiological signals and detection of changes in these signals that are observed as a consequence of hypoxic stimuli.

The identification of out-of-control wavelet coefficients are believed to indicate the time-frequency levels that undergo significant changes as a result of hypoxia. The results display that the methodology can successfully distinguish between before-hypoxia and after-hypoxia samples with low false alarm rates. The procedure is shown to perform better than the monitoring of channels separately.
5.2 Future Research

An important problem in the analysis of multichannel signals is the diagnosis of the simultaneous effects of certain variation sources on signals from different channels. It is clear that the analysis of multichannel data requires the identification of spatial variation patterns and the temporal behavior of these patterns. A natural motivation throughout this chapter is that understanding the nature of sources if variation that affect multichannel signals will help the identification and prevention of problems causing process variability. In particular, the results of the proposed methodology will be demonstrated to illuminate the dynamics of the underlying mechanisms that influence both respiratory and blood pressure signals under hypoxia conditions.

The next section explains the proposed method and its relation to Inter-Battery Factor Analysis [76, 78]. This model aims at the “blind” identification of a single set of unobservable sources that accounts for the correlation between observations from separate channels. The reason why the term “blind” is used here is the fact that the common variation sources are extracted completely from the data without specifying potential factors before the analysis is conducted. An advantage of this model is that it doesn’t require an equal number of features or variable from the signals. The following sections discuss the implications of modeling the cardiorespiratory interactions using the proposed model.

5.2.1 Factor Analysis

Factor analysis is a multivariate statistical analysis tool that to analyze the interrelationships among a large number of variables and to explain these variables in terms of their common underlying dimensions (factors) based on the internal structure of covariance or correlation matrices. Since its development by Spearman [59], it has been used in a range of studies in social sciences, mainly in the analysis of psychological or mental tests and economic quantities [60, 61, 62].
Factor analysis partitions a multivariate observation vector into an unobserved systematic part composed of a relatively small number of factors and an unobserved error part and separates the effect of factors from error. In contrast to principal component analysis which attempts at identifying uncorrelated linear combinations of the original variables that account for a major portion of the variability in the data, factor analysis aims to explain the covariance matrix by a “minimum” number of hypothetical variates or factors. In this sense, principal component analysis is considered to be variance-oriented whereas factor analysis is covariance-oriented.

Let \( x = [x_1, x_2, ..., x_n]^T \) be the \( n \)-dimensional vector of observations. The model for factor analysis is defined as

\[
x = \mu + \Lambda f + e
\]

where \( \mu \), and \( e \) are column vectors of \( n \) components, \( f \) is a column vector of \( k (\leq n) \) components, and \( \Lambda \) is a \( n \times k \) matrix. In model (5.1), the elements of the factor loading matrix described as \( \Lambda_{ij} \) simply represent the weight of the \( \text{ith} \) test-variate on the \( \text{th} \) hypothetical (latent) factor. It is assumed that \( e \) is distributed independently of \( f \) with mean \( E[e] = 0 \) and covariance matrix \( E[ee^T] = \Psi \), where \( \Psi \) is assumed to have a diagonal form. To understand the physical meaning of the components of the model in (5.1), consider the case of mental tests. In this context, each component of \( x \) denotes the score of the subject in a given test and the corresponding component of \( \mu \) denotes the mean score of this test in the population of subjects. It will be assumed, without loss of generality, that \( \mu = 0 \) (Note that this assumption can be easily satisfied by subtracting the mean of \( x \) from the data). When a battery of tests is given to a group of subjects, it is observed that the score of the subject in a given test is more related to his/her scores on the other tests rather than the scores of other individuals on other tests. Therefore, the motivation behind using a factor analysis model in the case of mental tests is to explain the interrelation between the tests of a particular subject through the separation of the effects acting
on a test score into two parts. The first part, contained in \( f \), consists of common factors that account for the correlations between scores in different scores and \( k \) denotes the number of unobserved ("latent") common factors. The vector \( f \) is assumed to be a zero-mean random vector with covariance matrix, \( \Phi \). The components of vector, \( e \), contains the factors that are specific to the corresponding test. In other words, \( e \) is the part of the test score not explained by the common factors and therefore, is often referred to as uniqueness.

Assuming that the model in (5.1) fits the data and the following set of assumptions are satisfied, a direct consequence will be that the covariance matrix of \( x \) can be expanded as

\[
\Sigma = E\left( (\Lambda f + e)(\Lambda f + e)\right) = \Lambda \Lambda^T + \Psi
\]

which is often called the fundamental theorem of factor analysis. Therefore, the problem of factor analysis becomes whether there exist a triplet \( \Lambda, \Phi, \) and positive definite and diagonal \( \Psi \) that satisfy (5.2) given a covariance matrix \( \Sigma \) and number of factors, \( k \). If a solution exists and is unique, the model is said to be identified. There is a fundamental indeterminacy in the factor analysis model due to the fact that any triplet can be transformed through multiplication by a nonsingular matrix \( C \) into an equivalent structure of \( \Lambda^* = \Lambda C \), \( \Phi^* = C^{-1} \Phi C^{T-1} \), and \( \Psi \) [63]. To rule out this indeterminacy problem, some \( k^2 \) identification conditions have to be imposed on \( \Lambda \) and \( \Psi \). A common approach is to add the restrictions that

\[
J = \Lambda \Psi^{-1} \Lambda
\]

is diagonal. If the diagonal elements of \( J \) are ordered and different from each other, then \( \Lambda \) is uniquely determined. Another set of restriction conditions imposes that the first \( k \) rows of \( \Lambda \) forms a lower triangular matrix.

A natural approach to determine the number of common factors in factor analysis has been the analysis of the eigenvalues \( \{ \lambda_i \}_{i=1}^{k} \) of the covariance matrix of the data, \( \Sigma \). A simple visual
tool is the scree plot [64] where eigenvalues are plotted against the number of factors and the number of factors at the cutoff point is taken as the optimal number of factors, \( \hat{k} \).

Alternative procedure based on maximum likelihood functions are Akaike information criterion (AIC) [65] and minimum description length (MDL) [66]. These criterion are based on a sequential Chi-square test of

\[
AIC(k) = N(n - k) \log \left( \frac{a_k}{g_k} \right) + k(2n - k) \quad (5.4)
\]

and

\[
MDL(k) = N(n - k) \log \left( \frac{a_k}{g_k} \right) + k(2n - k) \log(N) / 2 \quad (5.5)
\]

for \( k=0,...,n-1 \); where \( a_k \) and \( g_k \) denote the arithmetic and geometric means of the smallest \( n-k \) smallest eigenvalues of \( \Sigma \). In AIC and MDL methods, the optimal number of factors, \( \hat{k} \), is chosen as the \( k \) that minimizes (5.4) and (4.4) respectively. Bozdogan [67] proposed an alternative information complexity (ICOMP) criterion that it combines a badness-of-fit term (minus twice the maximum log likelihood) with a measure of complexity of a model by taking into account the interdependencies of the parameter estimates as well as the dependencies of the model residuals unlike AIC or MDL. A review of various model selection procedures is available in [68].

Once a set of factor loadings has been obtained, the next crucial step is the interpretation of the factor loadings in order to extract meaningful information about the analyzed data. This step requires the transformation of factor loadings through a rotation technique which is to be selected depending on the objectives and characteristics of the problem and data. In particular, for solutions with more than one factor (\( k > 1 \)) prior to rotation the first axis will lay in between the clusters of variables and the variables will not sort well on the factors obstructing their physical meaning. Rotation of the axes allows the factor loadings of each variable to be more clearly
differentiated by factor. For instance, in the case of psychological tests, each variable or factor in general has a positive direction such that more ability leads to higher test scores which implies that the factor loading should assume a positive value if it is non-zero.

Four different studies [69, 70, 71, 72] have independently proposed orthogonal rotation criteria based on different reasoning but yielding the same solution known as quartimax, which maximizes the variance of variables (rows) over all factors (columns). As a controversial method, Kaiser [73] proposed the popular varimax criterion which is an orthogonal rotation of the factor axes to maximize the variance of the squared loadings of a factor (column) on all the variables (rows) in a factor matrix. Therefore, varimax minimizes the number of variables which have high loadings on any one given factor and each factor consequently tends to have either large or small loadings of particular variables on it. Equimax rotation was proposed as a compromise between quartimax and varimax methods trying to accomplish the goals of both quartimax and varimax; however it has not gained much popularity.

![Figure 5.1 Plots of Unrotated and Varimax-Rotated Factor Loadings for a Sample Dataset (Note: The Numbers Indicate the Variable Index and All Numbers Are in Generic Values)](image)

47
As an alternative to orthogonal rotations, *direct oblimin* rotation and *promax* rotation [74] have been proposed as oblique (nonorthogonal) rotation methods. Promax rotation is known to be computationally faster than the direct oblimin method and therefore is sometimes used for very large datasets. Some of the other oblique factor rotation techniques are oblimax, quartimin, biquartimin, binormamin, and maxplane.

Figure 5.1 displays plots of loadings of variables from a sample dataset on unrotated and varimax-rotated factor axes. These plots show that varimax rotation has simplified the structure of the loadings with each variable depending on only one factor, which makes it possible to describe each factor in terms of the variables that it affects.

### 5.2.2 Multiple Battery Factor Analytic Model of Common Variation Patterns

Let \( \mathbf{x}_i : m \times 1 = \left[ \mathbf{x}'_i, \mathbf{x}''_i, ..., \mathbf{x}'_p \right] \) where \( \mathbf{x}_i : m \times 1 \) \((i = 1, ..., p)\) denotes the set of wavelet coefficients extracted from the \(i\)th channel, \(m_i\) denotes the number of variables obtained from channel \(i\), and \(m = \sum_{i=1}^{p} m_i\). The \(k\) \((i \in \{1, \ldots, m\})\) common variation sources are represented by the \(k\)-vector \(\mathbf{f}\). It is assumed that the observations obey the following model

\[
\begin{align*}
\mathbf{x}_1 &= \mathbf{\mu}_1 + \mathbf{\Lambda}_1 \mathbf{f} + \mathbf{\Gamma}_1 \mathbf{y}_1 + \mathbf{e}_1 \\
\mathbf{x}_2 &= \mathbf{\mu}_2 + \mathbf{\Lambda}_2 \mathbf{f} + \mathbf{\Gamma}_2 \mathbf{y}_2 + \mathbf{e}_2 \\
&\quad \vdots \\
\mathbf{x}_p &= \mathbf{\mu}_p + \mathbf{\Lambda}_p \mathbf{f} + \mathbf{\Gamma}_p \mathbf{y}_p + \mathbf{e}_p
\end{align*}
\]

where \(\mathbf{\Lambda}_i : p_i \times k\) contains the \(k\) *common loading vectors* describing the strength of common variation sources on the observations from the \(i\)th channel and \(\mathbf{\Gamma}_i\) contain *channel-specific loadings* indicating the significance of various sources acting on the \(i\)th channel only \((i = 1, ..., p)\).
Let \( y = [y'_1, y'_2, \ldots, y'_p] \), \( \mu = [\mu'_1, \mu'_2, \ldots, \mu'_p] \) and \( e = [e'_1, e'_2, \ldots, e'_p] \) contain channel-specific variation sources, mean vectors of the channel observations and residuals respectively. It is assumed that \( e \) is a zero-mean vector that is independent of \( f \), i.e. \( \text{Cov}(f, e') = 0 \). It is assumed that the mean vector is subtracted from the observation vectors, i.e. \( \mu = 0 \). It is also assumed that the common factors are independent and scaled to have unit variance, i.e. \( \text{Cov}(f, f') = I \). Therefore, the model in (5.6) attempts to model the effects of \( k \) independent variation sources that are common to all channels of data represented by \( x_i \) by separating them from channel-specific variation sources. The objective of model (5.6) then is the estimation of the matrices \( \Lambda_i \) \( (i = 1, \ldots, p) \) and vector \( f \). Since the elements of \( f \) are scaled to have unit variance, \( j \)th column of \( \Lambda_i \) indicates the strength of the \( j \)th common variation pattern on the observations from \( i \)th channel.

The structure and objective of model (5.6) are similar to the Multiple Battery Factor Analysis, where the elements of vector \( f \) are referred to as inter-battery factors and the matrices \( \Lambda_i \) \( (i = 1, \ldots, p) \) are referred to as inter-battery factor loadings. Multiple battery factor analysis is an extension of the two-battery case, Inter-Battery Factor Analysis, which was first proposed by Tucker [76] to provide information on the stability of factors over two batteries of tests. McDonald [77] generalized Tucker’s factor analysis method to multiple batteries. Based on Tucker’s work, Browne [78] derived the maximum likelihood estimates for the two-battery factor analysis and later generalized the maximum likelihood solutions to factor analysis with multiple batteries [78].

Note that the number of total parameters to be estimated in the proposed model is

\[
mk - \frac{1}{2}k(k - 1) + \frac{1}{2} \sum_{i=1}^{p} m_i (m_i - 1),
\]

which may result in a very large-size optimization problem depending on the method chosen. Therefore, the maximum likelihood estimates based on Gauss-Seidel algorithm proposed by Browne [78] will be used here in the estimation of factor loading.
matrices. It is assumed that the vector of inter-battery factors, \( f \), is normally distributed with zero-mean and covariance matrix \( \Phi = I \) implying that the common factors are assumed to be orthogonal. It is also assumed that \( y_{i} \sim N(0, \Theta_i) \) and \( e_{i} \sim N(0, Y_i) \), \((i = 1, \ldots, p)\). A consequence of this set of assumptions is that \( \mathbf{x} \) follows multivariate normal distribution with zero mean and \( m \times m \) covariance vector described as

\[
\Sigma = \Lambda \Lambda' + \Psi
\]  

(5.7)

where \( \Lambda = [\Lambda_1', \Lambda_2', \ldots, \Lambda_p'] \). This equation is the same as equation (2.5) except that the \( m \times m \) matrix \( \Psi \) is assumed to have the following block-diagonal form instead of the diagonal form in standard factor analysis:

\[
\Psi = \begin{bmatrix}
\Psi_{11} & 0 & \cdots & 0 \\
0 & \Psi_{22} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \Psi_{pp}
\end{bmatrix}
\]  

(5.8)

where \( \Psi_{ii} = \Gamma_i \Theta_i \Gamma_i' + Y_{ii}, \ (i = 1, \ldots, p) \). Due to the block-diagonal form of \( \Psi \), the identification problem of the standard factor analysis becomes more complicated in inter-battery factor analysis. A possible approach is to restrict the possible transformations of the loadings matrices to those that are permissible in standard factor analysis.

Let \( S \) be an unbiased estimate of the covariance matrix \( \Sigma \) obtained from \( N \) independent observations

\[
S = \begin{bmatrix}
S_{11} & S_{12} & \cdots & S_{1p} \\
S_{21} & S_{22} & \cdots & S_{2p} \\
\vdots & \vdots & \ddots & \vdots \\
S_{p1} & S_{p2} & \cdots & S_{pp}
\end{bmatrix}
\]  

(5.9)

Then \((N - 1)S\) has a Wishart distribution with \( N \)- degrees of freedom and parameter matrix \( \Sigma \).

The maximum likelihood estimates of \( \Lambda \) and \( \Psi \) may be derived by the minimization of the maximum likelihood discrepancy function.
\[ F(\Lambda, \Psi) = \ln|\Sigma| - \ln|\Sigma| + \text{tr}\left[\Sigma^{-1}\right] - p \]  

(5.5)

To minimize (5.5) the following equations must be satisfied [77]:

\[ \frac{\partial F}{\partial \Lambda} = 2 \Sigma^{-1} (S - \Sigma) \Sigma^{-1} \Lambda = 0 \]  

(5.10)

\[ \frac{\partial F}{\partial \Psi} = B_{\text{diag}} \left[ \Sigma^{-1} (S - \Sigma) \Sigma^{-1} \right] = 0 \]  

(5.11)

where \( B_{\text{diag}}[X] \) is the block diagonal matrix formed from the principal submatrices of \( X \) corresponding to \( \Psi_i \) (\( i = 1, \ldots, p \)).
REFERENCES


[58] Huang, Q. “Principal Curve Regression and Analysis of Multiple Curve Data”, 2004, *Unpublished*.


