Linköping Studies in Science and Technology
Thesis No. 1253

Processing of the Phonocardiographic Signal – Methods for the Intelligent Stethoscope

Christer Ahlström

Linköping University
INSTITUTE OF TECHNOLOGY

LiU-TEK-LIC-2006: 34

Department of Biomedical Engineering
Linköpings universitet, SE-58185 Linköping, Sweden
http://www.imt.liu.se

In cooperation with Biomedical Engineering,
Örebro County Council, Sweden

Linköping, May 2006

Processing of the Phonocardiographic Signal –
Methods for the Intelligent Stethoscope

Department of Biomedical Engineering
Linköpings universitet
SE-58185 Linköping
Sweden

ISBN: 91-85523-59-3 ISSN: 0280-7971

Printed in Sweden by LiU-Tryck, Linköping 2006

Abstract

Phonocardiographic signals contain bioacoustic information reflecting the operation of the heart. Normally there are two heart sounds, and additional sounds indicate disease. If a third heart sound is present it could be a sign of heart failure whereas a murmur indicates defective valves or an orifice in the septal wall. The primary aim of this thesis is to use signal processing tools to improve the diagnostic value of this information. More specifically, three different methods have been developed:

- A nonlinear change detection method has been applied to automatically detect heart sounds. The first and the second heart sounds can be found using recurrence times of the first kind while the third heart sound can be found using recurrence times of the second kind. Most third heart sound occurrences were detected (98 %), but the amount of false extra detections was rather high (7 % of the heart cycles).
- Heart sounds obscure the interpretation of lung sounds. A new method based on nonlinear prediction has been developed to remove this undesired disturbance. High similarity was obtained when comparing actual lung sounds with lung sounds after removal of heart sounds.
- Analysis methods such as Shannon energy, wavelets and recurrence quantification analysis were used to extract information from the phonocardiographic signal. The most prominent features, determined by a feature selection method, were used to create a new feature set for heart murmur classification. The classification result was 86 % when separating patients with aortic stenosis, mitral insufficiency and physiological murmurs.

The derived methods give reasonable results, and they all provide a step forward in the quest for an intelligent stethoscope, a universal phonocardiography tool able to enhance auscultation by improving sound quality, emphasizing abnormal events in the heart cycle and distinguishing different heart murmurs.

List of Publications

This thesis is based on three papers, which will be referred to in the text by their roman numerals.

I. Ahlstrom C, Liljefelt O, Hult P, Ask P: Heart Sound Cancellation from Lung Sound Recordings using Recurrence Time Statistics and Nonlinear Prediction. IEEE Signal Processing Letters. 2005. 12:812-815.

II. Ahlstrom C, Hult P, Ask P: Detection of the 3rd Heart Sound using Recurrence Time Statistics. Proc. 31st IEEE Int. Conf. on Acoustics, Speech and Signal Processing, Toulouse, France, 2006.

III. Ahlstrom C, Hult P, Rask P, Karlsson J-E, Nylander E, Dahlström U, Ask P: Feature Extraction for Systolic Heart Murmur Classification. Submitted.

Related publications not included in the thesis.

- Ahlstrom C, Johansson A, Hult P, Ask P: Chaotic Dynamics of Respiratory Sounds. Chaos, Solitons and Fractals. 2006. 29:1054-1062.
- Johansson A, Ahlstrom C, Länne T, Ask P: Pulse wave transit time for monitoring respiration rate. Accepted for publication in Medical & Biological Engineering & Computing. 2006.
- Ahlstrom C, Johansson A, Länne T, Ask P: Non-invasive Investigation of Blood Pressure Changes using Pulse Wave Transit Time: a novel approach in the monitoring of dialysis patients. Journal of Artificial Organs. 2005. 8:192-197.
- Ahlstrom C, Hult P, Ask P: Thresholding Distance Plots using True Recurrence Points. Proc. 31st IEEE Int. Conf. on Acoustics, Speech and Signal Processing, Toulouse, France, 2006.
- Ahlstrom C, Hult P, Ask P: Wheeze analysis and detection with non-linear phase space embedding. Proc. 13th Nordic Baltic Conf. in Biomedical Eng. and Med. Physics, Umeå, 2005.
- Hult P, Ahlstrom C, Rattfält L, Hagström C, Pettersson NE, Ask P: The intelligent stethoscope. 3rd European Med. Biol. Eng. Conf., Prague, Czech Republic, 2005

- Ahlstrom C, Johansson A, Länne T, Ask P: A Respiration Monitor Based on Electrocardiographic and Photoplethysmographic Sensor Fusion. Proc. 26th Ann. Int. Conf. IEEE Eng. Med. Biol., San Francisco, US, 2004.

The following M. Sc. theses have also contributed to this thesis.

- Hasfjord F: Heart Sound Analysis with Time-Dependent Fractal Dimensions.
  LiU-IMT—EX—358, 2004
- Nilsson E: Development of an Application for Visualising Heart Sounds.
  LiTH-IMT/FMT20-EX—04/397—SE, 2004
- Liljefeldt O: Heart Sound Cancellation from Lung Sounds using Non-linear Prediction.
  LiTH-IMT/FMT20-EX—05/388—SE, 2005

Preface

The intelligent stethoscope has occupied my mind for three years now, this dragon which consumes my time and drowns me in endless riddles. Many times have I looked at its mysteries in despair, but once and again the dusk disperses. Perhaps you can liken the process with the butterfly effect, where a butterfly flapping its wings over the beautiful island of Gotland can cause a hurricane in Canada. Similarly, the seed of an idea can be planted in the most unlikely ways; while hanging on the edge of a cliff or when biking along a swaying trail, while waiting in line at the local grocery store and sometimes even at work. A fragment of a thought suddenly starts to make sense, the idea tries to break free but gets lost and sinks into nothingness. Somewhere in a hidden corner the seed lay fallow, waiting for a new opportunity to rise. This could happen any day, any week or any year. In the end, you can only hope that the idea is unleashed while it still makes sense. After all, what would I write in my book if the seed decided not to grow?

Linköping, April 2006

Acknowledgements

To all of my friends, thank you for still being my friends. Especially Markus, who kept up with my complaints, and Jonas, for telling me when I was working too much (I believe the exact phrase was: “You’re about as funny as a genital lobotomy”).

My mother, father, brother and sister, the mainstay of my life, I’ll try to visit you more often now when this work is behind me.

All colleagues at the Department of Biomedical Engineering, particularly Amir, my gangsta brotha in arms, my faithful office-mate and my personal music provider. Your company has been most appreciated over the last couple of years. Linda, I owe you more than you’d like to admit, thanks again!

Anders Brun and Eva Nylander helped proof-reading the manuscript.

My supervisors; Per Ask, Peter Hult and Anders Johansson. Especially Per for having faith in my ideas, Peter for introducing me to the intelligent stethoscope and Anders for guiding me in scientific methodology and research ethics, for endless patience and most of all, for also being a friend.

Many colleagues have helped me in my works; your contributions will not be forgotten. To name but a few, Olle Liljefeldt, Fredrik Hasfjord, Erik Nilsson, Jan-Erik Karlsson, Peter Rask, Birgitta Schmekel, Christina Svensson, Björn Svensson, AnnSofie Sommer, Ulf Dahlström, January Gnitecki, Per Sveider, Bengt Ragnemalm, Solveig Carlsson, Susanne Skytt and Nils-Erik Pettersson with staff at Biomedical Engineering, Örebro County Council.

The artwork on the cover was kindly provided by Nancy Munford. The idea behind the image comes from the heart muscle being spiral in shape, but as it turns out, also the flow in the heart advances along spiral pathways (or rather, vortices or eddies). With a little good will it looks a bit like a signal moving around in its state space as well.

Finally, my beloved Anneli, thank you for letting me use your time.

This work was supported by grants from the Swedish Agency for Innovation Systems, the Health Research Council in the South-East of Sweden, the Swedish Research Council, the Swedish National Centre of Excellence for Non-invasive Medical Measurements and CORTECH (Swedish universities in cooperation for new cardiovascular technology).

Abbreviations

Abbreviations have been avoided as much as possible, but every now and then they tend to sneak in anyhow.

AR	Autoregressive
ARMA	Autoregressive moving average
AS	Aortic stenosis
MA	Moving average
MI	Mitral insufficiency
PM	Physiological murmur
ROC	Receiver operating curve
RP	Recurrence plot
RQA	Recurrence quantification analysis
S1	The first heart sound
S2	The second heart sound
S3	The third heart sound
S4	The fourth heart sound
ST	Stockwell transform
STFT	Short time Fourier transform
TFR	Time Frequency Representation
VFD	Variance fractal dimension
WT	Wavelet transform

Table of Contents

1. Introduction

“The way to the heart is through the ears.”
Katie Hurley

The history of auscultation, or listening to the sounds of the body, is easily described by a few evolutionary leaps. Hippocrates (460-377 BC) provided the foundation for auscultation when he put his ear against the chest of a patient and described the sounds he could hear from the heart. The next leap was made by Robert Hooke (1635-1703) who realized the diagnostic use of cardiac auscultation:

"I have been able to hear very plainly the beating of a man's heart…Who knows, I say, but that it may be possible to discover the motion of the internal parts of bodies…by the sound they make; one may discover the works performed in several offices and shops of a man's body and thereby discover what instrument is out of order."

The biggest breakthrough in auscultation came in 1816 when René Laennec (1781-1826) invented the stethoscope. Laennec was about to examine a woman with the symptoms of heart disease, but due to her sex and age, direct auscultation was inappropriate. Also percussion and palpation gave little information on account of the patient’s obesity [1]. Consequently, Laennec used a roll of paper to avoid physical contact during the examination. As a spin-off, he found that heart and lung sounds were amplified and previously unheard sounds emerged. The invention of the stethoscope resulted in, without precedent, the most widely spread diagnostic instrument in the history of biomedical engineering. The stethoscope has evolved over the years, but the underlying technology remains the same. Attempts have been made to take the stethoscope into the IT age, but the success has so far been limited. A selection of stethoscopes from different eras is presented in Figure 1.

Mechanical working processes in the body produce sounds which indicate the health status of the individual. This information is valuable in the diagnosis of patients, and it has been widely used since the days of Hippocrates. In modern health care, auscultation has found its primary role in primary or in home health

care, when deciding which patients need special care. The most important body sounds are heart sounds and lung sounds, but sounds from swallowing, micturition, muscles and arteries are also of clinical relevance. The main sources for production of body sounds are acceleration or deceleration of organs or fluids, friction rubs and turbulent flow of fluids or gases.

The auscultatory skills amongst physicians demonstrate a negative trend. The loss has occurred despite new teaching aids such as multimedia tutorials, and the reasons are the availability of new diagnostic tools such as echocardiography and magnetic resonance imaging, a lack of confidence and increased concern about litigations [2]. The art of auscultation is often described as quite difficult, partly because of the fact that only a portion of the cardiohemic vibrations are audible, see Figure 2.

Figure 1. Early monaural stethoscopes (top left), Cummanns and Allisons stethoscopes (lower left), a modern binaural stethoscope (middle) and a modern electronic stethoscope, Meditron M30 (right).

Acoustic stethoscopes transmit sound mechanically from a chest-piece via air-filled hollow tubes to the listener's ears. The diaphragm and the bell work as two filters, transmitting higher frequency sounds and lower frequency sounds, respectively. Electronic stethoscopes function in a similar way, but the sound is converted to an electronic signal which is transmitted to the listener by wire. Functionalities often included in electronic stethoscopes are amplification of the signal, filters imitating the function of the diaphragm and the bell and in some cases recording abilities to allow storage of data.

Heart sounds and murmurs are of relatively low intensity and are band limited to about 10–1000 Hz. Meanwhile, the human hearing is adapted to speech. This explains why physicians sometimes have easier to detect heart sounds by palpation than by auscultation. Part of this problem can be avoided by amplification of the sound, but much information contained in the phonocardiographic signal is hard to reach using acoustic stethoscopes as well as electronic stethoscopes. An intelligent stethoscope could make use of this extra information. Possible tasks for an intelligent stethoscope are to classify different murmurs, to detect differences in the heart sounds (such as the splitting of the second heart sound) and to detect additional sounds (such as the third heart sound and ejection clicks).

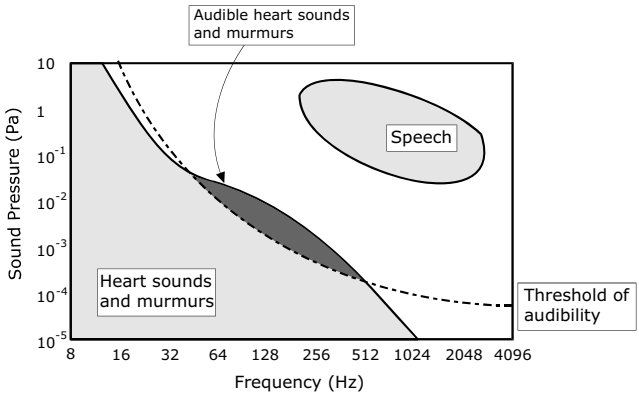

Figure 2. The frequency content of heart sounds and murmurs in relation to the human threshold of audibility. Drawn from [3]. Note that without amplification, the area representing the audible part of the phonocardiographic signal is very small.

50-80% of the population has murmurs during childhood, whereas only about 1% of the murmurs are pathological [2]. A simple tool able to screen murmurs would be both time- and cost-saving while relieving many patients from needless anxiety. The sensor technology in such a tool is basically simple and the parameters obtained are directly related to mechanical processes within the body (in contrast to ECG which measures the heart's electrical activity). In the new field of telemedicine and home care, bioacoustics is definitely a suitable method.

1.1. Aim of the Thesis

The all-embracing goal of bioacoustic research is to establish a relationship between mechanical events within the body and the sounds these events give rise to. The medical use of this knowledge is of course to link sounds that diverge from normality to certain pathological conditions. Clearly, there is valuable information hidden in the bioacoustic signal, and the primary aim of this thesis is to make use of signal processing tools to emphasize and extract this information. Clinical value has been a guiding-star throughout the work, and the goal has been to develop methods for an intelligent stethoscope. More specifically, the aims of this thesis were to:

- identify and compare signal analysis tools suitable for phonocardiographic signals.
- emphasize audibility of bioacoustic signals (Paper I emphasizes lung sounds by removal of heart sounds).
- extract specific components in the phonocardiographic signal (Paper II describes a method able to detect the third heart sound).
- extract information suitable for classification of heart diseases (Paper III presents a feature set for classification of systolic murmurs).

1.2. Thesis Outline

The thesis consists of two parts. The first part, chapter 1-7, contains an introduction while the second part consists of three papers.

Chapter 1 introduces the problem at hand and defines the aim of the thesis.

Chapter 2 contains a brief summary on the physiology of the heart and the origin of the phonocardiographic signal.

Chapter 3 describes methods for nonstationary and nonlinear time series analysis and introduces dynamical systems theory within the context of signal processing.

Chapter 4 covers the characteristics of phonocardiographic signals in different domains.

Chapter 5 discusses applications of phonocardiographic signal processing using the methods and results from chapter 3 and 4.

Chapter 6 contains discussion and conclusions, including notes on future work.

Chapter 7 is a review of the papers presented in the second part of the thesis.

2. Preliminaries on Heart Sounds and Heart Murmurs

"The heart is of such density that fire can scarcely damage it."
Leonardo da Vinci (1452-1519)

This chapter sets the scene for up-coming sections. The physics of sound is introduced followed by a review of the operation of the heart and the associated terminology. The genesis of heart sounds and heart murmurs is discussed and finally a short presentation of auscultation techniques and signal acquisition is given.

2.1. Physics of Sound

It would be a mistake for a study on heart sounds to leave out an introduction to the acoustic phenomena where everything actually starts. A sound is generated by a vibrating object and propagates as waves of alternating pressure. The vibrating source sets particles in motion, and if the sound is a pure tone, the individual particle moves back and forth with the frequency of that tone. Each particle is thus moving around its resting point, but as it pushes nearby particles they are also set in motion and this chain effect results in areas of compression and rarefactions. The alternating areas of compression and rarefaction constitute a pressure wave that moves away from the sound source, see Figure 3. These pressure variations can be detected via the mechanical effect they exert on some membrane (the diaphragm of the stethoscope, the tympanic membrane etc.). If the sound source vibrates in a more irregular manner, the resulting sound wave will be more complicated. Usually, sound is described by its intensity, duration, frequency and velocity [4]. If the sound is nonstationary, these measures have to be time varying to give relevant information. Time varying analysis techniques are described in section 3.1.

The number of vibrations per second, or frequency, is a physical entity. What humans perceive as frequency is however called pitch. The two are closely related, but the relationship is not linear. Up to 1 kHz, the measured frequency and the perceived pitch are fairly the same. Above 1 kHz, a larger increase in frequency is required to create an equal perceived change in pitch.

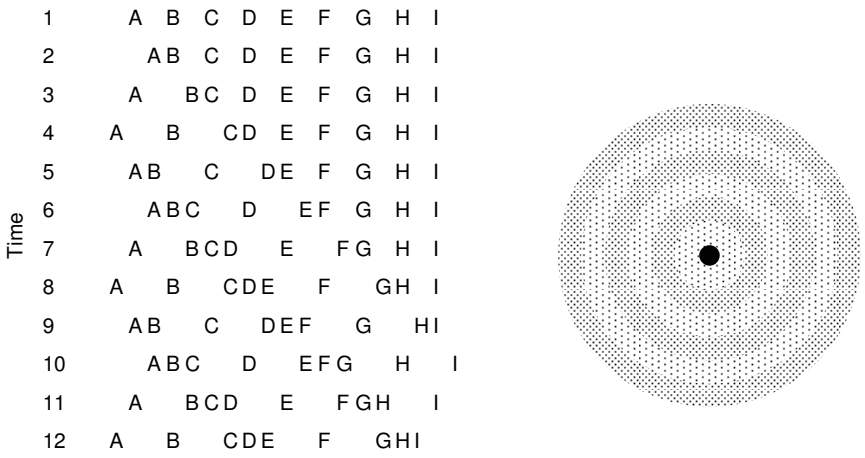

Figure 3. The left figure is a schematic drawing of nine particles in simple harmonic motion at twelwe moments in time. The sound source is located on the left side and the pressure wave, indicated by clustering of three adjacent particles, moves from left to right. Note that each particle moves relatively little around a rest position. In the right figure a pressure wave emanating from a sound source (black circle) is illustrated. Drawn from [4].

2.2. Physiology of the Heart

The primary task of the heart is to serve as a pump propelling blood around the circulatory system. When the heart contract, blood is forced through the valves, from the atria to the ventricles and eventually out through the body, see Figure 4.

There are four heart chambers; right and left atria and right and left ventricles. The two atria mainly act as collecting reservoirs for blood returning to the heart while the two ventricles act as pumps to eject the blood to the body. Four valves prevent backflow of blood; the atrioventricular valves (the mitral and tricuspid valve) prevent blood from flowing back from the ventricles to the atria and the semilunar valves (aortic and pulmonary valves) prevent blood from flowing back into the ventricles once being pumped into the aorta and the pulmonary artery. Deoxygenated blood from the body enters the right atrium, passes into the right ventricle and is ejected into the pulmonary artery on the way to the lungs. Oxygenated blood from the lungs re-enter the heart in the left atrium, passes into the left ventricle and is then ejected into the aorta.

The blood pressure within a chamber increases as the heart contracts, generating a flow from higher pressure areas towards lower pressure areas. During the rapid filling phase (atrial and ventricular diastole), venous blood from the body and from the lungs enters the atria and flows into the ventricles. As the pressure gradient between the atria and the ventricles level out (reduced filling phase), a final volume of blood is forced into the ventricles by atrial contraction (atrial systole). In the beginning of ventricular systole, all the valves are closed resulting in an isovolumic contraction. When the pressure in the ventricles exceeds the pressure

in the blood vessels, the semilunar valves open allowing blood to eject out through the aorta and the pulmonary trunk. As the ventricles relax the pressure gradient reverses, the semilunar valves close and a new heart cycle begins.

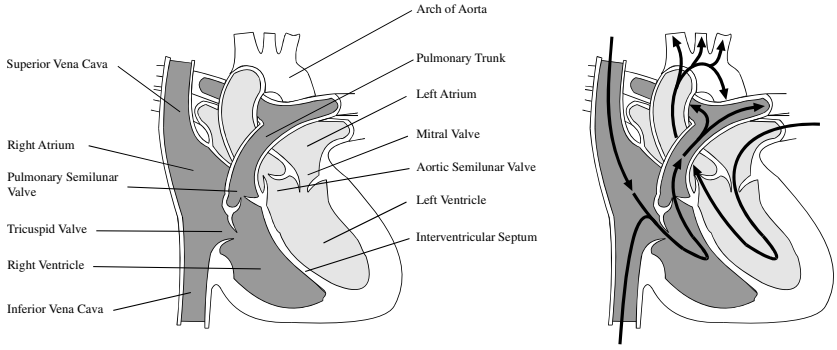

Figure 4. Anatomy of the heart (left) and the blood flow pathways through left and right heart (right).

2.3. Heart Sounds

The relationship between blood volumes, pressures and flows within the heart determines the opening and closing of the heart valves. Normal heart sounds occur during the closure of the valves, but how they are actually generated is still debated. The valvular theory states that heart sounds emanate from a point sources located near the valves, but this assumption is probably an oversimplification [5]. In the cardiohemic theory the heart and the blood represent an interdependent system that vibrates as a whole [5]. Both these theories originate from a time when the physiological picture was based on a one-dimensional conception of flow. Recent research provides means to visualize the actual three-dimensional flow patterns in the heart [6], and this new knowledge will probably clarify our view on the underlying mechanisms of heart sounds. An example of a visualisation technique called particle trace is shown in Figure 5. The blood's pathway through the heart is far from fully understood, but the induced vortices seem optimized to facilitate flow and thereby increase the efficiency of the heart as a pump. The impact of this new knowledge on the understanding of heart sounds and their origin is yet to be investigated. Awaiting this new insight, the cardiohemic theory will be assumed valid.

Normally, there are two heart sounds, see Figure 6. The first sound (S1) is heard in relation to the closing of the atrioventricular valves, and is believed to include four major components [3]. The initial vibrations occur when the first contraction of the ventricle move blood towards the atria, closing the AV-valves. The second component is caused by the abrupt tension of the closed AV-valves, decelerating the blood. The third component involves oscillation of blood between the root of the aorta and the ventricular walls, and the fourth component represents the

vibrations caused by turbulence in the ejected blood flowing into aorta. The second sound (S2) signals the end of systole and the beginning of diastole, and is heard at the time of the closing of the aortic and pulmonary valves [7]. S2 is probably the result of oscillations in the cardiohemic system caused by deceleration and reversal of flow into the aorta and the pulmonary artery [5].

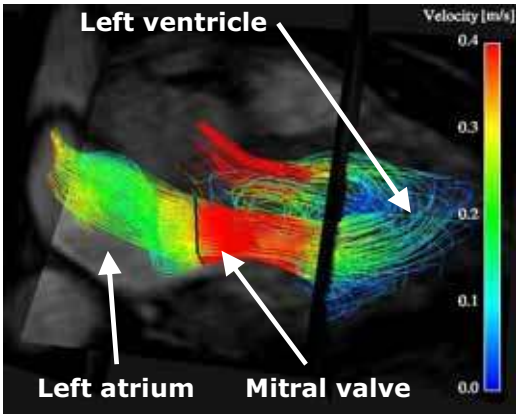

Figure 5. Particle trace (path line) visualization of intra-cardiac blood flow. The colour coding represents velocity according to the legend to the right. The image was adapted from [6].

There is also a third and a fourth heart sound (S3 and S4). They are both connected with the diastolic filling period. The rapid filling phase starts with the opening of the semilunar valves. Most investigators attribute S3 to the energy released with the sudden deceleration of blood that enters the ventricle throughout this period [8]. A fourth heart sound may occur during atrial systole where blood is forced into the ventricles. If the ventricle is stiff, the force of blood entering the ventricle is more vigorous, and the result is an impact sound in late diastole, S4 [7]. There are also sounds such as friction rubs and opening snaps, but they will not be treated further.

2.4. Heart Murmurs

Murmurs are produced by turbulent blood flow as a result of narrowing or leaking valves or from the presence of abnormal passages in the heart. More specifically, heart murmurs occur when the blood flow is accelerated above the Reynolds number. The resulting blood flow induces non-stationary random vibrations, which are transmitted through the cardiac and thoracic tissues up to the surface of the thorax. There are five main factors involved in the production of murmurs [7]:

- High rates of flow through the valves.
- Flow through a constricted valve (stenosis).

- Backward flow through an incompetent valve (insufficiency or regurgitation).
- Abnormal shunts between the left and right side of the heart (septal defects).
- Decreased viscosity, which causes increased turbulence.

Heart murmurs are graded by intensity from I to VI. Grade I is very faint and heard only with special effort while grade VI is extremely loud and accompanied by a palpable thrill. Grade VI murmurs are even heard with the stethoscope slightly removed from the chest. When the intensity of systolic murmurs is crescendo-decrescendo shaped and ends before one or both of the components of S2, it is assumed to be an ejection murmur (S2 is composed of two components, one from the aortic valve and one from the pulmonary valve). Murmurs due to backward flow across the atrioventricular valves are of more even intensity throughout systole and reach one or both components of S2. If the regurgitant systolic murmur starts with S1 it is called holosystolic and if it begins in mid- or late systole it is called a late systolic regurgitant murmur. Besides murmurs, ejection clicks might also be heard in systole. They are often caused by abnormalities in the pulmonary or aortic valves. Different murmurs, snaps, knocks and plops can also be heard in diastole, but such diastolic sounds are beyond the scope of this thesis.

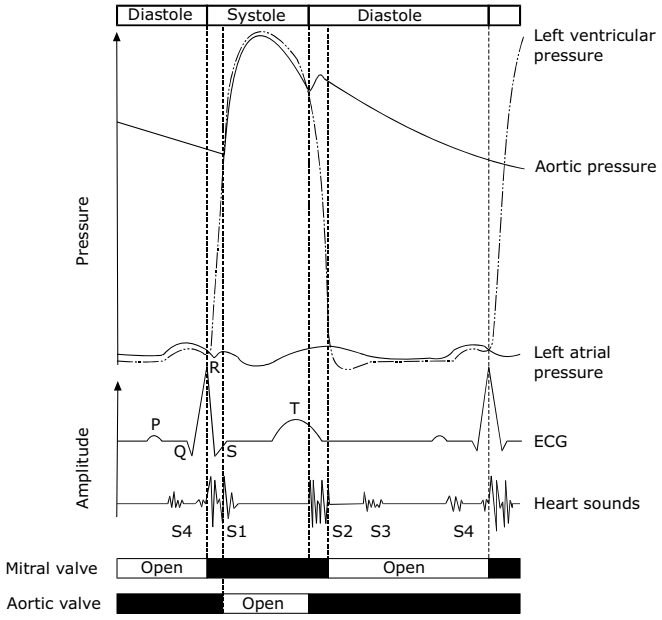

Figure 6. The four heart sounds in relation to various hemodynamic events and the ECG. All units are arbitrary.

2.5. Auscultation and the Phonocardiogram

Auscultation is the technical term for listening to the internal sounds of the body. The loudness of different components varies with the measurement location. For instance, when listening over the apex, S1 is louder than S2. Also, the location of a heart murmur often indicates its origin, e.g. mitral valve murmurs are usually loudest at the mitral auscultation area. The traditional areas of auscultation, see Figure 7, are defined as [7]:

- *Mitral area:* The cardiac apex.
- *Tricuspid area:* The fourth and fifth intercostal space along the left sternal border.
- *Aortic area:* The second intercostal space along the right sternal border.
- *Pulmonic area:* The second intercostal space along the left sternal border.

Even though the definition of these areas came to life long before we had much understanding of the physiology of the heart, they are still good starting points. Revised areas of auscultation, allowing more degrees of freedom, have however been adopted [7].

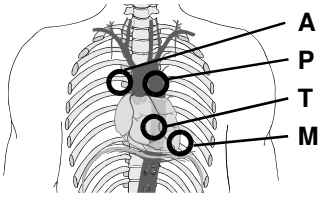

Figure 7. Traditional areas of auscultation (M refers to the mitral area, T the tricuspid area, P the pulmonic area, and A the aortic area).

A graphical printing of the waveform of cardiac sounds is called a phonocardiogram, PCG. An example of a phonocardiogram was shown in Figure 6. To obtain the phonocardiogram, a microphone is placed on the patient's chest and the signal is plotted on a printer, similar to the ones used in ECG recordings. This technique promotes visual analysis of cardiac sounds, thus allowing thorough investigation of temporal dependencies between mechanical processes of the heart and the sounds produced.

2.6. Acquisition of Phonocardiographic Signals

The audio recording chain involves a sequence of transformations of the signal: a sensor to convert sound or vibrations to electricity, a pre-amplifier to amplify the signal, a prefilter to avoid aliasing and an analogue to digital converter to convert the signal to digital form which can be stored permanently. In the setting of the intelligent stethoscope, this chain is complemented with an analysis step and an information presentation step.

2.6.1. Sensors

Microphones and accelerometers are the natural choice of sensor when recording sound. These sensors have a high-frequency response that is quite adequate for body sounds. Rather, it is the low-frequency region that might cause problems [9]. There are mainly two different kinds of sensors, microphones and accelerometers. The microphone is an air coupled sensor that measure pressure waves induced by chest-wall movements while accelerometers are contact sensors which directly measures chest-wall movements [10]. For recording of body sounds, both kinds can be used. More precisely, condenser microphones and piezoelectric accelerometers have been recommended [11].

Electronic stethoscopes make use of sensors specially designed to suit cardiac sounds. Compared to classic stethoscopes, electronic stethoscopes tries to make heart and lung sounds more clearly audible using different filters and amplifiers. Some also allow storage and the possibility to connect the stethoscope to a computer for further analysis of the recorded sounds. The leading suppliers of electronic stethoscopes are Thinklabs, Welch-Allyn and 3M. Thinklabs uses a novel electronic diaphragm detection system to directly convert sounds into electronic signals. Welch-Allyn Meditron uses a piezo-electric sensor on a metal shaft inside the chest piece, while 3M uses a conventional microphone.

The studies included in this thesis have used two different sensors; the Siemens Elema EMT25C contact accelerometer and an electronic stethoscope from Welch-Allyn Meditron (the Stethoscope, Meditron ASA, Oslo, Norway).

2.6.2. Pre-processing, digitalization and storage

The preamplifier amplifies low level signals from transducer to line level. This is important to be able to use the full range of the analogue to digital converter and thus minimizing quantization errors. Another matter concerning the digitalization of signals is aliasing which will occur unless the Nyquist-Shannon sampling theorem is fulfilled.

In this work, when using EMT25C, a custom-built replica of a phonocardiography amplifier (Siemens Elema E285E) was used. This amplifier includes a lowpass filter with a cut-off frequency of 2 kHz. The signal was digitized with 12-bits per sample using analogue to digital converters from National Instruments (see paper I and II for specific details). Acquisition of the data was conducted in a Labview-application (National Instruments, Austin, Texas, US) after which the data were stored on a personal computer.

For the electronic stethoscope, the matching acquisition equipment and software were used (Analyzer, Meditron ASA, Oslo, Norway). According to the manufacturer, the digital recordings are stored without pre-filtering. An excessive sampling frequency of 44.1 kHz was thus used to avoid aliasing (and with the idea of post-filtering in mind). The signals were stored in a database on a personal computer. This approach was used in paper III.

3. Signal Analysis Framework

"Calling a science nonlinear is like calling zoology the study of non-human animals."
Stanislaw Ulam

The underlying assumption of many signal processing tools is that the signals are Gaussian, stationary and linear. This chapter will introduce methods suitable for analysing signals that do not fall into these categories. Two short examples will precede the theoretical treatment to illustrate the problem at hand.

Example 1, Characteristics that vary with time: A sinusoid with changing mean, amplitude and frequency is shown in Figure 8. Using the Fourier transform to investigate the signal's frequency content, it can be seen that the signal contains frequencies up to about 35 Hz. However, much more information can be obtained by investigating how the frequency content varies over time. Methods able to investigate how a certain signal property varies over time are suitable for nonstationary signal analysis. A number of such methods are introduced in section 3.1.

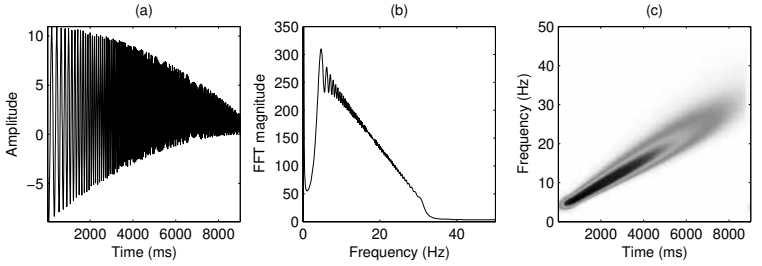

Figure 8. A sinusoid with changing mean, amplitude and frequency plotted as a waveform in the time domain (a), as a frequency spectrum in the frequency domain (b) and in a combined time-frequency domain (c).

Example 2, Distinguishing signals with similar spectra (example adapted from [12]): In many traditional linear methods it is assumed that the important signal characteristics are contained in the frequency power spectrum. From a stochastic process perspective, the first and second order statistics of the signal are

represented by this power spectral information. However, there are many types of signals, both theoretical and experimental, for which a frequency domain representation is insufficient to distinguish two signals from each other. For example, signals generated through nonlinear differential or difference equations typically exhibit broadband spectral characteristics that are difficult to interpret and compare. Two signals with indistinguishable power spectra are presented in Figure 9.

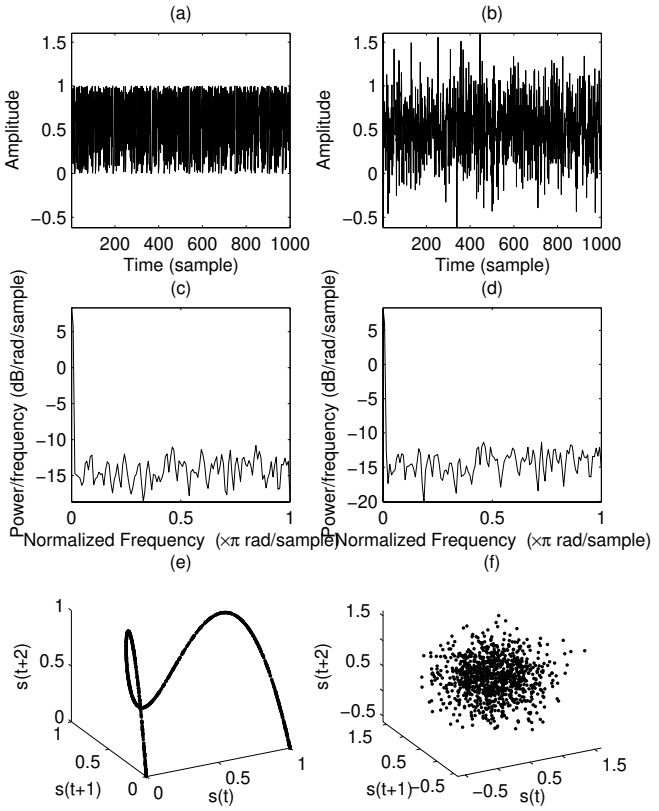

Figure 9. The logistic map, s(t+1) = c•s(t)[1-s(t)], where c is a constant, is often used as a model in population studies. Here a logistic map with c=4 and s(0) = 0.1 is presented in (a) and a phase randomized correspondence is shown in (b). Their respective frequency spectra, which are almost identical, are shown in (c) and (d). Finally, in (e) and (f) their corresponding phase portraits are shown.

The first signal is the logistic map and the second signal is its phase randomized correspondence. Even though they have the same (but rather obscure) frequency spectrum, the logistic map has structure in its phase portrait while the phase

randomized signal does not (the phase portrait is just the signal plotted against a time delayed version of itself). To distinguish between the two, or for that matter, to find the structure in the logistic map, it is obviously not enough to study their spectra. Methods for nonlinear signal analysis will be introduced in section 3.3.

3.1. Measuring Characteristics that Vary in Time

Much information can be gained by investigating how certain signal properties vary over time. Some properties possess natural time dependence, such as the envelope when calculated as the energy of a signal. In other cases, there is a typical trade-off between time resolution and accuracy in the sought parameter (the typical example being the short time Fourier transform, STFT). The property under investigation is then either a function of some quantity other than time (such as frequency) or a scalar value (such as the mean value of the data). In either way, time dependence can be introduced by calculating the wanted characteristic with a sliding window.

3.1.1. Intensity

The textbook approach to extract a signal's envelope, E(t), is via the analytic signal [13, 14]. The continuous analytic signal is composed by the original signal and its Hilbert transform according to equation (3.1) where *H(t)* is the Hilbert transform (equation (3.2)).

$$s_A(t) = s(t) + i \cdot s_H(t) \tag{3.1}$$

$$s_H(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{s(\tau)}{\tau - t} d\tau \tag{3.2}$$

The Hilbert transform can be interpreted as a convolution between the signal and -1/πt, or as a rotation of the argument with π/2 for positive frequencies and –π/2 for negative frequencies. Similarly, the analytic signal can be obtained by removing the negative frequencies and multiplying the positive frequencies by two [13]. Amongst several interesting properties of the analytic signal, the desired envelope can be found as:

$$\text{Analytic signal envelope: } s_A(t) = \sqrt{|s(t)|^2 + |s_H(t)|^2} \tag{3.3}$$

Other envelope-like measures are the absolute value or the square of the signal, see equations (3.4)-(3.5). The absolute value gives equal weight to all samples regardless of the signal's intensity. The energy (square) on the other hand colours the measure by emphasizing higher intensities compared to lower intensities. Two other envelope-like measures are the Shannon entropy and the Shannon energy [15], see equation (3.6)-(3.7). These measures give greater weight to medium intensity signal components, thus attenuating low intensity noise and high intensity disturbances. A practical issue with these approaches is that the envelope becomes rather jagged. This is usually dealt with by low-pass filtering *E(t)* [13, 15]. A

method developed by Teager, equation (3.8), results in an on-the-fly envelope-estimate that is very useful for analyzing signals from an energy point of view [16].

$$\text{Absolute value: } E(t) = \left|s(t)\right| \tag{3.4}$$

$$\text{Energy (square): } E(t) = s(t)^2 \tag{3.5}$$

$$\text{Shannon entropy: } E(t) = -\left|s(t)\right| \cdot \log\left|s(t)\right| \tag{3.6}$$

$$\text{Shannon energy: } E(t) = -s(t)^2 \cdot \log s(t)^2 \tag{3.7}$$

$$\text{Teager's energy operator: } E(t) = s(t)^2 - s(t-1)\,s(t+1) \tag{3.8}$$

A comparison of the methods can be seen in Figure 10. The test signal was a 200 Hz sinusoid with amplitude ranging from 0 to 1 and sampled with 10 kHz. In this comparison, all outputs (but Teager), were low pass filtered by a 5th order Butterworth filter with the cut off frequency 150 Hz.

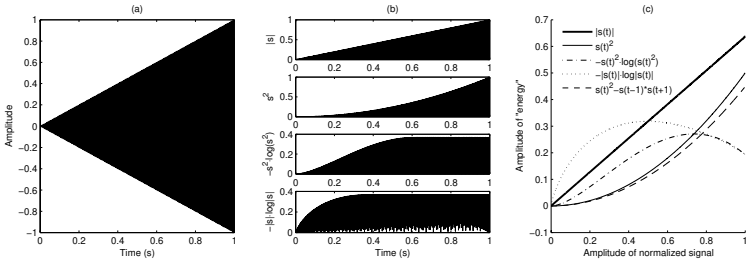

Figure 10. Comparison of different envelope estimation methods. The test signal is presented in (a) and the result of various envelope-measures are shown in (b). The smoothed envelope measures are presented in (c).

Another way to extract a signal's envelope is through homomorphic signal processing. Algebraically combined signals are here mapped into a domain where linear signal processing tools can be used [17]. The envelope can be seen as an amplitude modulation, where the slow varying envelope is multiplied with a higher frequency signal. By taking the logarithm of the signal, equation (3.9), the non-linear multiplication changes into a linear addition, equation (3.10). The low frequency contribution *l(t)* may then be extracted from the high frequency contribution *h(t)* by a low pass filter. Exponentiation takes the result back to the original signal domain, equation (3.11). If the low pass filter is properly chosen, *l(t)* will be a good estimate of the envelope.

$$s(t) = l(t)\,h(t) \tag{3.9}$$

$$\log s(t) = \log l(t) + \log h(t) \tag{3.10}$$

$$e^{LP\{\log l(t) + \log h(t)\}} = e^{LP\{\log l(t)\} + LP\{\log h(t)\}} \approx e^{LP\{\log l(t)\}} \approx l(t) \tag{3.11}$$

3.1.2. Frequency

The Fourier transform can be used to produce a time-averaged frequency spectrum. However, it is often desirable to study how the frequency content of a signal varies over time. There are many techniques available to perform such analyses, and they are generally named time-frequency representations (TFR). The simplest approach is probably the short time Fourier transform (STFT), which is merely a windowed Fourier transform:

$$STFT(m,k) = \sum_{t=1}^{N} s(t)\, w(t-m)\, e^{-2\pi i \frac{kt}{N}} \tag{3.12}$$

where w denotes the time window, m is the translation parameter and k the frequency parameter. If w is chosen as the Gaussian function in equation (3.13), the obtained TFR is called the Gabor transform [18].

$$g(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{t^2}{2\sigma^2}} \tag{3.13}$$

σ denotes the variance and is related to the width of the analyzing window. The STFT does, however, suffer from the uncertainty principle. This means that the frequency resolution decreases as the time resolution increases and the other way around (since time is the dual of frequency). One way to obtain better resolution is to use shorter windows for higher frequencies and longer windows for lower frequencies. One such approach is the wavelet transform (WT) [13]:

$$WT(m,a) = \frac{1}{\sqrt{|a|}} \sum_{t=1}^{N} s(t)\, w\left(\frac{t-m}{a}\right) \tag{3.14}$$

where m is a translation parameter and a is a scale parameter. The main idea is that any signal can be decomposed into a series of dilatations or compressions of a mother wavelet denoted $w(t)$. The mother wavelet should resemble interesting parts of the signal, and the choice is important for the results. An issue with wavelets is that the link to local frequency is lost (hence the term scale is preferred instead of frequency). A similar but phase corrected transform, able to maintain the notion of frequency, is the S-transform (ST) [19]:

$$ST(m,k) = \sum_{t=1}^{N} s(t) \frac{|k|}{\sqrt{2\pi}} e^{-\frac{(t-m)^2 k^2}{2}} e^{-2\pi i \frac{kt}{N}} \tag{3.15}$$

Compared to STFT, the window function is chosen as a Gaussian function where the variance is allowed to vary proportionally with the period of the analyzing sinusoid. When changing the variance, the width of the window is altered giving a multi-resolution description of the signal. An example comparing STFT, WT and ST is given in Figure 11.

There are many other approaches available for joint time-frequency analysis. The methods just described belong to the linear nonparametric group. The quadratic nonparametric group, the parametric group etc., will not be treated in this thesis.

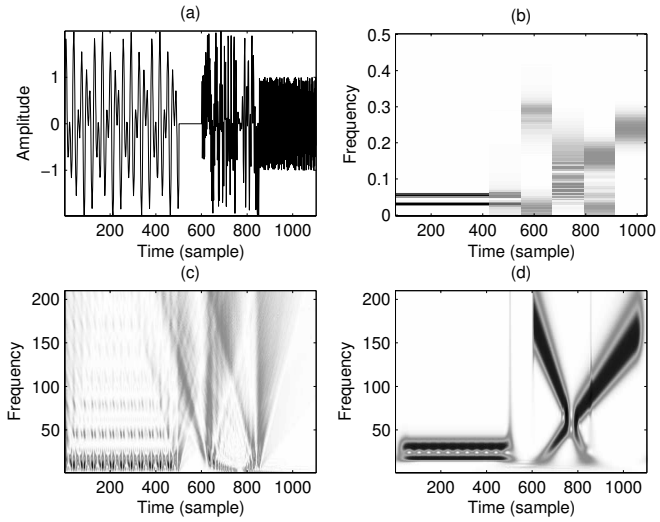

Figure 11. An example signal consisting of two well separated sinusoids (sample 1-500) and two chirps, one ascending and one descending (sample 600-1100) is given in (a). TFR-plots calculated with STFT, WT (with a Daubechie 2 mother wavelet) and ST are given in (b-d), respectively. The frequency axes are in arbitrary units.

3.2. Nonlinear Systems and Embedology

Dynamical systems theory is an important ingredient in nonlinear signal processing. A dynamical system is a system whose state changes with time. In continuous time, the system is described by differential equations and in discrete time by iterated maps. Finding explicit solutions for these equations is not the purpose of this chapter. Instead, the main goal is to improve knowledge about the system by deducing quantitative information. Since sampled data is used, only iterative maps will be considered.

The dynamics of a time discrete system is determined by its possible states in a multivariate vector space (called state space or phase space). The transitions between the states are described by vectors, and these vectors form a trajectory describing the time evolution of the system according to equation (3.16).

$$x(t+1) = \varphi(x(t)) \tag{3.16}$$

where $x(t)$ is the state of the system, t is the time index, φ is a mapping function such that $\varphi: M \rightarrow M$ and M is the true state space. A geometrical display of the trajectory, such as in Figure 12, provides a phase portrait of the system. If the

trajectory is drawn to a particular set this set is called an attractor. Examples of different attractors are given in Figure 12.

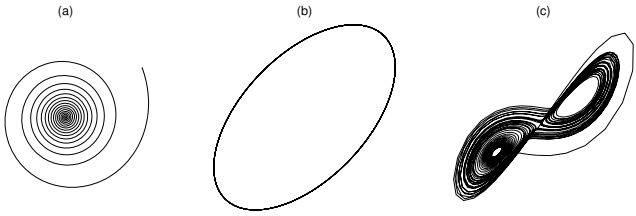

Figure 12. Examples of a fixed point attractor (a), a limit cycle (b) and a strange attractor from a Lorenz system (c). A physical example of a fix point attractor is a pendulum, where all initial states will converge to a single point. Modifying this example so that the pendulum has a driving force, thus creating a simple oscillation, a periodic attractor is obtained. Chaotic systems like the Lorenz system have been used to describe weather, and give rise to strange attractors, where the trajectories never cross or touch each other.

The true state space (*M*) thus contains the true states *x*, whose time evolution is described by the map φ, $x(t)= \varphi t(x(0))$. Now suppose that the only information available about this system is what we can find in a scalar measure $s(t)=h(x(t))$, where $h:M \to \mathbb{R}$. If *s(t)* is a projection from the true (multivariate) state space *M*, then it might be possible to undo the projection, see Figure 13. That is, given a measured signal *s(t)* in $\mathbb{R}$, is there a way to create a map from an unknown state *x(t)* in *M* to a corresponding point *y(t)* in a reconstructed state space in $\mathbb{R}^d$? Takens' theorem provides us with such a map [20]:

$$\begin{aligned} F: \quad M &\to \quad \mathbb{R}^d \\ x(t) &\to \quad y(t) = F\left(x(t)\right) = \left[s(t), s(t+\tau), ..., s\left(t+(d-1)\tau\right)\right] \end{aligned} \tag{3.17}$$

where τ is a delay parameter, *d* is the embedding dimension and *F* is the map from the true state space to the reconstructed state space. The selection of τ and *d* affects how accurately the embedding reconstructs the system's state space. These issues are important, but there are no bullet-proof ways to determine τ and *d*. In this thesis τ will be determined using average mutual information [21] and *d* will be chosen based on Cao's method [22].

What Takens actually proved was that the reconstructed state space $\mathbb{R}^d$ is a dynamical and topological equivalent to *M*. Since the dynamics of the reconstructed state space contains the same topological information as the original state space, characterization and prediction based on the reconstructed state space is as valid as if it was made in the true state space.

Table 1. Comparison of linear and nonlinear signal processing techniques. The table is adapted from [21].

Linear signal processing	Nonlinear signal processing
Finding the signal: Separate broadband noise from narrowband signal using spectral characteristics. Method: Matched filter in frequency domain.	*Finding the signal:* Separate broadband signal from broadband noise using the deterministic nature of the signal. Method: Manifold decomposition or statistics on the attractor.
Finding the space: Use Fourier space methods to turn difference equations into algebraic forms. $s(t)$ is observed and $S(f) = \sum s(t) e^{i 2\pi t f}$ is used.	*Finding the space:* Time lagged variables form coordinates for a reconstructed state space in d dimensions. $y(t) = \left[s(t), s(t+\tau), ..., s(t+(d-1)\tau) \right]$ where d and τ are determined by false nearest neighbours and mutual information.
Classify the signal: • Sharp spectral peaks • Resonant frequencies of the system	*Classify the signal:* • Lyapunov exponents • Fractal dimension measures • Unstable fixed points • Recurrence quantification • Statistical distribution of the attractor
Making models, predict: $s(t+1) = \sum \alpha_k s(t-k)$ Find parameters α_k consistent with invariant classifiers – location of spectral peaks.	*Making models, predict:* $\mathbf{y}(t) \rightarrow \mathbf{y}(t+1)$ $\mathbf{y}(t+1) = F\left[\mathbf{y}(t), a_1, a_2, \ldots, a_p \right]$ Find parameters a_j consistent with invariant classifier – Lyapunov exponents, fractal dimensions.

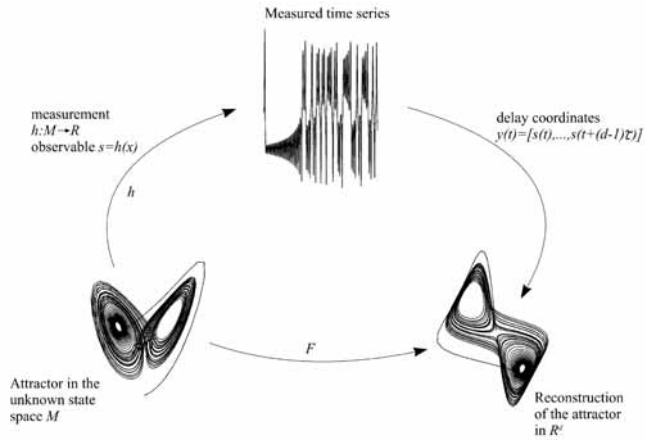

Figure 13. Delay reconstruction of states from a scalar time series (example using the Lorenz system). Redrawn from [20].

3.3. Nonlinear Analysis Tools

Nonlinear analysis tools are rather different from their linear analogue. A brief comparison between linear and nonlinear methods can be found in Table 1.

3.3.1. Non-integer dimensions

The concept of non-integer dimensions may sound abstract, but it can be intuitively motivated. For example, is the Henon map one or two dimensional? It seems two-dimensional when looking at it from a broad scale, and, it never breaks down into a one-dimensional line no matter how much it is magnified, see Figure 14. The answer to the question whether the Henon map is one or two-dimensional seems to be that it is "somewhere in between", i.e. it has a fractal dimension. Strange attractors are fractal, and their fractal dimension is less than the dimension of the state space it lives in.

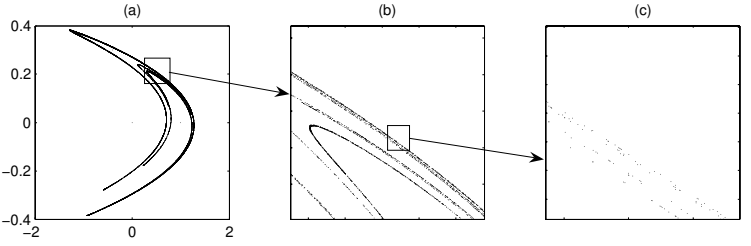

Figure 14. Zooming into the Henon map reveals new levels of complexity. No matter how much the figure is magnified it will never collapse into a one-dimensional line, nor does it fill the two-dimensional space in (a). Instead, the Henon map has a dimension somewhere in between one and two, i.e. a fractal dimension.

There are two types of approaches to estimate the fractal dimension; those that operate directly on the waveform and those that operate in the reconstructed state space. Note that the dimension of the attractor (measured in the reconstructed state space) is normally different from the waveform fractal dimension (measured on the projected signal *s(t)*, and thus limited to the range $1 \leq \dim \leq 2$. There are a number of problems when determining the fractal dimension of an attractor in state space, one being the computational burden [23]. For this reason, we will only consider waveform fractal dimensions. In this setting, the signal is looked upon as a planar set in $\mathbb{R}^2$, where the waveform is considered a geometric figure. Even though there are many ways to estimate the fractal dimension of a waveform, we focus on the variance fractal dimension (VFD) due to its robustness to noise [24]. The calculations are based on a power law relation between the variance of the amplitude increments of the signal and the corresponding time increments, see equation (3.18).

$$Var\left(s\left(t_2\right)-s\left(t_1\right)\right) \sim \left|t_2-t_1\right|^{2H} \tag{3.18}$$

where *H* is the Hurst exponent, a measure of the smoothness of a signal. The Hurst exponent can be calculated by taking the logarithm of (3.18):

$$\log\left(Var\left(s\left(t_2\right)-s\left(t_1\right)\right)\right) = 2H \cdot \log\left(\left|t_2-t_1\right|\right) \tag{3.19}$$

Plotting $log(Var(s(t_2)-s(t_1)))$ against $log(|t_2-t_1|)$ for different time increments in a log-log plot, *H* is determined as the slope of the linear regression line, see Figure 15.

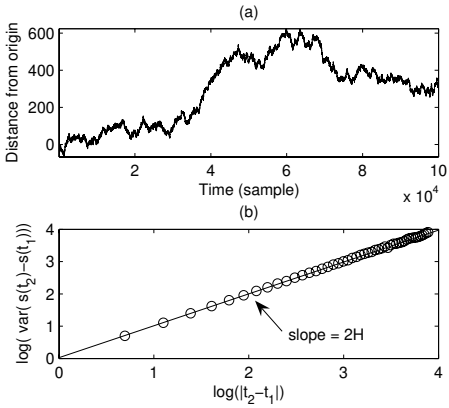

Figure 15. An example showing Brownian motion (a) and the corresponding log-log plot (b). *H* is calculated to 0.4983 which is very close to the theoretical answer 0.5. The variance fractal dimension is $VFD = E_d+1-H = 1+1-0.5 = 1.5$.

The variance fractal dimension is related to H as $VFD = E_d+1-H$, where E_d is the Euclidian dimension ($E_d = 1$ for time series). The choice of time increments is reflected in the VFD trajectory for the part of the signal where signal exists. If the time increments are chosen in a dyadic manner (1, 2, 4, 8,…), differences between various signal components are emphasized whereas if the time increments are chosen by unit decimation (1, 2, 3, 4,…), segmentation of signal from noise is favourable [24]. Since VFD is a scalar value calculated from the samples at hand, a sliding window approach has to be used to describe the complexity over time.

3.3.2. Recurrence quantification analysis

The state space of a system is often high-dimensional, especially when reconstructed from experimental data where noise tends to inflate the dimension. Its phase portrait can therefore only be visualized by projection into two or three dimensions. This operation does however fold the attractor, and by doing so, destroys its structure. A recurrence plot (RP) is a way to visually investigate the d-dimensional state space trajectory through a two-dimensional representation. RPs can be used on rather short time series and represent the recurrence of states of a system (i.e. how often a small region in phase space is visited). Unlike other methods such as Fourier, Wigner-Ville or wavelets, recurrence is a simple relation, which can be used for both linear and nonlinear data [25]. An RP is derived from the distance plot, which is a symmetric NxN matrix where a point (i, j) represents some distance between $y(i)$ and $y(j)$. Thresholding the distance plot at a certain cut-off value transforms it into an RP:

$$RP(i,j) = \Theta\left(\varepsilon - \left\| y(i) - y(j) \right\|\right) \qquad (3.20)$$

where $i,j=1,\ldots,N$, ε is a cut-off distance, $\|\bullet\|$ is some norm and $\Theta(\bullet)$ is the Heaviside function. An example of a recurrence plot is shown in Figure 16. States that are close to each other in the reconstructed state space are represented by black dots in the recurrence plot.

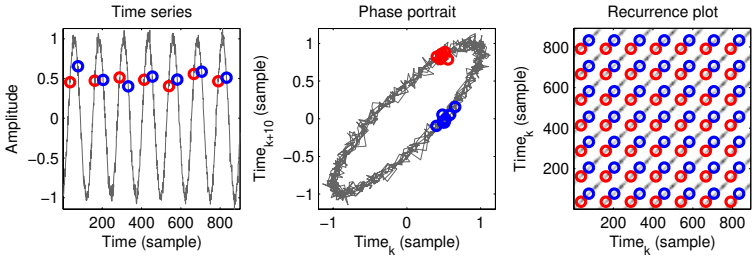

Figure 16. A noisy sinusoid represented with its waveform, its phase portrait (in a reconstructed state space) and by its recurrence plot. Dots are positioned on the waveform near amplitude values of 0.5, red dots for increasing amplitude and blue dots for decreasing amplitude. The recurrence plot shows clear diagonal lines which arise when trajectories in the state space run in parallel for some time period. The red and blue dots end up on these lines, and it can be seen that the distance between two diagonal lines is the period of the sinusoid.

There are seven parameters affecting the outcome of an RP; the embedding dimension d, the time delay τ, the range (or length) of the time series under investigation, the norm $\|\bullet\|$, the possibility to rescale the distance matrix, the cut-off distance ε and the minimal number of adjacent samples to be counted as a line (minimum line length) [26]. The last parameter is not used when creating RPs, but rather when trying to quantify them (recurrence quantification analysis, RQA). Measures used for RQA are often based on diagonal structures, vertical structures and time statistics. Isolated recurrence points occur if states are rare, if they do not persist for any time or if they fluctuate heavily. Diagonal lines occur when a segment of the trajectory runs in parallel with another segment, i.e. when the trajectory visits the same region of the phase space at different times, see Figure 16. Vertical (horizontal) lines mark a time length in which a state does not change or changes very slowly. The most common RQA-parameters are [27-29]:

- *Recurrence rate:* The percentage of recurrence points (black dots) in the recurrence matrix.
- *Determinism:* The percentage of the recurrence points that form diagonal lines. Diagonal lines are associated with deterministic patterns in the dynamic, hence determinism.
- L_{aver}*:* The average length of the diagonal lines.
- L_{max}*:* The length of the longest diagonal line. L_{max} is inversely proportional to the largest Lyapunov exponent which describes how fast trajectories diverge in the reconstructed state space.
- *Entropy:* The Shannon entropy of the distribution of the diagonal line lengths. Measures the complexity of the signal.
- *Laminarity:* The percentage of recurrence points which form vertical lines.
- *Trapping time:* The average length of the vertical lines.
- V_{max}*:* The length of the longest vertical line.
- *T1:* Recurrence time of the first kind, see below.
- *T2:* Recurrence time of the second kind, see below.

Detection of changes based on recurrence times

Detection of changes in signals is traditionally fitted into a residual based framework [30]. The main idea is to make a time varying model of the signal. If the model is correct, the residuals are expected to be white (or ideally zero). When the model and the signal no longer correspond to each other a change is indicated.

A change in the dynamics of a signal may also be detected as a change in the trajectories in the reconstructed state space. Distance measures of such changes often involve some count of nearest neighbours since the neighbours indicate recurrence of states. Actually, the RQA-parameters T1 and T2, recurrence times of the first and second kind, are such measures. Nearest neighbours in the reconstructed state space can be divided into true recurrence points and sojourn points [27], see Figure 17, where T1 is all the points and T2 is the black points.

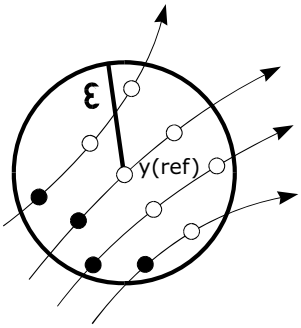

Figure 17. Recurrence points of the second kind (solid circles) and the sojourn points (open circles) in $B_\varepsilon(y(ref))$. Recurrence points of the first kind comprise all circles in the set.

More formally, an arbitrary state, y_{ref}, is chosen somewhere on the trajectory whereupon all neighbouring states within a hypersphere of radius ε are selected.

$$B_\varepsilon\left(y\left(ref\right)\right)=\left\{y\left(t\right):\left\|y\left(t\right)-y\left(ref\right)\right\|\le\varepsilon\right\}\quad\forall t \tag{3.21}$$

The recurrence points of the first kind (T1) are defined as all the points within the hypersphere (i.e. the entire set B_ε). Since the trajectory stays within the neighbourhood for a while (thus generating a whole sequence of points), T1 doesn't really reflect the recurrence of states. Therefore, the recurrence points of the second kind (T2) are defined as the first states entering the neighbourhood in each sequence (these points are commonly called true recurrence points). T2 is hence the set of points constituted by $B_\varepsilon(y(ref))$ excluding the sojourn points, see Figure 17. Both T1 and T2 are related to the information dimension via a power law, motivating their ability to detect weak signal transitions based on amplitude, period, dimension and complexity [31]. Specifically, T2 is able to detect very weak transitions with high accuracy, both in clean and noisy environments while T1 has the distinguished merit of being more robust to the noise level and not sensitive to the choice of ε. A mathematically more rigorous definition of T1 and T2 can be found in [31]. A sliding window approach is necessary to obtain time resolution.

3.3.3. Higher order statistics

Standard methods in signal processing are based on second-order statistics, but second-order measures contain no phase information. As a consequence, non-minimum phase signals and certain types of phase coupling (associated with nonlinearities) cannot be correctly identified. In contrast to second-order statistics, higher order statistics are based on averages over products of three or more samples of the signal, thus allowing nonlinear dependencies among multiple signal samples to be evaluated. Assuming zero mean signals and limiting the survey to order 4, the order moments and their corresponding cumulants are defined as [32]:

$$
\begin{aligned}
m_s^{(1)} &= c_s^{(1)} = E\{s(t)\} = 0 \\
m_s^{(2)}(\tau) &= c_s^{(2)}(\tau) = E\{s(t)s(t+\tau)\} \\
m_s^{(3)}(\tau_1,\tau_2) &= c_s^{(3)}(\tau_1,\tau_2) = E\{s(t)s(t+\tau_1)s(t+\tau_2)\} \\
m_s^{(4)}(\tau_1,\tau_2,\tau_3) &= E\{s(t)s(t+\tau)s(t+\tau_2)s(t+\tau_3)\} \\
c_s^{(4)}(\tau_1,\tau_2,\tau_3) &= E\{s(t)s(t+\tau)s(t+\tau_2)s(t+\tau_3)\} - 3\left[E\{s(t)s(t+\tau)\}\right]^2
\end{aligned}
\tag{3.22}
$$

where E represents the expected value. Interesting special cases are $c_s^{(1)}(0)$, $c_s^{(2)}(0,0)$ and $c_s^{(3)}(0,0,0)$ which represent the variance, skewness and kurtosis of $s(t)$. The Fourier transforms of cumulants are called polyspectra, and are defined according to equation (3.23). An example of a simple bispectrum, the Fourier transform of the third order cumulant, is shown in Figure 18.

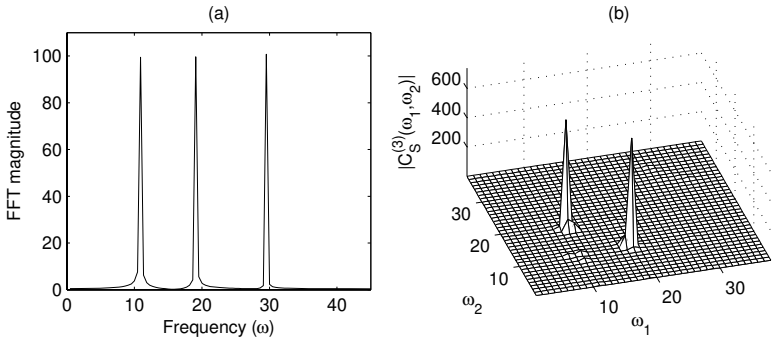

Figure 18. An example of phase coupling. The frequency spectrum of a signal composed of three sinusoids with frequencies λ_1, λ_2 and $\lambda_3 = \lambda_1 + \lambda_2$ is shown in (a). The corresponding bispectrum is shown in (b). Since λ_3 is caused by phase coupling between λ_1 and λ_2, a peak due to the phase coupling will appear in the bispectrum at $\omega_1 = \lambda_1$, $\omega_2 = \lambda_2$ (another peak will also emerge at $\omega_1 = \lambda_2$, $\omega_2 = \lambda_1$).

$$
\begin{aligned}
C_s^{(2)}(\omega) &= FT\{c_s^{(2)}(\tau)\} \quad \text{Power spectrum} \\
C_s^{(3)}(\omega_1,\omega_2) &= FT\{c_s^{(3)}(\tau_1,\tau_2)\} \quad \text{Bispectrum} \\
C_s^{(4)}(\omega_1,\omega_2,\omega_3) &= FT\{c_s^{(4)}(\tau_1,\tau_2,\tau_3)\} \quad \text{Trispectrum}
\end{aligned}
\tag{3.23}
$$

Complete characterisation of a stochastic process requires knowledge of all moments. Generally speaking, moments correspond to correlations and cumulants correspond to covariances. Even though both measures contain the same statistical information, cumulants are preferred in practice since [33]:

1. Cumulant spectra of order $n > 2$ are zero for Gaussian signals and their polyspectra provide a measure of the extent of non-Gaussianity.
2. The covariance function of white noise is an impulse function and its spectrum is flat. Similarly, cumulants of white noise are multidimensional impulse functions with multidimensionally flat polyspectra.

3. The cumulant of two independent random processes equals the sum of the cumulants of the individual random processes.

Higher order cumulants provide a measure of how much a random vector deviates from a Gaussian random vector with an identical mean and covariance matrix. This property can be used for extracting the nongaussian part of a signal (one application is removal of Gaussian noise). Other interesting properties are that the bispectrum is zero for a Gaussian signal and that the bicoherence (normalized bispectra) is constant for a linear signal.

3.4. Nonlinear Prediction

There are different sources of predictability in a time series. If the signal contain linear correlations in time, then linear models are suitable (moving average (MA) models, autoregressive (AR) models, autoregressive moving average (ARMA) models etc.). MA models can be used if the spectrum of the signal behaves like coloured noise while AR models are preferable if the spectrum is dominated by a few distinct peaks. The ARMA model is a natural extension of AR and MA. The AR model is obtained as a sum of weighted previous outputs *s(t-k)* and the innovation signal *e(t)*:

$$s(t) = \sum_{k=1}^{N} \alpha_k s(t-k) + e(t) \tag{3.24}$$

where α_k are the linear weights. For prediction, only the weighting coefficients are important and the prediction is obtained by ignoring the unknown innovation. This model can be expanded to allow nonlinear dependencies between previous outputs *s(t-k)*. Actually, a very general framework for predicting time series is given in Ljung [34] (φ may include all available signal samples including multivariate inputs and outputs):

$$\begin{aligned} \hat{s}(t|\theta) &= \sum_{k=1}^{N} \alpha_k g_k(\varphi) \\ \theta &= [\alpha_1, \alpha_2, ..., \alpha_n]^T \\ g_k(\varphi) &= \kappa(\beta_k(\varphi - \gamma_k)) \\ \varphi &= [s(t-k), ..., s(t-1)] \end{aligned} \tag{3.25}$$

where all the g_k are formed from dilated and translated versions of a mother basis function κ. θ is a vector of weights and φ is a vector of known signal samples. α are the coordinates or weights, β are the scale or dilation parameters and γ are the location or translation parameters. A few examples of how this model framework can be used are:

Autoregressive model: set most of the parameters to unity.
Sigmoid Neural Network: κ is a ridge construction such as the sigmoid function.
Radial basis networks: κ is a radial construction such as the Gaussian bell.

Turning back to the reconstructed state space setting, it can be seen that φ in equation (3.25) is very similar to a reconstructed coordinate. Toss in a delay parameter τ, or set $\tau = 1$, and φ turns into y (see equation (3.17)). A way to look at the model in (3.25) is thus as a function describing the whole attractor. Usually, all parameters but the α:s are design parameters that either vary in a predetermined way or are fixed. Inserted into a cost function, (3.25) leads to linear equations when estimating the α:s, thus simplifying their determination [23]. Since this modelling attempt tries to model the whole attractor, it is called a global model.

That being said about global models, we will abandon them altogether and focus on local methods working directly in the reconstructed state space. Similar trajectories in state space share the same waveform characteristics in time domain, and a way of predicting the future is thus to mimic the evolution of neighbouring trajectories, see Figure 19. If the data is sampled with high frequency, most of the discovered nearest neighbours will probably be samples adjacent to each other in the time series. A considerable improvement could thus be obtained by using nearest trajectories instead of nearest neighbours see Figure 20.

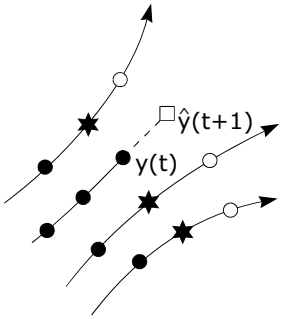

Figure 19. Three trajectory segments and a (forward) predicted trajectory in a two-dimensional phase space. The average change between the nearest neighbouring trajectory points (black stars) and their successors (white circles) are used to predict the next point (white square).

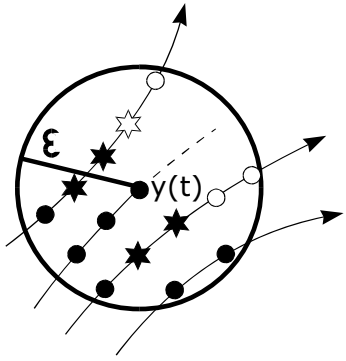

Figure 20. Many of the nearest neighbours to y(t), stars, are actually phoneys due to the high sampling rate. Using a nearest trajectory algorithm instead of a nearest neighbour algorithm is one solution.

There are two approaches to predict p steps ahead, either using iterated prediction or direct prediction. If the prediction is iterated, the algorithm predicts one step ahead p times (the predicted values will then be used as a starting point in the next iteration). In direct prediction, the evolutions of the nearest neighbours are modelled and the resulting function maps p steps into the future. It is empirically shown that iterated prediction is better on short term forecasts for a variety of non-linear models. However, iterated predictions do not take the accumulated errors in the input vector into account, and these errors grow exponentially [35].

More options are available. In averaged prediction, the average of the neighbours' successors (white circles) locations are chosen as the predicted value while in integrated prediction the next point is estimated as the current point plus the average change amongst the neighbours. If the trajectory that is to be predicted is an outlier; the mean of the nearest neighbours will always be misleading.

To conclude, local models can give excellent short-term prediction results, they are conceptually simple but may require a large computational burden due to the dependence of nearest neighbour calculations.

4. Properties of Phonocardiographic Signals

"Let your heart guide you. It whispers, so listen closely…"
The land before time (1988)

During auscultation, identification of heart sounds and murmurs is primarily based on pitch and timing of occurrences. In phonocardiography, information about morphology, and to some extent frequency content, is also included in the diagnosis. There is however more knowledge to extract. Physicians talk about auscultation as an art. The diagnosis is often based on a sensation that is hard to explain and even harder to implement in a computer. To automatically extract information that is even close to this sensation, every possible grain of information has to be exploited, even though it means going beyond the well-known properties of time and frequency.

Some examples in this chapter are based on data (used with permission) from FamilyPractice.com (http://www.familypractice.com/) and The Auscultation Assistant (http://www.med.ucla.edu/wilkes/). These data were chosen to give distinct and illustrative figures. The data were originally intended for educational purposes, and consists of text-book examples of recorded phonocardiographic signals and in some cases even simulated sounds. Other examples and all the presented results are however obtained using data from paper I-III. This means that noisy signals with, in some cases, large interpatient variability are used in the calculations.

4.1. Time and Frequency

Time and frequency properties are the most important features when physicians perform auscultation on a patient. In healthy subjects, the frequency spectrum of S1 contains a peak in the low frequency range (10-50 Hz) and in the medium frequency range (50-140 Hz) [36]. S2 contains peaks in low- (10-80 Hz), medium- (80-220 Hz) and high-frequency ranges (220-400 Hz) [37]. S2 is composed of two components, one originating from aortic valve closure and one originating from pulmonary valve closure. Normally, the aortic component (A2) is of higher frequency than the pulmonary component (P2) [38]. The peaks probably arise as a result of the elastic properties of the heart muscle and the dynamic events that

causes the various components of S1 and S2 [36, 37]. S3 and S4 are believed to originate from vibrations in the left ventricle and surrounding structures powered by the acceleration and deceleration of blood flow. 75 % of the total energy in S3 is contained below 60 Hz [39] while S4 mainly contain frequencies below 45 Hz [40]. The time and frequency properties of heart sounds are summarized in

Table 2 and examples of two heart sounds and their frequency spectra are illustrated in Figure 21. The different heart sounds are affected by various heart diseases, and the main changes are described in sections 4.1.1 - 4.1.3.

Table 2. Time and frequency properties for the heart sounds.

Sound	Location (ms)	Duration (ms)	Main frequency range (Hz)
S1	10-50 after R-peak in ECG	100-160	10-140
S2	280-360 after R-peak in ECG	80-140	10-400
S3	440-460 after R-peak in ECG or 120-180 after closure of semilunar valves	40-80	15-60
S4	40-120 after beginning of P-wave in ECG	30-60	15-45

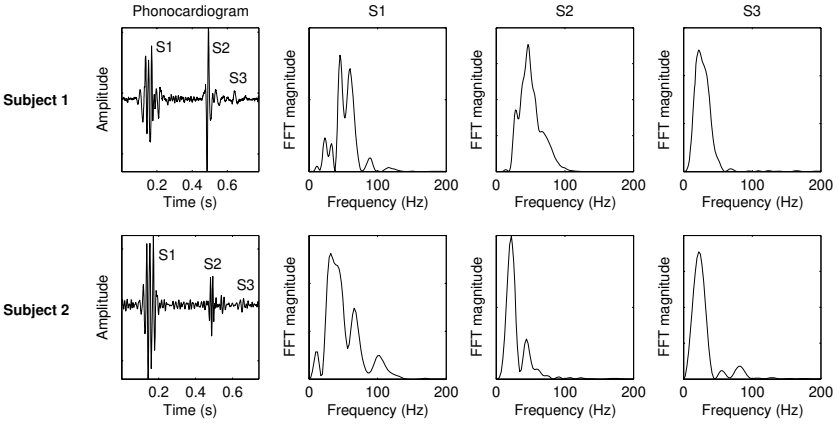

Figure 21. Heart sounds and their respective frequency spectra from a 13 year old girl (top row) and a 36 year old male (bottom row). Data obtained from paper I and II.

There is a small delay between the aortic component and the pulmonary component causing a splitting of S2 (since right ventricular ejection terminates after left ventricular ejection). Normally, the splitting increases with inspiration due to increased blood return to the right heart, increased vascular capacitance of the pulmonary bed and decreased blood return to the left heart [7]. That is, the

aortic component occurs earlier and the pulmonary component occurs later during inspiration. In certain heart diseases, this splitting can become wide, fixed or reversed (see sections 4.1.1 - 4.1.3). FFT analysis does not take timing into consideration, so it cannot reveal which of the two valves closes first. Meanwhile, it is hard to notice any difference between the two components in the time domain. A tool able to investigate how the signal's frequency content varies over time is thus called for. Such methods were introduced in section 3.1.2, and an example showing the four heart sounds is presented in Figure 22. Taking a closer look at S2 in Figure 22, it can be seen that the two components are merged together, but it is also clear that the higher frequency aortic component precede the lower frequency pulmonary component.

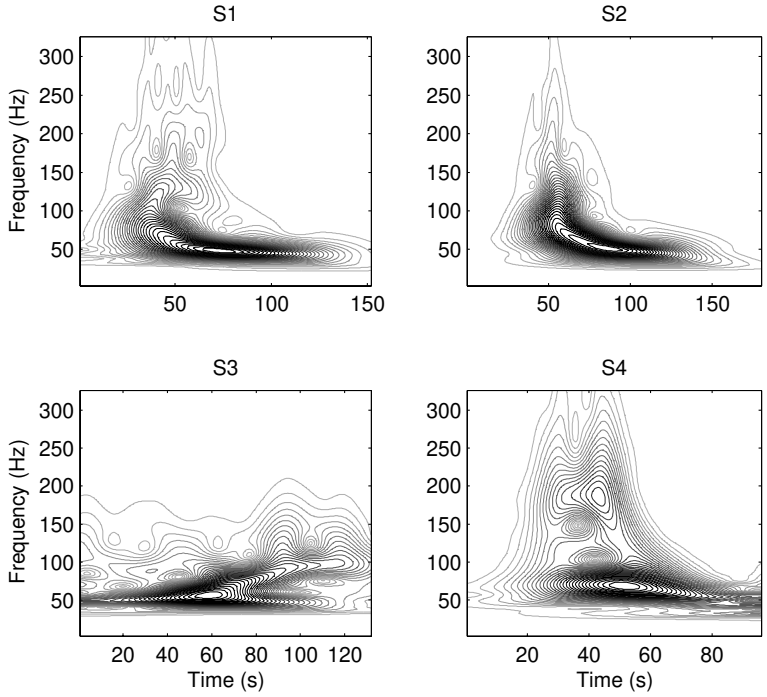

Figure 22. Example of TFR contour plots of S1, S2, S3 and S4 (note the different scaling of the x-axis). Stockwell's method was used to calculate the TFR. Data obtained from paper II.

An exposé of systolic murmurs is given in sections 4.1.1 - 4.1.3, where the emphasis is on the changes imposed by pathologies on the phonocardiogram. More information can be found in [7]. Basic layout sketches describe the time domain properties, while frequency domain information is illustrated with TFR-plots (since heart murmurs are nonstationary [41], useful frequency investigations have to be conducted using joint time frequency analysis tools).

4.1.1. Murmurs from stenotic semilunar valves

If the aortic or pulmonary valve becomes narrowed or constricted (stenotic), blood has to be forced through the valve opening. The arising turbulent blood flow causes vibrations in the cardiac structure which are transmitted through the tissue and perceived as a murmur. The murmur peaks in mid-systole at the time of maximal ejection and produces a crescendo-decrescendo shape in the phonocardiographic signal. The severity of the stenosis influences the shape of the murmur, where the intensity will increase and the peak will occur later in systole as the stenosis becomes more severe. An ejection click may occur if the valves are brought to an abrupt halt instead of opening completely when moving from their closed position in diastole to their open position in systole. Since a large boost from the atrium might be necessary to help build up the pressure in the ventricle, a fourth heart sound may be present. The appearance of a murmur caused by stenosis in the semilunar valves is illustrated in Figure 23.

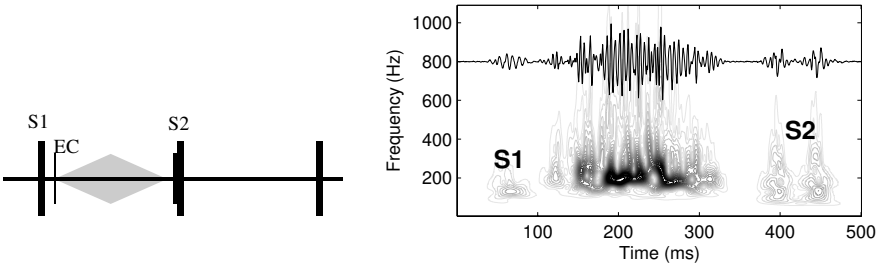

Figure 23. A basic layout sketch of the phonocardiographic signal from a murmur caused by stenosis in the semilunar valves is presented in the left plot while an example TFR (showing pulmonary stenosis, calculated by Stockwell's transform) is illustrated in the right plot. EC = Ejection click.

In aortic stenosis, there will be a paradoxical splitting of S2 that increases with expiration. If the stenosis is severe the aortic component is attenuated or even missing. Aortic stenosis is usually caused by congenital aortic valve disease, rheumatic fever or degenerative calcification.

In pulmonary stenosis, the splitting of S2 is caused by increased capacitance in the dilated pulmonary trunk. In severe stenosis, S2 is widely split but difficult to hear since the pulmonary component is faint and the aortic component is obscured by the murmur. The degree of stenosis correlates well with the width of S2. In mild stenosis the murmur ends before the aortic component of S2 while it is stretched up to the pulmonary component in severe cases. Pulmonary stenosis is usually caused by congenital fusion of the pulmonary valve cusps.

4.1.2. Murmurs from regurgitant atrioventricular valves

Backward flow through the mitral or tricuspid valves causes a murmur that begins as soon as the atrioventricular valves closes and continues up to the semilunar valve closure. Because the pressure gradient between ventricle and atrium is large throughout systole, the murmur tends to have a constant intensity throughout

systole (holosystolic). The appearance of a murmur caused by insufficient atrioventricular valves is illustrated in Figure 24.

In mitral insufficiency (MI), the murmur begins with systole and continues as long as left ventricular pressure exceeds that of the enlarged left atrium. The murmur engulfs the aortic component of S2 but stops before the pulmonary component. In acute mitral insufficiency the murmur is diamond-shaped, loud (grade IV or more) and does not necessarily extend to the end of systole. Due to the decrease in the rate of rise of the ventricular pressure, S1 is weak. A third heart sound is often present, although it doesn't necessarily imply systolic dysfunction or elevated filling pressure as it usually does. The mitral valve apparatus is composed of five different structures; mitral annulus, leaflets, chordae tendinae, papillary muscles and the free wall of the left ventricle. Malfunction in any of these may result in mitral insufficiency. Typical examples are rheumatic heart disease in the mitral valves, papillary muscle dysfunction or calcified mitral annulus.

In tricuspid insufficiency, the holosystolic murmur increases with inspiration. Unlike mitral insufficiency the murmur of tricuspid insufficiency persists through S2 and engulfs the pulmonary component. In mild cases, S4 may be present. Tricuspid insufficiency is often found in severe right heart failure.

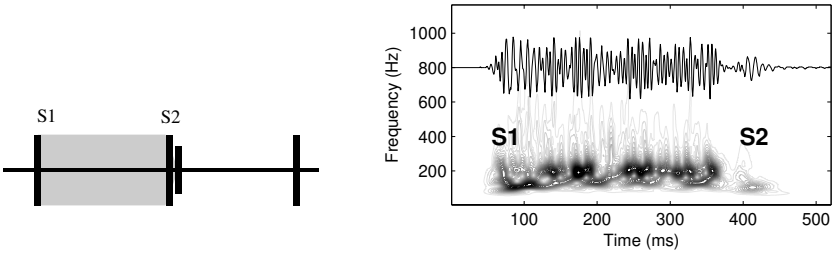

Figure 24. A basic layout sketch of the phonocardiogram from a murmur caused by a regurgitant atrioventricular valve is presented in the left plot while an example TFR (mitral insufficiency, calculated by Stockwell's transform) is illustrated in the right plot.

4.1.3. Murmurs caused by septal defects

A septal defect is an opening between the left and the right side of the heart. In atrial septal defect the opening connects the two atria and in ventricular septal defect the opening connects the two ventricles. Septal defects are often congenital, and ventricular septal defects are in fact the most common congenital heart defects.

In ventricular septal defect, already oxygen-rich blood is pumped back to the lungs, causing a dilation of the heart. During systole, a jet from the left to the right ventricle causes a holosystolic murmur that sometimes engulfs S2. The aortic component of S2 is early due to the short ejection period (caused by low impedance in the systemic circulation system). This results in an S2-split that increases with inspiration.

In atrial septal defect, recently oxygenated blood leaks back to the right atrium where it is, again, sent through the pulmonary circulation system. The increased flow through the pulmonary valve produces a soft mid-systolic ejection murmur. S2 has a large fix split caused by decreased resistance in the pulmonary aorta which delays the pulmonary component of S2.

4.1.4. Quantifying the results

It is a well known fact that systolic ejection murmurs like aortic stenosis are crescendo-decrescendo shaped while systolic murmurs due to backflow, like mitral insufficiency, are holosystolic (do not vary in intensity). This is not always obvious when looking at actual recorded signals. Using normalized Shannon energy as a measure of intensity, the shape of the murmur in aortic stenosis (AS), mitral insufficiency (MI) and physiological murmurs (PM) are shown in Figure 25. The data used is from the 36 patients in paper III. The classical shapes are indicated, but having the standard deviation in mind, this conclusion is not evident. In the AS case, the large standard deviations are probably due to the progress of the disease, which is ranging from moderate to severe.

The nine presented instants were selected at times before S1, peak S1, after S1, ¼ into systole, ½ into systole, ¾ into systole, before S2, peak S2 and finally after S2. The data were derived as mean values of each heart cycle in one patient. These mean-values were then used to create the plots in Figure 25. Similarly, the mean and standard deviation TFRs of AS, MI and PM were calculated. An average TFR was derived for each patient using all available heart cycles. These averaged TFRs were then used to create the mean and standard deviation TFRs in Figure 26. Distinct areas in the standard deviation plots are clearly areas where there are great differences between patients (this might be a hint on which time-frequency areas that are stable to use in a murmur classifier). From the figure, it is obvious that there is great interpatient variability. This is unfortunate when the aim is to classify patients into various categories; ideally all patients with the same disease behave similarly. Focusing on differences between diseases, it can be seen that the murmur in MI has frequency content up to about 160 Hz and that the murmur doesn't change much over systole (holosystolic). The murmur in AS on the other hand seems very unstable. Since the murmur in AS is crescendo-decrescendo shaped where the peak depends on the severity of the stenosis, it makes sense that the mean of several AS murmurs deviates between patients. The frequency content reaches up to about 180 Hz (as can be seen in the standard deviation plot). Finally, PM are known to be of low frequency, which is verified by the figure.

The TFR technique is a valuable tool when visually inspecting signals, but the amount of information is immense. In a classification setting we need to find a compact representation consisting of a manageable number of informative features. Paper III describes a number of methods able to quantify time-frequency information.

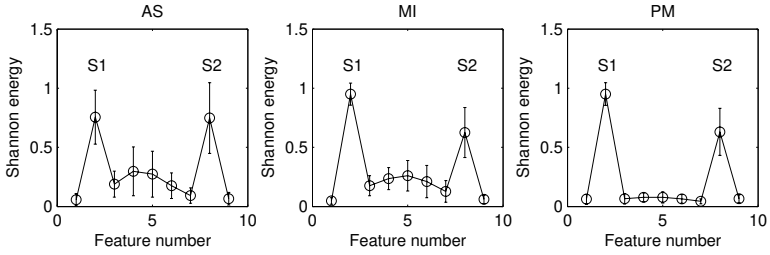

Figure 25. Mean value of the Shannon energy calculated at nine time instants in systole, the whiskers show the standard deviation. Data obtained from paper III.

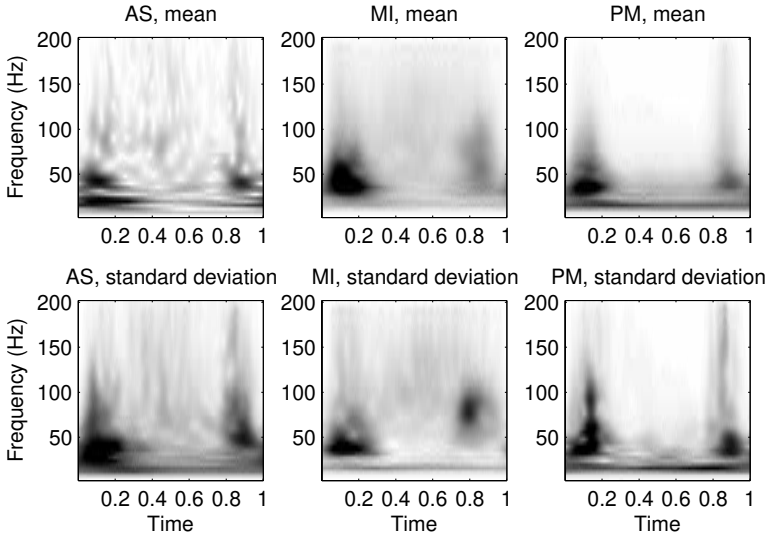

Figure 26. Mean (top) and standard deviation (bottom) TFRs (calculated by Stockwell's transform) of aortic stenosis, mitral insufficiency and physiological murmurs. The time scale was resampled to 2048 samples after calculating the TFR, and is here represented in arbitrary normalized units. Data obtained from paper III.

4.2. Higher Order Statistics

Since methods based on second-order statistics do not take nonlinearity and non Gaussianity into account, higher order statistics might provide more information about the phonocardiographic signal. Bispectra for signals from various systolic murmurs are presented in Figure 27. First of all it can be seen that the bispectrum is not zero, as it would be if the signals were Gaussian. Secondly, the main frequency content is well below 300 Hz and exhibits distinct peaks. Thirdly, considerable phase coupling exists between different frequencies. Finally it is also seen that the patterns revealed in the bispectra differ between various pathologies. It has previously been indicated that phonocardiographic signals are non Gaussian

[42, 43], but it has not been explicitly stated that this is the case. When performing Hinich's Gaussianity test on each heart cycle in the data from paper III, it turns out that each and every one of the 445 heart cycles have zero skewness with probability <0.05. This strongly suggests that the data are non Gaussian (nonzero skewness) and that investigations of the higher-order statistics of phonocardiographic signals are relevant. Similarly, a hypothesis regarding linearity could be rejected using Hinich's linearity test (for a nonlinear process, an estimated statistic may be expected to be much larger than a theoretical statistic, and in this case the estimated value is, on average, 3.4 times larger). This motivates the use of nonlinear techniques in the two subsequent sections, 0-4.3. Both Hinich's tests are described in [44].

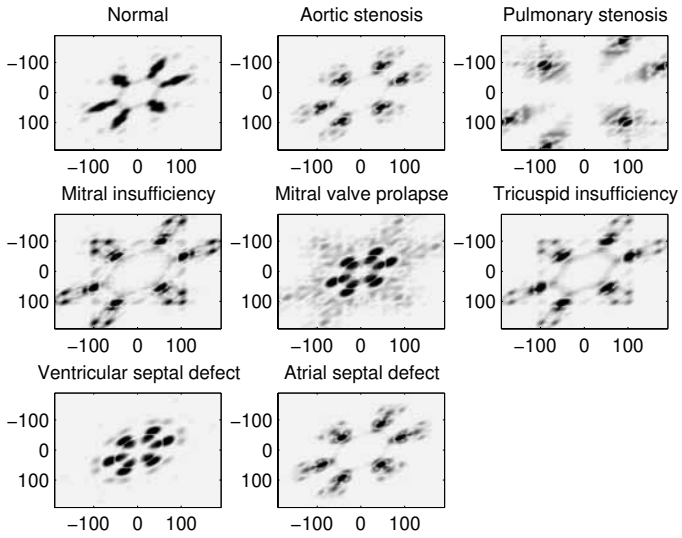

Figure 27. Examples of bispectra from one heart cycle in different heart diseases. One heart cycle here roughly corresponds start S1 to stop S2. All axes represent frequency in Hz.

The bispectra in Figure 27 are all very nice to look at, but we need to quantify them somehow. To make the number of quantifying units manageable, the bispectrum can be discretized [45], see Figure 28. Due to symmetry, it is enough to investigate the first nonredundant region [33]. Using data from paper III, box and whisker plots were derived for these 16 features, see Figure 29 are presented in. Unfortunately the features overlap and are more or less useless for classification purposes (t-tests show significant differences ($p<0.05$) for feature 2 between AS↔PM, feature 3 between AS↔MI and AS↔PM, feature 7-8, 11 between MI↔PM). The bispectrum is however a useful tool and Figure 27 does reveal a lot of information. If nothing more, it could be used as a visualisation technique to support the physician's decision.

There are distinct differences between the various heart valve diseases in Figure 27. Obviously these differences are lost in the discretization when attempting to reduce the information into a manageable feature set. A different approach is thus needed to extract this information. A few ideas are Gaussian mixture models or perhaps some parametric models like the non Gaussian AR model, but these issues are left for future studies.

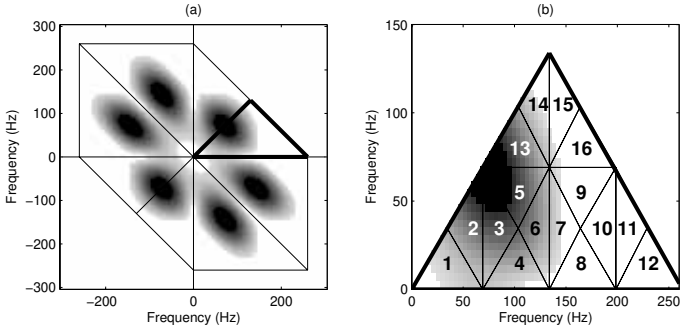

Figure 28. Example of bispectrum from a patient with aortic stenosis. The different regions of the bispectrum is plotted in (a) where the bold triangle shows the first non-redundant region. In (b) the region of interest is highlighted. The smaller triangles indicate the 16 features obtained from the bispectrum, where each feature is calculated as the mean intensity of each triangle.

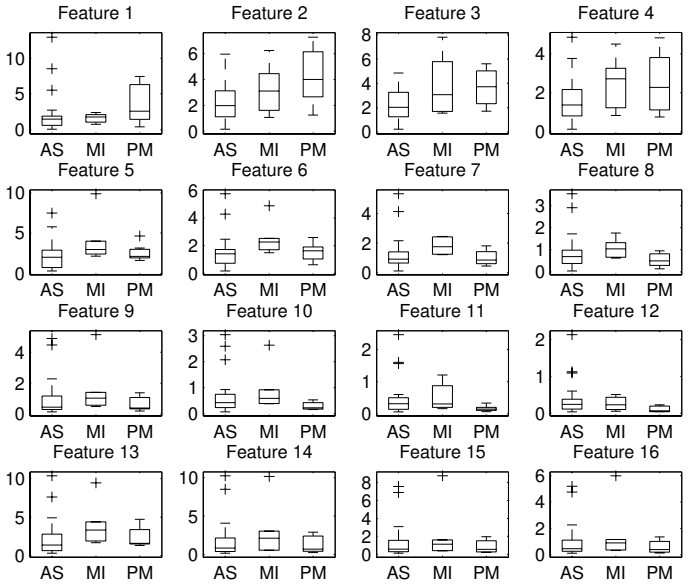

Figure 29. Box and whisker plots showing results from the bispectral analysis. The boxes have lines at the lower quartile, median, and upper quartile values. The whiskers show the extent of the data. Outliers (+) are data with values beyond the ends of the whiskers. Data obtained from paper III.

4.3. Reconstructed State Spaces

Phonocardiographic waveforms consist of a large variety of types, ranging from impulses (snaps and clicks) through turbulence induced sounds (murmurs) to nearly periodic oscillations (heart sounds). The transition between these types could be described by switching between different linear models, but using a nonlinear setting, such transitions occur naturally as bifurcations [46].

Hypothesising that the blood flow is a dynamic system which is observed via the recorded phonocardiographic signal, then the reconstructed state space would be an attempt to recreate the characteristics of the flow (compare with Figure 13). Since turbulence is a nonlinear phenomenon with strong interaction between the flow and the associated acoustic field [47], the theoretical foundation for the hypothesis seems valid.

Before pursuing any attempts to use nonlinear analysis tools, one should execute some tests to see whether the data really behave in a nonlinear fashion. Two such tests were performed on the data in paper III; Hinich's linearity test (see section 4.2) and phase randomized surrogate data [23]. Both tests indicated nonlinearity by rejecting the hypothesis of linearity.

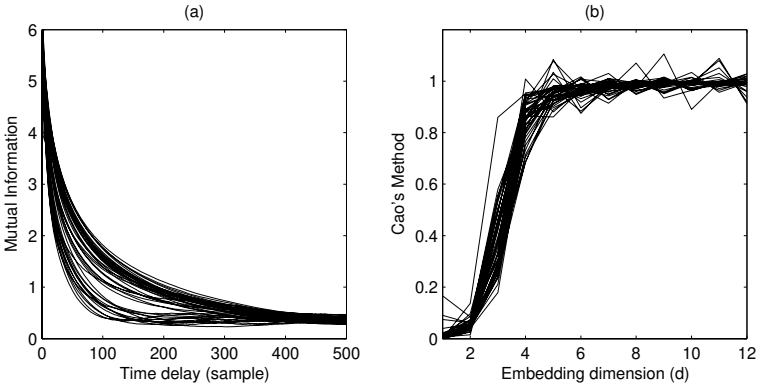

Figure 30. Average mutual information calculations are used to determine the time delay embedding parameter τ (a). The first minimum of the mutual information function indicates a delay where the signal contains little mutual information compared to a delayed version of itself (why the combination of the two provides as much information as possible). Cao's method (b) is used to determine the embedding dimension *d*. This method is similar to the common false nearest neighbour approach, which make use of the fact that points are moved closer together in a reconstructed state space, compared to the true state space, by folding. Data obtained from paper III.

Nonlinear analysis of measured data is generally based on the reconstruction of the signal in a multidimensional state space. Proper embedding parameters were calculated via mutual information and Cao's method. The embedding dimension was found to be $d = 4$, which can be seen by the clearly defined knee in Figure 30. Determination of the delay parameter was however less obvious. The mean value of first minima in the mutual information function was $\tau = 233 \pm 72$ samples.

Since roughly half of the patients had a minimum in the vicinity of $\tau = 150$, while the other half lacked an obvious minimum in the range $\tau = 1 \ldots 500$ samples, τ was set to 150. These routines should not be used on nonstationary data, and to minimize the damage, these values were determined on data consisting of only murmur data. En example of a phonocardiographic signal embedded in three dimensions is given in Figure 31. The heart sounds are clearly encircling the more complex murmur.

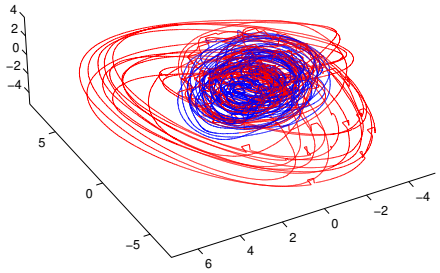

Figure 31. Example of embedded heart sound with d = 3 and τ = 200. Heart sounds (S1 and S2) are plotted in red and the murmur (AS) in blue. The small bumps in the trajectory are due to concatenation of segments. Data obtained from paper III.

4.3.1. Quantifying the reconstructed state space

Quantifying a four-dimensional phase portrait is not easy. There are a few common dynamical invariants that can be used. One of them is the fractal dimension which will be used in section 4.4. Another example of an invariant measure is the largest Lyapunov exponent. Disadvantages with invariant measures are that they are not sensitive to initial conditions or smooth transformations of the space. Another approach for quantification of the reconstructed state space is by direct modelling of its statistical distribution. A parametric model could be constituted by a Gaussian mixture model (GMM). In paper III, a GMM with five mixtures, see Figure 32, was fitted to the reconstructed state space using the Expectation-Maximization (EM) algorithm.

The centres of the mixtures and the eigenvalues of their covariance matrices can be used as a rather compact representation of the trajectory in the reconstructed state space. However, the five mixtures used here to describe the four-dimensional geometry require 40 parameters. Using only five mixtures to estimate the density function of the trajectory is probably far from enough, but the number of parameters increases rapidly with the number of mixtures. Due to the difficulty of giving an overall summary of three groups (AS, MI and PM) and 40 parameters, these results are omitted in this section. They are however used as features for murmur classification in section 5.4 and in paper III.

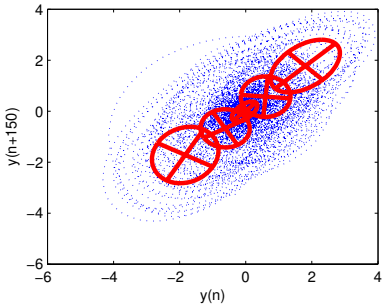

Figure 32. A reconstructed state space (d = 2, τ = 150) of the systolic period from a patient with aortic stenosis. The red ellipses symbolize a Gaussian mixture model with five mixtures. Note that d = 2 is not enough to unfold the trajectory.

4.3.2. Recurrence time statistics

In previous sections, higher dimensional state spaces have been visualized by projection into lower subspaces. Recurrence plots were introduced to avoid this procedure by visualizing high dimensional trajectories through a two-dimensional representation of its recurrences [28]. An example of recurrence plots is presented in Figure 33. Figure 34 shows the RQA results using the data in paper III. It is clear that several of these parameters are good at separating physiological murmurs from pathological murmurs (determinism, longest diagonal line, longest vertical line and trapping time). Unfortunately, the results obtained from AS and MI are heavily overlapping.

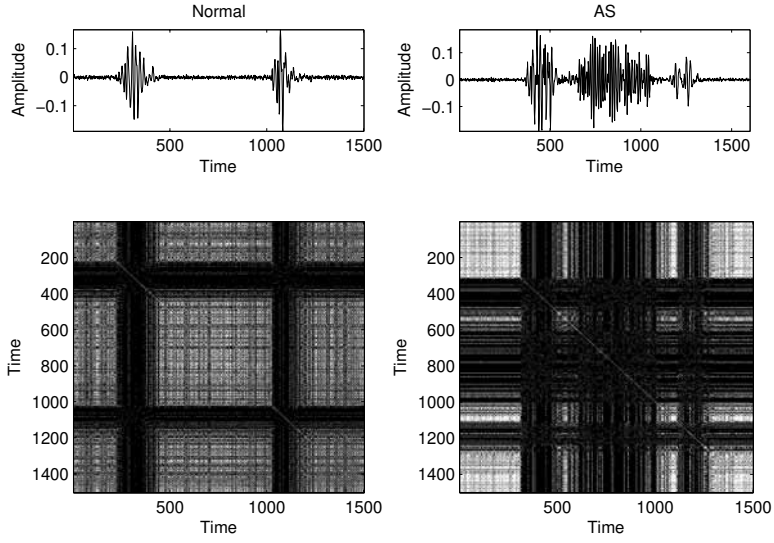

Figure 33. Example of recurrence plots for a normal phonocardiographic signal and for an AS case. The interpretation of recurrence plots was briefly explained in section 3.3.2.

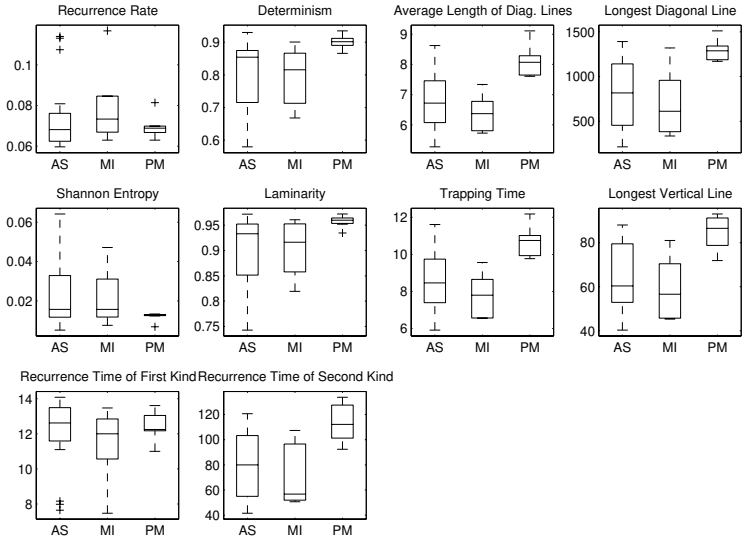

Figure 34. Box and whisker plots showing results from the recurrence quantification analysis. The boxes have lines at the lower quartile, median, and upper quartile values. The whiskers show the extent of the data. Outliers (+) are data with values beyond the ends of the whiskers. Data obtained from paper III.

4.4. Fractal Dimension

Theories attempting to explain turbulence forecast the existence of eddies (vortices with a characteristic size) at multiple scales [48], and this multiscale structure of turbulence can in some cases be quantified by fractals. It has therefore been suggested that turbulent flow is fractal in nature [48].

As illustrated in section 2.3, Figure 5, in vivo measures obtained with the particle trace technique show that blood flow through the normal heart contains large vortices. The resolution of the particle trace technique does not allow a thorough investigation on a wide range of scales, but model studies show that the interaction with the ventricle wall results in the generation of smaller secondary vortices which may in turn interact with the ventricle wall and thus creating even smaller vortices [49]. However, flow in the normal heart is predominantly laminar, even though it may involve vortices on a large scale. This behaviour fits the fractal theory with self similarity across scale, but only for large scales.

Stenosed or leaking valves give rise to turbulent flow in the heart or in the great vessels leaving the heart. In turbulent flow, unsteady vortices appear on many scales and interact with each other. Numerical studies on stenotic tube flow show that the positions at which vortices are initiated, their size, and their life span are a function of the Reynolds number [50]. There is thus a link connecting the

Reynolds number to the degree of turbulence which in turn is connected to fractal behaviour. It has been shown that the smallest length scales of turbulence is about three times larger than the size of a red blood cell [51], so self similarity can be found from very small scales up to large scales bounded by the size of the ventricle. Finally, the last piece of the puzzle is provided by the strong interaction between the flow and its induced sound field [47]. It is thus reasonable to believe that the turbulence can be quantified using the fractal dimension of the measured acoustic signal.

Considering the mere waveform of bioacoustic time series, it appears that these signals possess valid characteristics for pursuing fractal dimension calculations:

- The signals do not self-cross.
- The waveform is often self-affine, i.e. in order to scale the signal, a different scaling factor is required for each axis. In physical systems, this property is not strict but probabilistic, and there are minimum and maximum scaling limits (depending on the accuracy of the measurement, the sampling resolution etc.).
- The waveform exhibits clear quasiperiodicity (heart beats and breathing specifically).
- The power spectral density is broad-band.

An example is given in Figure 35, where the acoustic waveform from a patient with aortic stenosis is plotted along with its fractal dimension calculated over time. Comparing heart sounds, murmurs and background noise, HS have a curtain structure while murmurs are more complex and noise has no structure at all. It can also be seen that the fractal dimension of the murmur is rather constant despite the large amplitude variations (crescendo-decrescendo) in the time domain.

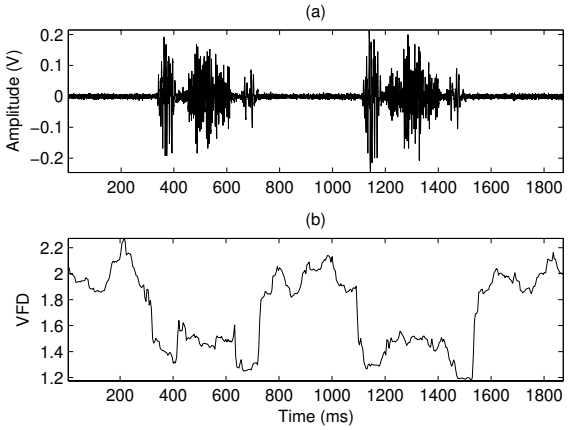

Figure 35. Example of aortic stenosis (a) showing the variance fractal dimension plotted over time (b).

Variance fractal dimension for the data in paper III is shown in Figure 36. The nine chosen instants were selected at times analogous to the Shannon energy plot in Figure 25. Again, the variance is rather large, especially in the AS case. VFD of the different murmurs are however quite well separated in their mean; AS = 1.202, MI = 1.037 and PM = 1.336 (calculated as mean values for feature number 4-6 in Figure 36, the values deviate from those in the figure due to normalization). Hypothesis testing for the difference in mean between the groups (t-test) shows a difference with significance p = 0.03, p = 0.07 and p = 0.01 when comparing AS↔MI, AS↔PM and MI↔PM, respectively. Boxplots of the same data are shown in Figure 37. Here the VFD was calculated using a concatenation of all S1 segments, all S2 segments and all murmur segments from each patient. Focusing on the murmur, the trend from Figure 36 is recognised; MI has lowest dimension, PM has highest dimension and AS is somewhere in between. The interpatient variability is still a problem, especially in the AS case (probably due to the wide range of mild to moderate AS).

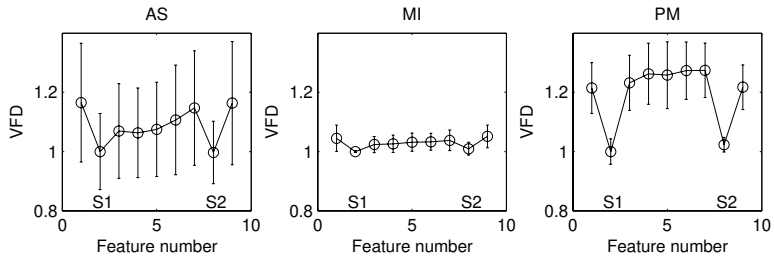

Figure 36. Mean values of the variance fractal dimension at nine time instants in systole, the whiskers show the standard deviation. The data were normalized so S1 had unit fractal dimension (for visual appearance). Data obtained from paper III.

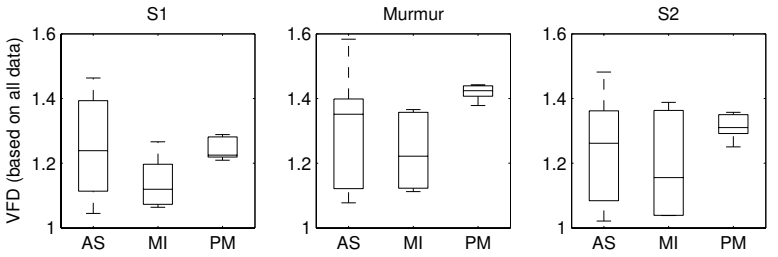

Figure 37. Boxplots of the VFD when calculated for S1, murmur and S2 when concatenating all S1 data, all S2 data and all murmur data, respectively, within each patient. The boxes have lines at the lower quartile, median, and upper quartile values. The whiskers show the extent of the data. Data obtained from paper III.

5. Applications in Phonocardiographic Signal Processing

"The modern age has a false sense of security because of the great mass of data at its disposal. But the valid issue is the extent to which people know how to form and master the material at their disposal."
Johann Wolfgang von Goethe (1832)

This chapter presents some applications that make use of the knowledge gained in previous chapters. Almost all phonocardiographic signal processing tasks are dependent on accurate segmentation of the signal. The segmentation algorithms that are described in 5.1 are later used in 5.2 - 5.4. Other treated applications include detection of the third heart sound, denoising of lung sound signals and classification or heart murmurs.

5.1. Segmentation of the Phonocardiographic Signal

Dividing the recorded signal into S1, systole, S2 and diastole is of great importance in all phonocardiographic signal processing tasks. While this is sometimes performed manually, several techniques are available to accomplish the task automatically. A quite simple and robust way to do this is by ECG gating. The ECG signal does usually have a better SNR than the sound signal, and numerous algorithms exist for automatic detection of the rather distinct R-peak. Since it is known that S1 follows shortly after the R wave in the ECG, segmentation into heart cycles is easily obtained.

Using additional sensors is not an optimal approach in the intelligent stethoscope setting, where simplicity is one of the key words. Present approaches for segmentation of the phonocardiographic signal are often based on peak picking after transformation into a domain where S1 and S2 are emphasized. Several choices of this transformation has been presented; coefficients from an 8th order AR model in a narrow sliding window [52], Shannon energy [15], wavelet decomposition plus Shannon energy [53], homomorphic filtering [17], matching pursuit [54] and fractal dimension trajectories [55, 56]. Typically, a threshold is used to locate the heart sounds, without classification into S1 and S2. The actual classification is done using interval statistics (the duration of systole is rather

constant and normally shorter than the duration of diastole, giving a binominal distribution, see Figure 38). A problem with all of these approaches is that it is hard to distinguish the heart sounds in noisy phonocardiographic signals. Noise in this case could be heart murmurs, lung sounds and/or background noise such as speech.

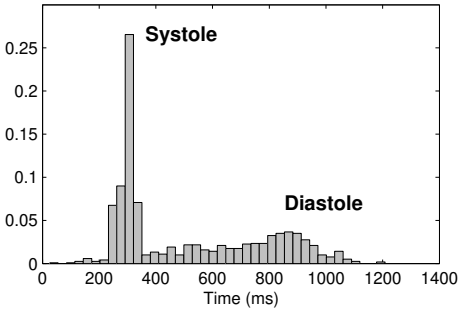

Figure 38. Histogram showing the distribution of systolic and diastolic duration in a heart cycle (based on 604 heart cycles from six subjects). It can be seen that the duration of systole is rather constant and normally shorter than the duration of diastole, giving a binominal distribution. Data obtained from paper I.

Different change detection methods have been employed for segmentation of heart sounds (unpublished work). One-model, two-model as well as multiple-model approaches were tried out, but neither of them was suitable for segmentation of heart sounds. A reason could be that they were all based on a linear regression framework, which was not able to separate changes due to heart sounds from changes due to noise. Parts of paper I was devoted to a nonlinear change detection method for the task of emphasizing S1 and S2 (as an alternative to the above mentioned transformations and change detection approaches). Heart sounds have a transient waveform that is superpositioned upon lung sounds and other disturbances. Since the heart sounds and the noise originate from different sources, they have different attractors, see Figure 39. These changes in signal dynamics can be detected with the change detection scheme in section 3.3.2 (based on recurrence times of the first kind, T1).

A sliding window was used to partition the phonocardiographic signal into overlapping segments to obtain time resolution. T1 is plotted in Figure 40. Since the application in paper I was to find and remove both S1 and S2, no attempts were made to actually classify the two sounds. This method was used in both paper I and paper II, and resulted in an error rate of 12.4 % and 0.5 %, respectively. The big difference in detection accuracy depends on heavy breathing in one of the provocation sequences in paper I.

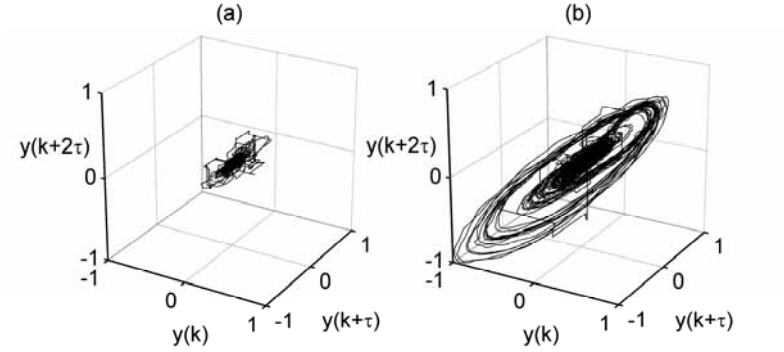

Figure 39. State space trajectories (d = 3, τ = 12) of a sound signal with S1 and S2 cut out (a). In (b) the whole signal including S1 and S2 is shown. The transition between the two attractors is reflected in the recurrence time statistic, hence indicating when a heart sound is present.

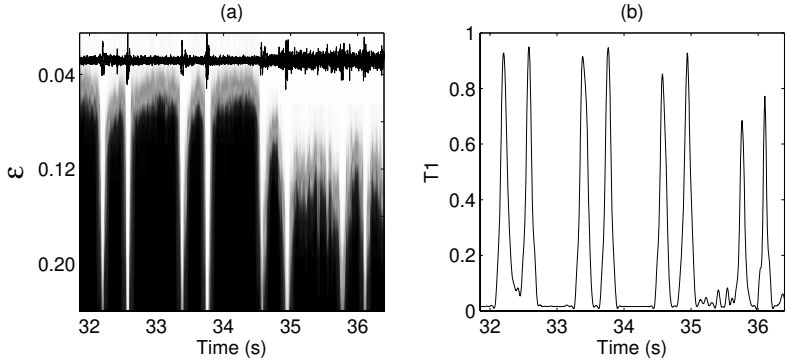

Figure 40. An example showing how the recurrence time statistic indicate the location of heart sounds. Note the obscuring noise (a deep breath) in the end of the signal. In (a) T1 is plotted over time for various ε-values where the grey scale indicates the strength of T1. Superimposed in the figure is the phonocardiographic signal (black waveform). T1(ε) for one fixed ε-value is plotted in (b).

5.2. Finding S3

The third heart sound occurs normally in children but disappears with increasing age. The sound can reappear in elderly persons and is clinically important because of its established connection with heart failure [8, 57]. Compared to the task of locating S1 and S2, finding S3 is harder due to its low amplitude, short duration and low frequency.

Previous methods to detect S3 is limited to a matched wavelet approach, where the mother wavelet was designed to have similar morphology as S3 [58, 59]. The idea was to divide the signal into four frequency bands; 17, 35, 60 and 160 Hz. S1 and

S2 are present in all frequency bands while a potential S3 should be represented in the three lower bands.

Paper II presents a novel technique based on the same method as the heart sound locator in paper I. Recurrence points of the first kind, T1, are used to locate S1 and S2 after which S3 is sought in time windows 100-300 ms after the two heart sounds. To avoid the problems involved in discriminating between S1 and S2, S3 was sought for within the predetermined time window following both heart sounds. As mentioned in section 3.3.2, T1 is more robust to noise while T2 is more sensitive to changes in the signal. S3 is a very weak signal, so T2 statistics is the better choice in this case.

Selecting the proper neighbourhood can be a problem. In Figure 41, T1 and T2 are plotted for a range of ε-values to visualize the dependence on the neighbourhood size. When looking for S1 and S2, a simple threshold can be used, but looking for S3 (in T2) is somewhat harder. The chosen approach was to use a whole range of ε-values to calculate a T2-matrix, see Figure 41d. The resulting 2D image can then be converted to 1D by an edge detection algorithm (here implemented by lowpass filtering and detection of the maximum value in each column). In the 1D signal, occurrences of S3 can be found by looking for a maximum within the previously defined time window. The detection rule could, for example, compare the amplitude of the maximum and the amplitude of the base line level. In paper II this rule states that the amplitude of the maximum should be one third larger than the base line level. Comparing the described method with the matched wavelet approach shows improved detection rate, 98 % compared to 93 %, at the expense of more false detections, 7 % compared to 2 %. The comparison was based on data from paper II consisting of ten children.

5.3. Filtering out Signal Components

Noise is a big problem in phonocardiography. The sensor, the sensor contact surface, the patient's position, the auscultation area, respiration phase and background noise all influence the quality of the sound. In practice this means that the recordings often contain noise such as friction rubs, rumbling sounds from the stomach, respiratory sounds from the lungs and background noise from the clinical environment. Most of these noise sources have frequency content in the same range as the signal of interest.

A special case of noise cancellation techniques deals with removal of transient noise. Potential use in a phonocardiographic setting is to remove disturbances such as friction rubs, but in paper I we focus on a related field in respiratory sound analysis, where the heart sounds themselves are the interfering noise.

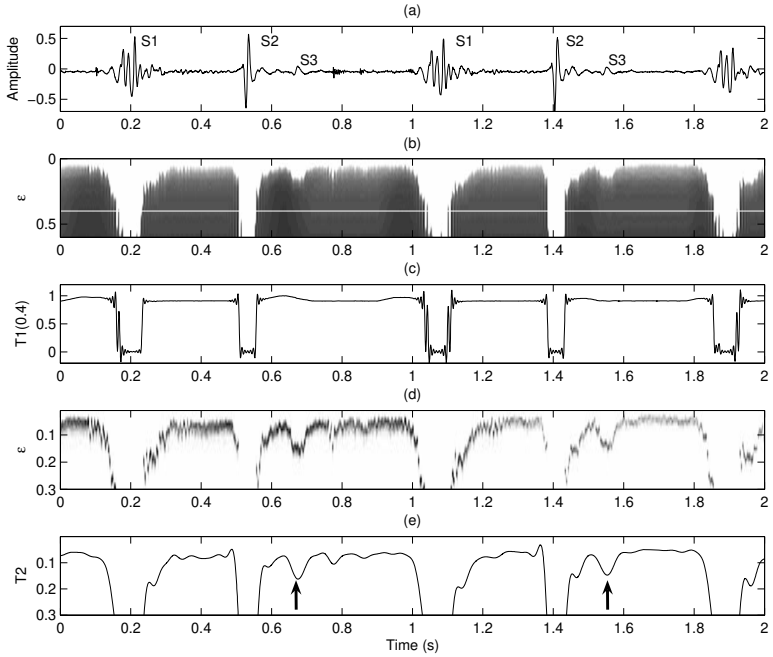

Figure 41. Example of a heart sound signal where S1, S2 and S3 are marked (a). T1, calculated for a range of ε-values, is shown in (b) while a single T1 is shown in (c) for ε = 0.4. T1(0.4) is used to find S1 and S2. T2, calculated for a whole range of ε-values is shown in (d). An edge detection algorithm is used to convert T2 to the 1D signal in (e) which is used to detect S3 (marked as arrows by the detection algorithm).

There are many different methods available for heart sound cancellation from lung sounds. Heart sounds and lung sounds have overlapping frequency spectra, and even though high pass filtering is often employed to reduce the influence of heart sounds, this results in loss of important signal information [60]. Previous approaches to heart sound cancellation include wavelet based methods [60], adaptive filtering techniques [61] and fourth-order statistics [62], all resulting in reduced but still audible HS. Recent studies indicate that cutting out segments containing HS followed by interpolation of the missing data yields promising results [63, 64]. The method developed in paper I is based on work by Thomas et al. [63, 64], but the used signal processing techniques are fundamentally different and allow nonlinear behaviour in the lung sound signal. This is an important difference since it has been indicated that lung sounds are indeed nonlinear [65-68].

The method suggested in paper I uses the heart sound locator presented in section 5.1. The detections are simply cut out and the resulting gaps are filled with predicted lung sound using the nonlinear prediction scheme described in section

3.4. Since the prediction error grows exponentially with prediction length [23], both forward and backward prediction was used (hence dividing the missing segment in two parts of half the size). To avoid discontinuities in the mid point, the number of predicted points was allowed to exceed past half of the segment. The two predictions were then merged in the time domain close to the midpoint at an intersection where the slopes were similar.

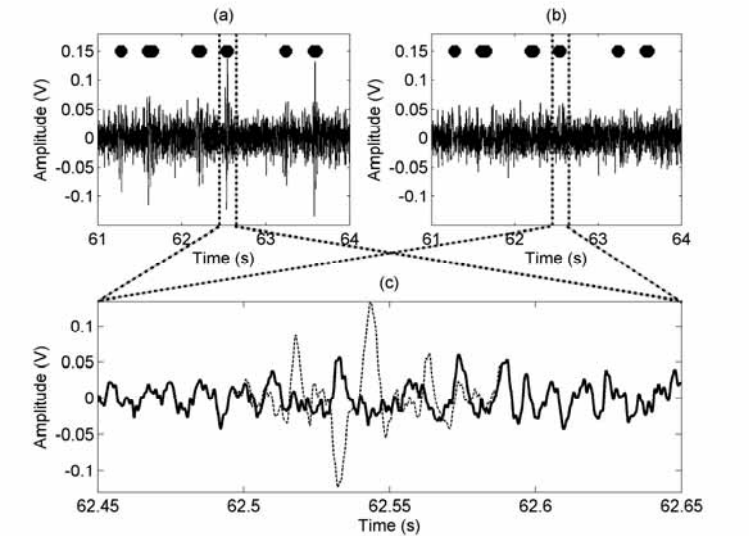

Figure 42. Example of a recorded lung sound signal with heart sounds present (a) and reconstructed lung sounds with heart sounds removed (b). The bars indicate heart sound detections. A zoomed in version showing the predicted lung sound (solid) and lung sound including heart sounds (dashed), is shown in (c).

The results are a bit hard to evaluate since the actual lung sound is unknown in the segments that are predicted. However, the waveform similarity between predicted segments and actual lung sound data is very high with a cross-correlation index of CCI = 0.997±0.004. The spectral difference was 0.34±0.25 dB/Hz, 0.50±0.33 dB/Hz, 0.46±0.35 dB/Hz and 0.94±0.64 dB/Hz in the frequency bands 20 – 40 Hz, 40 – 70 Hz, 70 – 150 Hz and 150 – 300 Hz, respectively. Since the main objective of the method was to give auditory high-quality results, a simple complementary listening test was performed by a skilled primary health care physician. The impression was that most heart sounds had been successfully replaced, but that some predictions had a slightly higher pitch than pure lung sounds. An example of heart sound cancellation is illustrated in Figure 42.

5.4. Classification of Murmurs

Common for all classification tasks is the importance of appropriate features, i.e. measures that retain similarities within classes while revealing differences between classes. Chapter 4 described characteristics of the phonocardiographic signal in

different domains. Although not mentioned explicitly, the underlying goal of the characterization was to find data representations with distinct differences between heart murmurs. This section is devoted to the problem of extracting features from these data representations and selecting those features with best discriminative power. Finally, the selected features are used to classify three different heart murmurs; aortic stenosis, mitral insufficiency and physiological murmurs.

5.4.1. Feature extraction

Feature extraction is about quantifying available information into a few descriptive measures. The different ways data were viewed upon in chapter 4 should hence be summarized in a few informative features. In phonocardiographic classification, the features are derived on a heart cycle basis. Knowledge about the accurate timing of events in the heart cycle is thus of great importance. Segmentation into the first heart sound (S1), systole, the second heart sound (S2) and diastole were discussed in section 5.1. In paper III the ECG gating technique was used, mostly because of its superior robustness when the phonocardiographic signal is noisy. Once the features have been extracted, averaging each feature over available heart cycles reduces the influence of noise.

A large amount of features were extracted in paper III. Some features provided information about the phonocardiographic signal's time varying behaviour. Shannon energy was used to measure intensity and a wavelet detail was used to measure intensity in a certain frequency interval. The variance fractal dimension was used to measure the time progress of the signal's complexity. An example of these features is shown in Figure 43.

Stockwell's TFR was used as a foundation when investigating how the frequency content varied over time. These representations are information rich, but they also contain a lot of data. Assuming a sample rate of 44.1 kHz and that one heart cycle is one second long, then each TFR matrix will contain $1.9448 \cdot 10^9$ samples. Clearly, considerable data reduction is required. The easiest way to do this is by down-sampling as in Figure 44. A perhaps more refined method for data reduction of TFR matrices use singular value decomposition. More details about this approach can be found in paper III.

These measures are based on second order statistics. Expanding the view to third order statistics, a measure of the non Gaussianity and nonlinearity of the signal can be obtained. Similarly to TFR down-sampling, the same technique can be applied to the bispectra, see Figure 28. Finally, features from the reconstructed state space can be derived. This procedure follows section 4.3 and uses the RQA measures and the GMM measures as features. These last additional features incorporate nonlinear behaviour into the description of the signal.

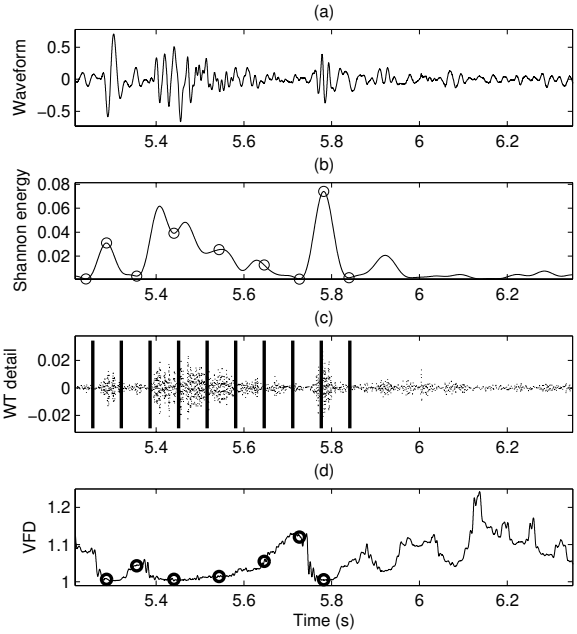

Figure 43. An example showing one heart cycle from a patient with aortic stenosis (a). In (b) the signal's envelope has been extracted (Shannon energy), the rings indicate the selected features. A wavelet detail is illustrated in (c), where the vertical lines are time markers equidistantly distributed over the region of interest. The absolute sum between each marker constitutes feature values. In (d) the variance fractal dimension trajectory is plotted together with the selected features marked as rings.

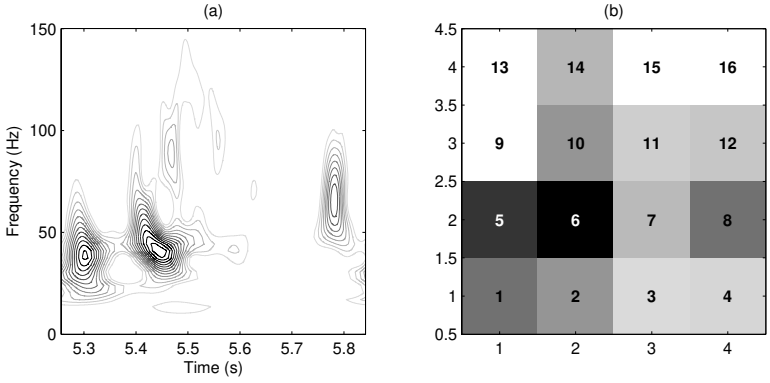

Figure 44. Time frequency representation of one systolic heart beat from a patient with aortic stenosis (a), S1 can be seen at 5.3 s and S2 at 5.8 seconds. In (b) the same data has been discretized into a 4x4 map of features.

5.4.2. Finding relevant features

A large number of features were described in the previous section. However, too many features often result in classifiers with low generality [69]. Actually, there are many potential benefits in reducing the number of features; facilitating data visualization, reducing the measurement and storage requirements, reducing training and utilization times and defying the curse of dimensionality [70].

Auscultation is an old science where a lot of information and experience have been gathered over the years. This domain knowledge should be incorporated in the classification task, and as a matter of fact, it is already included in the features. Many of the features in section 5.4.1 are based on the changes in timing, intensity and pitch that physicians use to separate normal from abnormal phonocardiographic signals. The first step in any feature selection process should be to incorporate domain knowledge, but as this was done in the feature extraction process, this important step can be omitted.

Scalar feature selection means that each feature is treated individually. A scoring function is defined, and its outcome indicates the predictive power of the feature. This way, all features can be ranked in decreasing order, and the best ones are selected for the classification task. The scoring function could, for example, be the distance from the feature to the centre of the distribution of the class it is supposed to belong to. Selecting the most (individually) relevant features is usually suboptimal for building a predictor though, particularly if the selected features are correlated and thus contain redundant information [70].

A problem with scalar feature selection is that is does not account for combinations of features that together have great predictive power. An optimal selection of features requires an exhaustive search over all features, but this is practically infeasible. Instead suboptimal search algorithms are employed, many of which uses greedy hill climbing (hill climbing is a search algorithm where the current path is extended with a successor node which is closer to the solution than the end of the current path). A possible subset of features is then evaluated, and other features are successively added or removed from this set to see if an improvement can be achieved [70]. A simple way to do this is to start with one feature (the one with highest ranking according to scalar feature selection), say x_1. Expand the set to contain two features by forming all possible pairs, say $\{x_1, x_2\}$, $\{x_1, x_3\}$, $\{x_1, x_4\}$. The pair that maximizes some class separability criterion is selected as the new feature subset. More features are then progressively added into larger and larger subsets until the desired number of features is reached. This method is often used when the size of the final subset is supposed to be small compared to the total amount of features. If the final subset is supposed to be large, then all features could be included in a preliminary subset which is progressively reduced. These methods are called sequential forward selection and sequential backward selection. A common drawback for both of these is that once a feature is included there is no way of getting rid of it (and vice versa in backward selection). Pudil's sequential floating forward selection is a workaround to this

problem, allowing features to be both included and excluded several times [71]. A flow chart describing the algorithm is presented in Figure 45.

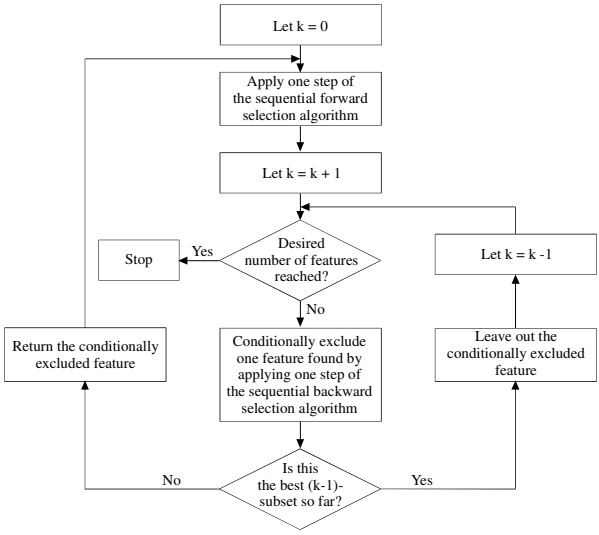

Figure 45. Flow chart of Pudil's sequential floating forward selection method, where *k* is the current number of features in the subset.

In paper III, Pudil's sequential floating forward selection method was used to reduce the number of features from 213 to 14. Inclusion or rejection of features was based on the error estimate of a 1-nearest neighbour leave-one-out classifier where the performance criterion equals 1 – the estimation error. The number of features used in the final set was chosen to maximize the performance criterion while keeping the number of features as low as possible.

14 features were selected, see Table 3, and the resulting set was denoted the SFFS subset. Bearing in mind that the investigated murmurs are aortic stenosis, mitral insufficiency and physiological murmurs, the selected features are actually very reasonable. A wavelet detail represents the end of systole, where it can be used to separate holosystolic mitral insufficiency murmurs from physiological murmurs and aortic stenosis murmurs which are of crescendo-decrescendo shape. Three Shannon energy measures represent the signal's intensity in mid systole, thereby describing the shape of the murmur in the time domain. A fractal dimension measure represents the complexity of the murmur in relation to the heart sounds. This measure can be seen as the amplitude normalized complexity of the murmur. Another fractal dimension measure, located at S1, represents the change of S1 that is associated with mitral insufficiency. Remaining features are a bit hard to explain in a physiologically meaningful way.

Table 3. A brief summary of the features selected by Pudil's sequential forward feature selection method, the SFFS subset.

Wavelet detail: One feature representing the end of systole
Wavelet entropy: One feature describing the information content in the high frequency range.
Shannon energy: Three features in mid systole able to describe the shape and intensity of the murmur and one feature after S2 revealing the noise level.
Stockwell's TFR: Two features giving a collected view of the low frequency content over the heart cycle.
Bispectrum: One feature indicating phase coupling and frequency content for low frequencies.
Reconstructed state space: Three features describing the width of the Gaussian mixture model (probably located in the part of state space where the murmur lives), two of these belong to the largest mixture.
Variance fractal dimension: Two features, one giving the amplitude normalized complexity of the murmur and the other describing S1.

5.4.3. Classifying murmurs

The work in this thesis does not intend to cover the classification step. However, to show the abilities of the feature set derived in the last section, a simple, off-the-shelf classifier was however employed. A fully connected feed-forward neural network was set up, with logarithmic sigmoid transfer functions and biased values throughout. The number of input units was set to the nearest larger integer of the square root of the number of features in the set, the number of units in the hidden layer was set to three and the number of output units was set to two. The target values were 00 (mitral insufficiency), 01 (aortic stenosis) or 10 (physiological murmur). Each output from the network was thresholded at 0.5 and compared to the results from a clinical echocardiography investigation. A leave-one-out approach was used for training and testing due to the limited amount of patients.

Three additional subsets were used for comparison with the SFFS subset. The first set consisted of the Shannon energy features, the second set consisted of the wavelet detail features and was adapted from [17, 72] and the third set consisted of a TFR decimated to 4x4 values (adapted from [73]).

Confusion matrices showing the classification results for the four tested feature subsets are presented in Table 4. The percentage of correct classifications was 58%, 44%, 53% and 86% for Shannon energy, Wavelet detail, TFR and the SFFS subset, respectively. If all pathological cases (mitral insufficiency and aortic stenosis) are treated as one group, sensitivity and specificity can be specified. In this setting, the sensitivity was 90%, 90%, 93% and 93% and the specificity was 28%, 14%, 57% and 100%, respectively. The number of patients with valve

pathology that were erroneously classified as physiological was comparable for all feature subsets; 10%, 10%, 7% and 7%, respectively.

Table 4. Confusion matrices showing the classification results from four different feature subsets. Target groups are presented horizontally while the predicted groups are presented vertically. Each number represents number of patients (total number of patients in the study is 36).

	Shannon Energy			Wavelet detail			TFR features			SFFS		
	AS	MI	PM	AS	MI	PM	AS	MI	PM	AS	MI	PM
AS	17	4	4	15	5	2	14	4	3	19	1	0
MI	3	2	1	6	0	4	8	1	0	2	5	0
PM	3	0	2	2	1	1	1	1	4	2	0	7

6. Discussion

"Doubt is not a pleasant condition,
but certainty is absurd."
Voltaire (1694 - 1778)

The focus of this thesis has been to investigate and develop new tools to facilitate physicians' daily work. Evaluation of patients with heart disease is a complex task, where auscultation provides one piece of the puzzle. Therefore, our intelligent stethoscope is not to be seen as a tool capable of replacing clinicians, but rather as a provider of quantitative decision support. The stethoscope's main usage will be in the primary health care, when deciding who requires special care. Consequently, it should neither be seen as a replacement of more advanced techniques such as echocardiography.

The main tasks for the intelligent stethoscope are to improve sound quality, to emphasize weak or abnormal events (such as reverse splitting of S2) and to distinguish different heart murmurs from each other. In a small pilot study from 2002, interviews with nine primary health care physicians revealed that the most interesting task for an intelligent stethoscope was classification of heart murmurs, especially to distinguish physiological murmurs from pathological murmurs.

A danger with projects such as the intelligent stethoscope is that technology is sometimes introduced for the sake of technology. Heart sound cancellation from lung sounds (paper I) tends in this direction, see section 6.1. Detection of the third heart sound is somewhat different since S3 can be very difficult to hear. Notifying the physician that a third heart sound is present could thus be of great value (paper II). When it comes to decision support and classification (paper III), the intended use of the system becomes an important issue, see section 6.2.

6.1. Context of the Papers

The research on signal processing of heart sound recordings has been extensive [74, 75], and several authors have investigated the possibility to automatically classify cardiac murmurs. The contribution of paper III is partly a thorough survey of features available for murmur classification and partly the addition of several

new features. The survey is based on features from the literature, ranging from time domain characteristics [76-78], spectral characteristics [79-81] and frequency representations with time resolution [17, 72, 73, 81-83]. The main contribution compared to previous works is the incorporation of nonlinear and chaos based features, a source of information that has not previously been explored in the context of heart murmur classification.

There are a number of methods available for heart sound cancellation from lung sound recordings. This is quite interesting since heart sound cancellation have limited clinical use (physicians are able to more or less ignore heart sounds while tuning in on lung sounds during auscultation). The problem at hand is a very intriguing engineering problem though, and this is probably one reason for its popularity. A justification of all these methods is that automatic classifiers seem to be confused by the heart sounds. When trying to separate different lung diseases based on lungs sounds, results tend to improve after removal of the heart sounds. Some of the previous approaches to heart sound cancellation include wavelet based methods [60], adaptive filtering techniques [61] and fourth-order statistics [62], all resulting in reduced but still audible HS. The contribution of paper I is that nonlinear behaviour in the lung sound signal is taken into account while cancelling out the heart sounds. This is an important extension since it has been indicted that lung sounds are indeed nonlinear [65-68].

In contrast to the other two papers, there is not much work available on detection of the third heart sound. A matched wavelet approach giving good results has previously been developed in our research group [58, 59]. The method presented in paper II is based on a change detection scheme developed to find weak transient components in signals. Compared to the wavelet approach, the change detection method finds more third heart sounds at the expense of more false detections. It is thus a complement rather than a replacement of the wavelet method, where the new approach could be used to find the third heart sounds while the wavelet approach could be used to exclude false detections.

6.2. Patients and Data Sets

Since this project is about developing of a new tool, a full clinical trial was not the goal when assessing the developed methods. Nevertheless, to emulate the environment where the intelligent stethoscope most likely will be used, most data has been recorded in a clinical environment (paper I and III). Below is a run-through of limitations implied by the selected study groups in paper I-III.

Paper I: This study was performed on six healthy male subjects (aged 28.5 ± 3.7 years). Limitations implied by the study population are that no adventitious lung sounds are present. Future work should thus contain patients with various pulmonary diseases. These limitations result in a few implications:

1. How would the heart sound detector perform in the presence of adventitious sounds? In paper I, it is indicated that the detector performs quite well even when the signal is obscured by heavy breathing. The

attractor of wheezing sounds has a similar morphology as the attractor of heart sounds, and it is not inconceivable that wheezes will disturb the detection algorithm. In paper I, the measurements contained several friction rubs and one test subject had a distinct third heart sound. Most of these friction rubs and third heart sounds were detected and removed along with the other heart sounds. Explosive lung sounds like crackles will probably be marked by the method as well. By including extra criteria, such as interval statistics or the degree of impulsiveness, it is possible that these false detections could be avoided.

2. How would the prediction scheme perform in the presence of adventitious lung sounds? Wheezing sounds have a periodic structure and their attractor is well defined (an example can be found in [84]). This would pose no problem to the predictor. Crackles on the other hand occurs rather stochastic (actually they come in avalanches following a fractal behaviour [85]), and are thus harder to predict. This will most likely fail.

Paper II: Ten healthy children were used in this study (male = 5, female = 5, aged 5-13 years), mostly because third heart sounds with high signal quality are common in this group. Investigation of pathological phonocardiograms is left for future studies. In a validation study, patients with heart failure should be chosen. After all, the intended use of the algorithm is to use S3 as a marker of heart failure.

Paper III: In total, 36 patients (19 male, 17 female, ages 69 ± 14 years, all with native heart valves) were enrolled in the study. The prerequisite for including a patient in the study was that a primary health care physician had identified a murmur and that the patient had been sent to the cardiology clinic for further investigations. As stated previously, physicians ask for a method able to separate physiological murmurs from pathological murmurs. The usage of such a tool is however more applicable in a young population. In a material with data from elderly patients, it is more interesting to determine the actual source of the murmur. This is exemplified in the following heart murmur classification scenarios, where the demands on the system turn out to be quite different from each other.

1. *Screening in a young population.* Physiological murmurs are very common in children, and methods able to separate physiological murmurs from pathological murmurs would be of great value. However, the performance requirements are extremely high. For example, physiological murmurs have to be separated from even very mild aortic stenosis due to the latter's implications on choice of profession, insurance issues and whether follow-ups of a possible disease are needed or not. Due to the consequences of an incorrect diagnosis, the tolerance for false positives and negatives is low. The performance requirements of the system are huge and the interesting question it all comes down to is: does the phonocardiographic signal contain enough information?

2. *Measuring the degree of stenosis in the elderly.* Pathological changes in the aortic valves are common in the elderly. Usually this change has little physiological importance since the stenosis is mild. However, it is important to find those patients who really have a significant narrowing of the valve opening; partly because surgical correction improves the prognosis of these patients and partly because, when surgery is out of the question, some medications should be avoided. The classification task would then be to measure the degree of the stenosis and decide whether the stenosis is mild or moderate/severe. This scenario is easier so solve compared to the previous scenario since the grey area between a physiological murmur and a mild stenosis is not of interest.

6.2.1. Measurement noise

The data sets used in this thesis where all more or less affected by noise, and as a matter of fact, the data in paper I and III were actually very noisy. In practice this means that the recordings contained friction rubs, rumbling sounds from the stomach, breathing sounds from the lungs (a necessity in paper I and an obstacle in paper III) and background noise from the clinical environment. All of these influence the recorded sound in a negative way, but a perhaps bigger problem is handling the stethoscope. This is a setback since the whole idea behind the intelligent stethoscope is that it should be easy to use (perhaps even in home care by the patient herself). Firm application of the sensor cannot be stressed enough to achieve high quality recordings.

6.3. Methodology

Most methods used in this thesis suffer from high computational burden. This is a problem since the software is supposed to be implemented in a portable stethoscope, preferably in real time. It is however difficult to assess the actual performance limitations of the used methods because they were never designed to be quick or efficient. A number of potential speed-ups come to mind.

- Matlab was used to implement all algorithms, but using a lower level language would increase performance.
- Fast nearest neighbour routines are available, but currently a very simple search routine is used.
- A sliding window approach is often used to gain time resolution. The reconstructed state space is nearly identical between iterations due to the overlap between segments, and this fact is not exploited.

A fundamentally different bottle-neck is the fact that the some calculations are non-causal. For instance, many of the features in paper III were derived as averages over all available heart cycles. In most cases this could be dealt with by only using old data. Accumulated statistics could then be used to increase the accuracy of the output as more data become available.

Significance tests could have been performed to statistically verify if there were any differences between the different groups in paper III. The number of tests would have been large, trying to separate three groups from each other in 213 cases (the total number of features). There are at least two reasons why these tests were not performed. Firstly, when performing a great number of statistical tests, the probability of getting significant differences by chance is rather high and secondly, variables that are useless by themselves might be useful in combination with others.

The greatest problems when using chaos based signal analysis tools are that the results are almost always open for interpretation, that nearly noise free data is required and that the amount of data should be large. Phonocardiographic data is rather cyclo-stationary than nonstationary, so by concatenating stationary segments, large data sets can be obtained. In this thesis, these segments were simply concatenated in the time domain, while a better approach would have been to append the reconstructed state space matrices to each other. An extra flag would then be appended to each coordinate, keeping track of the last coordinate in each segment. This way, false neighbours due to the concatenation procedure can be excluded from the calculations. This addendum would have impact on all methods where a reconstructed state space is constructed from concatenated data (prediction in paper I and certain features in paper III).

Estimation of fractal dimension characteristics should also be based on large enough data sets [86]. This implies a trade-off between time resolution and accuracy in the estimation of the time dependent fractal dimension (similar to the uncertainty principle when calculating time-frequency representations). As the investigated signal segment does not possess self-similarity over an infinite range of scales, the self-similar properties of the segment are lost if the sliding window is too short. Similarly, if the window size is set too long, the different characteristics of consecutive signal segments will be blurred together. Another reason for not using too short windows is that the number of signal amplitude increments used to calculate the variance in the variance fractal dimension algorithm must be greater than 30 to be statistically valid [24].

Ideally, four different data sets should have been used in paper III. One set for selecting analysis methods able to extract features, one set for selecting the most important features, one set for training the classifier and a final set for validation. This set-up requires an extensive database of phonocardiographic signals. Unfortunately, this was unattainable within the scope of this project. A good compromise when the number of data is limited is the leave-one-out approach. Here the training is performed on N-1 samples and the test is carried out on the excluded sample. This is repeated till all samples have been tested. The results when using separate training and testing sets compared to the leave-one-out approach are very similar [69]. However, the independence between the training and testing sets are compromised since the same sets are used for both feature selection and training of the classifier.

Selecting descriptive measures to demonstrate the performance of an algorithm can be cumbersome. In paper I and II, receiver operating curves (ROC) would have been informative to describe the dependence of thresholds. This was not feasible since all detections had to be confirmed or rejected manually.

In paper I, the accuracy of the predictions was difficult to verify since the true signal was not known. The spectral difference between the reconstructed lung sound signal and a lung sound signal with cut-out heart sounds was small. However, comparing spectral densities might be questionable as phase information is lost. Another attempt is to compare the waveform similarity between predicted segments and actual lung sound data. The results from this measure were very good, but this is not surprising since the prediction scheme exploits that trajectories in state space share the same waveform characteristics in time domain (i.e. the prediction tries to reproduce past parts of the time series). Since both of these measures are questionable, a simple complementary listening test was performed by a physician. However, such tests are certainly not quantitative.

6.4. Future Work

6.4.1. Clinical validation

The methods presented in paper I-III have not been clinically evaluated, validation on a large number of patients are thus necessary. The number of patients is especially low in paper III, where the methods used for feature selection as well as classification requires a lot of data to give reliable results. When recording a new material, the clinical need and the selected patient groups should be re-evaluated together with medical personnel. The two scenarios in section 6.2 are typical examples of interesting study groups.

6.4.2. Multi-sensor approach

The magnitude of different components in the phonocardiographic signal varies with the measurement location. For instance, listening over the apex, S1 is louder than S2. Also, the location where a heart murmur is best heard often indicates its origin. By using multiple sensors in parallel, this difference in intensity could be used as a parameter in a classification system. Further uses could be to derive time differences between the different signals and, using this information, calculate an estimate of the location of the event. For instance, using S1 and S2 as reference locations, the murmur location could be of diagnostic value. A third possible use of multiple sensors is to use several sources of the signal when creating the reconstructed state space. This would probably give better resistance to measurement noise and, above all, a better embedding of the signal.

The major drawback with multiple sensors is that the simplicity of the stethoscope would be afflicted. Nevertheless, if additional sensors are to be used, there are no reasons not to use conceptually different techniques such as ECG and Doppler ultrasound.

6.4.3. Dimension reduction

The recorded phonocardiographic signal is one-dimensional, and this signal can be unfolded into a high-dimensional space (section 3.2). This is all very nice, but how should the results be interpreted when they cannot even be visualized. One possibility is to use recurrence plots (section 3.3.2), but there might be better ways. Classical linear methods for dimension reduction include principal component analysis and multidimensional scaling. A drawback with linear methods is that data may not always be accurately summarized by linear combinations. An example is a helix in a three dimensional space, whose one dimensional structure cannot be discovered by linear methods.

Since the manifold on which the embedded phonocardiographic signal resides is non-linear, linear methods are not adequate. Instead of finding the most important linear subspace from a set of data points (like in principal component analysis), non-linear parameterizations can be sought. Many of these *manifold learning* techniques are based on a distance matrix (like multidimensional scaling, which tries to preserve the distance between each data point even after the dimensionality has been decreased). In principle, the difference compared to multidimensional scaling is the way that distances are calculated. Instead of measuring a global distance (a straight line through space), the distance is calculated by summing up local distances between states as we move from one point to another (i.e. we are only allowed to travel from point a to point b via other data points located in between).

A problem associated with these techniques is whether the embedded phonocardiogram really resides on a manifold and if so, whether the manifold is sampled dense enough to be able to draw any valid conclusions.

6.4.4. Choosing an appropriate classifier

A proper set-up of the classifier in paper III, and, for that matter, the choice of the actual classifier, was out of the scope of this thesis. The common picture is that the actual type of classifier is not extremely important, at least not in comparison with the importance of good features. Nevertheless, a good classifier should be used, and some effort should be spent on its design. Looking at it from a practical point of view, the chosen classifier might be adequate enough. After all, it works properly for the application in question.

7. Review of Papers

"There is a coherent plan in the universe,
though I don't know what it's a plan for."
Fred Hoyle (1915 - 2001)

This chapter introduces the papers which are included in the second part of this thesis.

7.1. Paper I, Heart Sound Cancellation

This paper presents a new method to detect and remove heart sounds from recorded lung sound signals. The idea is to locate the heart sounds, remove them altogether and predict what should have been in their place using surrounding data. The location of the heart sounds is found using a nonlinear change detection method based on recurrence time statistics. Once the heart sounds have been removed, the missing gap is filled in using a nonlinear prediction scheme.

The reason for using this rather odd change detection algorithm was that the data was very noisy (high flow in the respiratory part of the signal). The error rate was 4% false positives and 8% false negatives using this method. Results from other algorithms were not explicitly derived, but fast visual comparisons rejected methods like Shannon energy. Similarly, the choice of a nonlinear local prediction method was based on the fact that lung sounds are indeed nonlinear.

The proposed solution to the problem of heart sound cancellation was tested on six subjects. The spectral difference between the denoised lung sound signal and a lung sound signal with removed heart sounds was 0.34±0.25 dB/Hz, 0.50±0.33 dB/Hz, 0.46±0.35 dB/Hz and 0.94±0.64 dB/Hz in the frequency bands 20-40 Hz, 40-70 Hz, 70-150 Hz and 150-300 Hz, respectively. The cross-correlation index was found to be 99.7%, indicating excellent similarity between actual lung sound data and predicted lung sound data.

7.2. Paper II, Detection of the 3rd Heart Sound

An algorithm, actually based on the same recurrence statistics as the heart sound locator in paper I, was developed for detection of the third heart sound. Recurrence

points of the first kind (T1) are known to be rather noise insensitive and robust, while recurrence points of the second kind (T2) are better suited for finding very weak signals. The two could be used as a pair where T1 is used on a coarse scale to locate the first and second heart sounds, while T2 could be used on a finer scale to locate the third heart sound within a predetermined window succeeding the other heart sounds. The reason to look for S3 after both S1 and S2 was to avoid the problem of distinguishing the two (statistics about their timing is hard to use on a young population where the heart rate is high).

Since S3 is normally heard during auscultation of younger individuals, the method was tested on ten children. Most S3 occurrences were detected (98 %), but the amount of false extra detections was rather high (7% of the heart cycles).

7.3. Paper III, Feature Extraction from Systolic Murmurs

Available features for heart murmur classification were reviewed and new features, mostly inspired by research in speech processing, dynamical systems and chaos theory, were introduced. Techniques such as Shannon energy, wavelets, fractal dimensions and recurrence quantification analysis were used to extract 213 different features. 163 of these features have not previously been used for heart murmur classification. However, this many features often result in high computational complexity, mutual correlation and classifiers with low generality. Not to mention the curse of dimensionality. For this reason, a subset selection method was employed to reduce the number of features. The derived feature subset, maximizing a performance criterion while keeping the number of features low, resulted in a subset of 14 features.

Heart sound data from 36 patients with aortic valve stenosis, mitral insufficiency or physiological murmurs were used to test the new feature set. Using the results from a neural network classifier, this new feature set was compared with three other feature sets. The selected subset gave the best results with 86 % correct classifications, compared to 58 % for the first runner-up. In conclusion, the derived feature set was superior to those previously used, seems rather robust to noisy data sets and will be tested in more clinically oriented studies in the future.

References

[1] L. A. Geddes, "Birth of the stethoscope," *Engineering in Medicine and Biology Magazine, IEEE*, vol. 24, pp. 84, 2005.

[2] A. N. Pelech, "The physiology of cardiac auscultation," *Pediatr Clin North Am*, vol. 51, pp. 1515-1535, 2004.

[3] S. Persson and J. Engqvist, *Kardiologi: hjärtsjukdomar hos vuxna*, 5., rev. och utök. uppl. / ed. Lund: Studentlitteratur, 2003.

[4] G. J. Borden, K. S. Harris, and L. J. Raphael, *Speech science primer: physiology, acoustics, and perception of speech*, 3. rev. ed. Baltimore: Williams & Wilkins, 1994.

[5] D. Smith and E. Craige, "Heart Sounds: Toward a Consensus Regarding their Origin," *Am. J. Noninvas. Cardiol.*, vol. 2, pp. 169-179, 1988.

[6] L. Wigstrom, T. Ebbers, A. Fyrenius, M. Karlsson, J. Engvall, B. Wranne, and A. F. Bolger, "Particle trace visualization of intracardiac flow using time-resolved 3D phase contrast MRI," *Magn Reson Med*, vol. 41, pp. 793-799, 1999.

[7] A. G. Tilkian and M. B. Conover, *Understanding heart sounds and murmurs: with an introduction to lung sounds*, 4. ed. Philadelphia: Saunders, 2001.

[8] N. J. Mehta and I. A. Khan, "Third heart sound: genesis and clinical importance," *Int J Cardiol*, vol. 97, pp. 183-186, 2004.

[9] R. B. Northrop, *Noninvasive instrumentation and measurement in medical diagnosis*. Boca Raton, Fla., London: CRC; Chapman & Hall, 2002.

[10] H. Nygaard, "Evaluation of Heart Sounds and Murmurs - a Review with Special Reference to Aortic Valve Stenosis," Department of Electrical Engineering, College of Engineering, Aarhus, Denmark 1996.

[11] L. Vannuccini, J. E. Earis, P. Helistö, B. M. G. Cheetham, M. Rossi, A. R. A. Sovijärvi, and J. Vanderschoot, "Capturing and preprocessing of respiratory sounds," *Eur Respir Rev*, vol. 10, pp. 616-620, 2000.

[12] R. J. Povinelli, M. T. Johnson, A. C. Lindgren, F. M. Roberts, and J. Ye, "Statistical Models of Reconstructed Phase Spaces for Signal Classification," *IEEE Transactions on Signal Processing*, vol. In press, pp. -.

[13] R. L. Allen and D. W. Mills, *Signal analysis: time, frequency, scale and structure*. New York, Piscataway, N.J.: Wiley; IEEE Press, 2004.

[14] R. N. Bracewell, *The Fourier transform and its applications*, 3. ed. Boston: McGraw Hill, 2000.

[15] H. Liang, S. Lukkarinen, and I. Hartimo, "Heart sound segmentation algorithm based on heart sound envelogram," in *Computers in Cardiology*, Lund, Sweden, 1997, pp. 105-108.

[16] J. F. Kaiser, "Some useful properties of Teager's energy operators," in *ICASSP-93*, 1993, pp. 149-152.

[17] C. N. Gupta, R. Palaniappan, S. Swaminathan, and S. M. Krishnan, "Neural network classification of homomorphic segmented heart sounds," *Applied Soft Computing*, vol. In Press, Corrected Proof.

[18] G. Livanos, N. Ranganathan, and J. Jiang, "Heart sound analysis using the S transform," in *Computers in Cardiology*, 2000, pp. 587 - 590.

[19] R. G. Stockwell, L. Mansinha, and R. P. Lowe, "Localization of the complex spectrum: The S transform," *Ieee Transactions on Signal Processing*, vol. 44, pp. 998-1001, 1996.

[20] U. Parlitz, "Nonlinear time-series analysis," in *Nonlinear modeling: advanced black-box techniques*, J. A. K. Suykens and J. Vandewalle, Eds. Boston: Kluwer Academic Publishers, 1998, pp. 256.

[21] H. D. I. Abarbanel, *Analysis of observed chaotic data*. New York: Springer-Vlg, 1996.

[22] L. Y. Cao, "Practical method for determining the minimum embedding dimension of a scalar time series," *Physica D*, vol. 110, pp. 43-50, 1997.

[23] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*, 2. ed. Cambridge: Cambridge Univ. Press, 2004.

[24] W. Kinsner, "Batch and real-time computation of a fractal dimension based on variance of a time series," Dept. of Electrical & Computer Eng., University of Manitoba, Winnipeg, Canada DEL94-6, June 1994.

[25] J. P. Zbilut, N. Thomasson, and C. L. Webber, "Recurrence quantification analysis as a tool for nonlinear exploration of nonstationary cardiac signals," *Medical Engineering & Physics*, vol. 24, pp. 53-60, 2002.

[26] C. L. Webber and J. P. Zbilut, *"Recurrence quantification analysis of nonlinear dynamical systems." In: Tutorials in Contemporary Nonlinear Methods for the Behavioral Sciences*: National Science Foundation, 2005.

[27] J. B. Gao, "Recurrence time statistics for chaotic systems and their applications," *Phys. Rev. Lett.*, vol. 83, pp. 3178-3181, 1999.

[28] N. Marwan, N. Wessel, U. Meyerfeldt, A. Schirdewan, and J. Kurths, "Recurrence-plot-based measures of complexity and their application to heart-rate-variability data," *Phys. Rev. E.*, vol. 66, pp. 1-8, 2002.

[29] C. L. Webber and J. P. Zbilut, "Dynamical assessment of physiological systems and states using recurrence plot strategies," *J. Appl. Physiol.*, vol. 76, pp. 965-973, 1994.

[30] F. Gustafsson, *Adaptive filtering and change detection*. Chichester: Wiley, 2000.

[31] J. B. Gao, Y. H. Cao, L. Y. Gu, J. G. Harris, and J. C. Principe, "Detection of weak transitions in signal dynamics using recurrence time statistics," *Physics Letters A*, vol. 317, pp. 64-72, 2003.

[32] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent component analysis*. New York: Wiley, 2001.

[33] C. L. Nikias and J. M. Mendel, "Signal processing with higher-order spectra," *IEEE Signal Processing Magazine*, vol. 10, pp. 10 - 37, 1993.

[34] L. Ljung, *System identification: theory for the user*, 2. ed. Upper Saddle River, N.J.: Prentice Hall, 1999.

[35] J. McNames, "A nearest trajectory strategy for time series prediction," in *Proc. Int. Workshop on Advanced Black-Box Techniques for Nonlinear Modeling*, Leuven, Belgium, 1998, pp. 112-128.

[36] A. P. Yoganathan, R. Gupta, F. E. Udwadia, J. W. Miller, W. H. Corcoran, R. Sarma, J. L. Johnson, and R. J. Bing, "Use of the fast Fourier transform for frequency analysis of the first heart sound in normal man," *Med Biol Eng*, vol. 14, pp. 69-73, 1976.

[37] A. P. Yoganathan, R. Gupta, F. E. Udwadia, W. H. Corcoran, R. Sarma, and R. J. Bing, "Use of the fast Fourier transform in the frequency analysis of the second heart sound in normal man," *Med Biol Eng*, vol. 14, pp. 455-460, 1976.

[38] M. S. Obaidat, "Phonocardiogram signal analysis: techniques and performance comparison," *J Med Eng Technol*, vol. 17, pp. 221-227, 1993.

[39] C. Longhini, S. Aggio, E. Baracca, D. Mele, C. Fersini, and A. E. Aubert, "A mass-spring model hypothesis of the genesis of the physiological third heart sound," *Jpn Heart J*, vol. 30, pp. 265-273, 1989.

[40] E. Baracca, D. Scorzoni, M. C. Brunazzi, P. Sgobino, L. Longhini, D. Fratti, and C. Longhini, "Genesis and acoustic quality of the physiological fourth heart sound," *Acta Cardiol*, vol. 50, pp. 23-28, 1995.

[41] A. Haghighi-Mood and N. Torry, "Time-frequency analysis of systolic murmurs," in *Computers in Cardiology 1997*, 1997, pp. 113 - 116.

[42] L. J. Hadjileontiadis and S. M. Panas, "Discrimination of heart sounds using higher-order statistics," in *Proc. 19th Ann. Int. Conf. of the IEEE, EMBS*, 1997, pp. 1138-1141.

[43] S. Minfen and S. Lisha, "The analysis and classification of phonocardiogram based on higher-order spectra," in *Proc. of the IEEE-SP, Higher-Order Statistics.*, 1997, pp. 29-33.

[44] A. Swami, J. M. Mendel, and C. L. Nikias, "Higher-Order Spectral Analysis Toolbox Users Guide," 2001.

[45] Y. Xiang and S. K. Tso, "Detection and classification of flaws in concrete structure using bispectra and neural networks," *Ndt. & E. Int.*, vol. 35, pp. 19-27, 2002.

[46] G. Kubin, "Nonlinear processing of speech," in *Speech coding and synthesis*, W. B. Kleijn and K. K. Paliwal, Eds. Amsterdam: Elsevier, 1995, pp. 557-610.

[47] J. Whitmire and S. Sarkar, "Validation of acoustic-analogy predictions for sound radiated by turbulence," *Physics of Fluids*, vol. 12, pp. 381-391, 2000.

[48] P. Maragos and A. Potamianos, "Fractal dimensions of speech sounds: Computation and application to automatic speech recognition," *Journal of the Acoustical Society of America*, vol. 105, pp. 1925-1932, 1999.

[49] P. P. Vlachos, O. Pierrakos, A. Phillips, and D. P. Telionis, "Vorticity and turbulence characteristics inside a transparent flexible left ventricle," in *Bioengineering Conference, ASME*, 2001, pp. 493-494.

[50] D. Bluestein, C. Gutierrez, M. Londono, and R. T. Schoephoerster, "Vortex shedding in steady flow through a model of an arterial stenosis and its relevance to mural platelet deposition," *Ann Biomed Eng*, vol. 27, pp. 763-773, 1999.

[51] J. S. Liu, P. C. Lu, and S. H. Chu, "Turbulence characteristics downstream of bileaflet aortic valve prostheses," *J Biomech Eng*, vol. 122, pp. 118-124, 2000.

[52] A. Iwata, N. Ishii, N. Suzumura, and K. Ikegaya, "Algorithm for detecting the first and the second heart sounds by spectral tracking," *Med Biol Eng Comput*, vol. 18, pp. 19-26, 1980.

[53] H. Liang, L. Sakari, and H. Iiro, "A heart sound segmentation algorithm using wavelet decomposition and reconstruction," in *Proc. 19th Ann. Int. Conf. of the IEEE, EMBS*, 1997, pp. 1630 - 1633.

[54] H. Sava, P. Pibarot, and L. G. Durand, "Application of the matching pursuit method for structural decomposition and averaging of phonocardiographic signals," *Med Biol Eng Comput*, vol. 36, pp. 302-308, 1998.

[55] J. Gnitecki and Z. Moussavi, "Variance fractal dimension trajectory as a tool for hear sound localization in lung sounds recordings," in *Proc. 25th Ann. Int. Conf. IEEE EMBS*, 2003, pp. 2420-2423.

[56] V. Nigam and R. Priemer, "Accessing heart dynamics to estimate durations of heart sounds," *Physiological Measurement*, pp. 1005-1018, 2005.

[57] N. Joshi, "The third heart sound," *South Med J*, vol. 92, pp. 756-761, 1999.

[58] P. Hult, T. Fjallbrant, K. Hilden, U. Dahlstrom, B. Wranne, and P. Ask, "Detection of the third heart sound using a tailored wavelet approach: method verification," *Med Biol Eng Comput*, vol. 43, pp. 212-217, 2005.

[59] P. Hult, T. Fjallbrant, B. Wranne, and P. Ask, "Detection of the third heart sound using a tailored wavelet approach," *Med Biol Eng Comput*, vol. 42, pp. 253-258, 2004.

[60] S. Charleston, M. R. Azimi-Sadjadi, and R. Gonzalez-Camarena, "Interference cancellation in respiratory sounds via a multiresolution joint time-delay and signal-estimation scheme," *IEEE Trans Biomed Eng*, vol. 44, pp. 1006-1019, 1997.

[61] S. Charleston and M. R. Azimi-Sadjadi, "Reduced order Kalman filtering for the enhancement of respiratory sounds," *IEEE Trans Biomed Eng*, vol. 43, pp. 421-424, 1996.

[62] L. J. Hadjileontiadis and S. M. Panas, "Adaptive reduction of heart sounds from lung sounds using fourth-order statistics," *IEEE Trans Biomed Eng*, vol. 44, pp. 642, 1997.

[63] Z. K. Moussavi, D. Flores, and G. Thomas, "Heart sound cancellation based on multiscale products and linear prediction," in *Proc. 26th Annu. Int. Conf. IEEE Engineering in Medicine and Biology Society, EMBC'04*, San Francisco, USA, 2004, pp. 3840-3843.

[64] M. T. Pourazad, Z. K. Moussavi, and G. Thomas, "Heart sound cancellation from lung sound recordings using adaptive threshold and 2D interpolation in time-frequency domain," in *Proc. 25th Annu. Int. Conf. IEEE Engineering in Medicine and Biology Society, EMBC'03*, Cancun, Mexico, 2003, pp. 2586-2589.

[65] C. Ahlstrom, A. Johansson, P. Hult, and P. Ask, "Chaotic dynamics of respiratory sounds," *Chaos, Solitons & Fractals*, vol. In Press, Corrected Proof.

[66] J. Gnitecki and Z. Moussavi, "The fractality of lung sounds: A comparison of three waveform fractal dimension algorithms," *Chaos Solitons & Fractals*, vol. 26, pp. 1065-1072, 2005.

[67] J. Gnitecki, Z. Moussavi, and H. Pasterkamp, " Geometrical and Dynamical State Space Parameters of Lung Sounds," in *5th Int. Workshop on Biosignal Interpretation*, 2005, pp. 113-116.

[68] A. Vena, E. Conte, G. Perchiazzi, A. Federici, R. Giuliani, and J. P. Zbilut, "Detection of physiological singularities in respiratory dynamics analyzed by recurrence quantification analysis of tracheal sounds," *Chaos Solitons & Fractals*, vol. 22, pp. 869-881, 2004.

[69] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*, 2. ed. Amsterdam: Academic Press, 2003.
[70] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *J. Mach. Learn. Res.*, vol. 3, pp. 1157-1182, 2003.
[71] P. Pudil, J. Novovicova, and J. Kittler, "Floating search methods in feature-selection," *Patt. Recogn. Lett.*, vol. 15, pp. 1119-1125, 1994.
[72] T. Olmez and Z. Dokur, "Classification of heart sounds using an artificial neural network," *Pattern Recogn. Lett.*, vol. 24, pp. 617-629, 2003.
[73] T. S. Leung, P. R. White, W. B. Collis, E. Brown, and A. P. Salmon, "Analysing paediatric heart murmurs with discriminant analysis," in *Proc. 20th Ann. Int. Conf. of the IEEE, EMBS.*, 1998, pp. 1628 - 1631.
[74] L. G. Durand and P. Pibarot, "Digital signal processing of the phonocardiogram: review of the most recent advancements," *Crit. Rev. Biomed. Eng.*, vol. 23, pp. 163-219, 1995.
[75] R. M. Rangayyan and R. J. Lehner, "Phonocardiogram signal analysis: a review," *Crit Rev Biomed Eng*, vol. 15, pp. 211-236, 1987.
[76] I. Cathers, "Neural network assisted cardiac auscultation," *Artif. Intell. Med.*, vol. 7, pp. 53-66, 1995.
[77] S. A. Pavlopoulos, A. C. Stasis, and E. N. Loukis, "A decision tree--based method for the differential diagnosis of aortic stenosis from mitral regurgitation using heart sounds," *Biomed. Eng. Online*, vol. 3, pp. 21, 2004.
[78] Z. Sharif, M. S. Zainal, A. Z. Sha'ameri, and S. H. S. Salleh, "Analysis and classification of heart sounds and murmurs based on the instantaneous energy and frequency estimations," in *Proc. of TENCON*, Kuala Lumpur, 2000, pp. 130-134.
[79] C. G. DeGroff, S. Bhatikar, J. Hertzberg, R. Shandas, L. Valdes-Cruz, and R. L. Mahajan, "Artificial neural network-based method of screening heart murmurs in children," *Circulation*, vol. 103, pp. 2711-2716, 2001.
[80] H. Shino, H. Yoshida, K. Yana, K. Harada, J. Sudoh, and E. Harasewa, "Detection and classification of systolic murmur for phonocardiogram screening," in *Proc. 18th Ann. Int. Conf. of the IEEE, EMBS*, 1996, pp. 123-124.
[81] A. Voss, A. Mix, and T. Hubner, "Diagnosing aortic valve stenosis by parameter extraction of heart sound signals," *Ann. Biomed. Eng.*, vol. 33, pp. 1167-1174, 2005.
[82] H. Liang and I. Hartimo, "A heart sound feature extraction algorithm based on wavelet decomposition and reconstruction," in *Proc. 20th Ann. Int. Conf. of the IEEE, EMBS.*, 1998, pp. 1539 - 1542.
[83] I. Turkoglu, A. Arslan, and E. Ilkay, "An intelligent system for diagnosis of the heart valve diseases with wavelet packet neural networks," *Comput. Biol. Med.*, vol. 33, pp. 319-331, 2003.
[84] C. Ahlstrom, P. Hult, and P. Ask, "Wheeze Analysis and Detection with Non-linear Phase Space Embedding," in *13th Nordic Baltic Conference Biomedical Engineering and Medical Physics*, Umeå, Sweden, 2005, pp. 305-306.
[85] B. Suki, "Fluctuations and power laws in pulmonary physiology," *Am J Respir Crit Care Med*, vol. 166, pp. 133-137, 2002.
[86] R. Esteller, G. Vachtsevanos, J. Echauz, and B. Litt, "A comparison of waveform fractal dimension algorithms," *IEEE Transactions on Circuits and Systems I-Fundamental Theory and Applications*, vol. 48, pp. 177-183, 2001.

The phonocardiography factory.