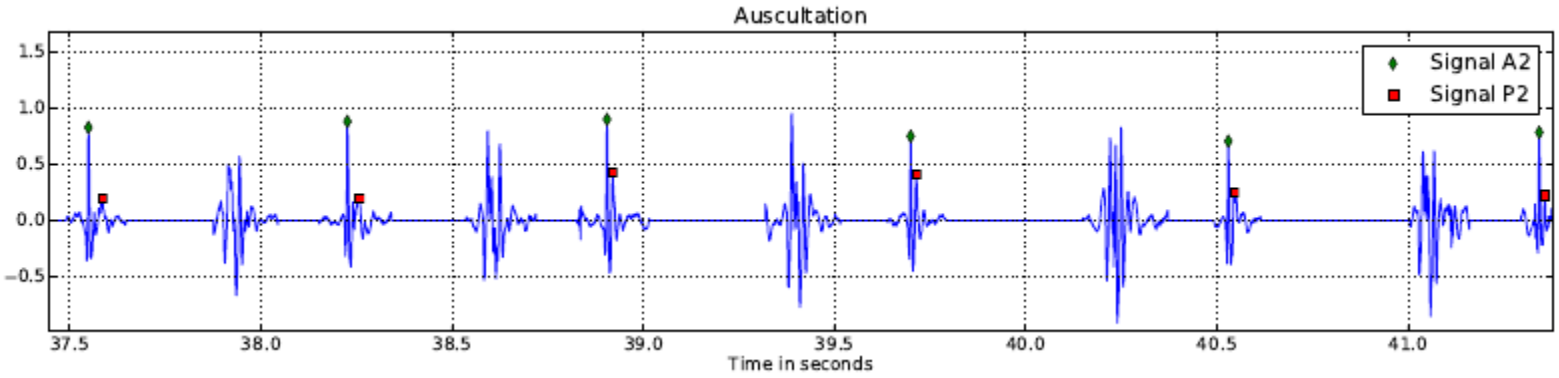


Looking at the second heart sound: A multi-facet study

Fabio de Lima Hedayioglu
Programa Doutoral em Informática
Departamento de Ciência de Computadores
2014

Orientador
Miguel Tavares Coimbra, Prof. Auxiliar, Faculdade de Ciências da Universidade do Porto

Coorientador
Sandra da Silva Mattos, UCMF - RHP

Contents

List of Figures

List of Tables

List of Algorithms

Acknowledgements

During the development of this thesis, a lot of people gave me invaluable input and did put effort into understanding the problem, suggesting solutions, and, also, being as puzzled as me with the riddles and facts this research posed. In fact, I am very grateful for sharing (and receiving help from) the team at the Centre for Digital Music (C4DM) group, at the Queen Mary University of London, who kindly helped to develop some ideas exposed in this thesis, specially Maria Jafari, and Mark Plumbley, who hosted me during my time at his research group. Back in Portugal, I am grateful to my supervisor, Miguel Coimbra for providing me with academic guidance and opportunities to meet amazing people and have such a diverse academic experiences, not only by going to international conferences but also by allowing me to take part in world-leading research groups, therefore, acquiring an amazing amount of technical and life experience.

I did get great input from Ana Castro who frequently told me to persevere and not to give up on my ideas. Even more input and sleepless nights were spent with Tiago Vinhoza, who not only did helped my with the technical and theoretical problems (through many, many discussions and debates), but also have shown how to do robust research.

I also would like to thank all my other colleagues from Porto and the staff of the Departamento de Ciência de Computadores for their great work environment they provided.

This acklowledgement wouldn't be complete without mentioning Bruno Lopes, who has been a great friend, who supported me whenever I needed, no matter if they where material (accomodation in Porto during the late stages of my Ph.D.), technical (by teaching me good programming practices) or in

a personal level (by giving me strengh to persevere and never give up).

My acknowledgements to my parents, who gave me all the necessary conditions to allow my intellectual, emotional and moral development.

I would like also, to express my gratitude to God for giving me the opportunity of having meeting people that helped me so much in my path of becoming a better human being (an endless task, I presume). Yet, I would like to thank Him for giving me wisdom to learn from the adversities and from the bad examples (and to give a chance to the ones who make those bad examples an opportunity to somehow help them).

I am also grateful to my understanding (and very patient) wife Julie and my stepdaughter Lauryn, who always believed in me and helped me persevere and survive the struggle of the PhD.

I would also like to acknowledge Instituto de Telecomunicações for being the host institution for this research, as well as the Real Hospital Português de Beneficiência em Pernambuco Brazil, specifically, the Unidade de Cardiologia e Medicina Fetal (UCMF) team for the effort in performing the data collection and the medical support necessary for this thesis. I could not fail to not acknowledge Dr. Sandra Mattos, who, since my undergrad believed in my idea of an electronic stethoscope and not only has been an academic corner stone for me, but also has been such a great friend, providing me with inspiration and wonderful insight not only in the academic level, but also throughout my life.

Last, but not least, I would like express my gratitude to my supervisor, the FCT - Fundação para a Ciência e a Tecnologia, who supported my word under the scholarship number SFRH / BD / 61655 / 2009, and the DigiScope project, under grant number PTDC/EIA-CCO/100844/2008.

Resumo

Nos ouvidos de um médico experiente, o estetoscópio produz informação clínica importante que pode ajudar a uma avaliação inicial da condição clínica do paciente e orientar a posterior necessidade de exames mais especializados. Isto é particularmente verdadeiro em Cardiologia e Pneumologia, sendo a razão pela qual o estetoscópio ainda mantém uma posição-chave em Medicina na era moderna.

O uso de um estetoscópio digital é adequado para a formação de médicos para melhorar suas habilidades básicas em diagnosticar e tratar doenças do coração, bem como uma ferramenta mais forte para a seleção mundial de patologias cardíacas específicas. Estes são alguns exemplos de como o estado da arte da tecnologia pode ser usado horizontalmente para beneficiar as pessoas em diferentes níveis econômicos, políticos ou geográficos.

Nesta tese estamos interessados na análise automática dos sinais acústicos adquiridos durante ausculta cardíaca, mais especificamente, a segunda bulha cardíaca. Analizaremos a segunda bulha a partir de três perspectivas diferentes: primeiramente a partir de uma perspectiva puramente de processamento de sinal, com poucas suposições sobre estes sinais; em seguida, criamos um modelo matemático dos componentes fisiológicos subjacentes do sinal e, por último, criamos um algoritmo que imita a maneira que os clínicos extraem algumas características fisiológicas do sinal.

Temos como objectivo a identificação de determinadas características importantes do som cardíaco, independentemente de estas corresponderem a uma ausculta saudável ou indiquem a existência de uma cardiopatia.

Dentre as contribuições desta tese, foi desenvolvida uma nova técnica para a utilização de um sensor em movimento para coletar sinais, simulando

a acquisição por meio de uma matriz de sensores. Também desenvolvemos um modelo matemático inspirado fisiologicamente para os componentes subjacentes da segunda bulha. O modelo matemático de segunda bulha foi utilizado como base para a criação de um novo simulador de som cardíaco que gera ausculta sintéticos com parâmetros clínicos e personalizável.

Também desenvolvemos um algoritmo que realiza a análise e identificação dos componentes fisiológicos da segunda bulha imitando o cardiologista quando interpretando um fonocardiograma.

Abstract

In the ears of an experienced physician, a stethoscope yields important clinical information which can help an initial assessment of a patient's clinical condition and guide the subsequent need for more specialized exams. This is particularly true in chest Medicine, i.e. Cardiology and Pneumology, which is the reason why the stethoscope still maintains a key position in Medicine in the modern era.

The use of a digitally enhanced stethoscope, adequate for training physicians to improve their basic skills in diagnosing and treating heart conditions, or as a stronger tool for world-wide screening of specific heart pathologies are some examples of how state of the art technology can be used horizontally to benefit people at different economical, political or geographical levels.

In this thesis, we are interested in the automatic analysis of the acoustic signal acquired during cardiac auscultation, more specifically, the second heart sound. We look at the second heart sound from three different perspectives: from a purely signal processing perspective, with very little assumptions about the underlying signals; then, we create a mathematical model of the underlying physiological components of the signal; and lastly, we mimic the way clinicians extract some physiological features of the signal.

We aim at identifying certain important features of the heart sound, regardless of whether these correspond to a healthy heart function or indicate the existence of a pathology.

Among the contributions of this thesis, we developed a novel technique for using a moving sensor to collect signals and simulate signal acquisition from an array or sensors. We also developed a physiologically inspired mathematical model for the underlying components of the second heart sound. The

mathematical model of the second heart sound was used as a base for the creation of a novel heart sound simulator that generates synthetic auscultations with customisable clinically meaningful parameters.

The algorithm that performs the analysis and identification of the physiological components of the second heart sound by mimicking the clinician when using a phonocardiogram was also developed.

Chapter 1

Introduction

1.1 Auscultation Through History

Since the early times, mankind knows that the body produces its own sounds and the importance of their interpretation have been recognised early by man. Between 460 B.C. and 370 B.C., Hippocrates mentions cardiac sounds [1], but it was only on the 17th Century that the clinician René Théophile Hyacinthe Laennec created a device that, later, would become the most remarkable symbols of the practitioner: the stethoscope.

At this time, in order to hear cardiac sounds, the clinician had to make use of a technique known as "direct auscultation", which consists in placing the ear directly on the patient's chest. This method is as uncomfortable for the clinician as it is for the patient, and a great inconvenience on women's examination. In hospitals, this method is impractical due to the great corporal contact, causing the risk of infection to increase significantly.

In 1816, Laennec (1781-1826) was requested to perform the examination of a woman with a high degree of obesity with general symptoms of cardiopathy. Direct auscultation was considered inappropriate by the woman's husband. Laennec, taking some sheets of paper, wrapping them up and putting one end of it on the patient's chest, and the other end on his ear, could hear the heart sounds clearer than using direct auscultation [2].

In 1819, Laennec published the result of his studies on the book: "Traité

de L'Auscultation Médiate, et des maladies des poumons et du coeur" (Treatise on Mediate Auscultation and Diseases of the Chest and Heart) [3]. In this work, he introduced new terms to describe more accurately the heart sounds and murmurs as well as the triple rhythm. Inadequate physiologic knowledge at that time led to a faulty interpretation of heart sounds. This was corrected, however, by the middle of the 19th century.

The first recording of heart sounds was made by Hurthle (1895), who connected a microphone to an inductorium, the secondary coil of which stimulated a frog nerve-muscle preparation. At about the same time, Willem Einthoven [4] recorded phonocardiograms (the graphic representation of the sounds which originate in the heart and great vessels), first by means of a capillary electrometer and then with a string galvanometer. The first monograph on phonocardiography was published by O. Weiss in 1909 [5].

Nowadays, stethoscopes are still widely used as the first cardiologic evaluation of patients. It is a non-invasive tool that detects a wide range of functional, hemodynamic and structural anomalies. This is an indicator of the effectiveness of the diagnostic capabilities of the stethoscope as a screening tool: the examination is fast, and can screen a broad range of cardiopathies.

1.2 Motivation for processing heart sounds

In this thesis, we focus on the processing of heart sounds acquired by auscultation, which has the advantage of being one of the simplest, quickest and most cost effective techniques to identify and diagnose a large number of heart conditions [6].

Cardiac auscultation is a difficult skill to master, which requires extensive training and years of experience [7]. This is because heart sounds are difficult to identify and analyse by the human listener as they are faint, significant events are closely spaced in time (less than 0.03 seconds), and their frequency content is at the lower end of the audible frequency range [8]. Auscultation is a skill that is also subjective to the hearing of the individual [9]. In fact, primary care physicians often never master auscultation [10] and studies have shown that they fail to correctly identify normal beats or benign murmurs

in as much as 80% of patients that they refer to cardiologists [10]. These unnecessary referrals represent a high financial cost for the medical system, including the visit to the cardiologist and any further tests that have to be performed [8]. Furthermore, a physician with poor auscultatory skills is also likely to fail to detect pathologies. Although there is no way to quantify how many problems go undiagnosed, it is obvious that this can have disastrous consequences for the individuals whose diseases remain undetected [11].

The difficulty in acquiring and maintaining auscultatory competence is compounded by the general decline in auscultation training in recent years, partly due to the introduction of new diagnostic technologies, and partly because of the scarcity of skilled instructors [6, 10]. Nonetheless, the need for a quick and inexpensive diagnosis method for cardiac disease remains, and this problem has attracted the attention of researchers who have been considering automating all or part of the auscultation procedure. This is an approach that is seen of particular importance for developing countries, where lack of resources may result in limited access to cardiologists, or in nurses taking on the role of primary carers in rural areas [12].

In this work, we are interested in the automatic analysis of the acoustic signal acquired during cardiac auscultation, more specifically, the second heart sound. We aim at identifying certain important features of the heart sound, regardless of whether these correspond to a healthy heart function or indicate the existence of a pathology.

1.3 The importance of the second heart sound

The second heart sound is closely related to the systemic circulation, pulmonary circulation and is greatly influenced by the respiratory cycle. Therefore, it is very important in the diagnosis of cardiac diseases.

The study of the main features of the second heart sound (amplitude of A2, amplitude of P2 and behaviour of the split), allows the clinician to detect a variety of cardiac conditions related to the pulmonary artery, aortic valve, systemic pressure, pulmonary artery pressure, right and left ventricle, and even some congenital heart defects, such as the Fallot Tetralogy [13, 14, 15,

16, 17, 18].

1.4 Objectives

In this thesis we want to look at the second heart sound from three different physiologically justifiable perspectives: from a purely signal processing perspectives, we assume very little about the signal components, and look at it as source separation problem; then we assume a mathematical structure to the underlying physiological components of the signal, creating a sparse representation; and then, buy trying to mimic the clinician, we extract some features of the physiological components of the signal.

1.5 Contributions

Throughout this thesis, some contributions were made. The main ones are listed below:

1. A novel technique using a moving sensor to collect data and simulate signal acquisition from an array of sensors was developed.

2. We developed a physiologically inspired mathematical model for the underlying components of the second heart sound.

3. A new algorithm that mimics the analysis and identification of A2 and P2 performed by the clinician when using a phonocardiogram was created.

4. A novel heart sound simulator that generates synthetic auscultations based on real ones, and with customisable clinically meaningful parameters was developed.

1.6 Publications

1. F. L. Hedayioglu; M. G. Jafari; S. S. Mattos; M. D. Plumbley; M. T. Coimbra, "Separating sources from sequentially acquired mixtures

of heart signals," Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on , vol., no., pp.653,656, 22-27 May 2011 doi: 10.1109/ICASSP.2011.5946488

2. M. G. Jafari; F. L. Hedayioglu; M. T. Coimbra; M. D. Plumbley, "Blind source separation of periodic sources from sequentially recorded instantaneous mixtures," Image and Signal Processing and Analysis (ISPA), 2011 7th International Symposium on , vol., no., pp.540-545, 4-6 Sept. 2011

3. F. L. Hedayioglu; M. G. Jafari; S. S. Mattos; M. D. Plumbley; M. T. Coimbra, "Denoising and segmentation of the second heart sound using matching pursuit," Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE (pp. 3440-3443). IEEE; (2012, August)

4. F. L. Hedayioglu; M.G. Jafari; S. S. Mattos; M. D. Plumbley; M. T. Coimbra, "An Exploratory Study On The Segmentation Of The Second Heart Sound", to be submitted

1.7 Thesis structure

Chapter 2 Introduces a brief summary on the physiology of the heart.

Chapter 3 Presents a study on the problem of performing source separation on heart signals.

Chapter 4 We describe an approach where we developed a physiologically inspired mathematical model for the underlying components of the second heart sound.

Chapter 5 In this chapter we postulate that the second heart sound components are peak-like, and the measures provided by these peaks also relates to the underlying components of S2.

Chapter 6 we presents a heart sound simulator program that generates synthetic auscultations based on real ones, and with customisable clinically meaningful parameters.

Chapter 7 Presents the main conclusions from this work, discuss some issues and presents possible future works on this field.

Chapter 2

Heart sounds

2.1 Physics of Sound

Sound may be described as the motion of waves of alternating pressure generated by a vibrating object [19]. This vibrating source sets particles in motion and in case of a sound with only one tone, the individual particles move around their resting point, with the same frequency of that tone.

In each movement, vibrating particles push others nearby, putting them in motion, therefore creating a chain effect, generating areas of high and low pressure. This alternation between low and high pressure moves away from the sound source and so does the sound wave. Usually those waves can be detected by their mechanical effect on a membrane (it could be a microphone's membrane or a stethoscope's diaphragm). A common way of describing a sound is by its intensity, frequency and duration [20].

Different materials conduct sound, and the nature of this conducting material defines the rate of propagation, varying directly with the elasticity of the conducting material (medium) and inversely with its density. On the human body, the transmission of sound waves is a very complex matter involving three different modes [21, 22, 23]:

- compression wave: the same speed of sound in water (with a speed about 1.5 Km/s)
- transverse shear wave: slower than the compression wave, propagates

most of the vibratory energy of the body (at about 20 m/s)

- surface wave: a mixture between transverse and compression waves, on the surface of the body (at about 20 m/s)

When the waves travel from one medium to the other, considerable distortion may be produced: waves may be reflected, refracted, among others, resulting in a loss of energy. Probably for these reasons, when the myocardial tissue is in a close position to the chest wall, heart sounds are well conducted to the body's surface, since they form a relatively homogeneous medium. However, when the air-filled lung is between the heart and the chest wall, waves must travel more mediums, therefore these waves are weakened considerably and the sound reaching the external surface becomes attenuated.

2.2 Physiology of the heart

The heart is an organ mostly constituted by striated cardiac muscle with two main functions: to collect oxygen-rich blood from the lungs and deliver it to all tissues of the body, and to collect blood rich with carbon dioxide from the tissues of the body and pump it to the lungs [24]. Its located inside the thorax, in the middle mediastinum, inside the pericardium membrane (the pericardial cavity), which involves the heart (Figure 2.1). This membrane has inner and outer layers, with a lubricating fluid in between (the pericardial fluid). The fluid allows the inner visceral pericardium to "glide" against the outer parietal pericardium [25].

The normal heart is composed by four chambers: the two upper chambers have the main function of collecting blood, injecting it into the ventricles, which are much stronger and work as a blood pump. The function of the right atrium and ventricle (right heart) is to collect blood rich in carbon dioxide from the body and pump it to the lungs. There is a one-way flow of blood through the heart, maintained by a set of four valves: the atrioventricular valves (tricuspid and bicuspid) allow blood to flow only from the atria to the ventricle; the semilunar valves (pulmonary and semilunar) allow blood to flow only from the ventricles out of the heart and through the great arteries,

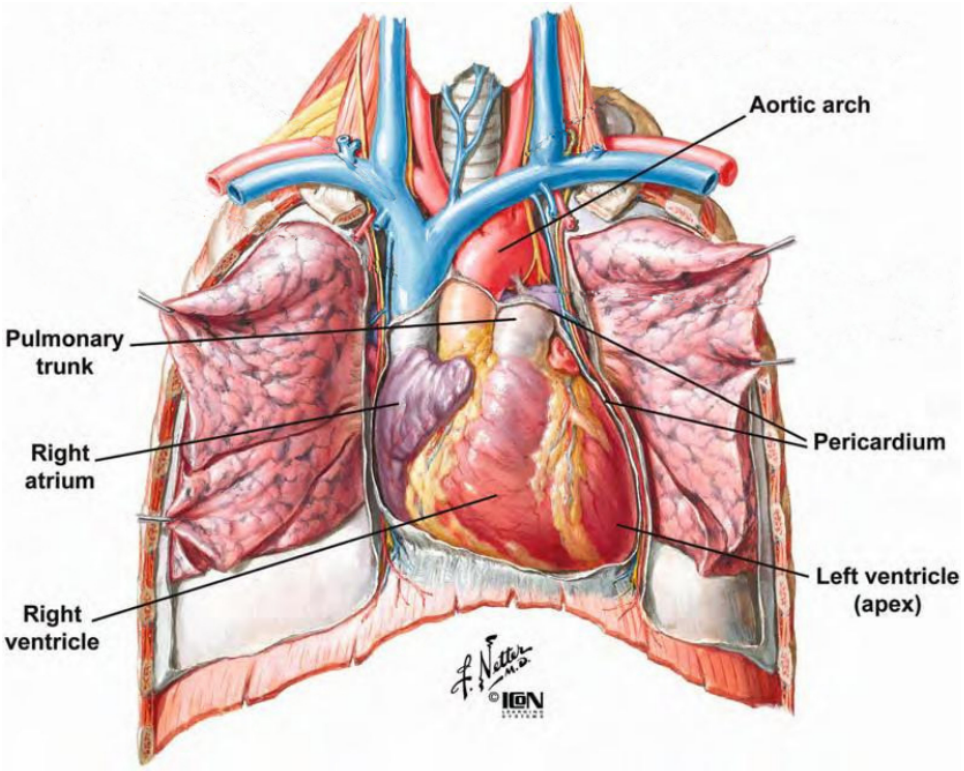


Figure 2.1: The position of the heart and the great vessels in the middle mediastinum [1]

as depicted in the Figure 2.2. In general, the anatomy of the right side of the heart (also known as right heart) is considerably different from that of the left heart, especially due to the effort the left heart has to make to pump blood to all tissues in the body; nonetheless the pumping principles of each are basically the same.

The cardiac valves passively open and close in response to the direction of the pressure gradient across them. The muscular cells (myocytes) of the ventricles are organised in a circumferential orientation; therefore, when they contract, the tension within the ventricular walls increases the pressure within the chamber. As the ventricular pressure exceeds the pressure in the pulmonary artery (on the right heart) and/or aorta artery (on the left heart), blood is forced out of the ventricular chamber. This contractive phase of the

[1] picture adapted from the book [25] page 37.

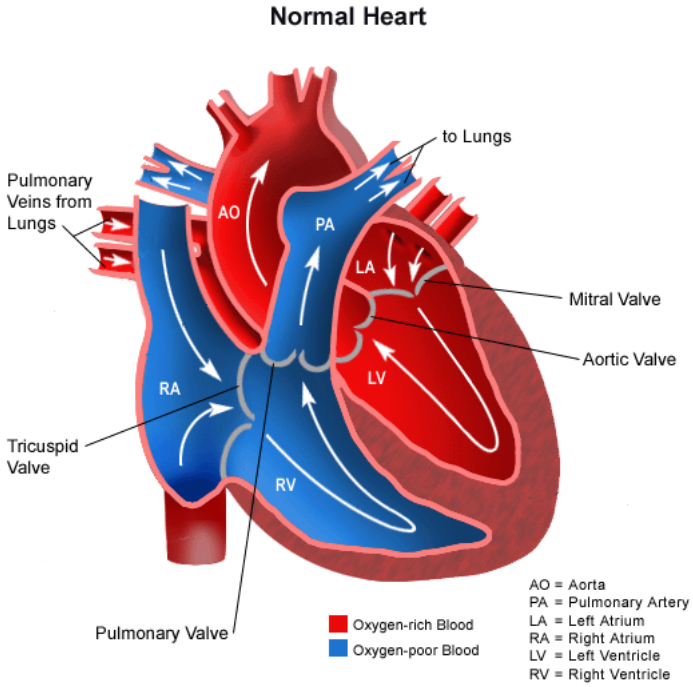


Figure 2.2: A Normal Heart [2]

cardiac cycle is known as systole. The pressures are higher in the ventricles than the atria during this systolic phase; hence, the atrioventricular (tricuspid and mitral) valves closes. When the ventricular muscles relax, the pressure inside them falls below those in the atria, the atrioventricular valves open and the ventricles are refilled; this is the diastolic phase. The aortic and pulmonary valves are closed during diastole because the arterial pressures (in the aorta and pulmonary artery) are greater than the intraventricular pressures [26]. This is depicted briefly on Figure 2.3

2.3 Heart Sounds

Heart sounds are caused by the dynamic events associated with the heartbeat and the blood flow. They are relatively brief and have different intensity

[2]picture adapted from the website: `http://healthinessbox.com/2012/01/27/coronary-heart-disease-chd/` accessed on 23/04/2013

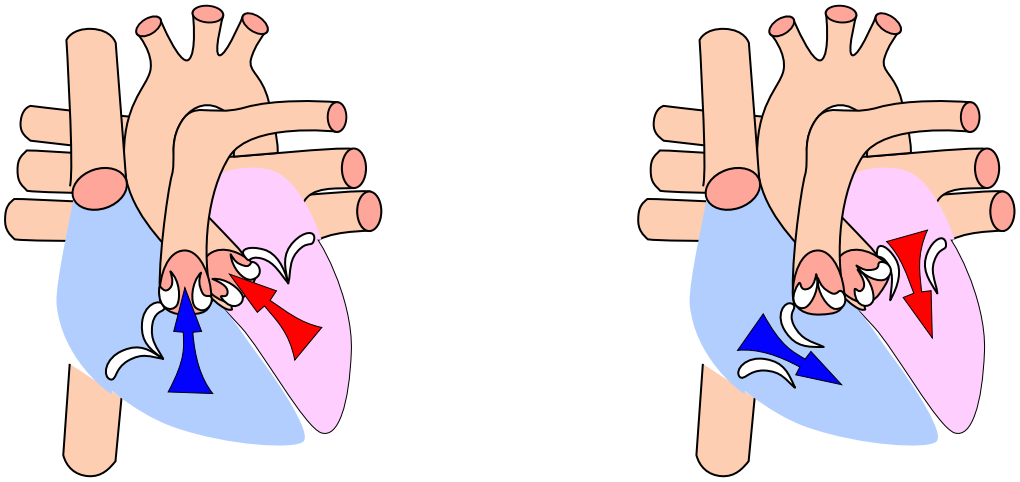


Figure 2.3: The Cardiac Cycle: the systole on the left and diastole on the right[3].

(loudness), frequency (pitch), clarity and duration. To better understand heart sounds, we need to know the physiology of the cardiac cycle.

2.3.1 The first heart sound

The first heart sound (S1) consists of several components, though only two of them are usually audible: the ones related to the closure of the mitral (M1) and tricuspid (T1) valves. These sounds are likely to be produced by the abrupt acceleration or deceleration of a mass of blood within the ventricles and by the sudden tensing of the entire atrioventricular (AV) valve structure that stretches the surrounding structures to their elastic limits. The greater these forces are, the louder and higher their frequencies are [27].

In the beginning of the diastole, the mitral and tricuspid valves open, making their leaflets widely separated. This allows the ventricular filling, where the leaflets of each of these valves begin their closure, until they become partially closed and the atrial systole begins, when they reopen, to be completely closed at the end of the atrial systole, when the ventricle recoil

[3]Picture acessed from the website: `http://commons.wikimedia.org/wiki/File:Heart_diasystole.svg` and `http://commons.wikimedia.org/wiki/File:Heart_systole.svg` on 23/04/2013

and the leaflets finally close. After the closure, the atrioventricular valves are stretched toward the atrium by the impulse of the ventricular blood. The closure of the AV valves, makes the blood volume to abruptly decelerate. The remaining vibrations of this heart-blood system generates sounds in the audible range (referred as S1) and is composed by the mitral closure sound (M1) and the tricuspid closure sound (T1) (Figure 2.4). In a normal auscultation, these sounds may be separated by 0.02 to 0.03 seconds. This splitting of S1 is referred to as normal or physiological splitting of the first heart sound. In the inspiration, the tricuspid component may become louder. The fact that S1 is split in normal subjects helps the detection of certain diseases: for instance, patients with Ebstein's abnormality may produce a loud tricuspid (T1) component; in mitral stenosis, T1 is very loud and may be heard throughout the precordium [27, 19, 13].

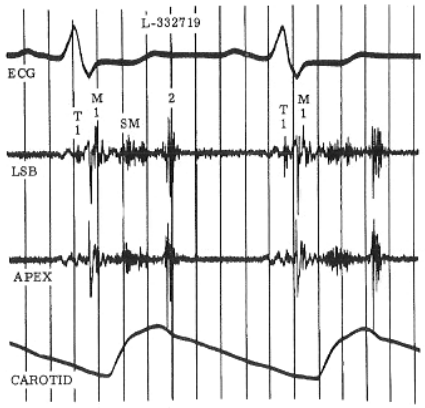


Figure 2.4: The M1 and T1 components of the first heart sound[4]

2.3.2 The second heart sound

At the end of the systole, the aortic and pulmonary valves cusps closes, and the elastic limits of the these tensed valve leaflets are met, then the blood flow suddenly decelerates and rebounds, setting these valves, the heart cavities and blood column into an oscillatory motion, that produces the second heart sound [27].

[4]picture adapted from Picture from [19]

The second heart sound has two components: the aortic component (A2) and the pulmonary component (P2). They are coincident with the incisurae of the aorta and pulmonary pressure curves, respectively, and mark the end of the left and right ventricular ejection periods (Figure 2.5). Right ventricular ejection begins before left ventricular ejection, has a slightly longer duration, and finishes after left ventricular ejection, as a result, P2 normally occurs after A2.

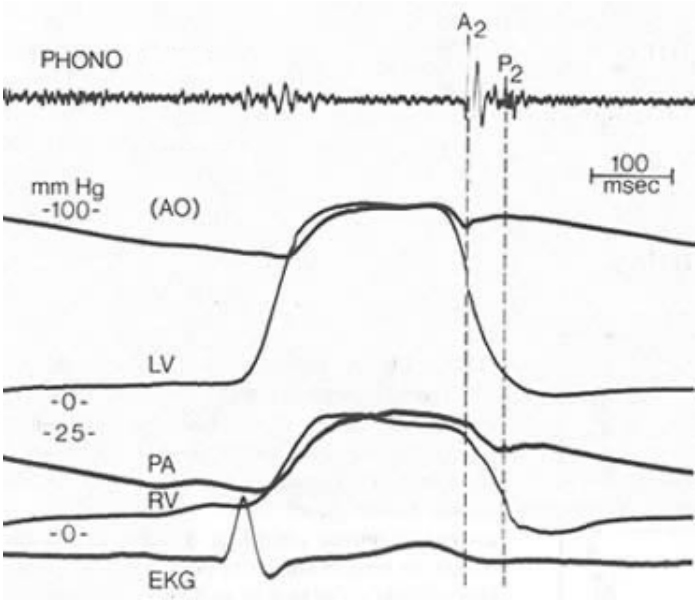


Figure 2.5: The pressure curves of the aorta (AO) and pulmonary artery (PA) [5]

To understand the effects of respiration on the splitting of S2, it is essential to understand the differences between the aortic and pulmonary artery vascular resistance: when we analyse the simultaneous pressure recordings of right ventricle and the pulmonary artery, the pulmonary artery pressure curves accompanies the left ventricle pressure curve until it descends, them the two pressure curves start to differ, with the left ventricle pressure curve decreasing before the pulmonary artery curve, therefore 'hanging out' for some milliseconds. The duration of this time is a measure of resistance in the pulmonary artery system [13, 26, 19]. In the left side of the heart, be-

[5] picture adapted from [28]

cause impedance is much greater, the hangout interval between the aorta and left ventricular pressure curves is negligible (less than or equal to 5 milliseconds). The hangout interval therefore correlates closely with impedance of the vascular bed into which blood is being ejected. Its duration appears to be inversely related to vascular impedance.

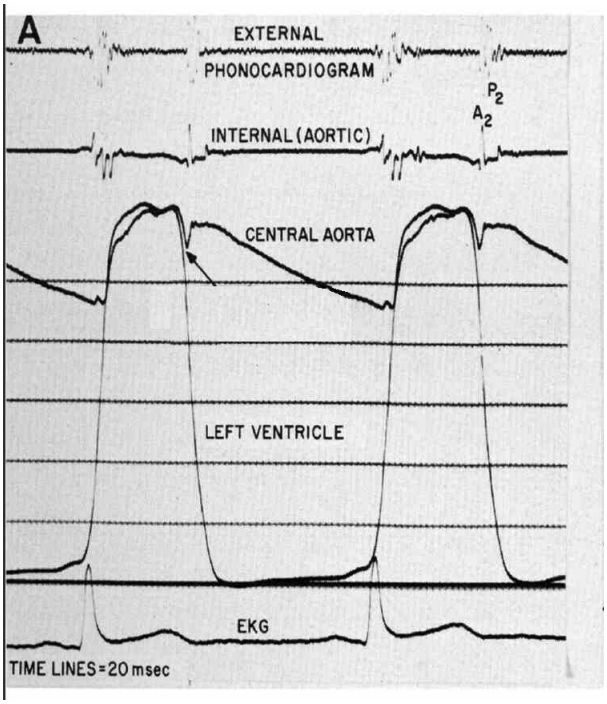


Figure 2.6: The pressure curves in normal subject [6]

Alterations in the resistance characteristics of the pulmonary vascular bed and the right-sided hangout interval are responsible for many of the observed abnormalities of S2. In a normal person, inspiration lowers the resistance of the pulmonary circuit, increases the hangout interval and delays pulmonary valve closure, therefore, producing an audible split of A2 and P2. At expiration, the reverse occurs: the pulmonary valve closes earlier, and the A2-P2 interval is shortened, being separated by less than 0.03 seconds

[6] picture adapted from [29]

and may sound single to the ear. Since the pulmonary circulation (the circulation between the heart and the lungs) has a much lower resistance than the systemic circulation (the circulation between the heart and the rest of the body), the blood flow through the pulmonary valve last longer than the blood flow through the aortic valve. The inspiratory split increases mainly because of the delay in the pulmonary component [19, 13].

2.3.3 Clinical assessment of the second heart sound

The clinical evaluation of S2 includes an assessment of its splitting and the determination of the relative intensities of A2 and P2. Usually the aortic closure sound (A2) occurs before the pulmonary closure sound (P2), and the interval between the two (splitting) increases on inspiration and decreases on expiration, as described earlier. With quiet respiration, the split time between A2 and P2 is from 0.02 to 0.08 seconds (with a mean of 0.03 to 0.04 seconds) in inspiration [19]. In younger subjects, the inspiratory splitting averages from 0.04 to 0.05 second during quiet respiration. With expiration, A2 and P2 may be superimposed and their split by more than 0.04 second [19]. If the second sound is split by more than 0.04 second on expiration, it is considered abnormal [19].

The respiratory variation of the second heart sound can be categorised as follows [27]:

1. Normal (physiologic) splitting: In normal individuals, during inspiration the splitting interval widens primarily due to the delayed P2, and during expiration, the A2-P2 interval narrows to the point that only a single sound is usually heard [19, 13, 27];

2. Persistent (audible expiratory) splitting, with normal respiratory variation: It suggests an audible expiratory interval (split greater than 0.03 seconds). Persistent splitting that is audible during both respiratory phases, with appropriate increase of split during inspiration and decrease of split in expiration, may occur in the recumbent position in normal children, teenagers, and young adults. However, in normal

adults, the expiratory split is not audible in the sitting or standing position. In almost all patients with heart disease and audible expiratory splitting in the recumbent position, the expiratory splitting is still audible even in the sitting or standing position. Thus, the audible expiratory splitting in both the recumbent and upright positions is a very sensitive screening test for heart disease [27];

3. Persistent splitting without respiratory variation (fixed splitting): Is the absence of significant variation of the splitting interval with respiration. The classic example of fixed splitting of the second heart sound is the atrial septal defect, where the expiratory splitting is caused by changes in the pulmonary vascular bed, that increases the time of the right ventricular systole. The fixed nature of the split is due to the two ventricles sharing a common reservoir (Figure 2.8), making the inspiratory delay of the aortic and pulmonary components almost the same as the expiratory time [27].The Valsalva maneuver may be used to exaggerate the effect of respiration and obtain clearer separation of the two components of the second sound . Patients with atrial septal defects show continuous splitting during the strain phase, and upon release the interval between the components increases by less than 0.02 second. In normal subjects, however, splitting is exaggerated during the release phase of the Valsalva maneuver. Variation of the cardiac cycle length may also be used to evaluate splitting of S2. During the longer cardiac cycle, patients with atrial septal defect may show greater splitting as a result of increased atrial shunting and greater disparity between stroke volume of the two ventricles. In normal subjects, there is no tendency to widen the splitting with longer cardiac cycles.

4. Reversed (paradoxical) splitting: Is the result of a delay in the A2 component. In the paradoxical splitting, P2 is produced before A2 and therefore, splitting is maximal on expiration, instead of inspiration (as it would be in a normal case). In inspiration, however, splitting may be minimal or even absent. This reversed behaviour and specially the narrowing or disappearance of the splitting on inspiration is an import-

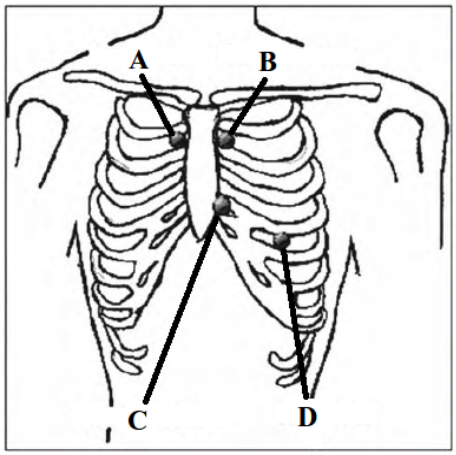


Figure 2.7: Auscultatory areas[7]: A) aortic, B) pulmonary, C) tricuspid and D) mitral

ant criteria for diagnosing reversed splitting by auscultation. Another way of identifying the paradoxical splitting of S2 is by identification of A2 and P2 by intensity and transmission: while A2 is audible in most areas of auscultation, P2 tend to be more clearly audible only at the pulmonary and aortic sites. This method, however, may fail in cases of pulmonary artery hypertension second to left ventricular failure, where P2 becomes as loud as A2. Paradoxical splitting indicates significant cardiovascular disease: it is usually caused by either prolongation of left ventricular activation or prolonged left ventricular emptying.

With the increase of age, the pulmonary vascular impedance increases and P2 may occur earlier, making normal patients over age 50 possibly exhibiting a single S2 or a narrow split on inspiration. However, normally a single S2 is due to a relatively soft pulmonary component that makes it difficult (or even unable) to auscultate. In healthy infants, children, and young adults the P2 component is not soft and P2 is rarely soft. In older persons under good auscultatory conditions, although soft, it can still usually be auscultated [27, 19, 13].

The inability of hearing P2 may suggest tetralogy of Fallot or pulmonary

[7]picture adapted from [30]

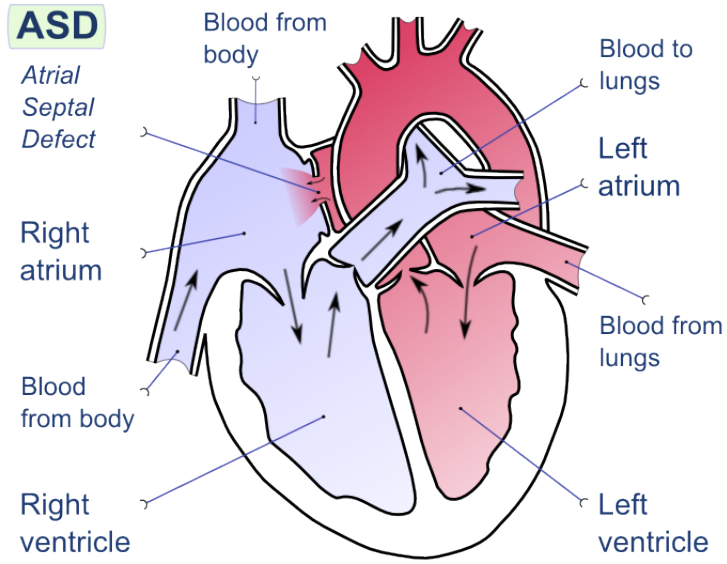


Figure 2.8: Heart with atrial septal defect.[8]

atresia. Other situations where P2 may be inaudible are either chronic right ventricular failure, or in cases where the aortic component may be masked by the systolic murmur, as it occurs in patients with aortic stenosis [27].

The amplitude (loudness) of each component of S2 is proportional to the respective pressures in the aorta and pulmonary artery at the moment of diastole. Another cause for an increase in the amplitude of A2 or P2 is the dilatation of the aorta or pulmonary artery, respectively. A2 usually presents greater amplitude than P2. The aortic component, therefore, radiates over the chest, whereas P2 is heard mainly in the second left intercostal space with some radiation down the left sternal border [27]. The higher pressure of the aorta is probably the reason why the aortic component has the greater radiation than P2. Since the pulmonary valve and the pulmonary arteries are closer to the chest wall, given the same level of pressure, the pulmonary component will be louder than the aortic component, therefore A2 must have a greater pressure level than P2 [27].

Reasons for a decreased intensity of A2 or P2 may be a stiff semilunar

[8]Picture acessed from the website: `http://en.wikipedia.org/wiki/File:Atrial_septal_defect-en.png` on 18/09/2013

valve, or low pressure beyond the semilunar valve, or deformity of the chest wall or lung. A low intensity of P2 is mostly common in patients with valvular pulmonic stenosis or chronic obstructive lung disease. Valvular aortic stenosis is another reason for a low intensity of the aortic component [27].

An overview of the changes in the volume and split of A2 and P2 can be seen in Table 2.1

2.4 Processing heart sounds

As we had seen in this chapter, the main constituents of a cardiac cycle are the first heart sound, the systolic period, the second heart sound (S2) and the diastolic period. Whenever a clinician is performing an auscultation, he tries to identify these individual components, and is trained to analyse related features such as rhythm, timing instants, intensity of heart sound components, and splitting of S2, among others [32, 7]. This analysis allows him to search for murmurs and sound abnormalities that might correspond to specific cardiac pathologies. From a signal processing perspective, processing of heart sounds is not only interesting by itself, but is also an essential first step for the subsequent task of automatic pathology classification. For sake of clarity, we will distinguish two sub-tasks of processing of heart sounds: heart sound segmentation and heart sound classification.

2.4.1 Heart sound segmentation

In heart sound segmentation we expect to identify and segment the four main constituents of a cardiac cycle. This is typically accomplished by identifying the position and duration of S1 and S2, using some sort of peak-picking methodology on a pre-processed signal.

Liang [33] has used discrete wavelet decomposition and reconstructed the signal using only the most relevant frequency bands. Peak-picking was performed by thresholding the normalised average Shannon energy, and discarding extra peaks via analysis of the mean and variance of peak intervals. Finally, they distinguish between S1 and S2 peaks (assuming that the dia-

stolic period is longer than the systolic one, and that the later has lower variation in duration), and estimate their durations. A classification accuracy of 93% was obtained on 515 periods of PCG signal recordings from 37 digital phonocardiographic recordings. The same authors further improved the statistical significance of their results by obtaining the same accuracy using 1165 cardiac periods from 77 recordings [32], and later attempted murmur classification based on these features and neural network classifiers, obtaining 74% accuracy [34]. Omran [35] has also studied this problem using normalised Shannon entropy after wavelet decomposition of the audio signal, but their experimental methodology is not so convincing.

Kumar [36] proposes a method for detection of the third heart sound (S3) that uses novel wavelet transform-simplicity filter, which separates S1, S2 and S3 from background noise and murmurs, then they used a technique developed earlier [37, 38], where S2 is assumed to have a high frequency signature, and S1 is detected by looking at the cardiac cycle. The third heart sound is detected by using temporal thresholds on the low frequency output of the simplicity filter. A sensitivity and specificity of 90.35% and 92.35% respectively was obtained. The same authors, later on also produced a paper on segmentation of cardiac murmur [39], where the simplicity filter is used in conjunction to adaptive thresholding in order to segment heart sounds in presence of murmurs.

Moukadem et al.[40] developed a robust heart sound segmentation algorithm by extracting the smoothed envelope of the Shannon energy of the local spectrum calculated by the S-transform [41] for each sample of the signal, for detection of presence of heart sound (S1 or S2), then a windowed version of the envelope is calculated, where the size of the window is changed in order to optimise the energy concentration and consequently, the boundaries of the heart sound. For S1 and S2 identification they use the techniques developed by Kumar et. al in [38, 37]. They also demonstrated that their approach is robust against additive Gaussian noise.

Besides the four main components of the cardiac cycle, there is a clinical interest in the analysis of some of its associated sub-components [42, 43]. It has been recognised that S1 may be composed of up to four components

produced during ventricular contraction [43], although the complexity of this task has been a very difficult hurdle for the signal processing community.

The S2 sound is more well known, being composed of an aortic component (A2), which is produced first during the closure and vibration of the aortic valve, the blood rebound and surrounding tissues, followed by the pulmonary component (P2) produced by a similar process associated with the pulmonary valve [42].

Xu [42] demonstrated a model where each component of S2 can be considered a narrow-band nonlinear chirp signal. Later [44] he adapted and validated this approach for the analysis and synthesis of overlapping A2 and P2 components of S2. To do so, the time-frequency representation of the signal is generated and then estimated and reconstructed using the highest instantaneous phase and amplitude of each component (A2 and P2). In this paper the accuracy evaluation was made by simulated A2 and P2 components having different overlapping factors. The reported error was between 1% and 6%, proportional to the duration of the overlapping interval.

Nigam [45] also presented a method for extracting A2 and P2 components by assuming them as statistically independent. To do so, four simultaneous auscultations are analysed using blind source separation. The main advantage of this method is the lower dependence on the A2-P2 time interval, although it needs a non-conventional 4-sensor stethoscope.

2.5 Heart sound classification

The vast majority of papers we have found regarding audio processing algorithms, concern the detection of specific heart pathologies. This highlights the interest of the scientific community on this topic but, there are still some major flaws in most of them such as the absence of a clinical validation step and unconvincing experimental methodologies.

Most papers use the well-established pattern recognition approach of feature extraction followed by a classifier. Bentley [46] uses Choi-Williams Distribution (CWD) as features, working with 45 normal/abnormal valve subjects. Some features were determined via visual inspection, others auto-

matically from the CWD by simple rule-based classification. Later [47], the authors show that CWD is a better method to represent the frequencies in PCG and to get heart sound descriptors, than other time- frequency (T-F) representations. According to them, a simple description of the T-F distribution allows an analysis of the heart valve's condition. However, they highlight the need of a more comprehensive evaluation using a larger population of test patients.

Wang [48] proposes a representation of heart sounds that is robust to noise levels of 20dB, using mel-scaled wavelet features. However, details regarding the used dataset are not clear enough for robust conclusions.

Liang [49] developed an interesting feature vector extraction algorithm where the systolic signal is decomposed by wavelets into subbands. Then, the best basis set is selected, and the average feature vector of each heart sound recording is calculated.

Neural Networks (NN) are used for classifying 20 samples after being trained with 65, obtaining an accuracy of 85%. Turkoglu [50], Ozgur [51] and El-Hanjouri [52] also used wavelets as feature vectors for classification, although they provide too few details regarding the used data sets. Trimmed mean spectrograms are used by Leung [53] to extract features of phonocardiograms. Together with the acoustic intensities in systole and diastole, the authors quantified the distinctive characteristics of different types of murmurs using NNs.

One of the few papers that is conscious about the important clinical validation step is from Kail [54]. The authors propose a novel sound representation (2D and 3D) and feature extraction algorithm using Morlet wavelet scalograms. After manual classification of the resulting graphs performed by two cardiologists on 773 subjects, they clinically validated the features as useful for sound and murmur extraction. Sharif [55] also proposes other features for classification systems based on central finite differences and zero crossing frequency estimation.

2.6 Challenges of processing the second heart sound

The S1 and S2 sounds can be robustly segmented and is a fundamental first step for most of the work on heart sound processing.

There is promising work regarding the extraction of secondary sounds such as A2 and P2, although this is still an open challenge, since there is still great limitations in several aspects of the research developed: there is still a great gap between the medical knowledge and the signal processing results in this area; and the use of some non-standard technological devices, such as a stethoscope with four sensors [45].

The automatic pathology classification scenario is not so evolved. Reviewing some of the papers and simply observing the disparity in the number of publications when compared with the other challenges, we conclude that there is a strong interesting in this topic.

The murmur detection and classification seems to be another area where robust results are being obtained, although there is still work to be done in creating techniques in order to produce a morphology-based classification of murmurs.

In our opinion, there is still a long way to go before we can have robust automatic classification systems that can be introduced in the clinical routine of hospitals.

In this thesis we worked mainly on the segmentation of the second heart sound into its two main components: A2 and P2. The second heart sound is quite important in the detection of several cardiac conditions, as seen in session 2.3.2.

The A2 is directly related to the aorta valve, the left heart and the systemic circulation, whilst P2, on the other hand is related to the pulmonary valve, the right heart and the pulmonary circulation. The behaviour of the split between A2 and P2 and how respiration change it during an auscultation is another important indicator of several cardiac conditions, linked to the hemodynamic of the heart, its ventricles and atria. In this way, looking at S2, A2 and P2 can provide a great overview of the heart and its hemody-

namic structure. The auscultation and interpretation of these components even by a clinician, however is of great difficulty, since they are quite short and close to each other in time.

Looking into the physiological background given in this chapter, we can see that A2 is produced by the blood rebound caused by the closure of the aortic valve, and that P2 is similarly produced by the closure of the pulmonary valve and its blood rebound. Therefore these components, are produced by independent mechanisms. We explore the possibility of separating these sound components by looking at S2 segmentation using a independent component analysis approach.

We also attacked the problem of extracting the subcomponents of S2 by postulating how these subcomponents could be estimated: Still keeping in mind the production of A2 and P2, we can also imagine that the initial nature of the vibratory motion set by their production would generate a high initial amplitude and frequency wave that fades with time (as suggested by Xu et al. [42]). This family of wave forms could be used as a model where A2 and P2 would be instances of: just like once one knows the letters of an alphabet, creating a word is just a matter of selecting the right words in the correct order and 'adding' them together; the problem of decomposing S2 into instances of waveforms becomes a problem of selecting the waveforms (letters) that best explains the word (S2).

Another challenge we tried to solve in this thesis is to lower the gap between the clinical knowledge of S2 and the knowledge produced by signal processing community: clinicians understand A2 and P2 as peaks in the phonocardiogram and most medical literature regarding phonocardiograms and A2 and P2 is based on the measurement of these peaks. This manual measurement, however has a precision problem and the quantification of amplitude of these peaks is still an open problem.

Table 2.1: Assesment of the second heart sound, adapted from [31]

Character	Clinical condition
Fixed A2-P2 interval	• Atrial septal defect (important clue)
Wide splitting, inspiratory increase in A2-P2	• Right-ventricular conduction delay • Idiopathic dilatation of pulmonary artery • Small atrial septal defect (unusual) • Pulmonic stenosis
Soft A2	• Aortic sclerosis or stenosis • Hypotension
Soft P2	• Pulmonic stenosis • Systemic hypertension
Loud A2	• Dilated Aorta • Systemic hypertension
Loud P2	• Pulmonary hypertension (important clue) • Dilated pulmonary artery

Chapter 3

Independent Component Analysis of S2

3.1 Introduction to source separation

In this chapter, we overview the independent component analysis technique for separation of simultaneously acquired signals, we propose a novel technique to perform source separation of a set of independent components assuming that we have access to one moving sensor and that the underlying source signals that we wish to separate are quasi-periodic, as it is the case with cardiac signals.

3.2 Blind source separation and independent component analysis

A classic problem in Blind Source Separation (BSS) is the cocktail party problem. The objective is to, given a mixture of sounds, with a given number of microphones (observations), separate each sound into a separate channel: the sources (Figure 3.1).

When two or more different signals are recorded by a single microphone its output is a signal mixture which is a simple weighted sum of the two signals. The relative proportion of each signal in the mixture (signal captured by

the microphone) depends on the loudness of each speech sound at its source, and the distance of each source from the microphone. The different distance of each source from the microphone ensures that each source contributes a different amount to the microphone's output, thus the mixture amplitude is the weighted sum of the source signals, namely:

$$x_1(t) = as_1(t) + bs_2(t) \tag{3.1}$$

$$x_2(t) = cs_2(t) + ds_2(t) \tag{3.2}$$

Where $x_1(t)$ and $x_2(t)$ are the mixtures, a, b, c, d are the relative proportion of each signal the microphones capture, and $s_1(t)$ and $s_2(t)$ are the source signals. If we know the values of $a; b; c; d$, then this problem could be solved by simply solving the above equations. However, these parameters are not always known a priori, which turns a simple problem into a complex one.

Among other complicating factors is the possibility of the sound of different sources reaching the microphones at different times; or the sounds being recorded in a reverberating environment, etc. These and other factors turns the BSS into quite a challenging problem.

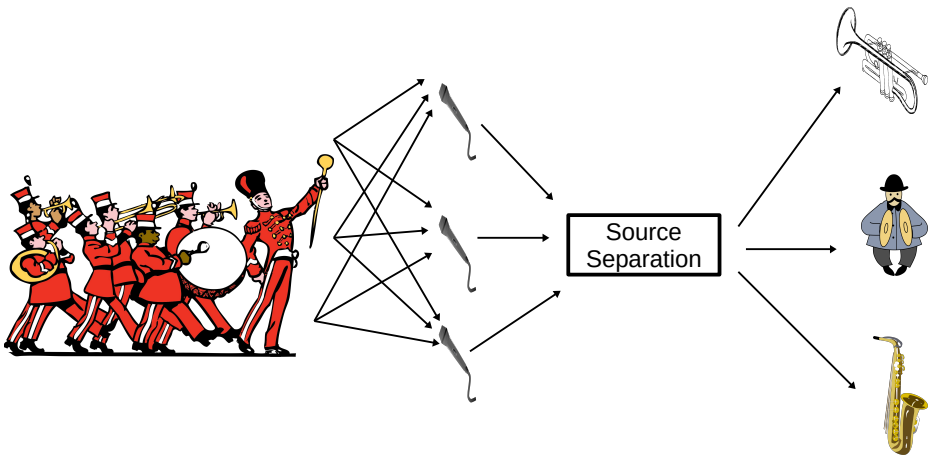


Figure 3.1: The classical problem of source separation

3.2.1 Historical background

The first papers on source separation date back to early 80's, with the works of Ans [56], Hérault [57] and Jutten [58] and in the context of motion coding by proprioceptive fibres.

Proprioception is one of the major modalities for somatic sensitivity [59, 60]. It is related to sensing both static position and movements of the limbs and body. Björklunds define it as "the perception of positions and movements of the body segments in relation to each other without the aid of vision, touch, or the organs of equilibrium [61].

The propioceptive sense comes from a combination of two afferent channels and several sensory structures: muscle spindle fibres, Golgi tendon organs, joint angle sensors, and cutaneous mechanoreceptors [62, 63]. When a joint is put into motion through muscle contraction, the muscle spindles detect changes in the length of these muscles, it encode this information by the rate of neuron firing, and sends this information to the central nervous system by two types of sensory afferent endings called primary and secondary endings. At the same time, the Golgi tendon organs detects the amount of stretch applied by the muscles, it also encode this information by the rate of neuron firing, sending the signal also using the same primary and secondary endings.

The primary and secondary endings transmit frequency-coded messages that are mixtures of information from both sensors to the central nervous system.

The simplified model of this transmission system is exemplified below:

$$\begin{aligned} f_1(t) &= a_{11}v(t) + a_{12}p(t) \\ f_2(t) &= a_{21}v(t) + a_{22}p(t) \end{aligned} \tag{3.3}$$

Where $v(t)$ is the muscle contraction, $p(t)$ is the muscle's stretch (position), and $a_{ij} = \{1, 2\}$ are all unknown. It is an undetermined system of equations, however, 'if spindle discharges are useful for kinesthetic sensations, the central nervous system must be able to distinguish which part is caused by fusimotor activity' [64]. Therefore, the central nervous system, somehow,

can solve this system.

By denoting $x(t) = (f_1(t), f_2(t))^T$, $s(t) = (v(t), p(t))^T$and A the matrix with entries a_{ij}, we get the classic model for instantaneous mixtures

$$x(t) = \mathbf{A}s(t) \tag{3.4}$$

3.2.2 Classical ICA

Blind Source Separation (BSS) is the process of separating a set of source signals from a set of mixtures of these sources without information about the source signals or the mixture process. There are several different methods for performing BSS, each one makes different assumptions: principal components analysis (PCA), assumes that the underlying signals are linearly uncorrelated to each other; projection pursuit searches for the sources with least Gaussian distribution, then subtract them from the mixture until all sources are found.

Independent component analysis (ICA) is a popular method for solving the BSS problem. ICA algorithms solve the model in Equation 3.4 using only the assumption that the sources are statistically independent. The $v(t)$ and $p(t)$ are regarded as sequences of samples from random variables v and p and these variables are assumed mutually independent.

Source separation methods typically assume independence of the source variables, therefore, implying independence of their physical processes represented by these individual sources: in simple terms, two variables are independent if knowing the value of one variable does not give any information about the value of the other. This statistical independence of the sources implies that these signals are produced by physically independent processes and the goal of the analysis is to separate such processes [65].

Note that independence of random variables is a stronger assumption than uncorrelation as it implies no correlation of any nonlinear transformations of variables. Independence is equivalent to uncorrelation only for Gaussian variables, but since there are infinitely many linear transformations providing uncorrelated sources, ICA is not possible for Gaussian variables.

ICA is based on the fundamental result about the separability of linear

mixtures [66], which states that by using the independence criteria it is possible to estimate sources among which there is at most one has a Gaussian distribution. However, there are two well-known ambiguities inherent to the process of ICA:

Scaling: the first one is that scale (or variance) of the components cannot be determined and therefore the variances of the sources are usually normalised to unity;

Permutation: the second one states that the order in which the independent components (e.g. $s_1(t)$, $s_2(t)$) are arranged in the output of the algorithm ($x_1(t)$, $x_2(t)$) cannot be determined.

There exist several approaches to solve the ICA problem. They typically estimate the sources using an unmixing matrix $\mathbf{W}$:

$$s(k) = \mathbf{W}x(k) \tag{3.5}$$

Perhaps the most rigorously justified approach to ICA is minimising the mutual information [67] as a measure of dependence between the sources. There are several algorithms based on different approximations of the mutual information, for example, order statistics [68] or using cumulants [66].

The minimisation of mutual information is essentially equivalent to maximising non-Gaussianity of the estimated sources [69]: this is a natural result which can be understood from the central limit theorem saying that under certain conditions a linear combination of independent random variables tends toward a Gaussian distribution. Thus, the distributions of the observations x_i should be closer to Gaussian compared to the original sources s_j and the goal of ICA is to maximise non-Gaussian components.

FastICA [70] is a popular algorithm based on optimising different measures of non-Gaussianity. Kurtosis is perhaps the simplest statistical quantity for measuring non-Gaussianity. It is defined as

$$kurt(s) = E\{s^4\} - 3(E\{s^3\})^2 \tag{3.6}$$

where $E\{\cdot\}$ denotes expectation. The kurtosis is zero for a Gaussian s and the more distant to zero the kurtosis is, the less Gaussian is the distribution. However, kurtosis is very sensitive to outliers and, therefore, other measures such as negentropy may be used. Negentropy is defined as

$$J(s) = H(s_{gauss}) - H(s) \tag{3.7}$$

where H denotes the differential entropy of s [67] and s_{gauss} is a Gaussian random variable with the same variance as s. A Gaussian variable has the maximum entropy among all random variables with the same variance, therefore, the larger $J(s)$ is, the "least Gaussian" s is. Estimating negentropy, however, is very difficult and it is usually approximated using higher-order moments or some appropriately chosen functions [69].

Another popular approach is the maximum likelihood estimation of the unmixing matrix $\mathbf{W}$ in Equation 3.5. In case the dimensionality N of x equals the dimensionality M of s, the corresponding log-likelihood [71] is given by

$$\mathcal{L} = \sum_{k=1}^{K} \sum_{j=1}^{N} \log p_j(w_j^T x(t)) + k \log |\det \mathbf{W}| \tag{3.8}$$

where K is the number of samples, w_j^T denotes the j-th row of matrix $\mathbf{W}$ and the functions p_j are the probability density functions of the sources s_j. The density functions p_j are not known and have to be estimated. It can be shown [72] that the maximum likelihood approach is closely related to the Infomax algorithm derived by Bell and Sejnowski [73] from the principle of maximising the output entropy of a neural network. In practice, the maximisation of the likelihood is considerably simplified using the concept of natural gradient, as introduced by Amari et al. [74].

3.2.3 Sequential Signal Acquisition

Blind source separation mainly exploits the spatial diversity of the system, that is, that different sensors in the measuring array receive different mixtures of the same sources [75]. This means that the mixtures must be recorded

simultaneously with an array of sensors at distinct positions for separation to be possible via classical BSS techniques. This set up is illustrated in Figure 3.2, for $m = 3$ sources and $n = 3$ sensors. In some settings, however, it may be impractical to use an array of sensors due, for instance, to space limitations or to common practice in the workplace.

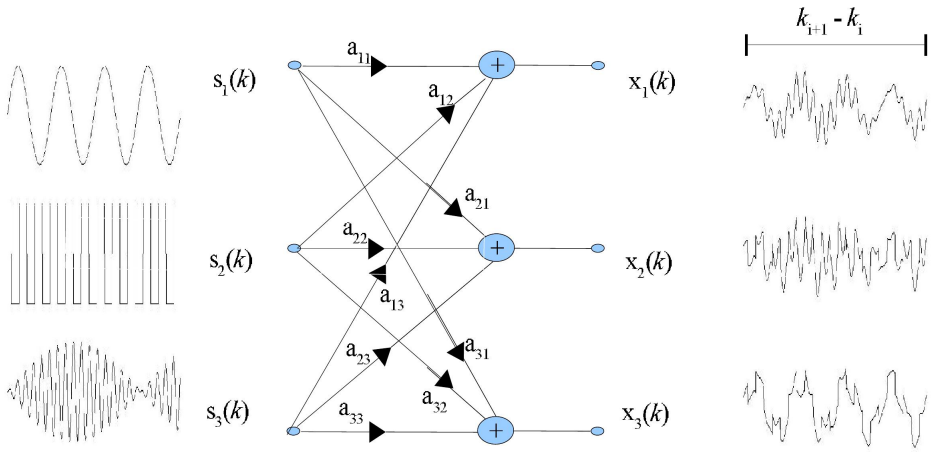


Figure 3.2: Conventional multichannel acquisition of a set of mixtures

An example of the latter is the examination performed by a medical practitioner to listen to the heart or other vital organs of a patient using a stethoscope. A physician listening to the heart will place the stethoscope head in several, pre-specified positions (as mentioned in chapter 2.6) to listen to the organ. In Figure 3.3, this is equivalent to placing a sensor first at location L_1 during time $k_1 < k < k_2$, then at L_2 for $k_3 < k < k_4$ and finally at L_3 for $k_5 < k < k_6$. If we were to record this signal, it would be a single time series containing all the measurement relating to the different positions on the chest (Figure 3.4). How do we recover the original sources for such signal?

At first sight this may appear to be a single sensor separation problem, such as discussed in [76]. However, that is not the case, because the mixture changes at each location, and if we were to treat it this way, we would fail to exploit the spatial diversity that the movable sensor is capturing. In addition, the availability of a moving sensor and the cyclic-stationary nature

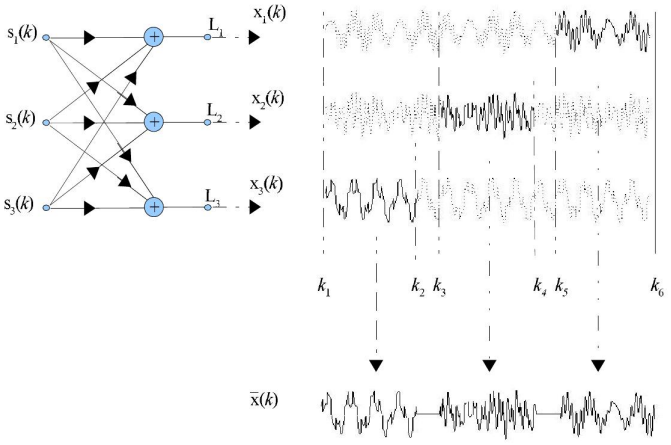


Figure 3.3: Mixed signals recorded sequentially

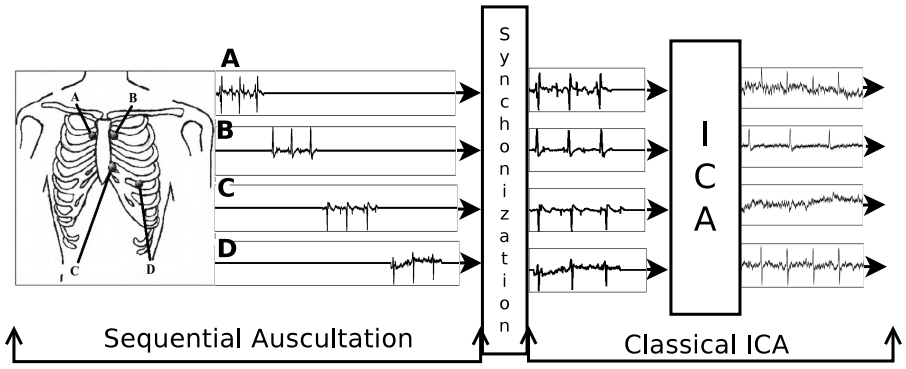


Figure 3.4: An overview of the proposed method

of the signal means that we can simulate an array of n spatially separated sensors which acquire m mixtures by time-shifting to align the mixtures. Thus, $x_1(k)$ will be recorded during the time interval $k_1 < k < k_2$, $x_2(k)$ during $k_2 < k < k_3$, etc., with each mixture being zero outside the time interval.

To overcome this problem, and apply BSS methods, researchers have proposed introducing novel ways of recording the mixtures using an array of sensors. Nigam and Priemer, for example, propose to use an array of stethoscopes to acquire simultaneous recordings of the heart sounds [45]. This has the drawback of needing to change an established procedure. This would be likely to encounter resistance from its potential users, particularly

in the medical profession, where some healthcare innovations have had to be withdrawn due to physicians' resistance [77]. In addition, using an array of sensors implies an individual adjustment for each patient depending on their age and biotype, making the process burdensome.

We propose to leave the measurement process unchanged and address this problem for a particular type of source signal, namely periodic or quasi-periodic sources. We exploit the periodicity property to align the different waveforms within the recorded mixture. Thus, the signal from the moving sensor is segmented to extract the separate mixture signals. These are then artificially synchronised and treated as if they were acquired simultaneously from an array of sensors. As a result, any pre-existing source separation algorithm can be used to extract the underlying sources. We make two main assumptions:

1. A single sensor is placed at several pre-defined locations and records the mixture at that location for a specified length of time, approximately equal for all locations

2. The source signals are periodic or quasi-periodic.

In our algorithm, the sensor is a stethoscope that is placed at particular locations in the thoracic region, during the routine listening of the heart sounds.

As result, the signal from the moving sensor can be segmented to extract the mixture signals, which are then artificially synchronised and treated as if they were acquired simultaneously from an array of sensors. Then, any ICA algorithm can be used to extract the underlying sources.

In this chapter we aim to describe this technique for acquiring signals using a moving sensor. We do this in a way that is inconspicuous to the clinician (in the case of auscultation), who remains free to perform a routine examination, including listening to the heart sounds sequentially at four standard sites, as shown in Figure 2.7.

In the scheme that we present, only one sensor is available, and it is firstly used to acquire a mixture of the sources from location A on the chest

(see Figure 3.4. The same sensor is then placed at the B location and so on, until all observations are obtained. The sensor signal, $x(k)$, contains the heart signal from the four locations, with periods of silence when the sensor is being relocated. Note that the timing between the source signals is not affected by the way the signals are acquired. In order to generate a mixture signal vector as in equation 3.9, we proposed to segment $\bar{x}(k)$ into four signals, so that the i-th heart signal is given by

$$\tilde{x}(k) = \begin{cases} \bar{x}_i(k) & \text{if } k_i < k < k_{i+1} \\ 0 & \text{otherwise} \end{cases} \tag{3.9}$$

where k_i represents the time at which recording at the next thoracic location begins (e.g. the sensor is placed at A at time k_1, at B at time k_2, and so on). We align the signals, so that they can be presented to the ICA algorithm as if they were acquired simultaneously. In doing this, we exploit the quasi-periodicity of the heart cycle. We seek to align the peaks of the mixtures in $\tilde{x}_i(k)$, according to

$$\hat{x}_i(k) = \begin{cases} \tilde{x}_i(k - k_i) & \text{if } i = 1 \\ \tilde{x}_i(k - (k_i + \delta_i)) & \text{otherwise } (i > 1) \end{cases} \tag{3.10}$$

where δ_i is the relative time shift to get the peaks to align. Since we are simulating several sensors capturing the heart sounds at the same time, we use the cross-correlation function (equation 3.11) in order to select the delay that maximises the similarity between the mixtures, therefore aligning the signals in a way that simulates an array of sensors capturing the heart sounds simultaneously.

$$\delta_{i,j} = \max(\textstyle\sum_{n=0}^{N-m-1} \tilde{x}_i(m) * \tilde{x}_j(n+m)) \quad \text{for } i \neq j \tag{3.11}$$

The final step in our proposed method is to perform blind source separation to recover the original source signals. To do this, we first form the mixture vector $\hat{x}(k) = [\hat{x}_1(k), ..., \hat{x}_n(k)]^T$ containing the aligned measurements $\hat{x}_i(k)$ from equation 3.11, at the desired sensor locations. This is now

the conventional observed vector as shown in equation 3.4. If we assume that there are as many sensors as sources, and that the mixing process is instantaneous, we can apply any BSS algorithm to estimate the source. Here, we select the FastICA algorithm [78], because it is well-established, has been extensively studied, and is routinely used in a variety of applications.

The algorithm is summarised below:

Algorithm 3.1 Separation via sequential signal synchronization

1. Segment the observed signal to extract the recordings a the different locations using:

$$\tilde{x}_i(k) = \begin{cases} \bar{x}_i(k) & \text{if } k_i < k < k_{i+1} \\ 0 & \text{otherwise} \end{cases}$$

   where k_i represents the time at which the next recording begins.
2. Align the mixture signals using

$$\hat{x}_i(k) = \begin{cases} \tilde{x}_i(k - k_i) & \text{if } i = 1 \\ \tilde{x}_i\left(k - (k_i + \delta_i)\right) & \text{otherwise } (i > 1) \end{cases}$$

3. Form the mixture vector

$$\hat{\mathbf{x}}(k) = [\hat{x}_1(k), \ldots, \hat{x}_n(k)]^T$$

4. Perform ICA.

3.3 Experiments

In this section we will test the development technique and compare it with classical ICA on different signals. The first experiment is performed with synthetic and noiseless periodic signals, as an initial proof of concept, where the sequential auscultation is also simulated. The remaining experiments explore the periodicity of different heart signals.

In the second experiment we tackle the biomedical engineering problem of separating the foetal electrocardiogram (FECG) from the maternal one. The electrical signal is collected through skin electrodes attached to the mother's body, therefore mixing the foetal electrocardiogram to the maternal electro-

cardiogram [79, 80].

The third experiment deals with the problem of separating heart sounds from the lungs sounds [81]. This is a common problem through auscultation, since, as it happens with the previous problem, due to the sensor's proximity to both heart and lungs, the heart sounds are mixed to the lung sounds, making the auscultatory procedure more difficult.

In the last experiment we try to separate the two subcomponents of the second heart sound. In order to do so, we first try to separate A2 and P2 using the synthetic model we developed in chapter 4, then we proceed to extract A2 and P2 in real signals.

3.3.1 Experiment 1: synthetic signals

In this section we investigate how well our approach performs in recovering the original sources in a number of settings. Firstly, we consider the separation of three known synthetic signals, to assess the separation performance that can be achieved using the method in Algorithm 3.1. We then evaluate the performance of the algorithm on two problems using heart-related signals acquired using a single sensor.

We begin by instantaneously mixing the three periodic signals shown defined by equation 3.13, where *square* function generates a square wave. The mixture is shown in the three upper plots of Figure 3.5, using the following mixing matrix:

$$A = \begin{pmatrix} 0.6948 & 0.0344 & 0.7655 \\ 0.3171 & 0.4387 & 0.7952 \\ 0.9502 & 0.3816 & 0.1869 \end{pmatrix} \tag{3.12}$$

$$\begin{aligned} s_1(k) &= \sin(0.01k\pi) \\ s_2(k) &= square(0.01k\pi) \\ s_3(k) &= \sin(0.05k\pi)\cos(0.0016k\pi) \end{aligned} \tag{3.13}$$

The resulting mixed signals are shown in the three lower plots of Figure 3.5. The mixtures were then randomly shifted, in order to simulate the

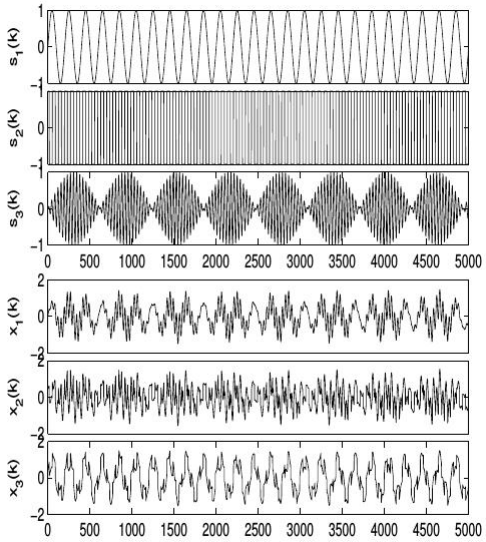


Figure 3.5: Three periodic source signals (upper three plots) and the mixtures obtained when a mixing matrix A is applied (lower three plots)

placing of the sensor at a particular location, which will occur randomly during the fundamental period of the mixture. Figure 3.6 shows an example of a single signal acquired in this case by relocating the sensor at three different positions. The three resulting delayed mixtures can be seen in the upper three plot of Figure 3.7, and it shows that the length of each mixture is now different. The lower three plots in the figure show the mixtures after alignment. Comparing these to the mixtures in Figure 3.5, we can see that the algorithm has successfully aligned the mixtures.

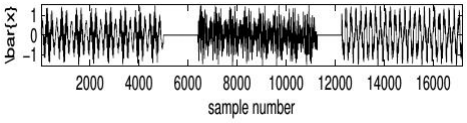


Figure 3.6: Example of single sensor signal acquired at three different locations. The regions between the mixtures, when the sensor is relocated, are set to zero

The separation was performed using FastICA. The upper three plots in Figure 3.8 show the sources recovered when separation was performed from the delayed mixtures. In this case the sources are not separated, because they

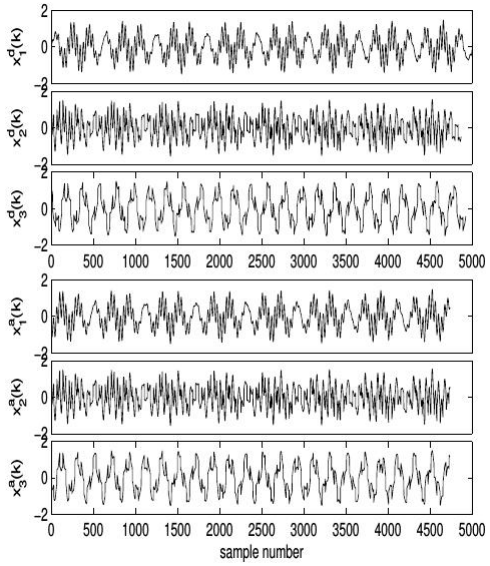


Figure 3.7: The mixtures in Figure 3.5 were delayed randomly, and the corresponding waveforms are shown in the upper three plots. The three lower plots show the signals aligned with the proposed approach. Looking at the peaks, we can see that the signals have been successfully aligned.

are effectively not recorded simultaneously. The three lower plots in the figure show performance following alignment of the mixtures; they show that the algorithm successfully recover the source signals from the aligned mixtures. Finally, Figure 3.9 shows the separation results from the mixtures that were recorded simultaneously (i.e. those in Figure 3.5). Comparing this to the results in Figure 3.8, it is clear that the performance of the proposed method is comparable to that of FastICA based of on the conventional recording set up. In addition to visual inspection, we can see in the following matrix that the Pearson coefficient between the real sources (s_1, s_2, s_3) and the estimated sources of the state of the art method (y_1, y_2, y_3) and the proposed method (y_4, y_5, y_6) are high:

	s_1	s_2	s_3
y_1	0.0007	0.0108	1
y_2	1	0.002	0.0007
y_3	0.0108	1	0.002
y_4	0.0085	0.9759	0.006
y_5	0.9730	0.0165	0.0034
y_6	0.0027	0.0061	0.9915

Table 3.1: The Pearson coefficient between real and estimated sources

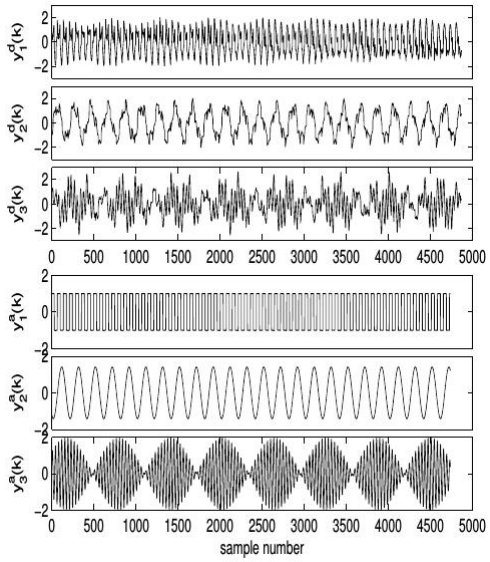


Figure 3.8: Separated sources when FastICA is applied to the delayed mixtures (upper three plots), and when it is applied to the aligned mixtures (lower three plots). The sources are incorrectly recovered from the delayed mixtures, while following alignment the correct sources are identified

3.3.2 Experiment 2: foetal and maternal electrocardiogram separation

Heart signals include the electrocardiogram (ECG) signal, and heart sounds. To further illustrate the proposed methodology, we decided to apply it to ECG signals, and separation of maternal and foetal ECG. The ECG is a

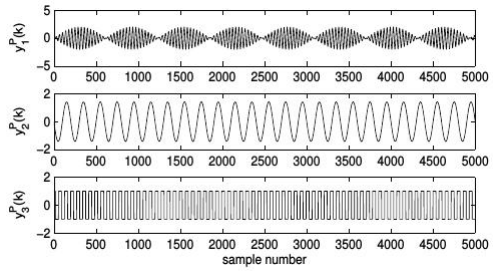
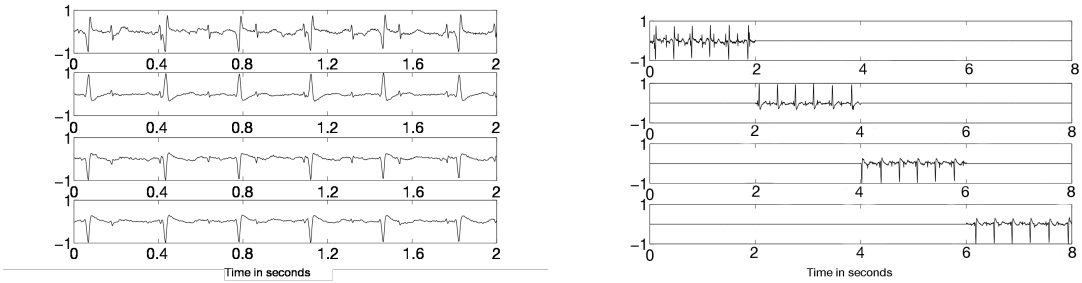


Figure 3.9: Sources separated using FastICA on the initial, non-delayed mixtures, which are recorded simultaneously

recording of the difference in potential between two electrodes during the cardiac cycle, and provides important information about the performance of the heart. ECG signal analysis typically entails removal of noise and interference. A related problem is the removal of maternal ECG (MECG) components from ECG signals recorded during pregnancy, also known as foetal ECG extraction. When risk factors are present during pregnancy, electrocardiograms, along with other measurement methods, may be of vital importance to both mother and child.

This problem has been studied at length [82], and several ICA algorithms have been used to extract the foetal heartbeat. Therefore, we use our algorithm to address this problem, as a step to establish the validity of the proposed method. We perform this experiment by simulating a relocated sensor, then aligning the mixtures, and compared the output of ICA applied on this set of mixtures, with the output of conventional ICA. We used the signals described in [77]. To simulate a relocating sensor, we selected random sections of the recorded signals from each of four sensors. We then aligned the sections as described previously and applied FastICA [83] to perform source separation. Figure 3.10a shows the mixture signals, while Figure 3.10b illustrates the simulated sequential signals, prior to alignment.

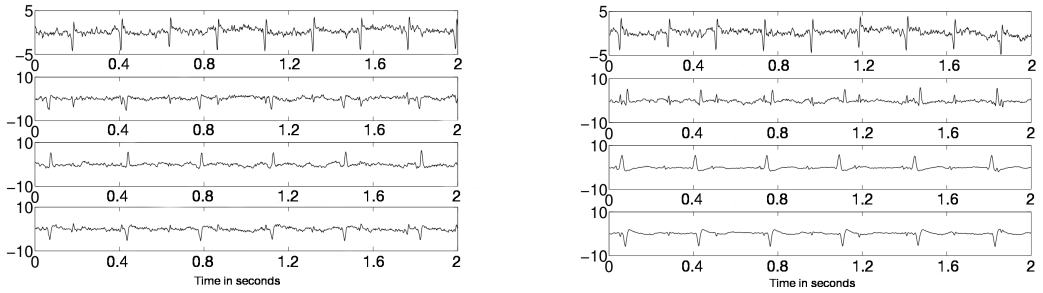


(a) (b)

Figure 3.10: Maternal ECG mixed with foetal ECG. a) is the classical input to ICA, and b) the input for the proposed method.

The separated signals are illustrated in Figures 3.11a and 3.11b, where the latter shows the output of the FastICA algorithm using the classical approach (note that in the first signal of both figures has the higher heart rate, from the foetus), while the former shows the output of the FastICA algorithm using the proposed method. The matrix 3.14 shows a high Pearson correlation between the outputs of the classical method (y_1, y_2, y_3, y_4) and the outputs of the proposed method (y_5, y_6, y_7). Comparing the two figures, we can see that the proposed algorithm successfully extract the foetal ECG (identified by the higher heart frequency), and generally recovers the same sources as conventional ICA. The scaling ambiguity of ICA causes the difference in the second source extracted by the two algorithms.

$$\rho = \begin{bmatrix} & y_1 & y_2 & y_3 & y_4 \\ y_5 & 0.7843 & 0.1164 & 0.2192 & 0.3768 \\ y_6 & 0.0948 & 0.8450 & 0.4698 & 0.1163 \\ y_7 & 0.1292 & 0.6185 & 0.7024 & 0.2608 \\ y_8 & 0.1827 & 0.111 & 0.6185 & 0.8463 \end{bmatrix} \tag{3.14}$$



(a) (b)

Figure 3.11: Output of ECG signals using: a) the proposed method and b) the classical ICA method. In both figures: the first signal is the foetal ECG. The second and forth signals are a mixture with maternal and foetal ECG, and the third signal is maternal ECG.

3.3.3 Experiment 3: heart sounds and lung sounds separation

When performing cardiac auscultation, it is common to hear heart sounds and lung sounds at the same time [81]. In addition, there is a significant overlap of the main energy content of both sounds [84, 85, 86]. This makes the problem of separating the heart sounds and the lung sounds a very interesting and important one. In this experiment we are using ICA as a way of separating the heart sounds from the lung sounds.

Experiment 3.A: heart sounds and lung sounds separation (sequential auscultation)

In this experiment we used one stethoscope recording over the classic auscultation sites (Figure 2.7) in a normal subject. The recordings have lung sounds superimposed to the S2 (Figure 3.12), in order to increase the sound similarities when the proposed technique and simulating four auscultations acquired at the same time.

	$mixture_1$	$mixture_2$	$mixture_3$	$mixture_4$
$mixture_1$		0.0092	0.1509	0.3935
$mixture_2$			0.5176	0.1467
$mixture_3$				0.5196

Table 3.2: The Pearson correlation between the mixtures

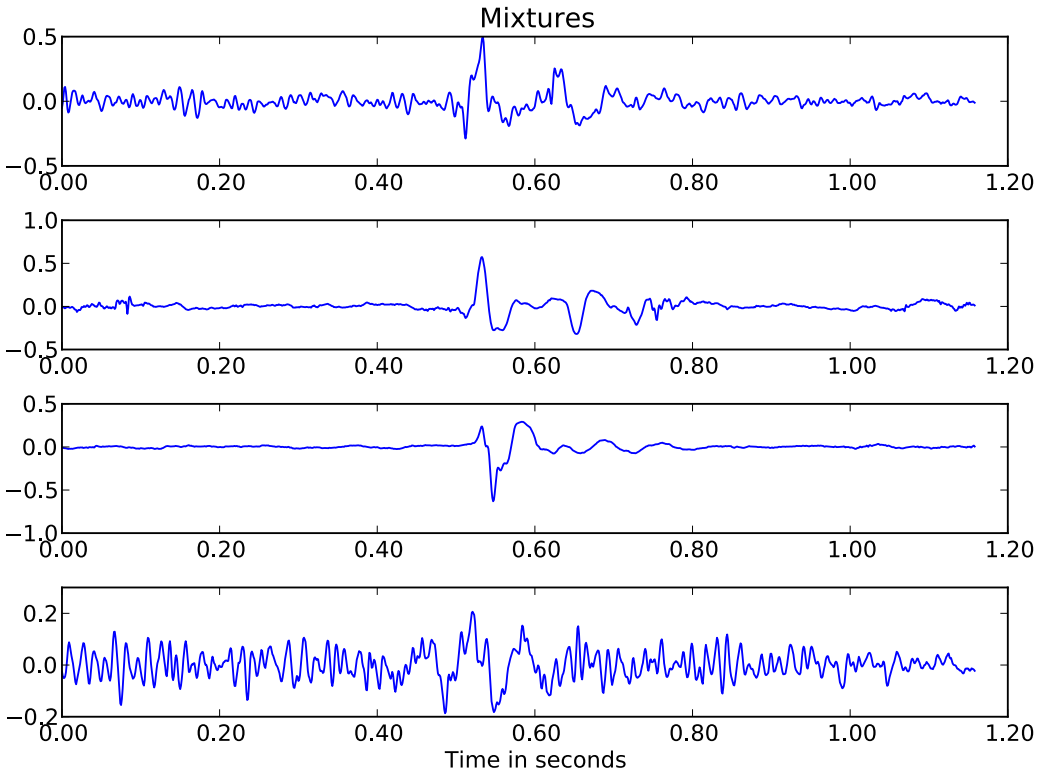


Figure 3.12: Sequential recordings of S2 and lung sounds

The Pearson coefficients between the mixtures shows they are not linear mixtures:

The eigenvalues of the autocorrelation matrix of the mixtures, on the other hand, give us some information about the total number of components present in the mixture:

$$S = \begin{bmatrix} 24.4912 & 0. & 0. & 0. \\ 0. & 20.4028 & 0. & 0. \\ 0. & 0. & 14.7631 & 0. \\ 0. & 0. & 0. & 9.8645 \end{bmatrix} \tag{3.15}$$

The high eigenvalues and low ρ values indicate that these mixtures are not linear and ICA sees them as having too many sources: there is a possibility that breath may be seen as a different source in each recording, or the sound of each lung is also seen as a different source? ICA detects more sources than sensors, therefore, it cannot really separate the sources:

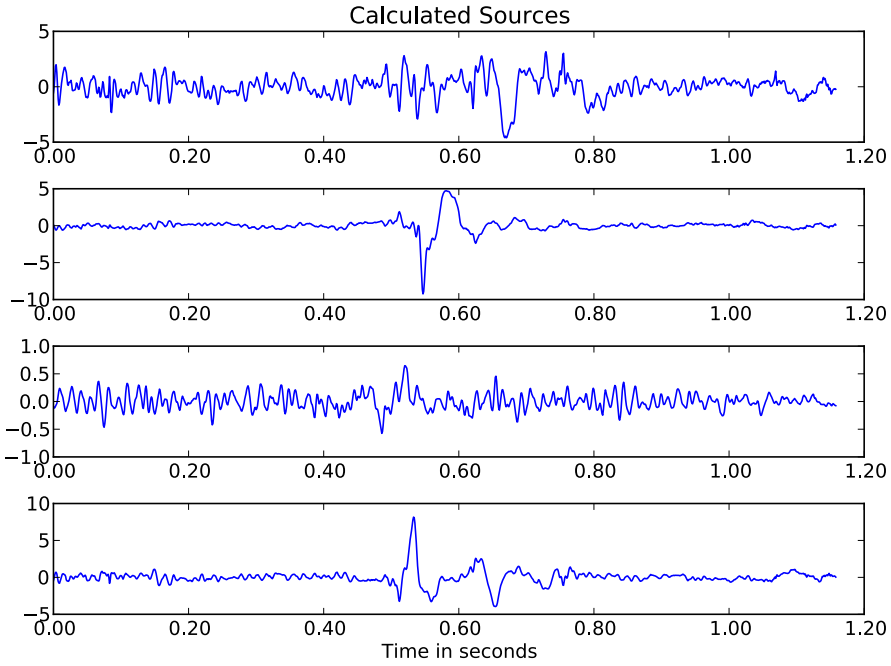


Figure 3.13: Estimated sources from Sequential recordings of S2 and lung sounds

In Figure 3.13, the first source may be seen as the noisy component of the fourth and first mixtures; the second source is the same as the third mixture; the third source is a slightly less noisy version of the fourth mixture; and the fourth source is a more noisy version of the first mixture and a less noisy version of the first mixture. See Figure 3.14

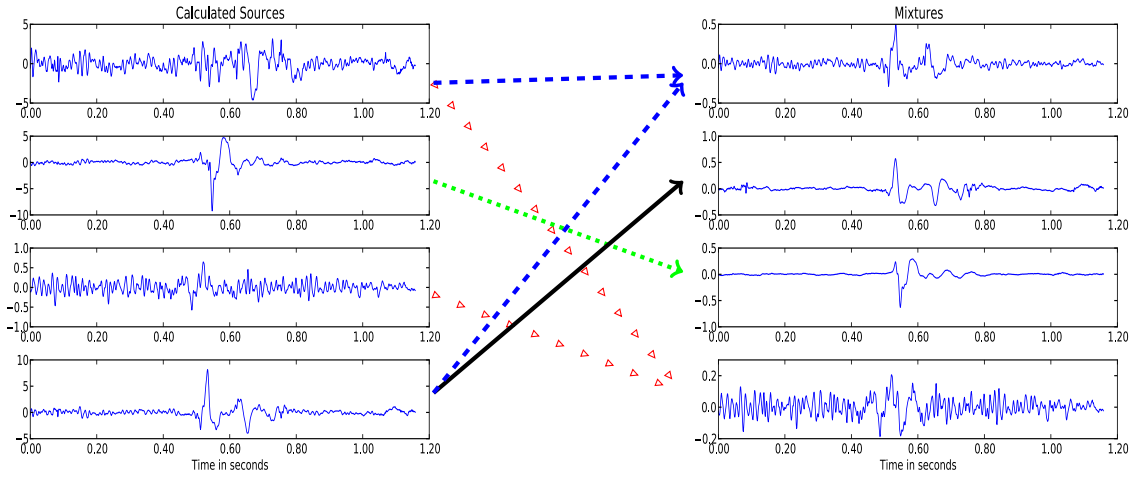


Figure 3.14: Mixtures explained

Experiment 3.B: heart sounds and lung sounds separation (parallel auscultation)

In this experiment we used four stethoscopes placed in the classic auscultation sites (Figure 2.7) in a normal subject. The recordings have lungs sounds over the whole S2, in order to get the most similarity among the recordings, as depicted in Figure 3.15

	$mixture_1$	$mixture_2$	$mixture_3$	$mixture_4$
$mixture_1$		0.3696	0.5966	0.3681
$mixture_2$			0.7511	0.4098
$mixture_3$				0.6127

Table 3.3: The Pearson correlation between the sources

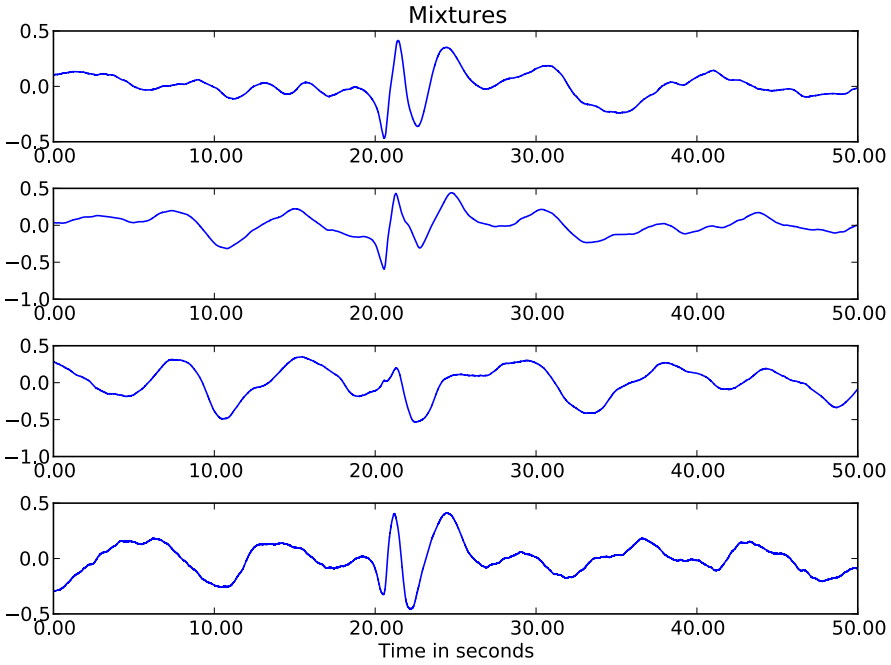


Figure 3.15: Simultaneous recordings of S2 and lung sounds

The eigenvalues of the correlation matrix, on the other hand, give us some information about the total number of components present in the mixture:

$$S = \begin{bmatrix} 16.7307 & 0 & 0 & 0 \\ 0 & 8.9188 & 0 & 0 \\ 0 & 0 & 6.9783 & 0 \\ 0 & 0 & 0 & 3.7382 \end{bmatrix} \tag{3.16}$$

The coefficient of linear correlation between the mixtures shows that these mixtures have low linearity:

The estimated sources and mixture can be seen at Figure 3.16. The

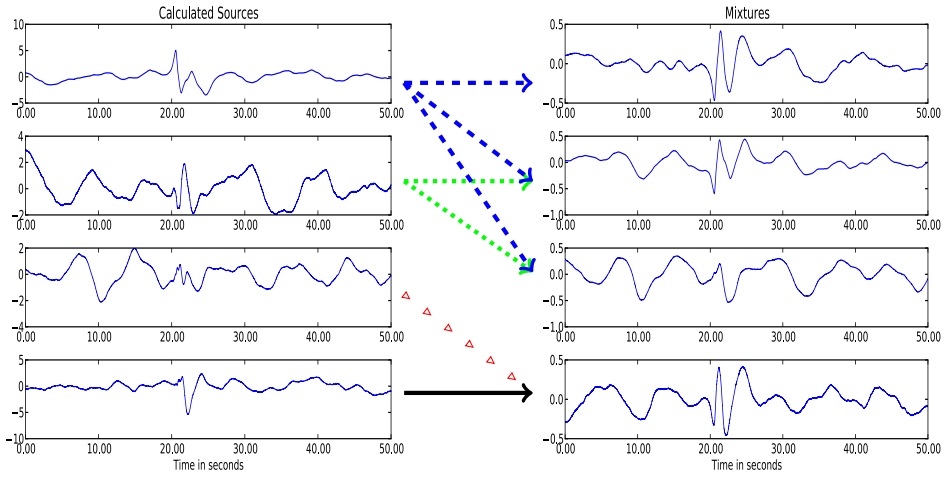


Figure 3.16: Mixtures explained by correlation (the sources with greatest absolute correlation with the mixture).

low eigenvalues and low ρ values (with the exception of $\rho_{(3,2)} = 0.75$) may indicate that ICA sees this mixture as being non-linearly dependent and, as it happened with the sequential case, with too many sources. A possible explanation is that the chest expansion and contraction caused by the respiratory process changes the mixing matrix through time, therefore the mixing coefficients are functions depending on the chest volume, as opposed to the case covered here where they are constants during all auscultation. Another possibility is that the different tissues within the chest produces a reverberation chamber: the different tissues in the human body transmits the sound with different speeds (as seen in chapter 2.6), creating different pathways where the sound can travel and reach the sensor at different times, therefore, producing a convolutive mixture.

3.3.4 Experiment 4: segmentation of the second heart sound

In this experiment we analyse the separation between A2 and P2, the two components of the second heart sound. We will use a synthesis-decomposition approach by using non-linear chirp signals as physiological components the

second heart sound.

Synthetic A2 and P2

As a model of A2 and P2 we are using a developed model based on non-linear chirp signals (further details on the development of this model on Chapter 4), expressed by equations 3.17 and 3.18:

$$g_{\gamma n}(s,u,f_1,f_2,t) = \frac{1}{\sqrt{s}}a(t)sin(\frac{f_1 t}{10} + 2f_2\sqrt{\frac{t}{10}} - 2f_2) \tag{3.17}$$

$$a(t) = (1 - e^{-\frac{t}{8}})e^{-\frac{t}{16}}sin(\frac{\pi t}{60}) \quad 0 \leq t \leq 60ms \tag{3.18}$$

Where s is the scale, u is the displacement (or shift), f_1 is the atom's highest frequency and f_2 is the atom's lowest frequency. The calculation of these parameters was performed using the Matching Pursuit technique, as described on Chapter 4.

The simulation mixing matrix (3.19) reproduces an auscultation in a normal subject. We used the clinical information about how well A2 and P2 can be heart throughout the auscultations sites (chapter 2.6): A2 is stronger than P2 in all auscultation sites; A2 is better heard in the aortic site (matrix element A_{11}); P2 is best heart in the pulmonary site (matrix element A_{22}); and in the other sites, P2 is quite faint (matrix elements$A_{32}$and A_{42}) and A2 is stronger (matrix elements$A_{31}$and A_{41}).

$$A = \begin{bmatrix} 60 & 20 \\ 50 & 25 \\ 12 & 5 \\ 10 & 5 \end{bmatrix} \tag{3.19}$$

Experiment 4.A: ICA with fixed split (simultaneous recording)

In the noiseless case, we mixed the A2 and P2 atoms, using the previous mixing matrix, producing four noiseless mixtures simulating a simultaneous recording:

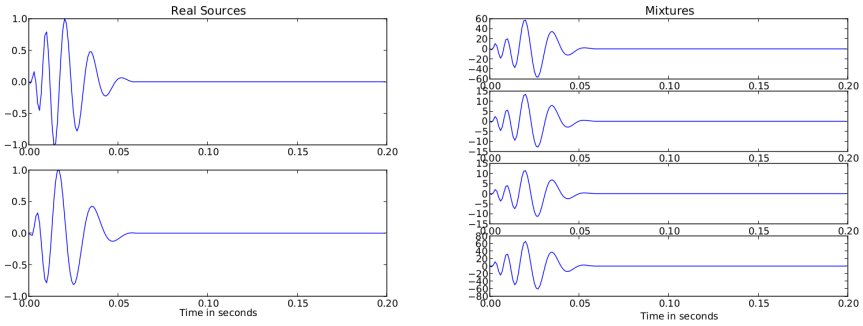


Figure 3.17: A2 and P2 (left) and their mixtures (right): simultaneous recordings and noiseless

As we can see in the correlation matrix R_x, although these mixtures are very similar, it is still possible to perform the separation, since we have more mixtures than sources.

$$R_x = \begin{bmatrix} 1 & 0.8589 & 0.9985 & 0.7441 \\ 0.8589 & 1 & 0.8853 & 0.9812 \\ 0.9985 & 0.8853 & 1 & 0.7790 \\ 0.7441 & 0.9812 & 0.7790 & 1 \end{bmatrix} \tag{3.20}$$

The eigenvalues of the correlation matrix, on the other hand, correctly shows that we have some information about the total number of components present in the mixture:

$$S = \begin{bmatrix} 3359282.95 & 0 & 0 & 0 \\ 0 & 1491118.7 & 0 & 0 \\ 0 & 0 & 1.0878 & 0 \\ 0 & 0 & 0 & 0.2698 \end{bmatrix} \tag{3.21}$$

We can clearly see two high eigenvalues and two quite low eigenvalues, as we only have two sources in our mixtures. As it is correctly separated (Figure 3.18). We can also see that both calculated sources are highly correlated to the original signals, as we can see in the correlation table below:

	x_1	x_2	y_1	y_2
x_1	1	−0.0091	0.1270	0.9918
x_2	−0.0091	1	0.9906	−0.1360
y_1	0.1270	0.9906	1	−0.0001
y_2	0.9918	−0.1360	−0.0001	1

Table 3.4: The Pearson correlation between the sources and estimated sources

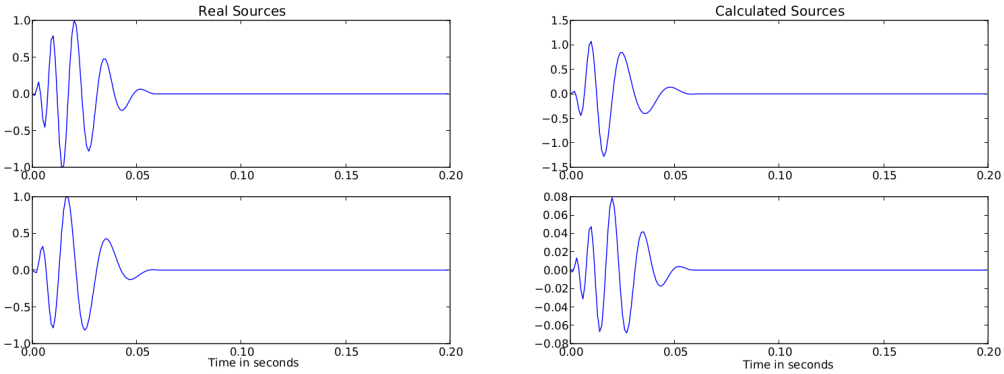


Figure 3.18: Left: A2 and P2 sources and Right: estimated sources

Experiment 4.B: ICA with variable split (sequential recording)

In this experiment, we mixed the A2 and P2 atoms, using the previous mixing matrix, and, for each mixture the delay P2 was of: $09, 14, 15, 17$ milliseconds - this is to simulate a real sequential recording where the cardiac cycle always has some small differences in the position of P2 (chapter 2.6, section 2.3.2). The four noiseless mixtures are sown in Figure 3.19:

The normalised correlation matrix R_x of the mixtures shows that although there is a variable shift in one of the components of the mixtures, the mixtures are still highly correlated, as it happens in the case with multiple sensors:

$$R_x = \begin{bmatrix} 1 & 0.9031 & 0.8404 & 0.7448 \\ 0.9031 & 1 & 0.9611 & 0.8487 \\ 0.8404 & 0.9611 & 1 & 0.9540 \\ 0.7448 & 0.8487 & 0.9540 & 1 \end{bmatrix} \tag{3.22}$$

The coefficient of linear correlation between the mixtures shows a lower

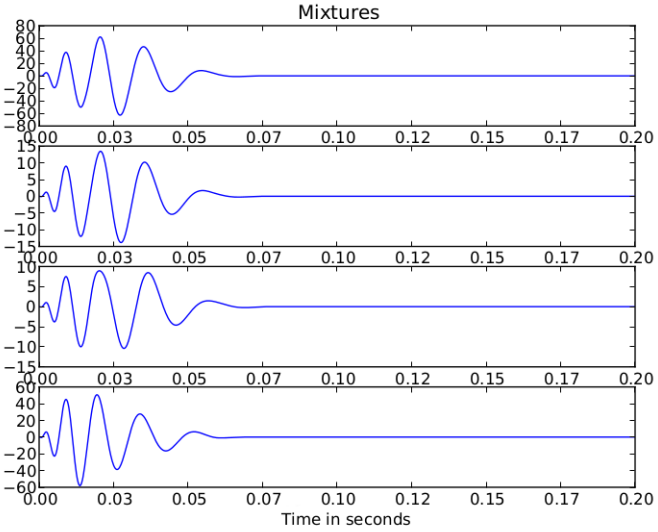


Figure 3.19: The recordings with P2 changing its delay to A2 (09, 14, 15, 17 milliseconds, respectively)

	$mixture_1$	$mixture_2$	$mixture_3$	$mixture_4$
$mixture_1$		-0.5715	0.8483	0.6293
$mixture_2$			-0.0708	-0.7357
$mixture_3$				0.2020

Table 3.5: The Pearson correlation between the mixtures

linear correlation (compared to the fixed split case), indicating some non-linearity in the mixing process introduced by the variable shift:

However, the eigenvalues still indicates the presence of two components:

$$S = \begin{bmatrix} 3.0963 & 0 & 0 & 0 \\ 0 & 0.1436 & 0 & 0 \\ 0 & 0 & 0.0224 & 0 \\ 0 & 0 & 0 & 0.0004 \end{bmatrix} \tag{3.23}$$

Two values are reasonably high if compared to each other, although in this case they are much lower than the case with multiple sensors, it is still enough to recover the original sources, as can be seen in Figure 3.20:

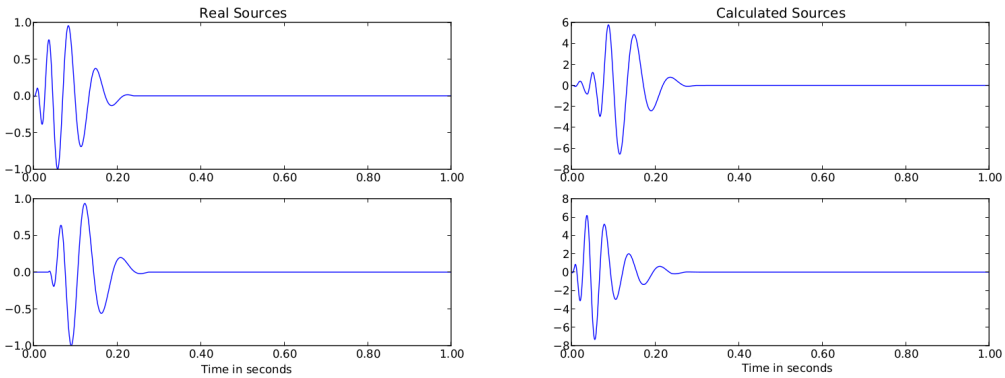


Figure 3.20: Left: A2 and P2 sources and Right: estimated sources

Experiment 4.C: ICA with recorded S2 with four sensors

In this experiment we used four stethoscopes on the areas depicted on Figure 3.4, selected the second heart sound as seen in Figure 3.21 and performed ICA on the mixtures.

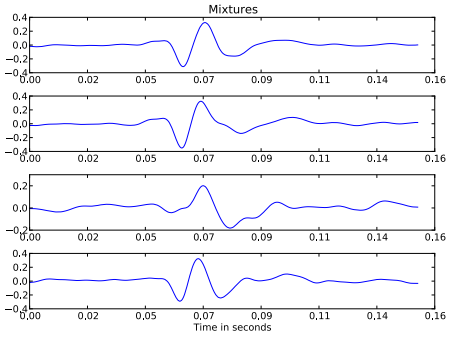


Figure 3.21: The second heart sound acquired using four stethoscopes

The eigenvalues of this mixture shows four components:

$$S = \begin{bmatrix} 42.1716 & 0 & 0 & 0 \\ 0 & 23.7645 & 0 & 0 \\ 0 & 0 & 18.3842 & 0 \\ 0 & 0 & 0 & 6.8032 \end{bmatrix} \tag{3.24}$$

In addition, the Pearson correlation between the mixtures shows they have low linearity:

	$mixture_1$	$mixture_2$	$mixture_3$	$mixture_4$
$mixture_1$		0.6702	0.7217	0.5835
$mixture_2$			0.9132	0.6978
$mixture_3$				0.6197

(3.25)

Table 3.6: The Pearson correlation between the mixtures

The exception here are $mixture_2$ and $mixture_3$, which corresponds to the Aortic and Pulmonary auscultation areas of auscultation. Although these mixtures have high linearity, they differ very little, and don't add much information to the system, since they are the two closest locations: most of the mixtures still have low linearity. This problem leads to a source extraction where although one estimated source (although still mixed) is calculated, and the other one is a mixture buried in noise:

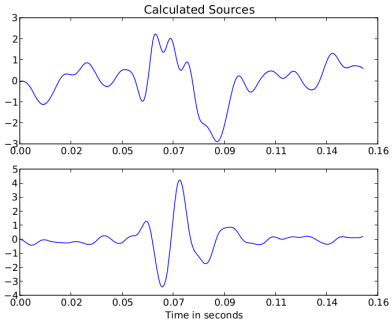


Figure 3.22: The calculated sources

In addition, the second calculated source of Figure 3.22 is similar to the first three mixtures and the first calculated source doesn't have a particular dominant signal.

3.4 Discussion and conclusions

We studied the problem of performing source separation on heart signals. A novel technique using a moving sensor to collect data and simulate signal acquisition from an array of sensors was developed. This technique only requires the sensor to move and the signal to be at least semi-cyclic. The

proposed method has the following advantages when compared with the use of an array of sensors:

1. An array of sensors for collecting auscultations would have to be developed.

2. The array of sensors would have to be adjusted for each patient depending on the size and shape of the chest.

3. The clinical routine would have to be adapted in order to use this array.

The proposed technique worked on synthetic data as well as in the problem of separating maternal ECG from the foetal ECG.

On the problem of separating heart sounds from lung sounds, the results suggest that some mixtures are non-linear and it appeared that the number of observations is less than the number of possible independent sources. This could be due to the changes in the chest volume caused by the respiration, making the mixing matrix change over time. Another possibility is that tissues on the chest make the mixture convolutive or the sensors are receiving different sources of noise. We believe that other techniques, such as convolutive independent analysis or even other techniques such as dependent component analysis could be more effective in solving this problem.

The problem of separating A2 and P2 from the second heart sound, however, proved to be much more harder in a number of ways:

Validation: A2 and P2 are very short bursts of sound, there is also a very short time interval between them (in some cases they can overlap completely). Hence, the location in time and validation of these underlying components has been shown very difficult. An extra and parallel measurement with great time resolution would have to be performed to locate in time these components.

ICA Ambiguities: ICA can identify the time where each underlying occurs. This information is important in identifying some cardiac diseases. However, due to the amplitude ambiguity, important information regarding the amplitude of each component is lost. This information is

particularly important in identifying some diseases such as pulmonary and systemic pressure levels.

Mixtures: Experiments have suggested that the mixing process on the human chest is likely to not be an instantaneous mixture: different tissues on the sound pathway may produce reverberation, leading to a convolutive mixing matrix; or the internal natural movement of the heart and lungs inside the middle mediastinum may lead to several mixing matrices that depends on the relative positions and instant volume of these organs.

To conclude, although we demonstrated the potential of the proposed technique on synthetic and real signals, we could not successfully segment components of the heart sounds using ICA. However, we created a technique that allows the use of ICA using only a single moving sensor (as long as the captured signal is a semi-cyclic one). In addition, a by-product of the research presented in this chapter was the creation of mathematical models of A2 and P2. These will be the discussed further in the next chapter along with their theoretical development.

3.5 Publications

The work developed in this chapter generated the following publications:

1. F. L. Hedayioglu; M. G. Jafari; S. S. Mattos; M. D. Plumbley; M. T. Coimbra, "Separating sources from sequentially acquired mixtures of heart signals," Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on , vol., no., pp.653,656, 22-27 May 2011 doi: 10.1109/ICASSP.2011.5946488 [87]

2. M. G. Jafari; F. L. Hedayioglu; M. T. Coimbra; M. D. Plumbley, "Blind source separation of periodic sources from sequentially recorded instantaneous mixtures," Image and Signal Processing and Analysis (ISPA), 2011 7th International Symposium on , vol., no., pp.540-545, 4-6 Sept. 2011 [88]

Chapter 4

Representation of S2 Using Matching Pursuit

4.1 Introduction

In the last chapter, we adopted an approach where we made very few assumptions regarding the signal components. We aimed to obtain underlying individual components, which when mixed resulted in the collected auscultation signal. This is theoretically possible since we are essentially facing a blind-source separation problem that explores that different auscultations sites have different mixtures of each signal. Results using ICA were interesting but somewhat limited, possibly due to the convoluted nature of the gathered mixtures.

In this chapter we describe an approach where we assume that the signal components can be modelled by a specific function such as non-linear chirp signals [42, 44]. This is a powerful simplification that transforms the problem into an optimisation one, in which we try to discover the component mixture that better explains the observed signal. From a physiological perspective this is quite reasonable since the mechanical process that produces heart sounds is well-known. It is limited, however, to patients that have 'standard' heart sounds and, depending on the chosen algorithm, might be affected by the presence of murmurs. However, since the search space is very large, sub-

optimal algorithms are needed in order to find the best combination of signal components that explain the observation. Matching pursuit is one of these sub-optimal algorithms that has the advantage of solving this problem with few components. In this chapter we show that matching pursuit using a dictionary of non-linear chirp signals can give us quite good results for this approach and has good physiological meaning.

4.2 Previous Work

T. Tran et al. in [89] describes a heart sound synthesiser that models the heart sound as produced exclusively by heart valves: mitral and tricuspid for S1, and aortic and pulmonary for S2.

The heart sound components are modelled using chirps with linear frequency decay on a normalised envelope (equation 4.3).

$$A(t, T, \alpha_A) = \left[(1 - \exp(-\frac{t}{0.2T}))\sin(\frac{\pi}{2T}T)\right]^{1-\frac{\alpha_A}{\pi}} \tag{4.1}$$

$$D(t, T, \alpha_D) = \left[(1 - \exp(-\frac{t}{0.4T}))\cos(\frac{\pi}{2T}t)\right]^{1-\frac{\alpha_D}{\pi}} \tag{4.2}$$

$$c(t, T, A, D, f_o, f_i) = \frac{A(t, T, \alpha_A)D(t, T, \alpha_D)}{\max(A(t, T, A)D(t, T, \alpha_D))}\sin((f_o + f_i)\pi t) \tag{4.3}$$

Where t is the time, T is the duration of the transient, A is the rate of attack, D is the rate of decay, f_o is the initial frequency and f_i is the final frequency. According to the authors, the component amplitude variation is calculated proportionally to that of real clinically recorded heart sounds. The authors, however, does not provide a comparison between the synthetic sounds produced by their equations and real recorded ones.

Xuan Zhang et. al. in [90] performed matching pursuit on 11 auscultation signals from an educational CD, containing some common pathological and normal auscultations recorded under an ideal and noiseless environment. They added Gaussian noise to the 11 auscultations, producing a dataset of 22 auscultations in total. Then, they applied matching pursuit using a re-

dundant and complete dictionary of Gabor atoms, (equation 4.4), where β_i is a normalizing factor, s_i is a scale factor to control the width of the envelope of $h_i(t)$, p_i controls the temporal placement of the envelope function g in the atom, f_i and ϕ_i are the frequency and phase of the atom, respectively. To successfully reproduce an auscultation (S1 and S2), 11 atoms where needed. However, these atoms do not represent the sub-components of the heart sounds, and the noiseless recording conditions of the dataset are not achievable on real clinical environment.

$$h_i(t) = \beta_i.g_i(t).u_i(t) \tag{4.4}$$

$$g_i(t) = g(\frac{t - p_i}{s_i}) \tag{4.5}$$

$$u_i(t) = \cos(2\pi f_i t + \phi_i) \tag{4.6}$$

Tang et. al. in [91] extends Koymen [92] work and models the aortic component of the second heart sound as exponentially damped sinusoids (4.7) using auscultations from six dogs. They use the mean filter of forward and backward predictor as means to calculate the parameters of the damped sinusoids. In this work, the aortic component was modelled using 6 to 20 components.

$$s(n) = \sum_{i=1}^{M} A_i e^{-n\alpha_i} \sin(2\pi n f_i + \varphi_i) \tag{4.7}$$

This model and technique, however, has problems representing transient components of the signal, and for this reason cannot be used to represent A2 and P2 as separated components.

Jingping Xu et. al [42, 44] postulated A2 and P2 components as nonlinear narrow-band chirp signals, with a fast decreasing instantaneous frequency over time. They performed a dechirping of the recorded S2 in order to generate a low-frequency estimate of the amplitude envelope of A2, then a 2D mask is designed by visual inspection and is applied on S2's Wigner-

Ville distribution, estimating the instantaneous frequency of the strongest component of that signal. Based on this frequency, they reconstructed A2 and subtracted it from the original signal, repeating this procedure in order to estimate the components of P2. This approach was applied to four pigs under anaesthesia with chemically induced severe pulmonary artery pressure (PAP). This approach, however, is sensitive to noise: the polynomial Wigner-Ville distribution's performance degrades, as well as the estimation of the instantaneous frequencies.

4.3 Methods

4.3.1 Matching pursuit

Matching pursuit (MP) is a greedy technique used to decompose a signal into a linear combination of simpler signals (atoms) that are selected from a dictionary of time-frequency functions in which the first atom always has the highest energy and highest dot product with the observed signal [93]. MP is a greedy method that tries to find a representation that is sparse in the dictionary, i.e. only a few atoms participate in the approximations. Therefore, MP can represent the underlying structures of the signal in a compact way.

MP, however, is one of the many solutions to the approximation problem: the problem of finding two atoms whose sum is the best approximation to S2 may not be found by the MP algorithm, as pointed by Rémi Gribonval [94], finding the best approximation of a signal by a linear combination of a dictionary D is an NP-hard problem, and matching pursuit does not provide such approximation. Other methods to solve the approximation problem include: Basis Pursuit [95], the Iterative Hard Thresholding Algorithm (IHT) [96]; and some others based on the MP, such as Stagewise Orthogonal Matching Pursuit (StOMP) [97], and Regularized Orthogonal Matching Pursuit (ROMP) [98].

Mathematically, the MP algorithm approximates the signal s as a sum of weighted atom $\psi_{\gamma k}$ (k-th atom with γ shift) from a dictionary D. The

approximation using $m-1$ atoms is given by:

$$s^{(m)} = \sum_{k=0}^{m-1} \alpha_k \psi_{\gamma k} \tag{4.8}$$

where α_k are the weights of each $\psi_{\gamma k}$ atom.

Let us define $r^{(m)} = s - s^{(m)}$ as the residual or the approximation error after s is approximated by $m-1$ atoms. The algorithm starts from an initial approximation $s^{(0)} = 0$ and a residual $r^{(0)} = s$. Then it searches for the atom in the dictionary that has the highest dot product with the current residual. Once this atom is found, a scalar multiple of that atom is added so that $s^{(k)} = s^{(k-1)} + \alpha_k \varphi_{\gamma k}$. The scalar $\alpha_k =< r^{(k-1)}, \varphi_{\gamma k} >$ where $< a, b >$ denotes the dot product between a and b. The MP procedure is summarised in the Algorithm 4.1:

Algorithm 4.1 The MP algorithm. The two stop conditions are: reaching maximum number of iterations *iteratLim*, or the minimum error threshold *errorThreshold*.

---

1. Initialize:
   $s^0 = 0$;
   $r^0 = s$;
   $m = 0$;
2. Repeat:
3. $m = m + 1$;
4. Find $\varphi_{\gamma k} \in D$ that $maximizes < r^{(k-1)}, \varphi_{\gamma k} >$
5. $\alpha_m =< r^{(m-1)}, \varphi_{\gamma k} >$;
6. $s^{(m)} = s^{(m-1)} + \alpha_k \varphi_{\gamma k}$; //update approximation
7. $r^{(m)} = s - s^{(m)}$; //update residual
8. Until $(m > iteratLim) \vee (r^{(m)} \leq errorThreshold)$

As a result, the signal s is decomposed into a series of time-frequency atoms 4.9 with decreased energy order.

$$s = \sum_{k=0}^{m-1} \alpha_k \varphi_{\gamma k} + r^{(m)} \tag{4.9}$$

4.3.2 The Model/Dictionary

We already know that S2 is composed by only two components: the sound following the closure of the aortic valve (A2) and the sound followed by the closure of the pulmonary valve (P2). Thus, we can postulate the production of two transient components as the components of the second heart sound. After the aortic valve closes, the blood column rebounds the closed valve, the arterial walls and the blood itself is put into an oscillatory state. The tension of the Aorta walls increases accordingly to Young-Laplace equation [99]. This wall will exert an increasing re-setting force back into the resting position, similar to a vibrating drum head. This will result in oscillation frequencies and amplitude proportional to blood pressure, followed by a decrease in oscillation frequency and amplitude, due to the system's natural loss of energy. The same mechanism happens to the pulmonary artery [100]. One consequence of this system is to consider the A2 and P2 components to be short-time signals with decreasing modular frequency, proportional to the pressure of their respective vessels. For such, a non-linear chirp model extended from Jingping Xu [42, 44] was developed.

We modelled A2 and P2 as narrow-band chirp signals defined as:

$$A_2(t, s_a, u_a, f_{oa}, \Delta f_a) = s_a Am_a(t + u_a) \sin(\varphi_a(t + u_a, f_{oa}, \Delta f_a)) \tag{4.10}$$

$$P_2(t, s_p, u_p, f_{op}, \Delta f_p) = s_p Am_p(t + u_p) \sin(\varphi_p(t + u_p, f_{op}, \Delta f_p)) \tag{4.11}$$

Where t is the time, s_a and s_p are the amplitudes, $u_a$and u_p are the displacements (or shift), $Am_a(t)$ and $Am_p(t)$ are the envelopes, and $\varphi_a(t, f_{oa}, \Delta f_a)$ and $\varphi_p(t, f_{op}, \Delta f_p)$ are the phase functions f_{oa}, and f_{op} are the initial frequencies, $\Delta f_a$and Δf_p are the decaying frequencies variations.

$$Am(t) = (1 - e^{-\frac{t}{8}})e^{-\frac{t}{16}} \sin(\frac{\pi t}{0.06}) \tag{4.12}$$

Since each component usually lasts for less than 0.05 seconds [42], we have

set the amplitude to zero after 0.06 seconds, allowing them to be localised in the time domain (Figure 4.1).

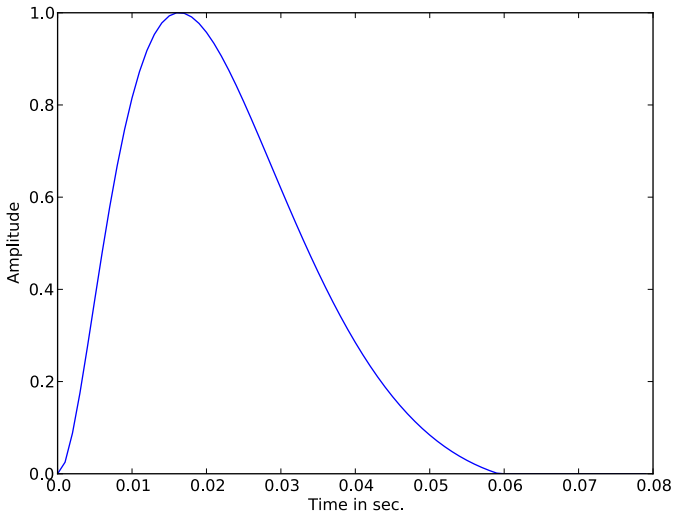


Figure 4.1: Graph of $Am(t)$ function with duration set to 60 milliseconds.

The phase function $\varphi_a(t)$ is estimated by integrating the instantaneous function $IF_a(t)$:

$$IF_a(t, f_{oa}, \Delta f_a) = f_o + \Delta f_a \sqrt{t} \tag{4.13}$$

therefore:

$$\varphi = \int_0^t IF_a(t, f_{oa}, \Delta f_a) dt \tag{4.14}$$

$$\varphi = f_{oa} t + \frac{2}{3} \Delta f_a t^{1.5} + f_{oa} \tag{4.15}$$

We now build a redundant and complete dictionary D by creating atoms with $u = 0$,$s = 1$, and all starting frequencies from 20 Hz up to 500 Hz, since this is the frequency range of the second heart sound [13, 19]. Since we are modelling the two components of the second heart sound, the matching pursuit code is set to stop within two interactions.

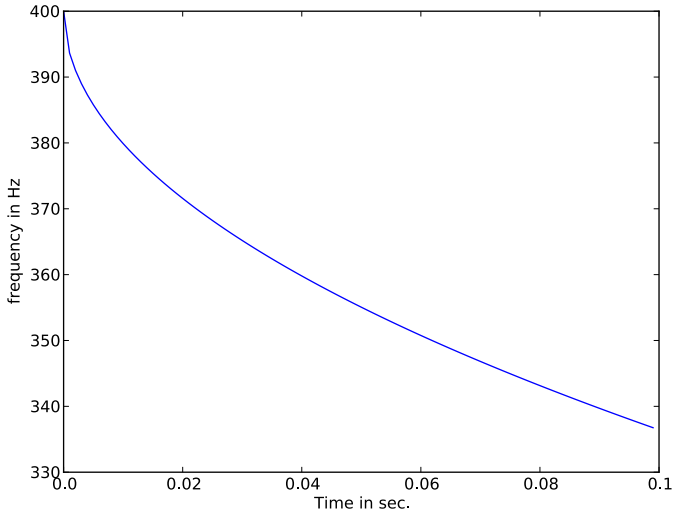


Figure 4.2: The frequency decay of an atom where $f_o = 400Hz$ and $f_f = 350Hz$ at 0.06 seconds.

4.4 Experiments

4.4.1 Dataset

The description of the dataset used in the experiments can be find in the Appendix A.

4.4.2 Experiment 1: denoising/sparse representation of S2

In this experiment we apply the matching pursuit using the atoms described earlier. The dictionary used for this experiment had all combination of atoms with frequencies ranging from 500 Hz down to 20 Hz, as show below:

$$\begin{bmatrix} f_0 & f_f \\ 500 & 499 \\ 500 & 498 \\ \vdots & \vdots \\ 22 & 21 \\ 22 & 20 \\ 21 & 20 \end{bmatrix} \tag{4.16}$$

We can easily see that the memory and speed demands for the MP algorithm can be enormous and saving memory and improving speed is a real problem. The size of the dictionary is great (over 115000 atoms). In order to optimise the memory usage of the dictionary, we opted to not build a dictionary with the atoms' signal, but only with their frequency parameters and generate the atoms on demand as we search the dictionary. We also parallelized the matching pursuit algorithm in order to improve the processing speed.

As a result, a noiseless signal that is an approximation of S2 was generated for each annotated S2 in our dataset (Figure 4.3).

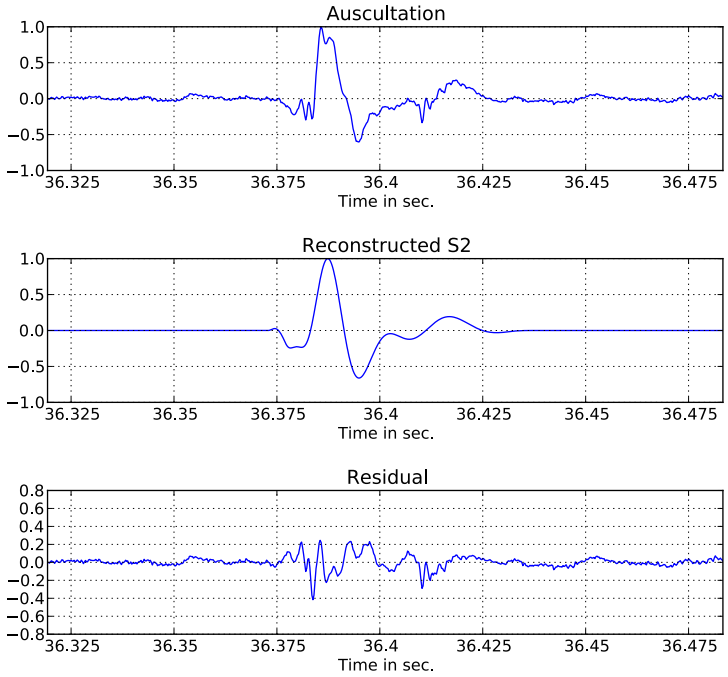


Figure 4.3: Recorded S2 (top), reconstructed S2 by MP (middle) and its residual (bottom)

The high correlation between the approximated S2 and the original one shows that this method can provide a good approximation of the real S2 (Figure 4.4). The example of this reconstructed S2's energy also captured 91.5% of the total energy of the recorded S2 (10.6), leaving a residual with only 8.5% of the total signal energy.

We now have a sparse representation of the second heart sound: this representation needs only 2 atoms and each atom (Equations 4.10 and 4.11) needs four parameters: the scale s; shift u; lowest frequency and frequency variation $(f_o, \Delta f)$ to represent each component of S2 component. In total only 8 parameters are required to fully reconstruct S2: $(s_a, u_a, f_{oa}, \Delta f_a, s_p, u_p, f_{op}, \Delta f_p)$.

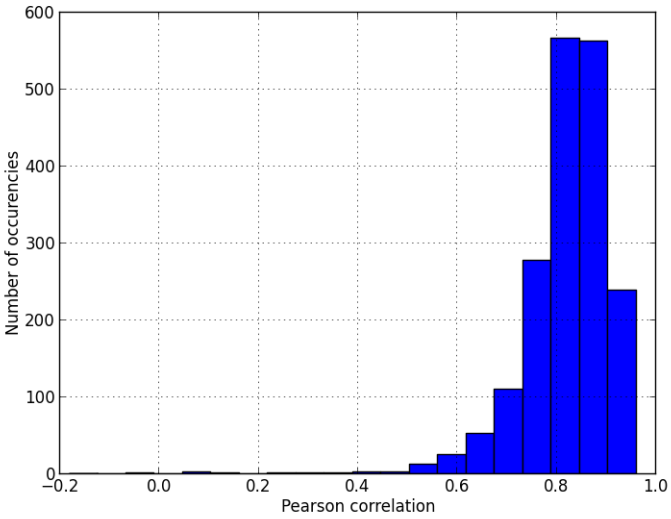


Figure 4.4: Histogram of the Pearson correlation between all recorded and reconstructed S2 from the dataset (n=1858).

4.4.3 Experiment 2: atoms and physiological features

The objective of this session is to show that since these atoms are physiologically inspired, they also have some information that may be useful to draw conclusions regarding the physiological components of the second heart sound.

The direct measurement of A2 and P2 from a normal auscultation has been proven quite difficult: these components have a very short duration, there is a significant overlap between them and auscultations may have a high level of noise. However, indirect inference of A2 and P2 is possible: if an auscultated patient has PAH, then the P2 component of S2 has an increased amplitude (and energy) if compared with a normal P2 (as seen in section 2.3.2). Therefore we use the total energy x^2 of the fully reconstructed S2 as a parameter. A second parameter used is the time difference between the starting time of A2 and P2 atom ($u_P - u_A$). The latter feature is based on the fact that commonly the PAH (with right heart failure) can impose a delay on P2, making the time difference between these components greater.

We used the Expectation Maximisation algorithm to cluster the data into two sets, as we can see in Figure 4.5: The cluster 0 (smaller, lower left of

the graph) is very compact and is coincident with the normal auscultations, showing that normality in terms of energy and split is well defined. The cluster 1, on the other hand, is much more disperse and covers most of the hyperphonetic heart sounds. The confusion matrix is shown in Table 4.1. We used the formulas in equations 4.17, 4.18 and 4.19, for calculating the accuracy, precision and balanced precision respectively, where TP is the number of true positives, TN is the number of true negatives, FP is the number of false positives and FN is the number of false negatives. For detecting normal auscultations (cluster 0) we found 70.27% accuracy, 91.30% precision and balanced accuracy of 70.71%.

$$Accuracy = 100 * \frac{TP + TN}{TP + FP + FN + TN} \tag{4.17}$$

$$Precision = 100 * \frac{TP}{TP + FP} \tag{4.18}$$

$$Balanced\_Accuracy = 100 * (\frac{0.5 * TP}{TP + FN} + \frac{0.5 * TN}{TN + FP}) \tag{4.19}$$

	Normal	Hp
Cluster 0	21	2
Cluster 1	5	9

Table 4.1: Confusion matrix: columns has the ground truth and rows has the cluster

(4.20)

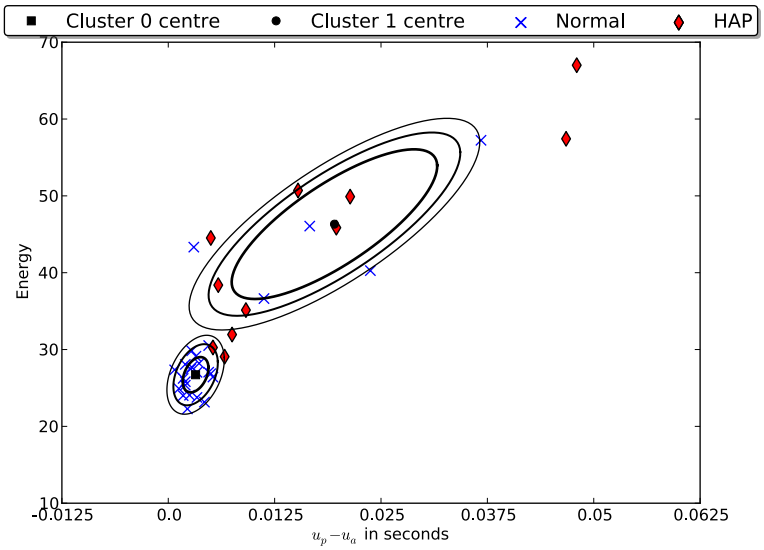


Figure 4.5: Scatter plot with the two Gaussian cluster's contour lines separating normal and hyperphonetic heart sounds. The x axis has the median of all differences between the starting time of A2 and P2 atom ($u_P - u_A$), the y axis has the median total energy (x^2) of the reconstructed S2. The legend displays the clinician classification of the auscultations.

4.5 Conclusions

In this chapter we developed an approach where we modelled the structure of A2 and P2 as non-linear chirp signals and applied matching pursuit in order to find the parameters of these components. This method has the advantage of generating a sparse representation for the second heart sound and to create a noiseless representation of the original heart sound.

In the first study our results show a high correlation between the recorded and reconstructed S2. It can be seen that just two atoms are enough to capture the main features of the signal and produce noiseless reconstructed version of the original S2. It is also seen that although in general, matching pursuit produces reconstructed signals with high correlation with the original ones (Figure 4.4), the cases where a lower correlation is obtained can be explained by the greedy nature of the MP technique used, where a sub-optimal solution is found. This may be further improved by developing other

search strategies, for instance: maximising the energy of each first interaction of the MP algorithm may lead in some cases to a sub-optimal solution (having the most energy of the S2 captured by only one component, leaving the second component to capture noise), a more sensible strategy could favour the selection of atoms with a more balanced contribution to S2 (therefore, they would tend to represent more relevant features of the signal).

In the second study we demonstrated that since the atoms are physiologically inspired, they also express information that may be related to physiological components of S2. After using the Expectation Maximisation algorithm we have shown that the total energy x^2 of the fully reconstructed S2 and the time difference between the starting time of A2 and P2 atom ($u_P - u_A$) can be used to cluster the normal heart sounds, achieving 70.27% accuracy, 91.13% precision and balanced accuracy of 70.71%.

4.6 Publications

The work developed in this chapter generated the following publications:

1. F. L. Hedayioglu; M. G. Jafari; S. S. Mattos; M. D. Plumbley; M. T. Coimbra, "Denoising and segmentation of the second heart sound using matching pursuit," Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE (pp. 3440-3443). IEEE; (2012, August) [101]

2. F. L. Hedayioglu; M.G. Jafari; S. S. Mattos; M. D. Plumbley; M. T. Coimbra, "An Exploratory Study On The Segmentation Of The Second Heart Sound", to be submitted

Chapter 5

Energy Based Segmentation of S2

5.1 Introduction

In the previous chapter we modelled the second heart sound as the sum of non-linear decaying chirp atoms and showed that features produced by these atoms can be used to extract information about the underlying physiological components of S2.

In this chapter we postulate that the second heart sound components are peak-like, and the measurement of their amplitude and delay can also relate to the underlying physiological components of S2. In fact, this is the very same procedure carried out by doctors when reading a phonocardiogram: they consider the positive amplitude of the peaks of each component (Figure 5.1). However, phonocardiograms require an acoustically isolated room [13, 19], and calibration is required on every new patient [13, 19, 102, 103, 104, 105]. In addition, each phonocardiogram manufacturer produces his own set of filters, making standardisation and calibration a complex problem [19, 102, 103, 105].

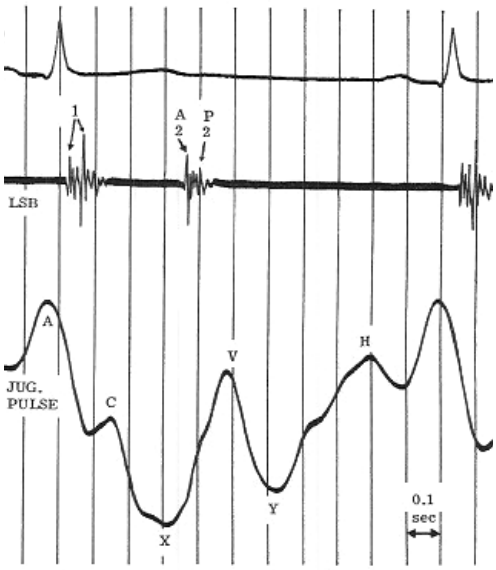


Figure 5.1: Phonocardiogram with A2 and P2 depicted in the middle trace.

Two evident advantages of the proposed approach in this chapter are the direct correlation it yields with the medical procedure and its computational simplicity. In addition to that, a threshold where an auscultation can be considered normal or hyperphonetic was also inferred. A second advantage is its low memory and low processing demands, allowing it to be easily incorporated into portable devices.

5.2 Characterisation of S2

Before going into details about the this method, a small review of some facts about the second heart sound and phonocardiographic measurements is in place in order to better understand the rationale behind the procedure proposed in this chapter. For more details, please refer to chapter 2.6.

The second heart sound is the sound generated by the closure of the semilunar valves [106]. The closure of these valves, however is usually not synchronous, causing S2 to split. The A2 component originates in the aorta and is well transmitted to the pulmonary artery, whereas the P2 component

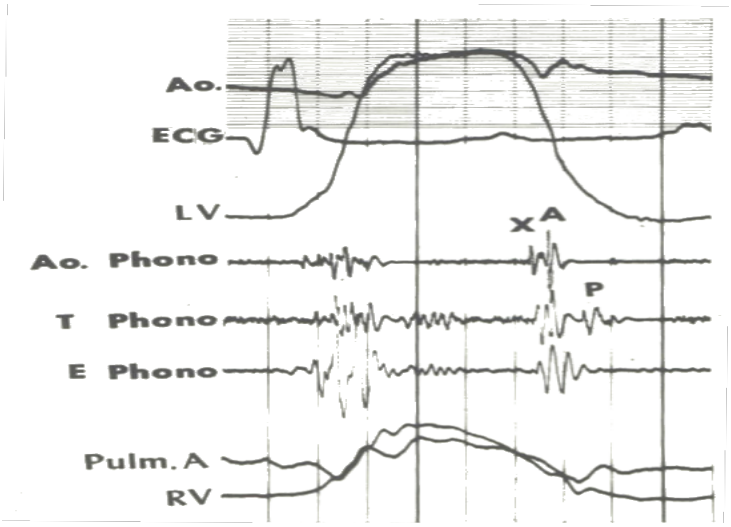


Figure 5.2: A phonocardiogram of a human subject depicting the X wave[1].

originates in the pulmonary artery and is poorly transmitted to the aorta because of its smaller magnitude and the thicker wall of the aorta. Both components are generated by the sharp wave of back blood flow produced by the valves closure.

In addition to these components, sometimes there is a third component, namely the X component preceding A2 by an average of 7 milliseconds [13, 107, 16]. This small X component can be found also in the external phonocardiogram and within the left ventricle: It may be explained by a vibration of the left ventricular wall when systolic contraction ceases and the pressure drops rapidly (Figure 5.3 and 5.2).

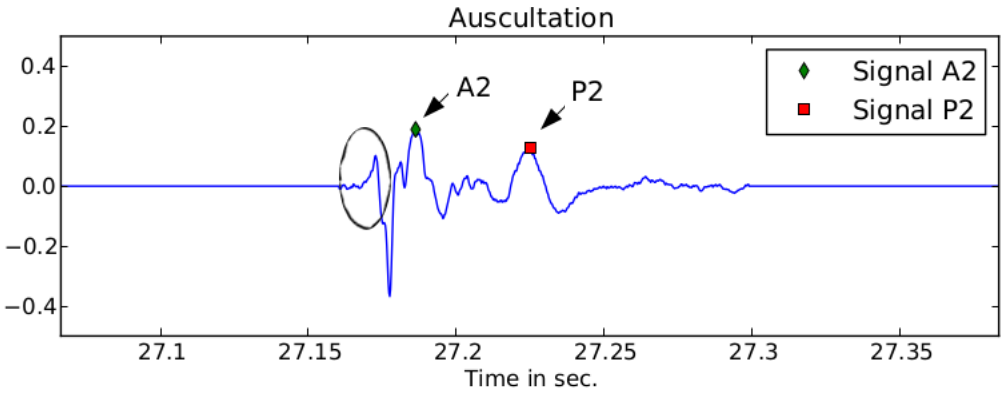


Figure 5.3: The second heart sound from an auscultation from in data set showing the X wave right in the beginning of S2 (circled), followed by A2 and P2.

Normal respiration causes A2 and P2 to be closer in expiration and more widely separated in expiration (Figure 5.4). This is produced by the delay on P2 because of the transient prolongation of right ventricle contraction, and the shortened left-ventricular systole, and earlier A2.

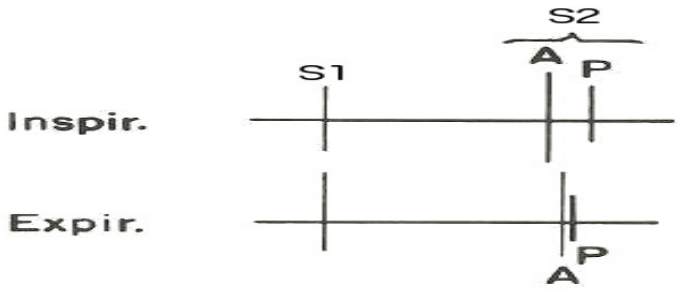


Figure 5.4: Physiological splitting of S2 (adapted from [13], p. 23).

In order to be differentiated and heard as two distinct sounds, the A2-P2 interval should be greater than 0.03 seconds. In normal adult patients, during expiration, A2-P2 is typically less than 0.03 seconds, being considered superimposed in 90% of normal persons [108, 109], therefore it may sound as a single beat to the ear. If the split is greater than 0.04 seconds in expiration, then it is usually abnormal (split in expiration is an important clue to

[1]picture adapted from [13]

abnormality). In children or adolescents, however, the splitting is audible in inspiration.

In normal individuals A2 exceeds P2 in intensity in 90% at the pulmonary auscultation site [109, 108]. The intensity of S2 is related to the inertial energy involved in deceleration of the blood volume back flow, determined by: wall elasticity, systolic runoff of preceding stroke volume and rapidity of ventricular relaxation [110].

5.2.1 Energy of S2

Since the total intensity of S2 is related to the mechanical energy involved in the generation of A2 and P2 sounds [110], this will result in the S2 waveform having energy peaks. The proposed method tries to emulate the way a clinician measures A2 and P2 in a phonocardiogram: by measuring only the positive part of peaks of the signal. In order to look for peaks on the positive side of the signal, we propose to to discard the negative part of the signal:

$$x[t] = \begin{cases} x[t] & \text{if } x[t] > 0 \\ 0 & \text{otherwise} \end{cases} \tag{5.1}$$

Then, in order to enhance the amplitude difference between the A2 and P2, we squared the signal and applied a sliding Hanning window to calculate the energy of this positive part of the signal. This will facilitate the peak selection procedure and makes your energy measure less sensitive to high frequency noise (Figure 5.5).

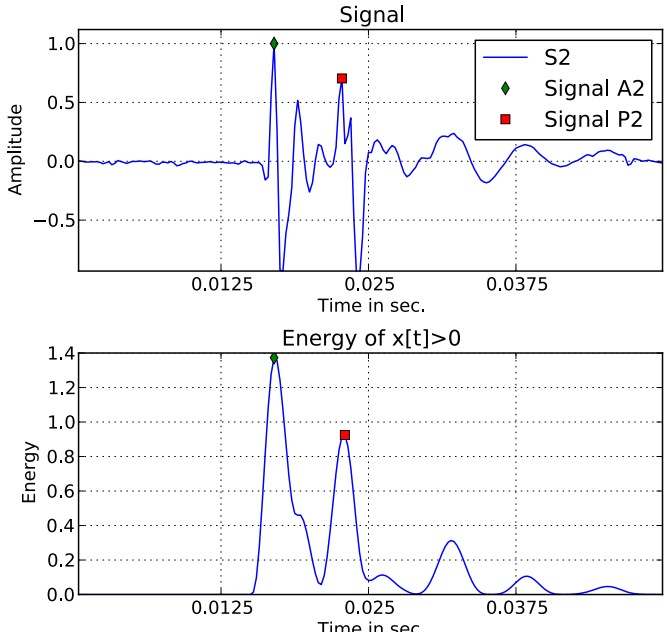


Figure 5.5: Recorded S2 (top), its calculated energy (bottom) and the detected A2 and P2 components.

5.2.2 Algorithm

The algorithm works as as follows. After calculating the energy of the positive part of the signal, we apply peak-picking to select the candidate peaks. Now their eligibility is to be verified. Since the auscultations are usually very noisy, it is common that the noise addition to the neighbourhood of a peak would create another fake peak. Therefore to rule out these high-energy peaks created by additive noise, we need a decreasing threshold so that it would ignore these noisy peaks that happen to be too close to each other, but would consider sufficiently strong peaks further away from the each other. A function that captures well this behaviour is the exponential decreasing threshold given by (5.2) and, depicted in Figure 5.6.

$$th(\Delta t) = e^{-\lambda \Delta t} \tag{5.2}$$

The threshold is parametrized by the rate of decay λ and the time difference Δt between the candidate peaks. To avoid the wrong detection of noise

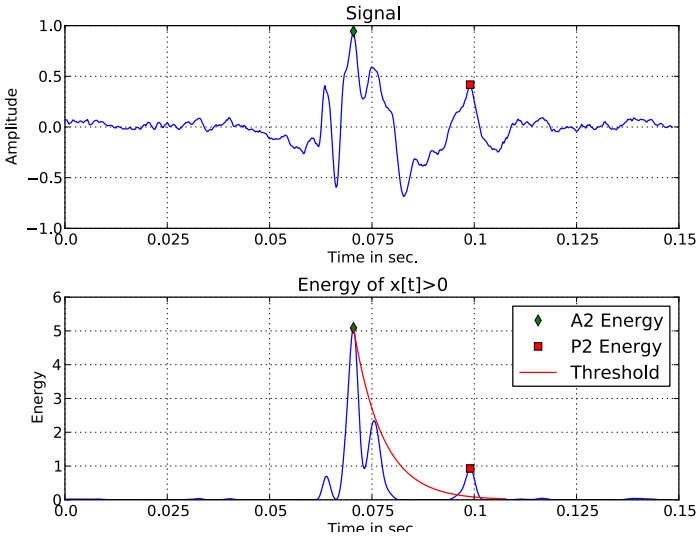


Figure 5.6: The non-linear threshold for A2 and P2 detection

or artefacts as another peak, λ is chosen empirically. The peaks are ordered in decreasing energy order. After selecting the highest energy peak ($peak_1$), we select the one with the second highest energy $peak_2$. On these peaks we perform two tests:

1. If the time difference between $peak1$ and $peak2$ is 7 milliseconds or less, then we are dealing with the initial low frequency vibrations of the aorta's leaflets: the X wave [13, 107, 16]; the earlier peak should be ignored [13, 107, 16]; and the next high-energy peak is selected as $peak1$ or $peak2$.

2. If the energy of the positive signal (Equation 5.1) is lower than the threshold (Equation 5.2), then the latter peak should be discarded and the next high-energy peak should be selected as $peak1$ or $peak2$.

These tests are repeated until we can find $peak1$ and $peak2$ that can pass both tests. If we can't find such suitable peaks, then the split is considered too small (single S2), and therefore we can't detect any component in this S2. We name A2 the peak that comes first in time. The latter peak is named P2.

The algorithm's pseudocode is outlined below:

Algorithm 5.2 Peak Detection algorithm

1. Calculate signal's positive energy:

$$x_p(t) = \begin{cases} x[t] & \text{if } x[t] > 0 \\ 0 & \text{otherwise} \end{cases}$$

$$E_+(\tau) = \int w(t+\tau) * x_p(t)^2 dt$$

where w represents the Hanning window and τ represents the shift over time.

2. detect peaks in the energy

$$candidates = list\ of\ peaks\ in\ E_+$$

3. Initialize peak search

```
Let Peaks_by_E=candidates by decreasing energy level
Let Peaks_by_t=candidates by decreasing time
```

4. Discard x-waves

```
P1 = Peaks_by_t[0]
P2 = Peaks_by_t[1]
while |time(P1)-time(P2)|$\leq 0.007$
        Remove earlier peak from Peaks_by_t and Peaks_by_E
        P1 = Peaks_by_t[0]
        P2 = Peaks_by_t[1]
endwhile
```

5. Search for valid signal peaks

```
P1 = Peaks_by_E[0]
P2 = Peaks_by_E[1]
Δt = |time(P1) − time(P2)|
while th(Δt) > Energy(P2)
        remove earlier peak from Peaks_by_t and Peaks_by_E
        P1 = Peaks_by_E[0]
        P2 = Peaks_by_E[1]
        Δt = |time(P1) − time(P2)|
endwhile
a2 = P1
p2 = P2
return a2, p2                    93
```

Here $time(P1)$ is the time of $P1$, and $th()$ is the non-linear threshold of equation 5.2

5.3 Experiments

The description of the dataset used in the experiments can be find in the Appendix A. In some auscultations, the signal's quality did change and the annotation of these heart sounds under such conditions was not done, generating periods without heart sounds (Figure 5.7).

In the following subsections we will perform experiments to indirectly validate the A2 and P2 detected: in subjects with hyperphonesis, the P2 component tends to be louder than in normal auscultations; another indirect validation used is the the analysis of the split (time difference between P2 and A2) and their expected values reported in the medical literature.

5.3.1 Normal Auscultations

In normal auscultations, we measured the amplitude values and the split (time interval between the detected P2 and A2) to compare against the literature findings.

Taking a normal auscultation as an example (Figure 5.7)

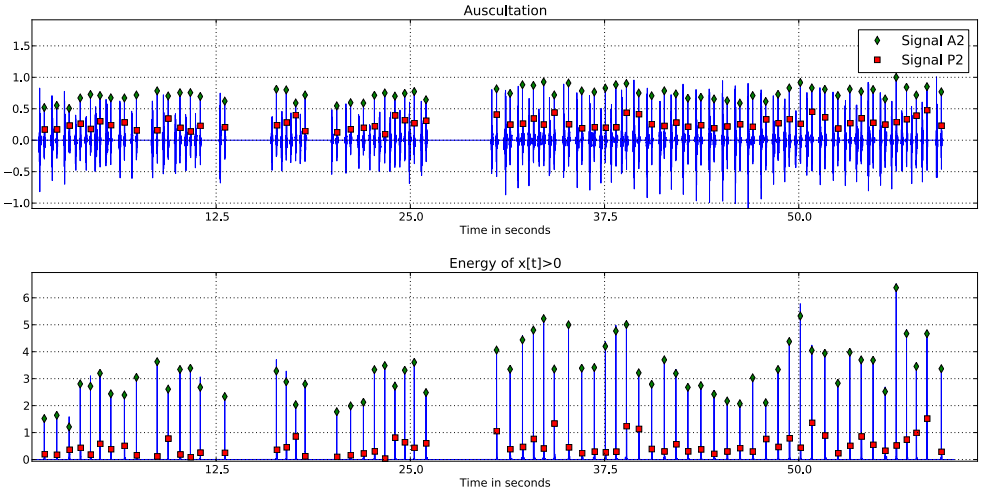


Figure 5.7: The detected A2 and P2 in a normal auscultation (top), and its energy (bottom)

On Figure 5.8 and 5.9 we can notice the variation in amplitude of A2 and

P2 through the auscultation: Though A2 has a small variation in relation to the mean, P2 on the other hand varies greatly in amplitude relative to its mean. This could be due to the changes in pressure in the Pulmonary Artery [13, 19, 100, 110].

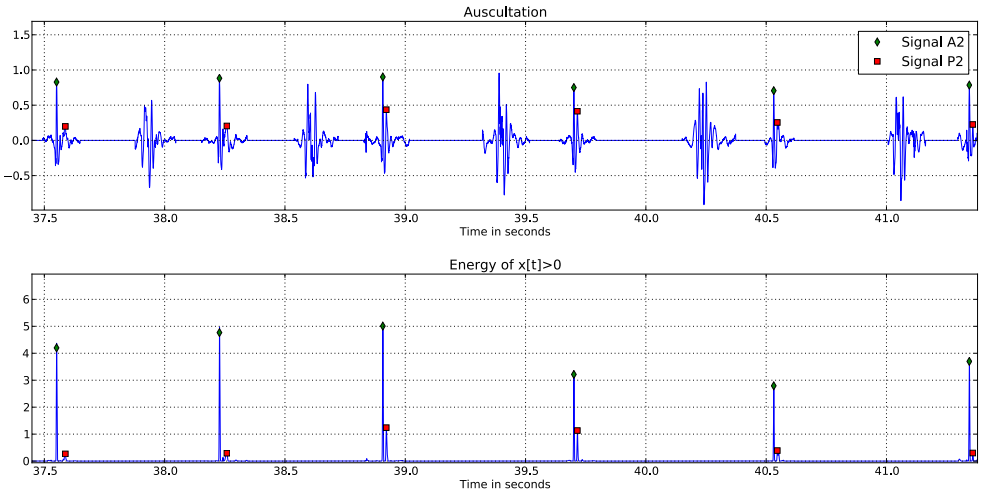


Figure 5.8: Variation in amplitude of A2 and P2: notice how P2 (red square) changes greatly during an auscultation, and A2 (green diamond) changes much less than P2

During the auscultations, we could not record the respiratory cycle, but since the patients where in relatively normal breathing, the respiratory cycle was estimated by calculating all the (P2-A2 split time) and taking the median (Figure 5.10) time.

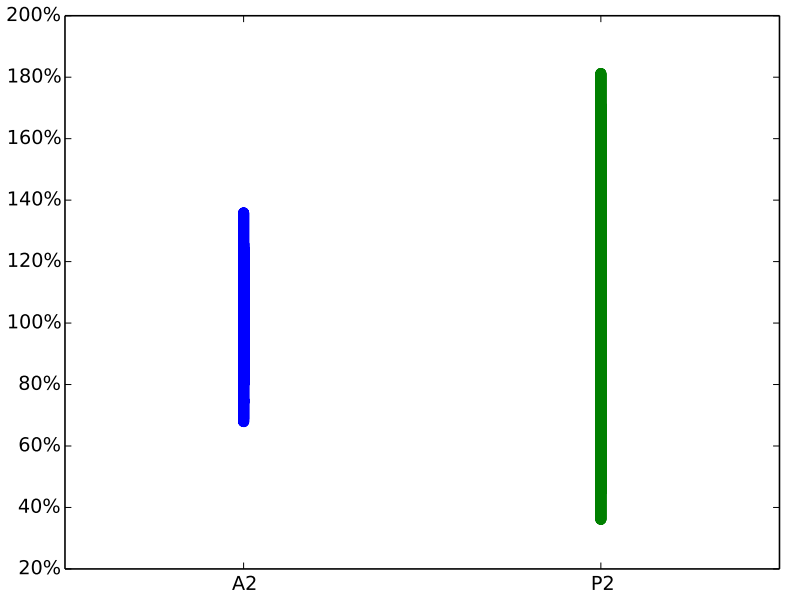


Figure 5.9: The amplitudes of each A2 and P2 relative to the mean (100%)

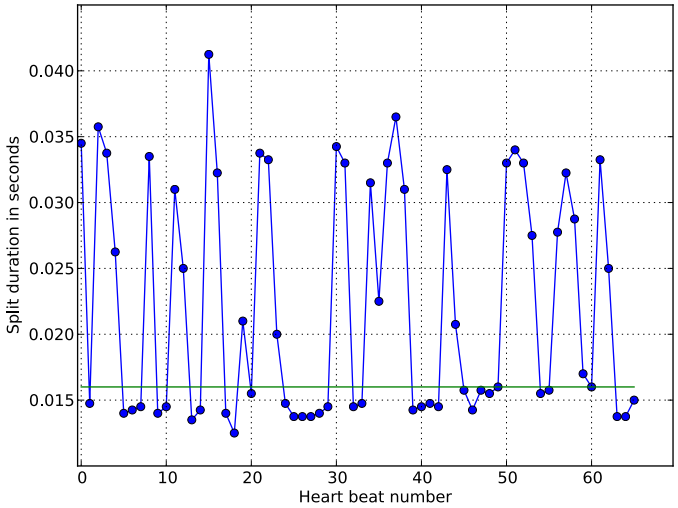


Figure 5.10: P2-A2 split time recorded on every S2 of a normal auscultation. The green horizontal line represents the median. Values below this line expected to be part of the expiration cycle whilst values above the line are considered part of the inspiration cycle.

As we can see in Table 5.1, the mean inspiratory split of 0.0302 seconds is within the literature range of a mean inspiratory split between 0.03 to 0.04 seconds. The mean expiratory value of 0.0145 seconds, with a minimum value of 0.0125 seconds and maximum value of 0.016 seconds means that on expiration, S2 is heard as a 'single' one and no split is heard, again, in accordance with the literature.

Estimated Inspiration

Mean split	0.0302
Median split	0.0323
Min split	0.0170
Max split	0.0412
Number of S2	32

a)

Estimated Expiration

Mean split	0.0145
Median split	0.0145
Min split	0.0125
Max split	0.016
Number of S2	34

b)

Table 5.1: Values of the split time (P2-A2) between inspiration a) and expiration b). All values are in seconds

5.3.2 Hyperphonetic Auscultations

The same analysis was made on hyperphonetic auscultations (auscultations where the P2 component is louder than the normal case): we measured the amplitude values and the split (time interval between the detected P2 and A2) to compare against the literature findings. A hyperphonetic auscultation can be seen at Figure 5.11:

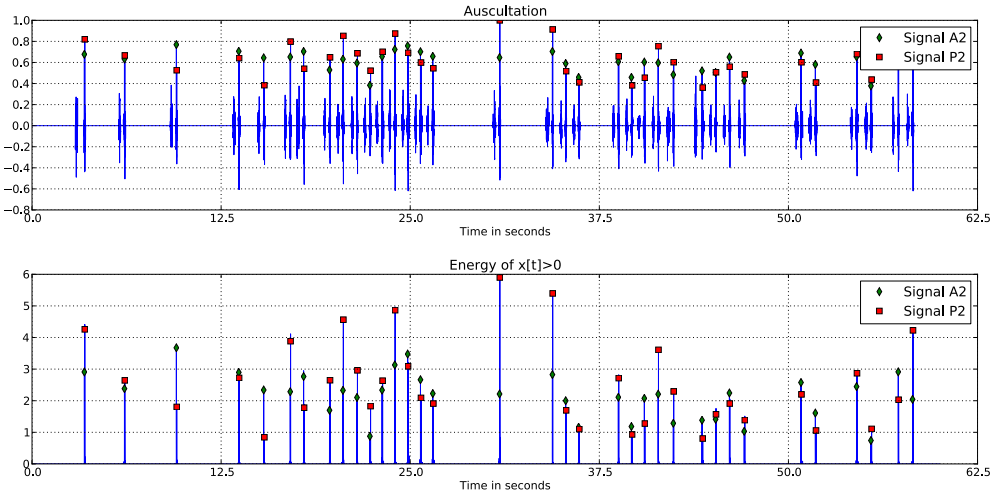


Figure 5.11: The detected A2 and P2 in a hyperphonetic auscultation (top), and its energy (bottom)

On Figure 5.12 we can notice that P2's amplitude is greater in overall if compared to Figure 5.8. Another point worth noting is the variation of amplitude and split of S2: this shows the influence of respiration on the split and amplitude of both components.

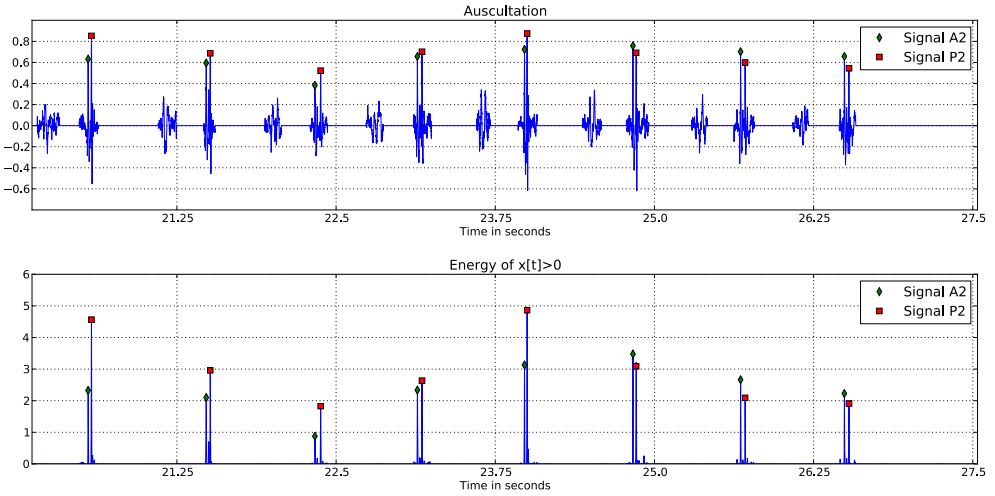


Figure 5.12: Variation in amplitude of A2 and P2: notice how the amplitude of P2 (red square) is generally greater than P2 on Figure 5.8, sometimes it reaches a slightly greater amplitude than A2, as it can be seen at 21.5 seconds.

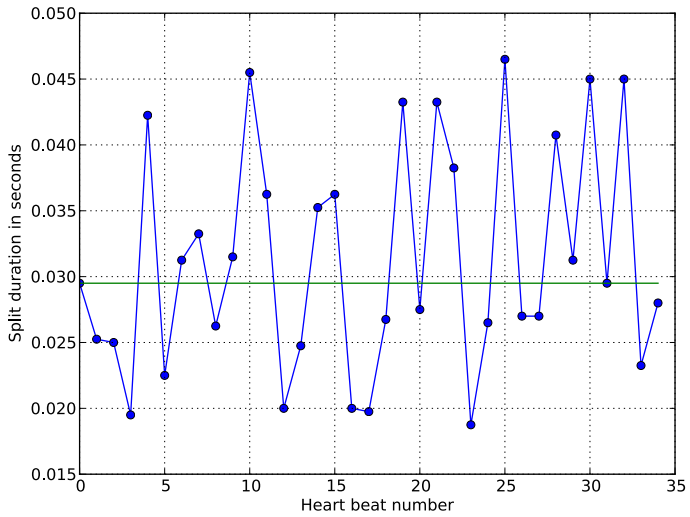


Figure 5.13: P2-A2 split time recorded on every S2 of a hyperphonetic auscultation. The green horizontal line represents the median.

Figure 5.13 shows the beat by beat values of the split of S2. We can notice the larger split on estimated inspiration (points above the green line). Also, compared with the normal case (Figure 5.10), the inspiratory values are higher, varies more, and the same happens with the expiratory values.

Table 5.2 contains the values for the estimated inspiration and estimated expiration. Although these values are overall higher than the normal auscultation (Table 5.1), the estimated inspiration mean values are within the upper literature limit for normality [108, 109, 13], but the estimated expiration mean values are higher than normal: since the split is greater than 0.020 seconds, an experienced cardiologist may hear a discrete split of the second heart sound.

Estimated Inspiration

Mean split	0.0390
Median split	0.0395
Min split	0.0312
Max split	0.0465
Number of S2	16

a)

Estimated Expiration

Mean split	0.0245
Median split	0.0252
Min split	0.0187
Max split	0.0295
Number of S2	19

b)

Table 5.2: Values of the split time (P2-A2) between inspiration a) and expiration b). All values are in seconds

5.3.3 S2 components by auscultation group

As seen in section 2.3.2, and in the section 5.2, besides the split information, another important feature on A2 and P2 components is their amplitude: a number of papers have shown that the amplitude of A2 and P2 are correlated to the pressure in their respective arteries [108, 16, 19, 13, 15, 111].

This information about the amplitude of these components is paramount on the diagnosis of heart conditions. Such as, the pulmonary artery hypertension (PAP). As described on sections 4.4.3 and 5.2, the main indicator of PAP is an hyperphonetic P2.

The hyperphonesis of P2 is not only defined by the absolute amplitude of P2. The clinician looks at the overall amplitude of the second heart sound, and apply his experience and common sense to judge if the given P2 component is hyperphonetic or normal. Besides, the amplitude of A2 and P2 also depends on a number of factors, such as: placement of stethoscope, patient's body mass index, etc. Therefore, the absolute amplitude of A2 and P2 may vary in such a way that measuring P2 alone is not enough to quantify P2 hyperphonesis. (Figure 5.14).

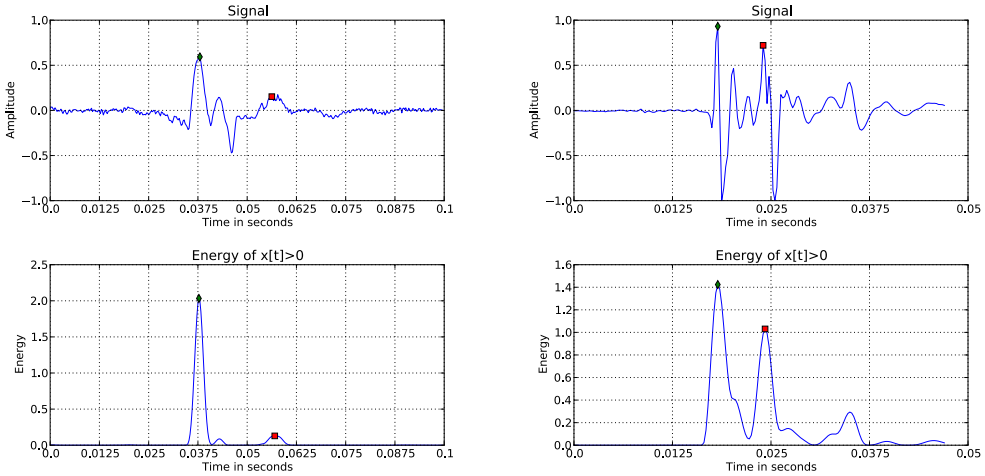


Figure 5.14: Detected A2 (diamond) and P2 (square): in a normal auscultation (top left: signal, bottom left: energy), and in a hyperphonetic auscultation (top right: signal, bottom right: energy).

In this experiment, we looked into the amplitudes of A2 and P2 of auscultations diagnosed as normal and hyperphonetic in order to build a feature that separates normal and hyperphonetic auscultations. We intend to turn the clinician's subjective judgement of hyperphonesis into an objective one.

We found out that, as expected, the absolute value value of the amplitude of P2 is not a feature capable of separating normal from hyperphonetic auscultations, but the amplitude of P2 divided the amplitude of A2 do provide us with a feature that is good enough to separate these two groups. For each auscultation, in order to have one feature value that is more resilient to outliers, we calculate the median $P2/A2$, as we can see in Equation 5.3, where $x[P2_1]$ and $x[A2_i]$ is the amplitude of the first P2 and A2 in the $i-th$ S2 of an auscultation with n second heart sounds. The groups can be seen at Figure 5.15.

$$median(\frac{x[P2_1]}{x[A2_1]}, \frac{x[P2_2]}{x[A2_2]}, \cdots, \frac{x[P2_{n-1}]}{x[A2_{n-1}]}, \frac{x[P2_n]}{x[A2_n]}) \tag{5.3}$$

	Normal	Hyperphonetic
n	26	11
Mean ranks	14.3	30.2
U	266	20

Table 5.3: Mann-whitney U test

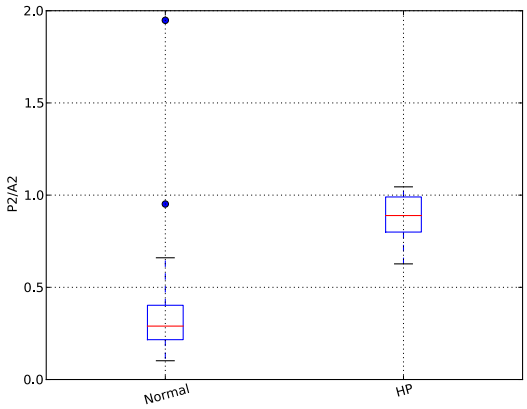


Figure 5.15: Separation between normal and hyperphonetic S2. The outlier is due to an auscultation were the detection of A2 and P2 failed due to high level of noise.

The Mann-Whitney U test [112] reported these two groups statistically different on the median $P2/A2$ of each auscultation with $P < 0.0001$ and mean ranks of 14.3 and 30.2 for the normal and hyperphonetic groups respectively (Table 5.3). It is also observed that with a threshold of $\frac{P2}{A2} = 0.7$, the two groups can be separated with 89.7% accuracy, 75% precision and balanced accuracy of 92.59%.

5.4 Conclusions

In this chapter, we assumed that A2 and P2 are energy peaks produced by the positive values of the recorded second heart sound. This method mimics the analysis and identification of A2 and P2 performed by the clinician when using a phonocardiogram. Although at first glance this may seem an

oversimplification of the components of the second heart sound, the measures produced by this method have direct clinical usage: they are the automatisation of the measures of the phonocardiogram, therefore, the clinical knowledge derived from phonocardiograms are directly applicable to these measurements.

We performed two experiments to indirectly validate the detected A2 and P2. In the first experiment we analysed the time-difference (split) between P2 and A2 in a normal subject and in a subject with hyperphonetic auscultation. The respiratory cycle was estimated by calculating the median of the splits and assuming the expiration to be the splits smaller than this median, since according to the literature, the splits are shorter during the expiratory phase, if compared to the splits on the inspiratory phase. In both subjects, the values found are in accordance to the literature. However, in some auscultations, P2 was too faint and was wrongly detected (Figure 5.16). This may occasionally happens, but should not change the overall analysis, since the auscultations used in our dataset where also the ones a clinician hear and give a diagnosis.

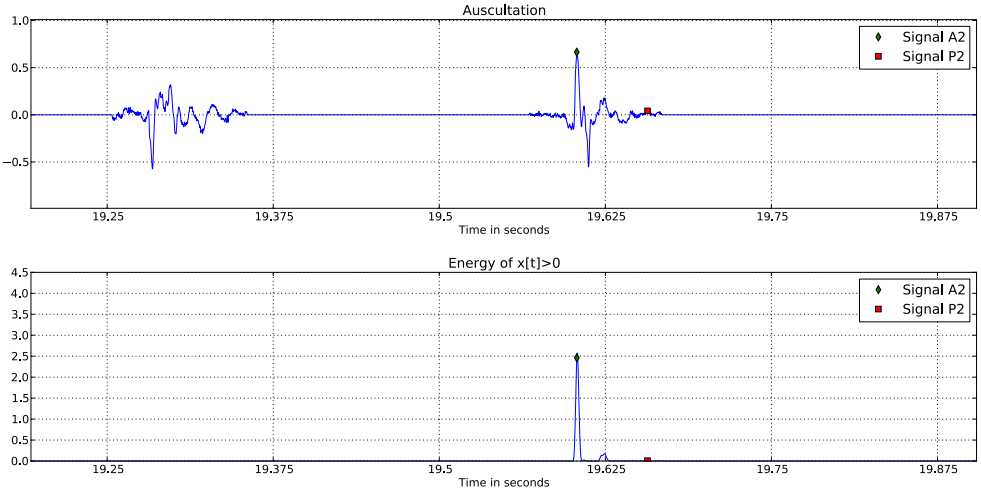


Figure 5.16: Miss-identification of P2 due to very low energy intensity.

In the second experiment, we used the amplitude information to separate the normal auscultations from the hyperphonetic ones. In comparison with the matching pursuit technique, the energy technique had the best performance in separating these groups.

The downside of the energy method lies in the peak selection procedure, specially in cases where both A2 and P2 have similar energy and lasts longer than usual. This results in shift information error (Figure 5.17). This kind of situation is more likely to happen if S2 is buried in noise or murmur. This situation, however, can be improved, by using consistency analysis and/or application of a strong filter. A clinician, however, would not be able to listen to the components of the second heart sound in this situation, rendering the validation of this correction quite difficult.

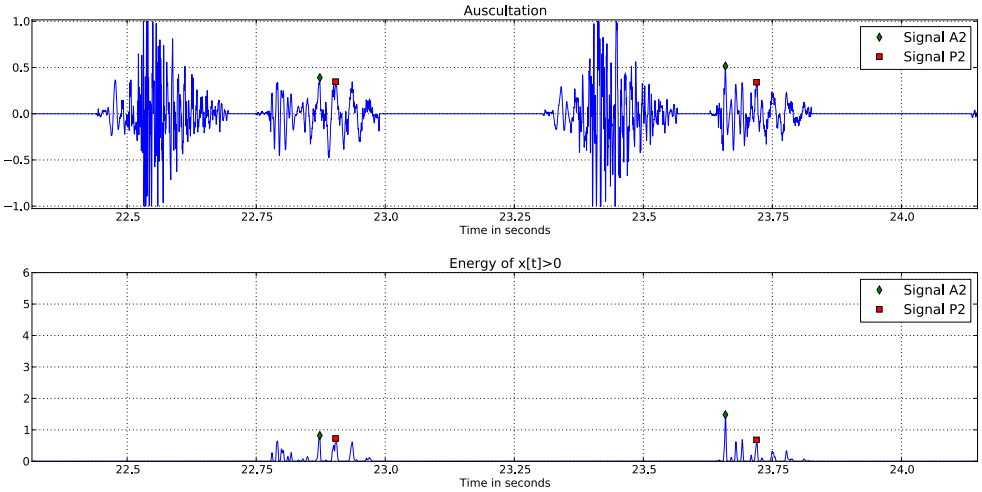


Figure 5.17: Incorrectly detected A2: on the first S2, the wrongly detected A2 component has a slightly greater amplitude than the real one, which comes before - compare with the following S2, where the components where detected correctly.

Another problem with the energy method is the occurrence of miss-identifications of A2 and P2 particularly when auscultations have too low volume and therefore levels of noise are comparable to the heartbeat itself. There are two possible solutions: creating an auditory threshold where detected peaks below that threshold are ignored, or using consistency checks and filtering, which would diminish the problem, however, this solution is beyond the scope of this work and should be explored in the future.

5.5 Publications

1. F. Hedayioglu; M. G. Jafari; S. S. Mattos; M. D. Plumbley,; M. T. Coimbra, "An Exploratory Study On The Segmentation Of The Second Heart Sound", to be submitted

Chapter 6

Synthetic auscultation generator

6.1 Introduction

During the development of this thesis, specifically during the tests of the ICA algorithms in chapter 3, it was clear that some way of validating the underlying components of S2 was necessary. It was particularly difficult to test the validity of our results using the ICA techniques: how do we know that the obtained signals are really the A2 and P2 waves and not something else that looks like them? More knowledge about the structure of A2 and P2 was necessary, and from this initial necessity, the physiologically inspired atoms described on chapter 4 were developed.

In this chapter we use the theory and techniques developed on chapter 4 as basis to create a heart sound simulator that has as its main emphasis, the capability of generating synthetic auscultations where S2 is not only realistic but its clinically meaningful parameters are customisable (e.g. split variation, A2 and P2 amplitude). In addition to a customisable S2, this simulator can also generate auscultations with realistic systolic and diastolic times. Besides generating an audio file, an annotation file compatible with the open source audio editor Audacity [113] is also generated.

6.2 State of the art

One of the most important and difficult parts of the auscultation process to a clinician is the training and the subjectivity aspects related to auscultation interpretation. The heart sound simulators where first used as a tool in the teaching of auscultation to medicine students. Due to this market opportunity, there is a good number of heart sound simulators in the market. There is not, however, a rich literature in this area.

Most commercial simulators use recorded heart sounds together with some audio processing techniques for changing some features of these sounds such as intensity and speed.

Among the papers describing the generation of auscultations, Almasi in [114], uses a model based in dynamic ECG generation [115], where three differential equations are used to generate different morphologies (systolic and diastolic times) of normal auscultations. The first and second heart sound are modelled by using two Gabor kernels per heart sound. These heart sounds are repeated throughout the auscultation and no information about how their parameters are calculated is given.

In [89], Tran et al. developed a simulator with emphasis on the generation of customisable first and second heart sounds. It uses linear chirp signals to model the heart sounds components. In the first heart sound, it models them M (mitral), T (tricuspid) components, and in the second heart sound, it models the A2 and P2 components. In total, 14 parameters are used to generate each heart sound. The authors, however, do not mention how the systolic or diastolic times are calculated.

6.3 Synthesis of auscultation

6.3.1 First heart sound generation

To represent the first heart sound, we used the works of Damin Chen [116, 117], where they created a parametric model of S1 that is composed by the vibrations of two valves and the myocardium.

The valves are modelled by exponentially decaying sinusoids [118, 91]:

$$S_v(t) = \sum_{i=1}^{2} A_i e^{-k_i t} \sin(2\pi f_i t + \phi_i) \tag{6.1}$$

where A_i is the amplitude, k_i is the damping factor, f_i is the frequency, andϕ_i is the phase.

The myocardial component is represented by:

$$S_m(t) = A_m(t) \sin(2\pi(f_o + f_m(t)) + \phi_m(t)) \tag{6.2}$$

Where A_m is the amplitude modulating function, f_o is the carrier frequency, $f_m(t)$ is the frequency modulating wave and $\phi_m(t)$ is the phase function.

The whole S1 is represented as:

$$S1(t) = S_m(t) + \begin{cases} 0 & if\ 0 < t \leq t_0 \\ S_v(t - t_0) & \text{otherwise} \end{cases} \tag{6.3}$$

Where $t_0 = 0.01$ is the time between the start of the myocardial component and the start of the S_v components.

The parameters used in the simulator where calculated by [107, 119] and used by Chen in [116, 117]:

i	f_i	A_i	ϕ_i	k_i
1	50Hz	1	$-\pi$	60
2	150Hz	0.5	$-\pi$	60

Table 6.1: Sv parameters

$$A_m = \begin{cases} 0.275(1.1 - 0.9\cos(1.4\pi t)) & 0 < t \leq 0.012 \\ 0.55 & 0.012 < t \leq 0.03 \\ 0.275(1 - \cos(2\pi t)) & 0.03 < t \leq 0.06 \\ 0 & t > 0.06 \end{cases} \tag{6.4}$$

$$\begin{aligned} f_m(t) = -40\cos(2\pi t),\ f_0 = 60Hz,\ \phi_m(t) = 0 \qquad & 0 < t \leq 0.03 \\ f_m(t) = 0,\ f_0(t) = 100Hz,\ \phi_m(t) = -0.4\pi \qquad & 0.03 < t < 0.06 \end{aligned} \tag{6.5}$$

These parameters are calculated by Chen from real auscultations and the functions are parametric, which creates the possibility of creating more realistic S1 sounds in the future. In this simulator S1 only changes in amplitude in accordance to the second heart sound.

6.3.2 Second heart sound generation

For the generation of the second heart sound, we used the atoms described in chapter 4. Their parameters were calculated by performing matching pursuit in auscultations with good audio quality during the whole recording. In total 7 normal auscultations and 4 hyperphonetic auscultations were selected. Then, matching pursuit was performed on them and their parameters saved in a file. These parameters are later used by the heart sound simulator as basis to, according to the user input, generate customisable second heart sounds.

6.3.3 Systole and diastole calculation

There are three types of systoles and diastoles: Electric, ventricular and atrial [24]. In this chapter we are dealing only on the ventricular systolic and diastolic times. We can calculate the instantaneous heart rate of the $i-th$ heart cycle (ihr) using Equation 6.6. The auscultations used in this dataset had a mean instantaneous heart rate ranging from 62 bpm to 98 bpm, with an average of 80 bpm. The minimum instantaneous heart rate of 55 bpm and maximum instantaneous heart rate of 109 bpm.

$$ihr(i) = \frac{60}{systole_i + diastole_i} \tag{6.6}$$

Although [120] suggests a linear equation to calculate systolic times from heart rate, from a visual inspection, our data suggests that an exponential

relationship between diastolic time and heart rate is a better fit (Figure 6.1).

We then tested two functions in order to fit this data: exponential and linear

$$HR(d) = ad + b \tag{6.7}$$

$$HR(d) = b(d)^a \tag{6.8}$$

Where d is the diastole in seconds, $HR(d)$ is the heart rate in beats per minutes as function of the diastolic time d. The systole is calculated by $S(d) = \frac{60}{HR(d)} - d$.

Analysing the linear case first, we define D as the observed diastolic time, X as the function parameters, and HR as as the observed heart rate:

$$D = \begin{bmatrix} d1 & 1 \\ d2 & 1 \\ \vdots & \vdots \\ dn & 1 \end{bmatrix} \quad X = \begin{bmatrix} a \\ b \end{bmatrix} \quad HR = \begin{bmatrix} hr_1 \\ hr_2 \\ \vdots \\ hr_n \end{bmatrix} \tag{6.9}$$

Where dn is the n-th observed diastolic time and hr_n is the n-th instantaneous heart rate. The problem of minimising the sum of squared error (Equation 6.10) becomes one where we want to find the optimum X (Equation 6.11.

$$E = ||DX - HR||^2 \tag{6.10}$$

$$X = (D^T D)^{-1} D^T HR \tag{6.11}$$

To find the optimum a, b on equation 6.8, however a linearisation process must be performed:

$$D_{log} = \begin{bmatrix} \log(d1) & 1 \\ \log(d2) & 1 \\ \vdots & \vdots \\ \log(dn) & 1 \end{bmatrix} \quad X_{log} = \begin{bmatrix} a \\ \log(b) \end{bmatrix} \quad HR_{log} = \begin{bmatrix} \log(hr_1) \\ \log(hr_2) \\ \vdots \\ \log(hr_n) \end{bmatrix} \tag{6.12}$$

then we need to find the optimum X:

$$X = (D_{log}^T D)^{-1} D_{log}^T HR_{log} \tag{6.13}$$

An example of the linear and exponential fit functions on one auscultation is given in Figure 6.1. As we can see in Table 6.2, the exponential function has a lower mean squared error, although, in some cases it is similar to the linear function.

Auscultation	Number of heart cycles	MSE(exponential)	MSE(linear)
133	65	0.9544	1.0411
109	69	0.6349	0.7859
124	57	19.6016	19.1021
114	62	6.4597	6.3914
113	59	2.2782	2.9792
111	43	1.2179	1.2965
115	54	0.4805	1.1389
131	49	0.1617	0.2875
112	52	20.3785	18.9880
105	83	1.9416	2.3981
137	49	1.9443	2.0486
108	47	1.9450	2.1203

Table 6.2: The mean squared errors between the linear and exponential functions

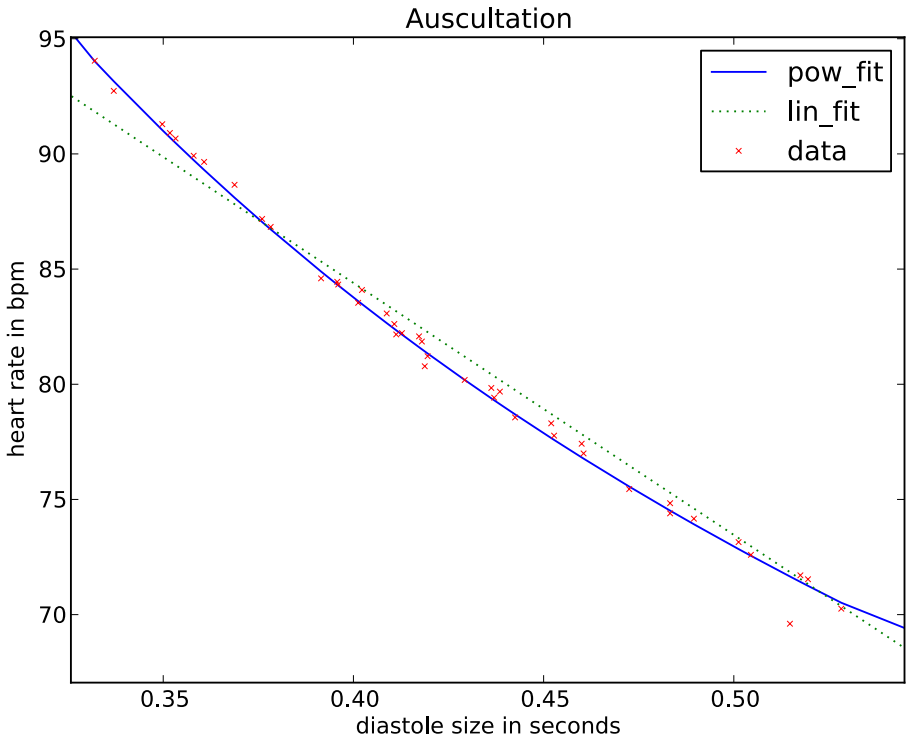


Figure 6.1: Here we can see the two fit function used in a normal auscultation: lin_fit is the function on equation 6.7, where pow_fit is the exponential function on equation 6.8. Although the difference between them is relatively small, the exponential function clearly provides a better fit.

6.3.4 Generating an auscultation

To generate an auscultation, the first step is to use a real auscultation as a 'model' auscultation. Suppose this model auscultation has n annotated S2. From this model, it will be extracted:

1. Atoms parameters: $s_a, u_a, f_{oa}, \Delta f_a$ for the first atom, and $s_p, u_p, f_{op}, \Delta f_p$ for each one of the n annotated S2.

2. All instantaneous heart rates ihr (Equation 6.6).

3. The parameters for equation 6.8

The second step is to get the parameters for the heart sounds to be produced. In our model, we do not change the first heart sound, since it is beyond the scope of this thesis. We use the parameters from the works in [116, 117].

The atoms used for the generated auscultations will have the same parameters as the ones obtained by the matching pursuit on the selected model auscultation. The user has the freedom to alter the parameters s_a, s_p and u_p to produce a proportional amplitude change and increase or decrease in the splits in comparison to the original recording.

If the generated auscultation needs more S2 than the available n, the vector of atom parameters will be read in reversal order (e.g.. $0, 1, 2, ..., n-2, n-1, n-2, ..., 1, 0, 1, 2...$). This is done by:

$$index = |x \mod (N-1) - [x \mod 2(N-1) - x \mod (N-1)]| \quad (6.14)$$

Where $x \mod y$ is the remainder of x/y.

The third step is to calculate the size of the systoles and diastoles. During an auscultation, the instantaneous heart frequency changes, as we can see in Figure 6.1.

Looking at the heart rate variation of the i-th heart cycle is defined by:

$$\Delta_{hr}(i) = hr(i) - hr(i-1) \quad (6.15)$$

where $hr(i)$ is the observed heart rate of the i-th heart cycle, we can see that the instantaneous heart rate variation on a beat by beat basis is small,

smooth and independent of the mean heart frequency (Table 6.3).

Auscultation	$mean(hr)$	$mean(\Delta_{hr})$	$std(\Delta_{hr})$
131	78.2355	0.1218	5.0851
124	69.2848	0.0129	10.1821
111	83.2123	-0.1053	2.7264
114	66.6770	0.0089	6.0314
137	84.5143	0.0965	5.3136
210	81.7844	0.3061	4.9001
112	71.9386	0.0212	5.3857
142	83.2452	-0.0045	2.5998
129	84.7339	-0.4331	7.6970
140	70.2664	-0.0403	5.0126

Table 6.3: Variation of heart rate on different auscultations.

This small beat by beat variation was implemented in our simulator by calculating all instantaneous heart rates of the model auscultation, and then subtracting the average. This leaves us with just these small fluctuations of the heart rate throughout the auscultation. This vector is then added to the mean heart rate frequency informed by the user. This way we can have some variation on the heart rate during the generated auscultation. If the vector is too small to produce an auscultation of the desired duration, then the vector will be read in reversed order following Equation 6.14.

6.4 Experiments

Here we show the interface and output of the simulator. The simulator was made in such a way that, being supplied by an auscultation's systoles and diastoles times, and annotated first and second heart sounds, it can extract automatically all the other parameters in order to use this auscultation as a model to generate new synthetic auscultations in one auscultation site,

together with their respective annotations.

6.4.1 Auscultations generation

In Figure 6.2, we can see the GUI of the simulator. The input fields are as follows:

a Duration of the generated auscultation in seconds

b Auscultation's approximated average heart rate: varies between 60 bpm to 99 bpm

c First atom's amplitude factor: This parameter is multiplied by the A2's atom's amplitude s_a: accept values between 0.1 to 2.0

d Seconds atom's amplitude factor: This parameter is multiplied by the P2's atom's amplitude s_p: accept values between 0.1 to 2.0

e Split variation factor: specify how bigger or shorter the split would be compared to the original auscultation: values range between 0.1 to 2.0

f Auscultation to be selected as model: it will be used to extract the parameters to generate the baseline atoms for S2, heart rate variation, and derive the function for mapping heart rate into diastole time. Currently with 7 normal auscultation and 4 hyperphonetic ones.

g Display the annotations on the generated auscultation: S1, S2, A2 and P2.

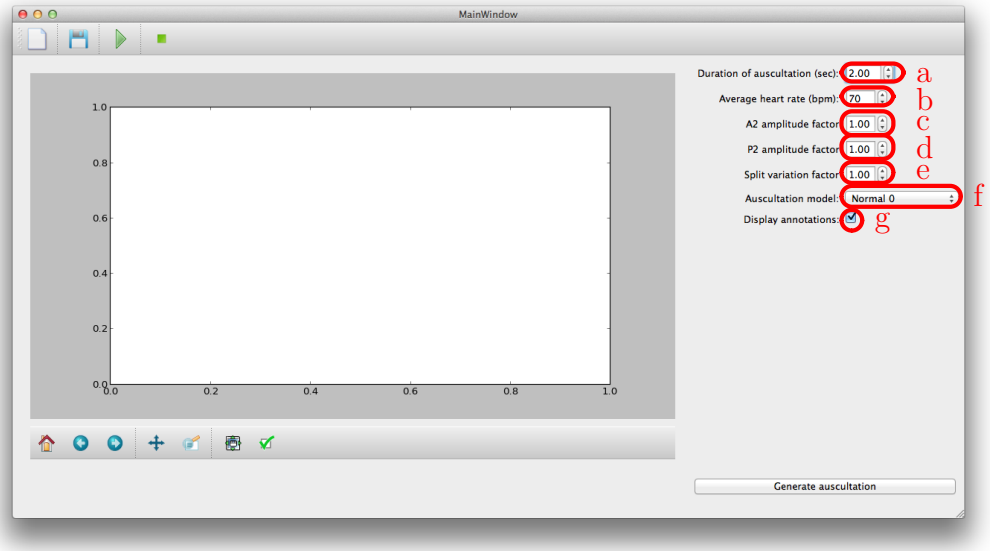
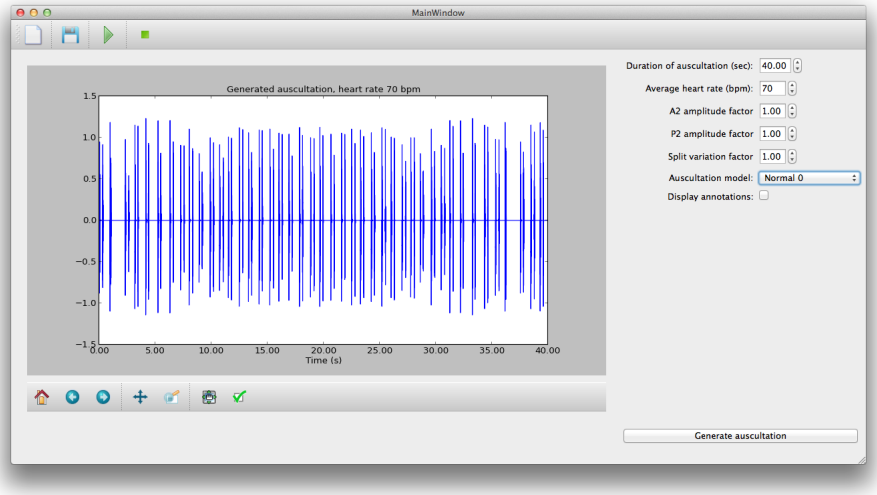


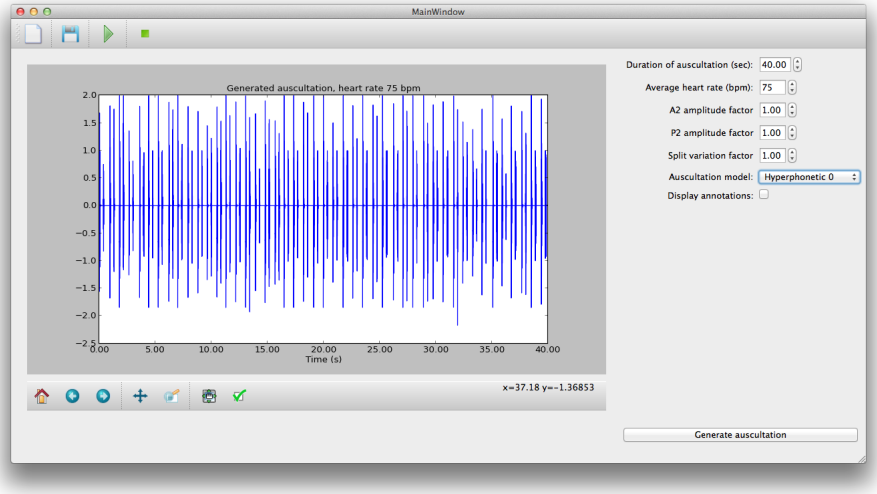
Figure 6.2: The GUI of the simulator with the input fields.

The white area displays the generated auscultation and allows the user to change its zoom, pan, in a way similar to the Audacity visual editor[113].

In Figure 6.3a and Figure 6.3b, we can see the amplitude variation throughout the auscultation in a normal and hyperphonetic auscultation, respectively. Since the amplitude and the split vectors are read in the same order, they are in sync with the respiration cycle. However, the volume of the chest in a given respiratory cycle is rarely the same. This will make the heart sound have small differences in their amplitude, caused by this change in the transmission medium. We accounted for this effect by introducing a small random variation (between 0% and 0.1% of the amplitude) in the amplitude of S1 and S2.



(a) Normal auscultation

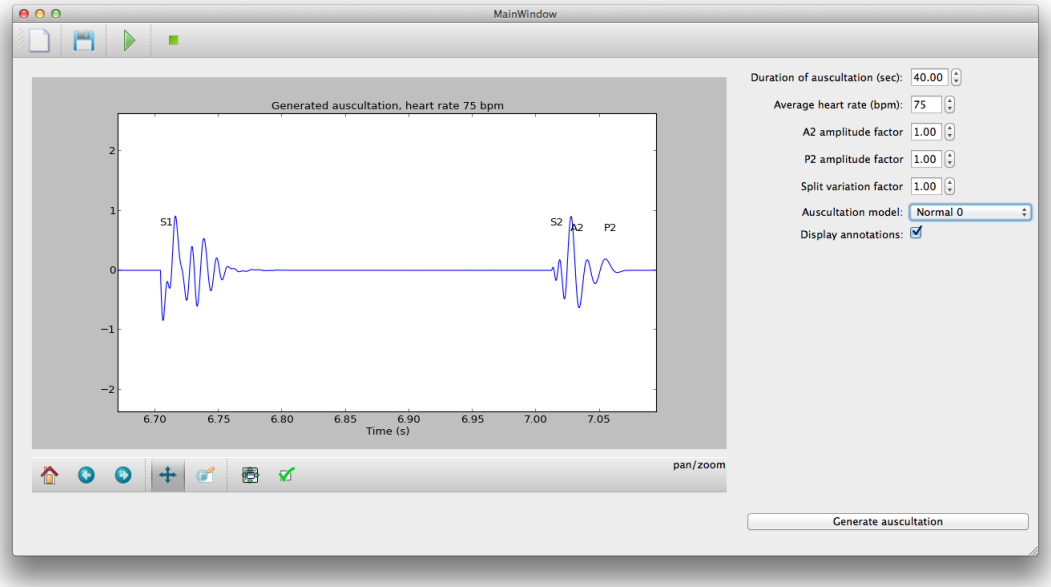


(b) A hyperphonetic auscultation

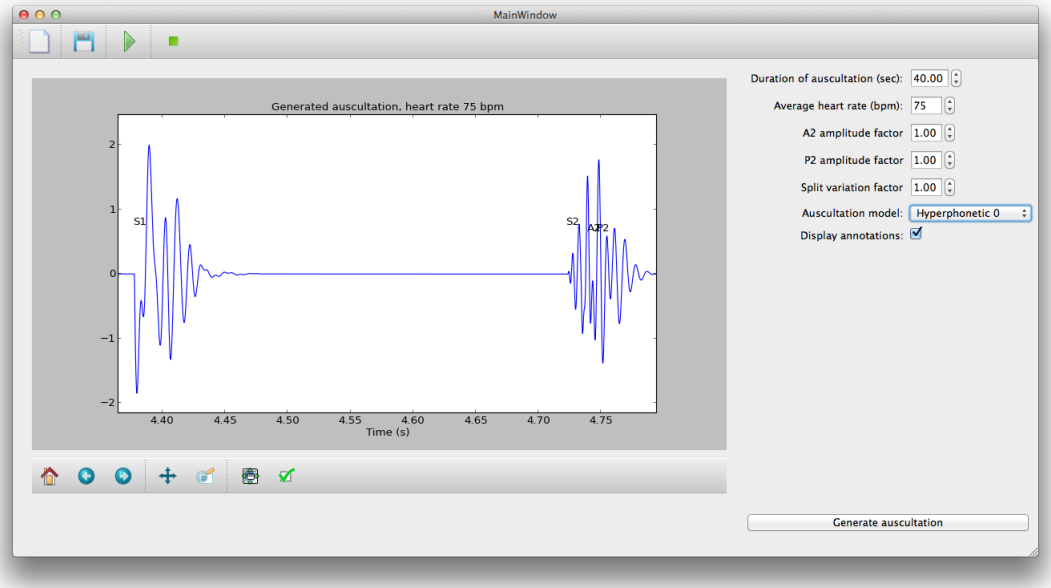
Figure 6.3

We can see the first and second heart sound in Figures 6.4a and 6.4b. Since these atoms parameters are calculated based on the matching pursuit algorithm described on chapter 4, they are closely related to the real second heart sounds present in the original auscultation. The visual annotation of A2 and P2, however is related to where the atoms begin, not where the clinician would annotate them, which would be the same as the energy segmentation

method, from chapter 5 would detect.



(a) Normal auscultation, with S1, S2 and A2 and P2 atoms annotated.



(b) An hyperphonetic auscultation with zoom on the heart sounds and their annotations.

Figure 6.4

One of the great strengths of this heart sound simulator is the capability of customisation: the second heart sound can be changed to exhibit a determined feature: may it be a frequency range or its position in the second heart sound. This was paramount for us when generating a 'custom' second heart sound to test algorithms. One interesting experiment we performed was the transformation of a normal auscultation into a hyperphonetic one: this was achieved by increasing the P2 amplitude factor and the split factor, as we can see in Figure 6.5 (compare with S2 in a hyperphonetic auscultation in Figure 6.4b or even with Figure 5.14). This is confirmed by the energy method of chapter 5 reported an $A2/P2 = 0.536$ for the normal auscultation, where the auscultation made hyperphonetic reported an $A2/P2 = 1.131$, a value within the range of hyperphonetic class. We used the method of chapter 4 to generate a scatter plot with the two synthetic auscultations (Figure 6.6). It can be seen that the auscultation generated by the normal model is correctly localised in the lower left cluster that represents normal auscultations. The upper right cluster represents the hyperphonetic auscultations and has the auscultation generated by the same normal model but with the amplitude and split parameters changed to make it an hyperphonetic auscultation.

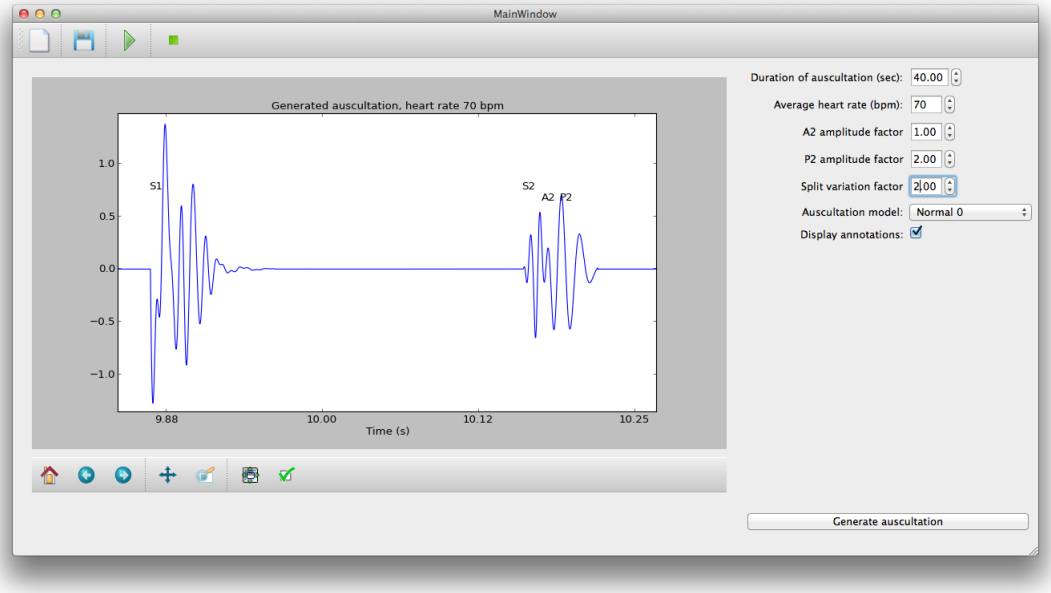


Figure 6.5: Normal auscultation with P2 and Split factor changed in order to make it hyperphonetic.

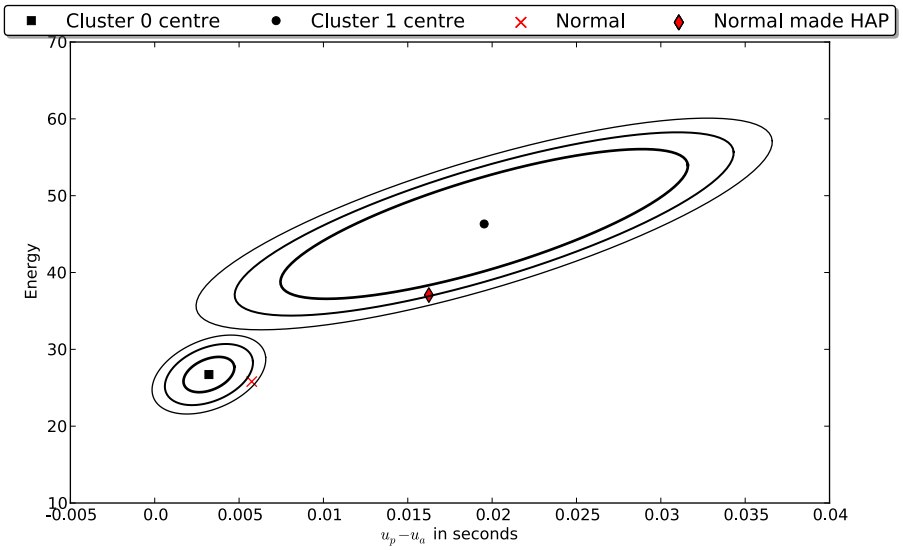


Figure 6.6: Scatter plot with the two synthetic auscultations: the lower left cluster is the normal cluster; and the hyperphonetic cluster is located upper right

In Figure 6.7, we can see the differences between the model auscultation and its annotation and the generated auscultation and its annotation: The original auscultation have an average heart rate of 81 bpm, while the generated one has 87 bpm. This change in the heart rate still maintained the natural features of the synthetic heart sound, and, allied with the atoms produced by the matching pursuit method, the synthetic auscultation reproduces the main features of the second heart sound (as can be seen in the same Figure).

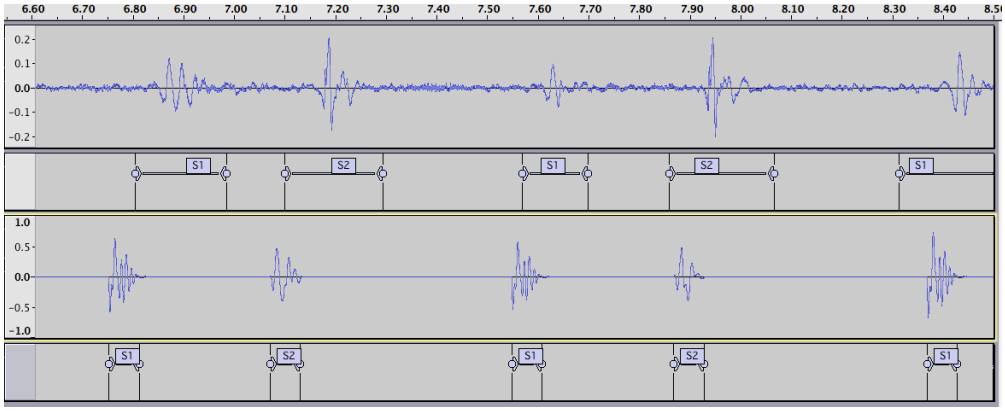


Figure 6.7: From top: the original auscultation, its manual annotation, the generated auscultation and the automatically generated auscultation.

6.5 Conclusions

In this chapter we developed a heart sound simulator program that generates synthetic auscultations based on real ones, and with customisable clinically meaningful parameters.

To do so, we also performed a study on the relationship between ventricular diastolic time and instantaneous heart rate and found that a better fit was achieved by using an exponential function. For each auscultation we calculated the parameters of this function and, based on them, we generated auscultations with customisable average heart rate and realistic instantaneous heart rate and systolic and diastolic times throughout the auscultation.

We also suggested a way of changing the parameters of S2 to generate

auscultations where the second heart sound can change throughout the auscultation in a somewhat natural way and yet, customisable by the user. The S2 parameters that can be changed are: split variation, and A2 and P2 amplitude variation.

The heart sound simulator not only generates auscultations but also can visualise, reproduce and save them and their annotations. This simulator can be important to test heart sound segmentation algorithms and, in our case, it was used to test ICA algorithms in separating the components of the second heart sound.

Chapter 7

Conclusions and future work

During this thesis we tried to understand the second heart sound through three different approaches, all of them based in the physiology of S2.

In the first approach, we looked to the second heart sound as a source separation problem: it was assumed that S2 is the set of physiological and independent events, such as the closure of the aorta and pulmonary artery, followed by the blood rebound. We made attempts to extract the signal based on the four auscultation sites using only one sensor: the stethoscope. To do so, we explored the quasi-periodicity of the heart signals and the fact that auscultation is performed in these auscultation sites, in order to simulate several sensors and perform ICA. This method differs from the single sensor separation problem because the mixture in this case differs at each location.

It is important to note that this is not a single sensor separation problem, where the additional sources are obtained by manipulating the frequencies sub-bands of the mixed signal. Our method explores the quasi-periodicity nature of the mixed signal and the moving sensor in order to simulate an array of spatially separated sensors acquiring mixtures. The source separation results did not produce the expected results, however the developed technique for sequentially acquiring signals did perform well for other types of signals, such as ECG and periodic synthetic signals.

As further research in the problem of source separation on S2, we could probably get better results by assuming a convoluted mixture instead of

instantaneous mixtures, as it was done in this thesis. Another possibility we believe would be interesting to study is the possibility of the mixing matrix be changing through time, since the patient chest is always changing its volume.

In the second approach, we assumed that the second heart sound is due to the blood rebound produced by the closure of the aortic and pulmonary arteries. These signals were modeled as non-linear chirps, inspired by the work of Xu [42]. We suggested that the second heart sound would be composed by the summation of these underlying physiological components (atoms), therefore, making the problem of calculating the parameters of these components into an optimization problem. In order to calculate the parameters of these two components, matching pursuit was performed with good results. Another result of this work was the generation of a sparse representation of the second heart sound where some parameters could convey important physiological information. This method, however, takes a considerable amount of time to calculate parameters of the atoms, since the search space is quite large and, due to the greedy nature of the matching pursuit algorithm, the selected parameters are not always the best ones. In order to improve this technique on those two points, we believe that the insertion of some parameters into the matching pursuit algorithm, or even the use of other algorithms could lead to better calculation of these parameters.

The increase in the size and diversity of our database with auscultations from both normal and pathologic subjects would also help us to characterize better the atoms of such cardiac conditions. Another possible point to explore with this approach is the use of such atoms as means to transmit, store and denoise the second heart sound.

As the third approach, we emulate what the clinicians do when reading a phonocardiogram: we assumed A2 and P2 to be energy peaks produced by the positive part of the auscultation signal. This approach takes the advantage of producing data about A2 and P2 that are similar to the ones reported on the medical literature - namely phonocardiogram literature. We performed an indirect validation of this method using patients with normal auscultation and patients exhibiting a hyperphonetic P2 component. We observed that a $P2 \geq 0.7A2$ do separates normal auscultations from hyper-

phonetic ones with high accuracy. This is a more precise measurement than the clinical literature that states $P2 > A2$ as a measure to separate normal from hyperphonetic auscultations.

This method performs poorly when auscultation is noisy, therefore the application of filtering or multiresolution analysis could improve its performance in such cases.

Another point of improvement is the introduction of some consistency check in order to detect sporadic wrong detections of A2 and P2 in auscultations.

In the future, it would also be interesting to monitor the respiratory cycle during the auscultation, since this could give some interesting improvements in the estimation of the respiratory cycle throughout the auscultation.

In addition to those points of views, we realized that the development of a physiologically inspired model of A2 and P2 opens the possibility of creating S2 with some physiological features customisable. This was used several times during the thesis to validate some of the techniques developed.

The creation of a heart sound simulator where physiological features could be changed was, a natural consequence of this development. In order to calculate the relationship between diastolic times and heart rate, we found that an exponential function is a better approximation than the state of the art linear function.

One important point of improvement in the simulator is the creation of a S1 that changes realistically throughout the auscultation and the addition of more normal and abnormal auscultations to the simulator's database: that would increase the flexibility and capabilities of the simulator in generating a wider variety of auscultations. Another interesting improvement of the simulator is the addition of several types of noise: synthetic, recorded, white, colored, constant, or transient.

Appendix A
Dataset

The auscultations used in this chapter were collected in the UCMF clinic at the Real Hospital Português de Beneficência in Recife and in the Procape clinic at the Oswaldo Cruz hospital, both in Recife - Pernambuco, Brazil. In total 47 auscultation from adults were collected as part of the regular clinical procedure. All recordings have a length of 60 seconds and were done in the regular clinical environment. This database is composed by 36 normal auscultations and 11 auscultations collected from patients with pulmonary artery hypertension. The Digiscope prototype (digiscope.up.pt) was used in conjunction with a Littmann 3100 stethoscope with a sampling rate of 4000Hz, 8 bits resolution and all filters disabled. Given the nature of the clinical environment, the collected auscultations had a wide range of background noise and the quality of the recorded heart sounds did vary greatly. Therefore we choose to include in our study only auscultations where the heart sound components could be heard, leaving 26 normal auscultations and 11 auscultations from patients with pulmonary hypertension. After recording the auscultations, the first and second heart sounds intervals where annotated using the Audacity software [113].

Bibliography

[1] V. A. McKusick, *Cardiovascular sound in Heart and Disease.* Baltimore: Wiliams & Wilkins, Jan. 1958, the History of Cardiovascular Sound.

[2] J. Forbes, *A Treatise on the Disease of the Chest.* Underwood, 1821.

[3] R. T. H. Laennec and M. Laennec, *Traité de l'auscultation médiate, et des maladies des poumons et du coeur.* Société Typographique Belge, 1837.

[4] B. G. E. and D. Pasquale, *A history of electrocardiography.* Norman Publishing, 1990.

[5] O. Weiss and G. Joachin, "Registrierung von hertzönen und herzgeräuschen mittels des phonoskops und ihre beziehungen zum elektrokardiogramm," *Ztschr. klin. Med*, vol. 73, p. 240, 1911.

[6] S. Yuenyong, W. Kongpravechnon, K. Tungpimolrut, and A. Nishihara, "Automatic heart sound analysis for tele-cardiac auscultation," in *Proc. ICROS-SICE Int. Joint Conference*, 2009, pp. 1599–1604.

[7] U. Alam, O. Asghar, S. Khan, S. Hayat, and R. Malik, "Cardiac auscultation: an essential clinical skill in decline," *British Journal of Cardiology*, vol. 17, pp. 8–10, 2010.

[8] A. Pease, "If the heart could speak," *Pictures of the Future*, pp. 60–61, 2001.

[9] E. Pretorius, M. Cronje, and O. Strydom, "Development of a pediatric cardiac computer aided auscultation decision support system," in *Annual Int. Conference of the IEEE Eng. in Medicine and Biology Society*, 2010, pp. 6078–6082.

[10] S. Mangione and L. Nieman, "Cardiac auscultatory skills of internal medicine and family practice trainees," *Journal of the American Medical Association*, vol. 278, pp. 717–722, 1997.

[11] M. J. Green and R. Rieck, "Missed it," *Annals of internal medicine*, vol. 158, no. 5_Part_1, pp. 357–361, 2013.

[12] J. de Vos and M. Blanckenberg, "Automated pediatric cardiac auscultation," *IEEE Trans. on Biomedical Engineering*, vol. 54, pp. 244–252, 2007.

[13] A. A. Luisada, *From Auscultation to phonocardiograpy.* The C. V. Mosby Company, 1965.

[14] M. M. AYGEN and E. BRAUNWALD, "The splitting of the second heart sound in normal subjects and in patients with congenital heart disease," *Circulation*, vol. 25, no. 2, pp. 328–345, 1962.

[15] R. Kusukawa, D. W. Bruce, T. Sakamoto, D. M. MacCanon, and A. A. Luisada, "Hemodynamic determinants of the amplitude of the second heart sound." *J Appl Physiol*, vol. 21, pp. 938–946, 1966.

[16] M. Mori, P. Shah, D. MacCanon, and A. Luisada, "Hemodynamic correlates of the various components of the second heart sound," *Cardiology*, vol. 44, no. 2, pp. 65–77, 1964.

[17] L. Vogelpoel and V. SCHRIRE, "The role of auscultation in the differentiation of fallot's tetralogy from severe pulmonary stenosis with intact ventricular septum and right-to-left interatrial shunt," *Circulation*, vol. 11, no. 5, pp. 714–732, 1955.

[18] A. Leatham and D. Weitzman, "Auscultatory and phonocardiographic signs of pulmonary stenosis," *British Heart Journal*, vol. 19, no. 3, p. 303, 1957.

[19] M. E. Tavel, *Clinical phonocardiography and external pulse recording*. Year Book Medical Publishers, 1972.

[20] G. Borden and K. Harris, *Speech science primer: physiology, acoustics, and perception of speech*. Williams & Wilkins, 1994.

[21] H. Von Gierke, "Transmission of vibratory energy through human body tissue," in *Proc. of First Nat. Biophysics Conf*, 1957.

[22] J. Faber and A. Burton, "Spread of heart sounds over chest wall." *Circulation research*, vol. 11, p. 96, 1962.

[23] J. Verburg, "Transmission of vibrations of the heart to the chest wall," *Adv. Cardiovasc. Phys*, vol. 5, no. Part III, pp. 84–103, 1983.

[24] J. E. Hall, *Guyton and Hall Textbook of Medical Physiology: Enhanced E-book*. Elsevier Health Sciences, 2010.

[25] e. a. Paul A. Iazzo, *Handbook of Cardiac Anatomy, Physiology, and Devices*, P. A. Iazzo, Ed. Humana Press Totowa, New Jersey, 2005.

[26] W. G. C. S. Germann, *Principles of Human Physiology*. Pearson Education/Benjamin Cummings, 2002.

[27] H. K. Walker, W. D. Hall, J. W. Hurst, and J. M. Felner, *Clinical methods*. Butterworths, 1990.

[28] (2009, 9). [Online]. Available: http://lessons4medicos.blogspot.co.uk/2009/07/splitting-of-second-heart-sound.html

[29] E. Curtiss, R. Matthews, and J. Shaver, "Mechanism of normal splitting of the second heart sound." *Circulation*, vol. 51, no. 1, pp. 157–164, 1975.

[30] P. A. Iaizzo, *Handbook of cardiac anatomy, physiology, and devices*. Springer, 2009.

[31] J. Abrams, "Physical examination of the heart and circulation," in *Essential Cardiology*. Springer, 2006, pp. 99–115.

[32] S. L. H. Liang and I. Hartimo, "A heart sound segmentation algorithm using wavelet decomposition and reconstruction," in *19th International Conference - IEEE/EMBS*, Chicago, IL, USA, Oct., Nov. 1997.

[33] H. Liang, S. Lukkarinen, and I. Hartimo, "Heart sound segmentation algorithm based on heart sound envelogram," *Computers in Cardiology*, vol. 24, 1997.

[34] H. Liang and I. Hartimo, "A heart sound feature extraction algorithm based on wavelet decomposition and reconstruction," in *Proc. IEEE EMBS*, vol. 20, no. 3, 1998.

[35] M. T. Sherif Omran, "A heart sound segmentation and feature extraction algorithm using wavelet," in *Proc. of IEEE MWSCAS '03*, vol. 1, 2003, pp. 27–30.

[36] D. Kumar, P. Carvalho, R. Couceiro, M. Antunes, R. P. Paiva, and J. Henriques, "Heart murmur classification using complexity signatures," in *Pattern Recognition (ICPR), 2010 20th International Conference on*. IEEE, 2010, pp. 2564–2567.

[37] D. Kumar, P. Carvalho, M. Antunes, J. Henriques, L. Eugenio, R. Schmidt, and J. Habetha, "Detection of s1 and s2 heart sounds by high frequency signatures," in *Engineering in Medicine and Biology Society, 2006. EMBS'06. 28th Annual International Conference of the IEEE*. IEEE, 2006, pp. 1410–1416.

[38] D. Kumar, P. d. Carvalho, M. Antunes, J. Henriques, M. Maldonado, R. Schmidt, and J. Habetha, "Wavelet transform and simplicity based heart murmur segmentation," in *Computers in Cardiology, 2006*. IEEE, 2006, pp. 173–176.

[39] D. Kumar, P. Carvalho, M. Antunes, R. P. Paiva, and J. Henriques, "An adaptive approach to abnormal heart sound segmentation," in *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on.* IEEE, 2011, pp. 661–664.

[40] A. Moukadem, A. Dieterlen, N. Hueber, and C. Brandt, "A robust heart sounds segmentation module based on s-transform," *Biomedical Signal Processing and Control*, 2013.

[41] R. Stockwell, L. Mansinha, and R. Lowe, "Localisation of the complex spectrum: the s transform," *Journal of Association of Exploration Geophysicists*, vol. 17, no. 3, pp. 99–114, 1996.

[42] P. P. JingPing Xu, L.G. Durand, "Nonlinear transient chirp signal modeling of the aortic and pulmonary components of the secound heart sound," *IEEE Transactions on Biomedical Engineering*, vol. 47, no. 7, March 2000.

[43] L.-G. Durand and P. Pibarot, "Digital signal processing of the phonocardiogram: Review of the most recent advancements," *Critical Reviews in Biomedical Engineering*, 1995.

[44] J. Xu, L. Durand, and P. Pibarot, "Extraction of the aortic and pulmonary components of the secound heart sound using nonlinear transient chirp signal model," *IEEE Transactions on Biomedical Engineering*, vol. 48, no. 3, March 2001.

[45] V. Nigam and R. Priemer, "A procedure to extract the aortic and the pulmonary sounds from the phonocardiogram," in *Proceedings of the 28th IEEE EMBS Annual International Conference*, New York City, USA, Aug 30-Sep 3 2006.

[46] P. M. Bentley, J. T. E. McDonnell, and P. M. Grant, "Classification of native heart valve sounds using the choi-williams time-frequency distribution," *Biomedical Engineering, IEEE Transactions on*, 1995.

[47] P. M. Bentley, P. M. Grant, and J. T. E. McDonnell, "Time-frequency and time-scale techniques for the classification of native and bioprosthetic heart valve sounds," *IEEE Transactions on Biomedical Engineering*, vol. 45, no. 1, January 1998.

[48] P. Wang, Y. Kim, and C. B. Soh, "Feature extraction based on mel-scaled wavelet transform for heart sound analysis," in *Engineering in Medicine and Biology Society*, 2005.

[49] H. Liang and I. Hartimo, "A feature extraction algorithm based on wavelet packet decomposition for heart sound signals," in *Proceedings of the IEEE-SP International Symposium*, October 1998.

[50] I. Turkoglu and A. Arslan, "An intelligent pattern recognition system based on neural network and wavelet decomposition for interpretation of heart sounds," in *Proceedings of the 23rd Annual International Conference of the IEEE*, vol. 2, October 2001, pp. 25–28.

[51] Z. D. Ozgur Say and T. Olmez, "Classification of heart sounds by using wavelet transform," in *Proceedings of the Second Joint EMBS/BMES Conference*, vol. 1, 2002.

[52] M. El-Hanjouri, W. Alkhaldi, N. Hamdy, and A. Alim, "Heart diseases diagnosis using hmm," in *Electrotechnical Conference -MELECON*, 2002.

[53] T. Leung, P. White, W. B. Collins, E. Brown, and A. P. Salmon, "Classification of heart sounds using time-frequency method and artificial neural networks," in *Proceedings of the 22nd Annual International Conference of the IEEE*, vol. 2, July 2000.

[54] E. Kail, S. Khoor, B. Kail, K. Fugedi, and F. Balazs, "Internet digital phonocardiography in clinical settings and in population screening," *Computers in Cardiology*, 2004.

[55] 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 instantan-

eous energy and frequency estimations," in *Proceedings of TENCON*, vol. 2, 2000.

[56] B. Ans, J. Hérault, and C. Jutten, "Adaptive neural architectures: detection of primitives," *Proc. of COGNITIVA'85*, pp. 593–597, 1985.

[57] J. Herault and B. Ans, "Circuits neuronaux synapses modifiables : dÃ©codage de messages composites par apprentissage non supervisÃ©," in *C.R. de l'Acadmie des Science*, 1984.

[58] J. Herault, C. Jutten, and B. Ans, "Detection de grandeurs primitives dans un message composite par une architecture de calcul neuromimetique en apprentissage non supervise," in *Actes du Xe colloque GRETSI*, vol. 2, 1985, pp. 1017–1022.

[59] D. Barrett, A. Cobb, and G. Bentley, "Joint proprioception in normal, osteoarthritic and replaced knees," *Journal of Bone & Joint Surgery, British Volume*, vol. 73, no. 1, pp. 53–56, 1991.

[60] S. M. Lephart, D. M. Pincivero, J. L. Giraido, and F. H. Fu, "The role of proprioception in the management and rehabilitation of athletic injuries," *The American journal of sports medicine*, vol. 25, no. 1, pp. 130–137, 1997.

[61] M. Björklund, "Effects of repetitive work on proprioception and of stretching on sensory mechanisms: implications for work-related neuromuscular disorders," Ph.D. dissertation, Umeå University, 2004.

[62] D. Collins and A. Prochazka, "Movement illusions evoked by ensemble cutaneous input from the dorsum of the human hand." *The Journal of physiology*, vol. 496, no. Pt 3, pp. 857–871, 1996.

[63] D. F. Collins, K. M. Refshauge, G. Todd, and S. C. Gandevia, "Cutaneous receptors contribute to kinesthesia at the index finger, elbow, and knee," *Journal of Neurophysiology*, vol. 94, no. 3, pp. 1699–1706, 2005.

[64] D. McCloskey, "Kinesthetic sensibility." *Physiological Reviews*, vol. 58, no. 4, pp. 763–820, 1978.

[65] J. Stone, *Independent component analysis: a tutorial introduction*, J. Stone, Ed. The MIT Press, CambridgeMassachusetts, 2004.

[66] P. Comon, "Independent component analysis, a new concept?" *Signal processing*, vol. 36, no. 3, pp. 287–314, 1994.

[67] T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley & Sons, 2012.

[68] D.-T. Pham, "Blind separation of instantaneous mixture of sources based on order statistics," *Signal Processing, IEEE Transactions on*, vol. 48, no. 2, pp. 363–375, 2000.

[69] A. Hyvarinen, "Blind source separation by nonstationarity of variance: A cumulant-based approach," *Neural Networks, IEEE Transactions on*, vol. 12, no. 6, pp. 1471–1474, 2001.

[70] A. Hyvärinen and E. Oja, "A fast fixed-point algorithm for independent component analysis," *Neural computation*, vol. 9, no. 7, pp. 1483–1492, 1997.

[71] D. T. Pham and P. Garat, "Blind separation of mixture of independent sources through a maximum likelihood approach," in *In Proc. EUSIPCO*. Citeseer, 1997.

[72] J.-F. Cardoso, "Infomax and maximum likelihood for blind source separation," *Signal Processing Letters, IEEE*, vol. 4, no. 4, pp. 112–114, 1997.

[73] A. J. Bell and T. J. Sejnowski, "An information-maximization approach to blind separation and blind deconvolution," *Neural computation*, vol. 7, no. 6, pp. 1129–1159, 1995.

[74] S.-i. Amari, A. Cichocki, H. H. Yang *et al.*, "A new learning algorithm for blind signal separation," *Advances in neural information processing systems*, pp. 757–763, 1996.

[75] J.-F. Cardoso, "Blind signal separation: statistical principles," *Proceedings of the IEEE*, vol. 86, no. 10, pp. 2009–2025, 1998.

[76] S. T. Roweis, "One microphone source separation," in *NIPS*, 2000, pp. 793–799.

[77] A. Bhattacherjee and N. Hikmet, "Physicians' resistance toward healthcare information technology: a theoretical model and empirical test," *European Journal of Information Systems*, vol. 16, no. 6, pp. 725–737, 2007.

[78] A. Hyvärinen and E. Oja, "Independent component analysis: algorithms and applications," *Neural networks*, vol. 13, no. 4, pp. 411–430, 2000.

[79] V. Zarzoso, A. Nandi, and E. Bacharakis, "Maternal and foetal ecg separation using blind source separation methods," *Mathematical Medicine and Biology*, vol. 14, no. 3, pp. 207–225, 1997.

[80] V. Zarzoso and A. K. Nandi, "Noninvasive fetal electrocardiogram extraction: blind separation versus adaptive noise cancellation," *Biomedical Engineering, IEEE Transactions on*, vol. 48, no. 1, pp. 12–18, 2001.

[81] J. Gnitecki and Z. M. Moussavi, "Separating heart sounds from lung sounds-accurate diagnosis of respiratory disease depends on understanding noises," *Engineering in Medicine and Biology Magazine, IEEE*, vol. 26, no. 1, pp. 20–29, 2007.

[82] M. Jafari, "Novel sequential algorithms for blind source separation of instantaneous mixtures," Ph.D. dissertation, King's College London, September 2002.

[83] L. De Lathauwer, B. De Moor, and J. Vandewalle, "Fetal electrocardiogram extraction by blind source subspace separation," *Biomedical Engineering, IEEE Transactions on*, vol. 47, no. 5, pp. 567–572, 2000.

[84] A. Sovijarvi, L. Malmberg, G. Charbonneau, J. Vanderschoot, F. Dalmasso, C. Sacco, M. Rossi, and J. Earis, "Characteristics of breath sounds and adventitious respiratory sounds," *European Respiratory Review*, vol. 10, no. 77, pp. 591–596, 2000.

[85] P. Arnott, G. Pfeiffer, and M. Tavel, "Spectral analysis of heart sounds: relationships between some physical characteristics and frequency spectra of first and second heart sounds in normals and hypertensives," *Journal of biomedical engineering*, vol. 6, no. 2, pp. 121–128, 1984.

[86] F. Ghaderi, H. R. Mohseni, and S. Sanei, "Localizing heart sounds in respiratory signals using singular spectrum analysis," *Biomedical Engineering, IEEE Transactions on*, vol. 58, no. 12, pp. 3360–3367, 2011.

[87] F. L. Hedayioglu, M. G. Jafari, S. S. Mattos, M. D. Plumbley, and M. T. Coimbra, "Separating sources from sequentially acquired mixtures of heart signals," in *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*. IEEE, 2011, pp. 653–656.

[88] M. G. Jafari, F. L. Hedayioglu, M. T. Coimbra, and M. D. Plumbley, "Blind source separation of periodic sources from sequentially recorded instantaneous mixtures," in *Image and Signal Processing and Analysis (ISPA), 2011 7th International Symposium on*. IEEE, 2011, pp. 540–545.

[89] T. Tran, N. Jones, and J. Fothergill, "Heart sound simulator," *Medical and Biological Engineering and Computing*, vol. 33, no. 3, pp. 357–359, 1995.

[90] X. Zhang, L. Durand, L. Senhadji, H. C. Lee, and J.-L. Coatrieux, "Analysis-synthesis of the phonocardiogram based on the matching

pursuit method," *Biomedical Engineering, IEEE Transactions on*, vol. 45, no. 8, pp. 962–971, 1998.

[91] Y. Tang, C. Danmin, and L. Durand, "The synthesis of the aortic valve closure sound of the dog by the mean filter of forward and backward predictor," *Biomedical Engineering, IEEE Transactions on*, vol. 39, no. 1, pp. 1–8, 1992.

[92] H. Koymen, B. K. Altay, and Y. Z. Ider, "A study of prosthetic heart valve sounds," *Biomedical Engineering, IEEE Transactions on*, no. 11, pp. 853–863, 1987.

[93] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *Signal Processing, IEEE Transactions on*, vol. 41, no. 12, pp. 3397–3415, 1993.

[94] R. Gribonval, "Fast matching pursuit with a multiscale dictionary of gaussian chirps," *Signal Processing, IEEE Transactions on*, vol. 49, no. 5, pp. 994–1001, 2001.

[95] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM journal on scientific computing*, vol. 20, no. 1, pp. 33–61, 1998.

[96] T. Blumensath and M. E. Davies, "Iterative hard thresholding for compressed sensing," *Applied and Computational Harmonic Analysis*, vol. 27, no. 3, pp. 265–274, 2009.

[97] D. L. Donoho, Y. Tsaig, I. Drori, and J.-L. Starck, "Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit," *Information Theory, IEEE Transactions on*, vol. 58, no. 2, pp. 1094–1121, 2012.

[98] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *Information Theory, IEEE Transactions on*, vol. 53, no. 12, pp. 4655–4666, 2007.

[99] H. Åstrand, "Regulation of aortic wall mechanics and stress: An experimental study in man," Ph.D. dissertation, Linköping, 2008.

[100] A. Bartels and D. Harder, "Non-invasive determination of systolic blood pressure by heart sound pattern analysis," *Clinical Physics and Physiological Measurement*, vol. 13, no. 3, p. 249, 1992.

[101] F. L. Hedayioglu, M. Jafari, S. S. Mattos, M. D. Plumbley, and M. Coimbra, "Denoising and segmentation of the second heart sound using matching pursuit," in *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*. IEEE, 2012, pp. 3440–3443.

[102] K. Holldack, A. A. Luisada, and H. Ueda, "Standardization of phonocardiography," *The American Journal of Cardiology*, vol. 15, no. 3, pp. 419–421, 1965.

[103] A. A. Luisada and G. Gamna, "Clinical calibration in phonocardiography," *American Heart Journal*, vol. 48, no. 6, pp. 826–834, 1954.

[104] A. Luisada, L. Richmond, and C. Aravanis, "Selective phonocardiography," *American Heart Journal*, vol. 51, no. 2, pp. 221–233, 1956.

[105] A. Sloan and J. Greer, "Calibration of an electronic phonocardiograph," *British heart journal*, vol. 17, no. 2, p. 138, 1955.

[106] J. Rouanet, "Analyse des bruits du coeur.(cand. j. rouanet.)." Ph.D. dissertation, 1832.

[107] 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," *Medical and biological engineering*, vol. 14, no. 1, pp. 69–73, 1976.

[108] G. Sutton, A. Harris, and A. Leatham, "Second heart sound in pulmonary hypertension." *British heart journal*, vol. 30, no. 6, p. 743, 1968.

[109] A. Harris and G. Sutton, "Second heart sound in normal subjects." *British Heart Journal*, vol. 30, no. 6, p. 739, 1968.

[110] T. Sakamoto, R. Kusukawa, D. M. Maccanon, A. A. Luisada, and I. Harvey, "Hemodynamic determinants of the amplitude of the first heart sound," *Circulation research*, vol. 16, no. 1, pp. 45–57, 1965.

[111] V. A. Mckusick, "Cardiovascular sound in health and disease," *The American Journal of the Medical Sciences*, vol. 238, no. 1, p. 128, 1959.

[112] E. Lehmann, *Nonparametrics: statistical methods based on ranks (POD)*, 2006.

[113] (2009, 09) Audacity: Free audio editor and recorder. [Online]. Available: http://audacity.sourceforge.net/

[114] A. Almasi, M. B. Shamsollahi, and L. Senhadji, "A dynamical model for generating synthetic phonocardiogram signals," in *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*. IEEE, 2011, pp. 5686–5689.

[115] P. E. McSharry, G. D. Clifford, L. Tarassenko, and L. A. Smith, "A dynamical model for generating synthetic electrocardiogram signals," *Biomedical Engineering, IEEE Transactions on*, vol. 50, no. 3, pp. 289–294, 2003.

[116] C. Danmin, "Application of time-frequency representations to study the genesis of the first heart sound," Master's thesis, McGill University, March 1995.

[117] D. Chen, L.-G. Durand, and H. Lee, "Time-frequency analysis of the first heart sound. part 1: Simulation and analysis," *Medical and Biological Engineering and Computing*, vol. 35, no. 4, pp. 306–310, 1997.

[118] G. Cloutier, R. Guardo, and L.-G. Durand, "Spectral analysis of closing sounds produced by lonescu-shiley bioprosthetic aortic heart valves," *Medical and Biological Engineering and Computing*, vol. 25, no. 5, pp. 497–503, 1987.

[119] R. Rangayyan and R. J. Lehner, "Phonocardiogram signal analysis: a review." *Critical reviews in biomedical engineering*, vol. 15, no. 3, p. 211, 1987.

[120] M. El-Segaier, O. Lilja, S. Lukkarinen, L. Sörnmo, R. Sepponen, and E. Pesonen, "Computer-based detection and analysis of heart sound and murmur," *Annals of biomedical engineering*, vol. 33, no. 7, pp. 937–942, 2005.