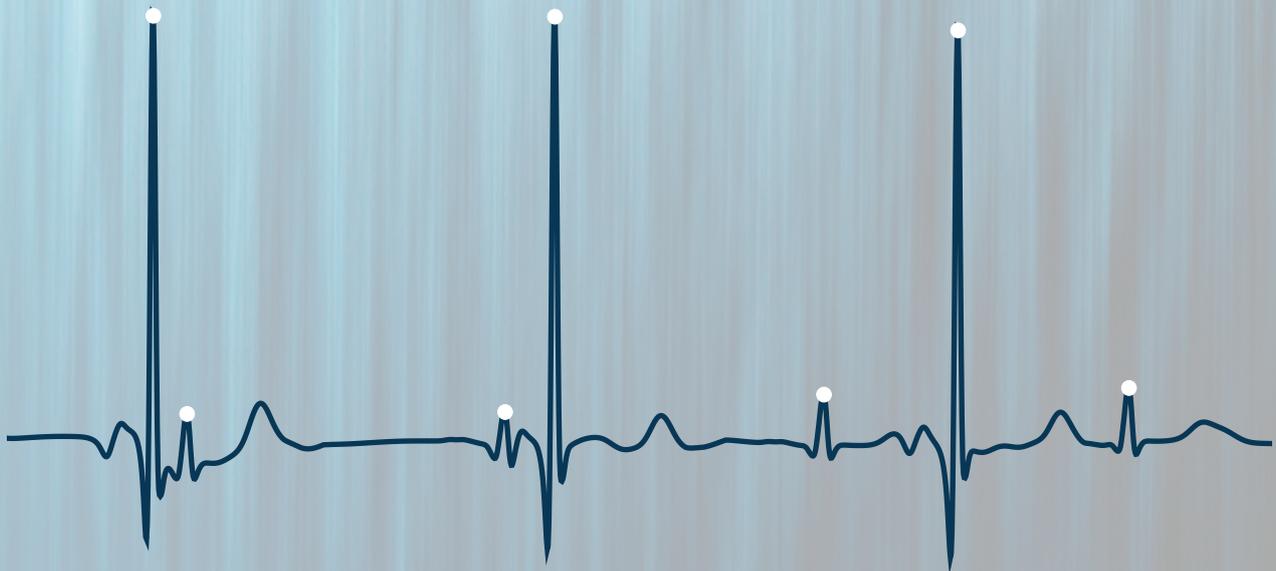


SIGNAL PROCESSING & COMPUTING

# NON-INVASIVE FETAL ECG USING CONSTRAINED ICA



Rasmus Gundorff Sæderup



MASTER'S THESIS

JUNE 2018

Copyright © Aalborg University 2018

This report is compiled in L<sup>A</sup>T<sub>E</sub>X, originally developed by Leslie Lamport, based on Donald Knuth's T<sub>E</sub>X. The main text is written in *Computer Modern* pt 11, designed by Donald Knuth. Flowcharts and diagrams are made using draw.io, a free online diagram software.



Institute of Electronic Systems  
Fredrik Bajers Vej 7  
9220 Aalborg Øst

## AALBORG UNIVERSITY

### STUDENT REPORT

**Title:**

Non-invasive fetal ECG using  
constrained ICA

**Master's Programme:**

Signal Processing and Computing

**Project Period:**

February 2018 - June 2018

**Project Group:**

SPC10-1071

**Participants:**

Rasmus Gundorff Sæderup

**Supervisors:**

Søren Holdt Jensen  
Samuel Schmidt

**Number of Pages:**

147 incl. appendices

**Date of Completion:**

June 15, 2018

**Abstract:**

Non-invasive methods for estimating the fetal electrocardiogram (FECCG) by using ECG recordings from the abdomen of the mother are of great interest in order to carry out less complicated and safer fetal monitoring compared to the invasive methods. In this thesis, independent component analysis (ICA) is used to extract morphologically accurate fetal ECG, by estimating the maximally independent sources from the observed mixture.

A constrained ICA (cICA) algorithm is derived, which is based on constraining the ICA optimization problem such that the estimated sources must be correlated with some reference signals.

Different reference signals are generated, ranging from pulse signals at the QRS locations to templates generated from e.g. maternal ECGs. A hyper-parameter search is conducted, where different step-sizes, correlations thresholds and combinations of maternal and fetal reference signals are tested. Using the parameters giving the best results, the cICA algorithm is tested on a synthetic dataset, where its performance is compared to other extraction algorithms such as template subtraction PCA as well as other ICA methods such as FastICA and Infomax. From this test, it is clear that cICA can do accurate extraction if the true FECCGs are used as reference, but performs less well if other references are used, and is comparable to classical ICA methods in these cases. None of the tested algorithms were able to extract morphologic features (QT-interval and T/QRS ratio) from real ECG mixtures. It was found that cICA converges very fast, but to a local minimum, which is due to the non-convex nature of the problem. It can therefore be concluded that constrained ICA will not provide morphologically accurate FECCG if the true FECCGs are not provided.

# Nomenclature

$\alpha$	Step-size/learning rate
$\mathbf{\Lambda}$	Diagonal matrix of eigenvalues
$\epsilon$	Correlation threshold
$\epsilon_f$	Fetal correlation threshold
$\epsilon_m$	Maternal correlation threshold
$\epsilon_W$	Convergence threshold for cICA
$\eta$	Momentum factor
$\gamma$	A penalty parameter
$\mathbf{A}$	The mixing matrix, $\mathbf{A} \in \mathbb{R}^{M \times K}$
$\mathbf{B}$	The de-mixing matrix, $\mathbf{B} \in \mathbb{R}^{K \times M}$
$\mathbf{I}$	Identity matrix
$\mathbf{P}$	The permutation matrix, $\mathbf{P} \in \mathbb{R}^{K \times K}$
$\mathbf{R}$	Covariance matrix
$\mathbf{U}$	Matrix of stacked eigenvectors
$\mathbf{V}$	Whitening matrix, $\mathbf{V} \in \mathbb{R}^{M \times M}$
$\mathbf{W}$	The de-mixing matrix applied to whitened data, $\mathbf{W} \in \mathbb{R}^{K \times M}$
$\mathbf{X}$	Observation matrix, $\mathbf{X} \in \mathbb{R}^{M \times N}$
$\nabla_x$	The gradient w.r.t. $x$
$\hat{\mathbf{s}}$	Estimated source signal vector, $\hat{\mathbf{s}} \in \mathbb{R}^{K \times 1}$
$\hat{\mathbf{s}}[n]$	Estimated source signal sequence, $\hat{\mathbf{s}} \in \mathbb{R}^{K \times N}$
$\mathbf{b}$	A single row of the de-mixing matrix $\mathbf{B}$
$\lambda$	A vector of Lagrange multipliers for equality constraints
$\mu$	A vector of Lagrange multipliers for inequality constraints
$\mathbf{r}$	Reference signal vector as a random variable, $\mathbf{r} \in \mathbb{R}^{K \times 1}$
$\mathbf{s}$	Source signal vector as a random variable, $\mathbf{s} \in \mathbb{R}^{K \times 1}$
$\mathbf{s}[n]$	Source signal sequence, $\hat{\mathbf{s}} \in \mathbb{R}^{K \times N}$
$\mathbf{v}$	A vector of slack-variables
$\mathbf{w}$	A noise vector, $\mathbf{w} \in \mathbb{R}^{M \times 1}$
$\mathbf{w}$	A single row of the de-mixing matrix $\mathbf{W}$

$\mathbf{x}$	Observed signal vector as a random variable, $\mathbf{x} \in \mathbb{R}^{M \times 1}$
$\mathbf{x}[n]$	Observed signal sequence, $\mathbf{x} \in \mathbb{R}^{M \times N}$
$\mathbf{z}$	Whitened signal vector as a random variable, $\mathbf{z} \in \mathbb{R}^{M \times 1}$
$K$	Number of signals/sources
$k$	Source index
$M$	Number of sensors
$m$	Observation index
$N$	Number of samples
$n$	Time sample index
$p$	Indexes fetal and maternal sources, $p \in m, f$
cICA	Constrained ICA
ECG	Electrocardiogram
EKF	Extended Kalman Filter
FEKG	Fetal electrocardiogram
FQRS	Fetal QRS complex
HPF	High-pass Filter
ICA	Independent Component Analysis
KF	Kalman Filter
LPF	Low-pass Filter
MECG	Maternal electrocardiogram
MQRS	Maternal QRS complex
PCA	Principal Component Analysis
PF	Particle Filter
QRS	The QRS-complex of the ECG
SNR	Signal to noise ratio
TS	Template-subtraction
VCG	Vector-cardiogram

# Contents

<b>Preface</b>	<b>ix</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Background and Motivation . . . . .	1
1.2 Aim of Thesis . . . . .	2
<b>2 Clinical Background</b>	<b>3</b>
2.1 Cardiac Physiology . . . . .	3
2.2 Fetal Circulation . . . . .	5
2.3 The Electrocardiogram . . . . .	6
2.4 Fetal ECG Monitoring . . . . .	8
<b>3 Non-Invasive Fetal Heart Rate Estimation</b>	<b>13</b>
3.1 Databases of Fetal ECG Recordings . . . . .	13
3.2 Overview of Current Methods . . . . .	16
3.3 Comparison of Current Methods . . . . .	20
<b>4 Non-Invasive Fetal ECG Morphology Estimation</b>	<b>27</b>
4.1 The Challenges of Fetal ECG Morphology . . . . .	27
4.2 Overview of Current Methods . . . . .	29
4.3 Comparison of Current Methods . . . . .	32
<b>5 Extraction Methods Considerations</b>	<b>35</b>
5.1 Problems with Existing Methods . . . . .	35
5.2 Potential Extraction Methods . . . . .	37
5.3 VCG-ICA + EKF . . . . .	43
5.4 Constrained ICA . . . . .	46
5.5 Preprocessing . . . . .	49
<b>6 Independent Component Analysis</b>	<b>55</b>
6.1 General ICA Concepts . . . . .	55
6.2 ICA by Maximization of Non-Gaussianity . . . . .	62
6.3 FastICA . . . . .	66
6.4 An Example of ICA . . . . .	71
6.5 Infomax . . . . .	73
<b>7 ICA with Reference</b>	<b>79</b>
7.1 Deriving the cICA Algorithm . . . . .	79
7.2 Summary of cICA Algorithm . . . . .	84
7.3 Generating Reference Signals . . . . .	86
<b>8 Test and Performance Results</b>	<b>93</b>
8.1 Test Measures . . . . .	93
8.2 Test Plan . . . . .	98
8.3 Test Results 1A - ICA on Non-ECG Signals . . . . .	102

8.4	Test Results 1B - ICA on ECG signals . . . . .	105
8.5	Test Results 2A - Morphology of Artificial Signals . . . . .	113
8.6	Test Results 2B - Morphology of Real Signals . . . . .	115
<b>9</b>	<b>Discussion</b>	<b>117</b>
<b>10</b>	<b>Conclusion</b>	<b>121</b>
<b>11</b>	<b>Future Work</b>	<b>123</b>
 <b>Bibliography</b>		 <b>125</b>
 <b>Appendices</b>		 <b>133</b>
A	Electrocardiographic Leads . . . . .	133
B	Beamforming Methods . . . . .	136
C	Annotation GUI . . . . .	141
D	Results of ICA on Non-ECG Signals . . . . .	143
E	Results of ICA Comparison on ECG Signals . . . . .	145
F	Morphology Test Results on Artificial Signals . . . . .	147

# List of Algorithms

1	FastICA . . . . .	70
2	Infomax . . . . .	75
3	Extended Infomax . . . . .	77
4	cICA . . . . .	85

# Preface

This master's thesis in Signal Processing & Computing has been carried out during the spring of 2018, from February to June, at Aalborg University (AAU).

The thesis is on the topic of non-invasive fetal electrocardiogram (ECG) extraction, obtained from a series of abdominal sensors, and is structured as follows: First, an introduction to the topic of fetal ECG monitoring is given, together with an overview of the necessary clinical background on ECGs and fetal monitoring. Hereafter follows an in-depth literature review of the current methods used for non-invasive fetal ECG monitoring, and what their limitations are. Then, a series of new methods are proposed and briefly explained, after which one of them, constrained independent component analysis (cICA), is chosen and described in detail. An algorithm for extracting fetal ECG is developed, and its results are presented in chapter 8. Finally, a discussion and conclusion follows.

The reader of this thesis is assumed to have an understanding of various signal processing techniques and concepts, such as constrained optimization, probability theory/stochastics, adaptive filtering, as well as more classical signal processing methods such as digital filters and multirate systems.

References to sources are of the APA-style, i.e. of the type [*source name/author's surname, year, optional page number*]. Sources are listed in a bibliography at the end of the report. Figures without a source are of own production.

This thesis would not have taken this form, was it not for the help of some highly skilled and helpful people. I would first and foremost like to thank Viewcare A/S for proposing the topic of non-invasive fetal ECG, and for providing data collected with their Centaflow device.

Henrik Zimmermann, R&D engineer at Viewcare A/S, is thanked for his sincere interest in the progress of the project, and for discussing ideas, problems and possible solutions that have come up throughout the thesis.

A great thank you goes to Professor MSO Jan Østergaard from the Signal & Information Processing (SIP) section at AAU, for his assistance discussing information theory and optimization problems.

The staff at the Medical Informatics Group at Aalborg University are thanked for always having open office doors and being willing to discuss complicated topics with a signal processing student with no prior biomedical experience. In this regard, a special thanks sounds to Associate Professors John Hansen and Claus Graff as well as Professor Johannes Struijk.

Finally, a huge thanks to my fellow SPC student and partner in crime through the last two years, Poul Hoang. You've been a great help discussing ideas and understanding difficult topics, as well as helping making every day in the group room a bit more fun.

Rasmus Gundorff Sæderup,

Aalborg University, June 15, 2018



# 1 | Introduction

## 1.1 Background and Motivation

The fetal electrocardiogram (FECG) provides information about the wellbeing of the fetal heart, and thus the fetus in general. This is especially useful during birth, where monitoring the fetal heart rate is used as a tool to evaluate fetal oxygenation: If the oxygen levels are too low, medical intervention such as performing a caesarean section might be necessary. [Freeman et al., 2012, p. 8]

The ability to extract and analyze the FECG is thus of great interest as it is a highly valuable tool for monitoring the fetal wellbeing, and therefore also in ensuring a higher infant survival rate.

The death of a child is probably the most traumatic event any parent can experience, and every single child saved due to fetal monitoring is of great importance for the family having the baby.

However, from a higher, more social perspective, it is in fact also of great interest to increase the infant survival rate: The world population is growing in an exponential way: 200 years ago there were less than one billion humans living on earth. Today, there are over 7 billion humans on earth [Roser and Ortiz-Ospina, 2018]. Although not entirely accepted, overpopulation is thought to be "a main threat to the planet" according to some experts [Connor, 2006; Moody, 2017]. On the contrary, it can be argued that overpopulation will not become a problem, as theories such as *Demographic Transition*<sup>1</sup> predict that the population growth will decrease and eventually reverse at a certain point [Deonandan, 2017]. And in fact, the growth rate has decreased in recent years [Roser and Ortiz-Ospina, 2018], which raises the question of why this decrease has been occurring. One of the reasons is reduced child mortality:

In 2016, 2.6 million newborns died globally according to [UN, 2017, p. 1], equivalent to 7000 every day. Neonatal<sup>2</sup> deaths accounted for 46% of all deaths of children under the age of five years [UN, 2017, p. 1]. One may initially think that decreasing child mortality will increase the population, but in fact, the opposite is true:

*"A very cynical view is that a decrease in child mortality is bad for the world since it would contribute to the overpopulation of the planet (...): When more infants survive, fertility goes down and the temporary population growth comes to an end. If we want to ensure that the world's population increase comes to an end soon we must work to increase child survival."* [Roser, 2018]

More precisely, countries with an infant mortality rate (IMR)<sup>3</sup> of less than 20 have an average total fertility rate (TFR)<sup>4</sup> of 1.7 children. Countries with an IMR of over 100 have an average TFR of 6.2 children [Sachs and WHO, 2001, p. 36]. The ability to monitor the fetus and intervene if necessary, may be one of the reasons the global child mortality fell from 18.2% in 1960 to 4.3% in 2015 [Roser, 2018].

In fact, "children face the highest risk of dying in their first month of life, at a rate of 19 deaths per 1,000 live births" [UN, 2017, p. 1], so monitoring the fetuses to ensure a birth with a positive clinical outcome is of great importance.

---

<sup>1</sup>The demographic transition model describes how populations and societies in general change over time, due to changes in birth rates and death rates [Bongaarts, 2009].

<sup>2</sup>Neonatal death is defined as death during the first month of life [Pathirana et al., 2016]

<sup>3</sup>According to [Sachs and WHO, 2001, p. 118], IMR is the number of deaths under age 1 per 1000 live births

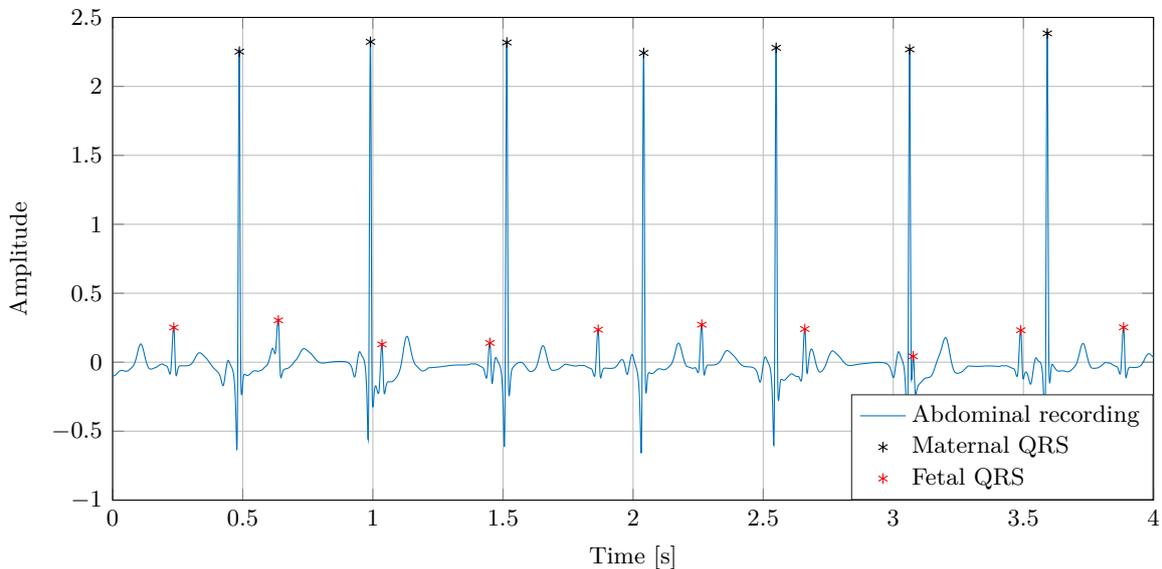
<sup>4</sup>The TFR is, roughly, the average number of children per women during her reproductive lifetime [Sachs and WHO, 2001, p. 118]

## 1.2 Aim of Thesis

While techniques such as a scalp electrode are able to obtain the FECG in an invasive matter, non-invasive methods are of great interest as they are both less risky and less complicated, compared to invasive methods [Freeman et al., 2012, p. 3-6]. Non-invasive methods build on the idea of recording the electric activity on the abdomen of the mother, which means both the mother's (maternal) and fetal ECGs are obtained simultaneously, see Figure 1.1. The difficulty with non-invasive methods is however that since both the maternal and FECGs are recorded, the FECG must somehow be extracted from the abdominal mixture. This is not a straight-forward task as the maternal ECG is much more powerful than the FECG, which is why the topic of non-invasive fetal ECG (NI-FECG) estimation is a hot signal processing research field. This work aims at developing a new method for non-invasive FECG, which is able to extract morphologically accurate FECG, as opposed to the current methods, which focus more on obtaining a good estimate of the fetal heart rate. The question which this thesis aims to solve is therefore:

*How can a morphologically accurate fetal ECG be extracted non-invasively from a set of abdominal ECG recordings?*

Emphasis will be on describing the theoretical, mathematical foundation behind the algorithms, instead of using more heuristic methods which may not be mathematically well-founded. The considerations going into choosing which extraction methods to use will also be described, in order to make this work helpful for anyone who might have an interest in designing a system to perform morphologically accurate fetal ECG.



**Figure 1.1:** A single (simulated) abdominal ECG recording, from which the maternal and fetal QRS complexes are marked with black and red crosses, respectively. For now, the QRS complex can be considered as the biggest peak of the ECG.

## 2 | Clinical Background

This chapter provides the clinical background on the topic of fetal electrocardiogram (FECG) monitoring. This includes an overview of the physiology of the adult and fetal heart, and a description of the ECG. Finally information on the topic of FECG monitoring is provided, together with a motivation for electronic fetal heart rate monitoring.

### 2.1 Cardiac Physiology

An overview of the cardiac physiology is provided in this section, with a focus on the electric part of the heart.

#### 2.1.1 The anatomy of the heart

The heart consists of four chambers: The right and left atria and the right and left ventricle. The right side of the heart collects oxygen-poor blood from the body from the superior and inferior vena cava, leading it into the right atrium. From here, it flows through into the right ventricle, and then into the lungs through the pulmonary artery. The left side of the heart handles the oxygen-rich blood - the left atrium collects oxygenated blood from the lungs, from which it flows into left ventricle and then into the aorta, supplying blood to the body [VanPutte et al., 2016, p. 323]. The chambers of the heart can be seen in Figure 2.1.

#### 2.1.2 The cardiac cells

The cells that make up the heart are called cardiac cells, which are a sub-group of muscle cells. A key property is thus their ability to contract and to conduct electric signals, signifying that other cells should contract. Cardiac cells are polarized in their resting state, meaning that their outsides are positively charged relative to the inside, by around 90 mV. In the heart, electrical signals are propagated through a process called *depolarization*, which is a process in which the cell loses the charge differential between its outside and inside. This depolarization can propagate from cell to cell, as a depolarization of a cell will trigger a depolarization of its neighboring cell, thus producing a wave of depolarization propagating through the heart. After depolarization, the reverse process occurs, called repolarization, in which the cardiac cells restore their charge differential. [Thaler, 2007, p. 10-16]

From an electrical perspective, the heart consists of three different types of cells [Thaler, 2007, p. 10-16]:

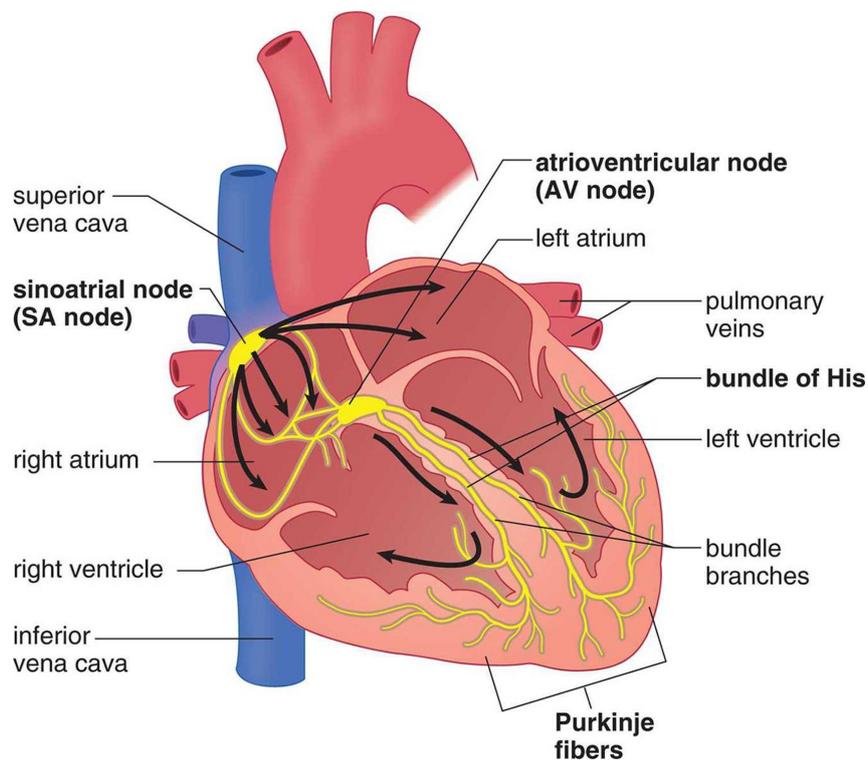
- *Pacemaker cells*: These cells depolarize (spontaneously) by themselves after a certain timespan, and act as the source of the depolarization wave.
- *Electrical conducting cells*: These cells are good (i.e. fast) at depolarizing meaning they are good electrical conductors
- *Myocardial cells*: These cells contract when depolarized, and constitute the largest part of the heart tissue.

### 2.1.3 The cardiac conduction system

A diagram showing the electric conduction system of the heart can be seen on Figure 2.1, and will be described in the following. The "pacemaker" of the heart, i.e. the part that triggers the contraction of the heart, is the sinoatrial node (also called SA node), and is located at the junction of the superior vena cava and the right atrium [Fuster et al., 2007, p. 92]. Consisting primarily of pacemaker cells, the SA node fires at a typical<sup>1</sup> rate of 60-100 times per minute, resulting in a pulse of 60-100 bpm. [Thaler, 2007, p. 14]

From the SA node, the depolarization wave propagates to the myocardial cells of the right atrium, and through Bachman's bundle to left atria, making both atria contract. The wave propagates through the atrial conducting system to the atrioventricular node (AV node), located in the posterior wall of the right atrium [Hall, 2016, p. 125]. The AV node serves the purpose of delaying the electric signal by about 130 ms, such that the atria have time to contract before the electric signal propagates to the ventricles [Hall, 2016, p. 125].

From the AV node, the depolarization wave propagates into the ventricular conducting system, by first entering the bundle of His which then splits up into the left and right bundle branches. The bundle branches terminate in the Purkinje fibers located around the ventricles, which deliver the electric signal to the myocardial fibers of the ventricles, making them contract. [Hall, 2016, p. 126]

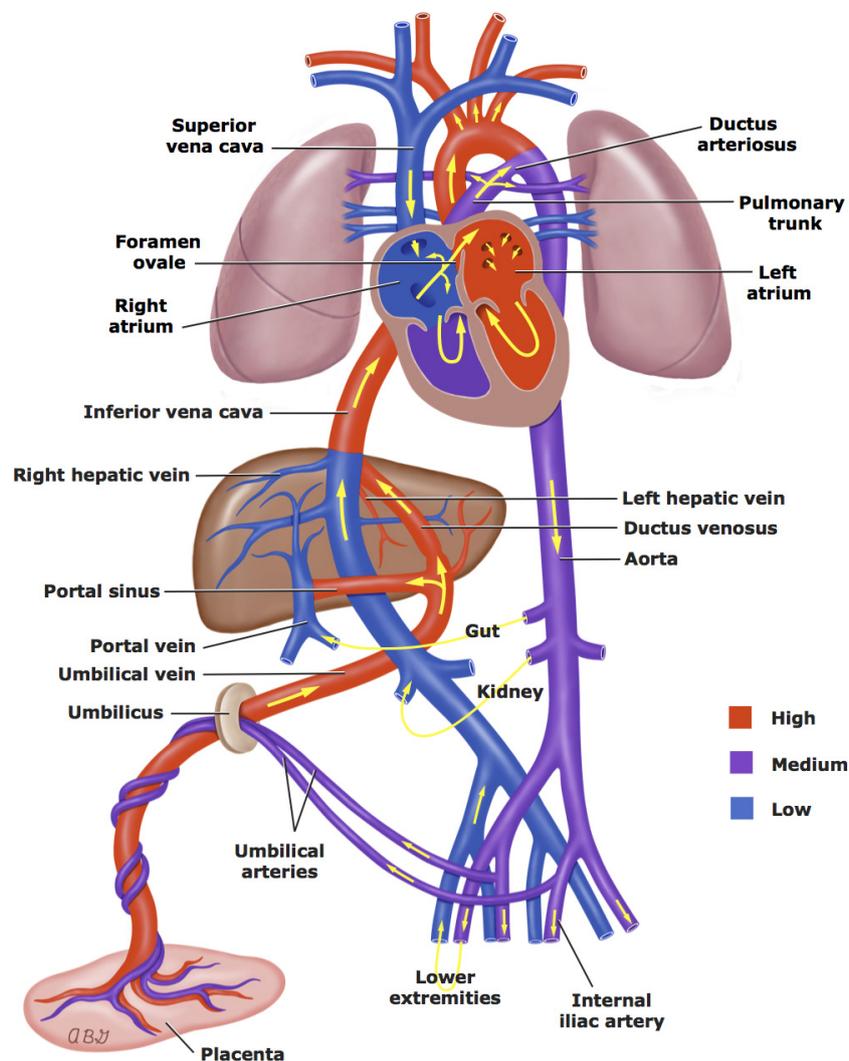


**Figure 2.1:** A diagram of the cardiac conduction system. [Anatomybodysystem, 2017]

<sup>1</sup>The rate varies depending on the activity of the autonomic nervous system, as e.g. adrenaline increases the heart rate.

## 2.2 Fetal Circulation

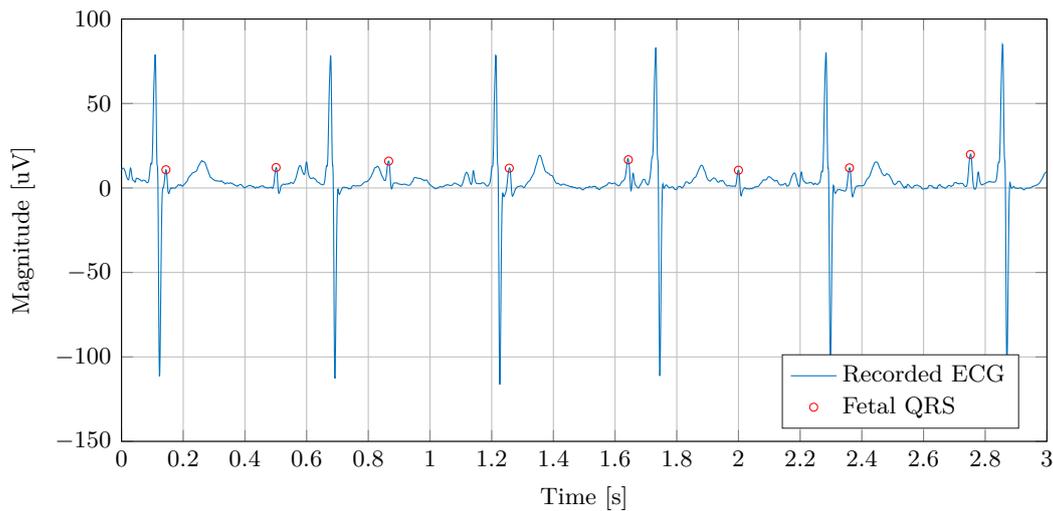
The fetal circulation can be seen in Figure 2.2, and is slightly different than that of adults, as the fetus cannot use the lungs to oxygenate the blood. Instead, it receives oxygenated blood from the placenta, which, through the umbilical vein and ductus venosus, transports oxygenated blood to the liver and via the inferior vena cava to the right atrium [Fuster et al., 2007, p. 1856]. The foramen ovale is an opening between the right and left atria, which makes it possible for some the oxygenated blood to pass into the left atrium, from which it flows into the left ventricle and into the aorta. Approx. 63% of the total ventricular output comes from the right ventricle, of which 90% goes through the ductus arteriosus and into the aorta. That way, the lungs have been bypassed, and oxygenated blood is delivered to the fetal body. To complete the circulation, part of the deoxygenated blood is returned to the placenta through the umbilical arteries. [Fuster et al., 2007, p. 1856]



**Figure 2.2:** The fetal circulation system. Red indicates high oxygen content blood, purple is medium and blue is low oxygen concentration. [Sepulveda W, 2013]

### 2.2.1 Correlation between mother and fetus

An key factor to consider in the topic of FECG is the relation between the fetus and the mother. As described in Section 2.4.2, the fetal heart rate (FHR) is around 40% higher than the maternal, meaning they are not identical to each other. Some correlation is however present, at least in the extreme cases: If the mother dies, so does the fetus, unless intervention is done in order to save the fetus. Stress responses are also transmitted from mother to fetus, affecting the fetal heart rate as shown by [Monk et al., 2000]. Likewise, it was found by [DiPietro et al., 2003] that "there is no evidence of a protective effect of diminished maternal sensitivity to stress on the fetus". One can therefore not assume the maternal and FECGs independent, however their correlation changes over time, depending on the maternal stress response. The power of the FECG signal is about 15-30 dB *lower* than the maternal one, as can also be seen from Figure 2.3, while the signal to noise ration (SNR) between FECG and the noise is -5 to -15 dB [Martens et al., 2007, p. 385]. From around week 28-32 the vernix caseosa forms, a fatty layer surrounding the fetus, which dissolves again around week 37-38. This makes the amplitude of the FECG dramatically reduce, making it almost immeasurable around week 30 [Stinstra, 2001, p. 85]. This is due to the insulating effects of the layer, making it hard for the electric currents to reach the abdomen for recording.



**Figure 2.3:** A 3-second ECG recording from the abdomen of the mother, showing both fetal and maternal ECG. Fetal QRS complexes are marked with red. Data obtained from [Physionet, 2017]

With the fetal circulation described, it is now possible to investigate why fetal monitoring is needed, as this relates to the fetal circulation.

## 2.3 The Electrocardiogram

The electrocardiogram (ECG) is "the most commonly used laboratory procedure for the diagnosis of heart disease" [Fuster et al., 2007, p. 295]. Understanding how the ECG is obtained, and what it measures, is therefore crucial. The purpose of this section is to provide an explanation of the components of the normal ECG, followed by a description of how ECGs are recorded.

### 2.3.1 The components of the ECG

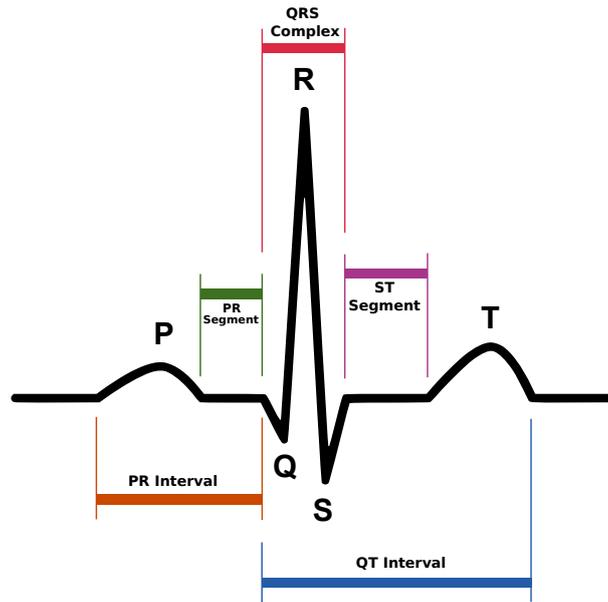
The normal ECG consists of a three main components, namely a P-wave, a QRS complex and a T wave, as can be seen in Figure 2.4. The components are briefly described in the following [Hall, 2016, p. 131]:

- *P-wave*: The P-wave comes from the depolarization of the atria. It thus signifies atrial contraction. The first part of the P-wave is related to the right atrium (as this is where the SA node is located), while the second part relates to the left atrium [Thaler, 2007, p. 20]
- *QRS-complex*: In most cases, the QRS complex consists of three separate waves, namely the Q-wave, R-wave and S-wave. The QRS complex is caused by the depolarization of the ventricles before they contract. If the first deflection is downwards, it is called the Q wave, while the upward deflection is called the R-wave. The QRS complex has a duration ranging from 6-100 ms [Thaler, 2007, p. 55], while the QRS complex of the fetal ECG has a typical duration is 20-50 ms, depending on where in the literature one looks [Allan et al., 2001, p.18-20].
- *T-wave*: The T-wave originates from the repolarization of the ventricles, and is therefore also known as the repolarization wave.

Note that a repolarization wave of the atria also occurs, but since this happens during ventricular depolarization, the wave coming from it is hidden by the QRS complex. As repolarization (the T-wave) occurs in the opposite direction of depolarization, but with a reversed electrical sign, the repolarization wave will also appear as a positive deflection on the ECG.

The flat lines between waves are called segments, and signify electric silence. One such segment is the ST segment, which is the line between the S-wave and the T-wave. This measures the time interval from the end of ventricular depolarization to the start of ventricular repolarization. [Thaler, 2007, p. 29]

Intervals denote the time between events, and include at least one wave and the connecting straight line/segment. The PR interval measures the time from the start of atrial depolarization to the start of ventricular depolarization, and has a duration of about 160 ms [Hall, 2016, p. 125]. The QT interval is the time interval between the onset of the QRS complex and the offset of the T-wave, and measures the time from beginning of ventricular depolarization to the end of ventricular repolarization [Thaler, 2007, p. 29]. It has a normal duration of 350 ms. [Hall, 2016, p. 126]



**Figure 2.4:** The normal sinus rhythm for the human heart [Atkielski, 2007; Fuster et al., 2007]

In Appendix A, a description of the electrocardiographic leads can be found, which are the standard way of obtaining an ECG. With the normal ECG of an adult described, the fetal ECG can now be described, which is done in the following section.

## 2.4 Fetal ECG Monitoring

The fetal electrocardiogram can be used as a tool to evaluate the well-being of the fetus. In this section, a brief overview of the diseases related to fetal distress (and which might be prevented using fetal ECG monitoring) is given, together with a description of the two classical types of fetal ECG monitoring, namely the fetal heart rate (and its variability) and the morphology of the ECG.

### 2.4.1 Motivation for non-invasive fetal monitoring

As mentioned in Section 2.2, the fetus gets its supply of oxygen from the mother, through the placenta. During labour, strong uterine contractions may stop the flow of oxygen from the placenta, thus reducing the oxygen flow to the fetus [Neilson, 2015]. A reduction of oxygen flow can lead to hypoxia, which is the medical term for a lack of oxygenation at the tissue level [Samuel and Franklin, 2008]. The umbilical cord may also be compressed during labour, which may stress the fetus [Neilson, 2015].

According to [Freeman et al., 2012, p. 28-31], fetal hypoxia may lead to a wide variety of diseases: Cerebral Palsy (CP)<sup>2</sup>, mental retardation, learning disorders, respiratory distress syndrome, kidney damage and damage to the digestive system.

Having a method to monitor the fetus in order to detect fetal hypoxia is therefore of great interest, such that medical staff can intervene and improve the clinical outcome.

---

<sup>2</sup>CP is a neurological disorder caused by brain malformation during the development of the child's brain, or a low level of oxygen (hypoxia) in the brain during its development. CP primarily affects body movement and muscle coordination [MedlinePlus, 2017]

Around the 1840's it was proposed that fetal heart rates (FHR) might be an indicator of the wellbeing of the fetus [Freeman et al., 2012, p. 1]. However, it was later found that the term "fetal distress" was an inappropriate term to describe FHR patterns, and FHR patterns thought to be associated with fetal hypoxia were actually found not to lead to newborns with signs of fetal hypoxia. [Freeman et al., 2012, p. 2]

One reason for this might be the uncertainty related to obtaining the FHR: A simple experiment in which 15 obstetricians were to count the FHR of an artificial fetus showed a wide range of counts between the obstetricians, with error rates up to 30% [Jenkins, 1989, p. 210].

Based on this, Hon [1958] suggested in 1958 that the use of electronic techniques for the continuous evaluation of FHR in labour would permit a more accurate indication of fetal distress than clinical methods (of more sporadic nature).

Building on this, the fetal scalp electrode was invented in 1963, making direct ECG monitoring of the fetus possible, but at the expense of having to insert the electrode into the vagina and cervix against the fetus' scalp, thus being more risky and complicated than non-invasive methods [Freeman et al., 2012, p. 3-6]. Another problem with the scalp electrode is that it can only be used during labour, as it requires a rupture of membranes (i.e. a punctured amniotic sac).

For these reasons, it would be advantageous to have non-invasive monitoring methods. The two most popular types of monitoring are described in the following, namely monitoring of the fetal heart rate and fetal morphology.

## 2.4.2 Fetal heart rate (FHR) & FHR variability

By analyzing 1.5 billion individual fetal heart rate (FHR) measurements from 78852 CTG<sup>3</sup> tracings, [Pildner von Steinburg et al., 2013] found that the normal<sup>4</sup> FHR range is 120 to 160 bpm. This fits well with the baseline FHR mentioned in [Macones et al., 2008, p. 663]: "Abnormal baseline is termed *bradycardia* when the baseline FHR is 110 bpm; it is termed *tachycardia* when the baseline FHR is 160 bpm."

Baseline FHR *variability* is defined as "fluctuations in the baseline FHR that are irregular in amplitude and frequency" [Macones et al., 2008, p. 663], and is an important measure of the well-being of the fetus.

As described in Section 2.4.1, fetal hypoxia and hypoxemia are critical diagnoses which can greatly affect the life of the newborn. For this reason, it is of great interest to monitor the fetal heart rate (FHR), as fetal hypoxia is related to the FHR variability: "Moderate FHR variability reliably predicts the absence of fetal metabolic acidemia<sup>5</sup> at the time it is observed" [Macones et al., 2008, p. 663].

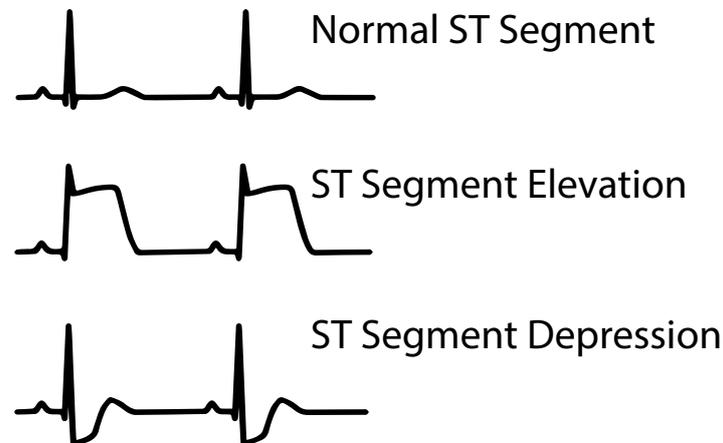
If the FHR variability is absent and the FHR is bradycardia, the FHR tracing is called *abnormal*, calling for immediate action, such as providing the mother with excess oxygen [Macones et al., 2008, p. 664].

---

<sup>3</sup>Cardiotocography is described in greater detail in Section 2.4.4

<sup>4</sup>Normality is here used in a statistical sense, defined as the 5th to 95th percentile

<sup>5</sup>Acidemia and acidosis means there is a high hydrogen ion concentration in the tissues, i.e. the body is making too much acid or cannot get rid of it. In fetuses, hypoxia may lead to acidosis. [Bobov and Soothill, 1999]



**Figure 2.5:** A normal ST segment as well as an elevated and depressed ST segment. Figure modified from [OPENPediatrics Staff, 2015].

### 2.4.3 ECG morphology

Looking at the shape of the ECG (morphology) is also a useful diagnosis tool for fetal hypoxia: "During severe intrapartum<sup>6</sup> hypoxia and metabolic acidosis, there is a significant shortening of the QT interval. The intrapartum QT interval may therefore provide additional information on the condition of the fetus" [Oudijk et al., 2004]. Another interesting finding by [Oudijk et al., 2004] is that the QT interval is not dependent on the FHR or the general stress of the fetus. This might make it a better indicator of fetal acidosis than the FHR: "Fetuses without metabolic acidosis did not show a QT shortening despite fetal heart rate changes indicative of such events."

It was shown by [Pardi et al., 1971] that ST segment depression (found by averaging multiple ST segments) is related to fetal hypoxia, although it has not been possible to develop the technique enough to make it clinically applicable [Freeman et al., 2012, p. 2]. An example of ST depression (and elevation) can be seen in Figure 2.5. From the literature, it is not fully clear how useful ST elevation monitoring is for ensuring the well-being of the fetus: "In comparison to continuous electronic fetal heart rate monitoring alone, the use of adjunctive ST waveform analysis made no obvious difference to primary outcomes" [Neilson, 2015]. However, experiments on fetal sheep show that hypoxia is linked to ST elevation [Sameni and Clifford, 2010, p. 7], indicating that the morphology of the fetal ST analysis may be of interest.

Apart from the QT-interval and the ST-segment, the PR and RR intervals are also of interest: "Shortening of the PR interval, and an inversion of the normal positive relationship between the PR and RR intervals, have been found to be early indicators of fetal stress, with ST segment and T-wave changes being later indications of fetal suffocation" [Jenkins, 1989, p. 213].

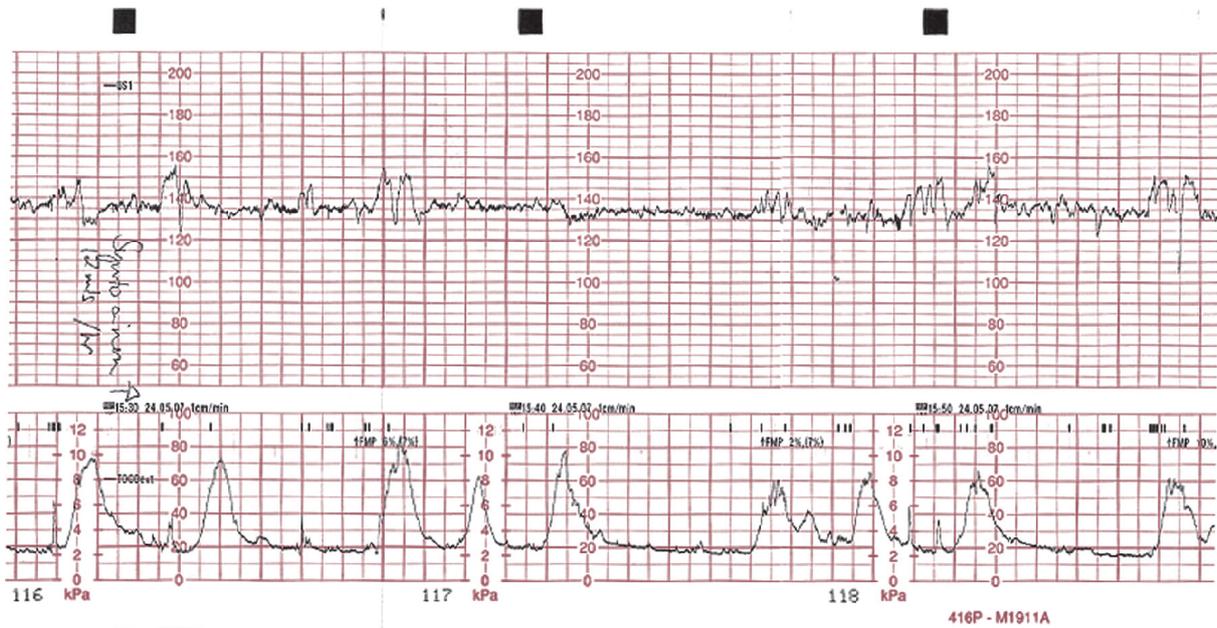
It can thus be concluded that by obtaining a full fetal ECG is of great interest, as it is a helpful diagnosis tool. It is now interesting to have a brief look at Cardiotocography, one of the most commonly used methods for fetal monitoring.

---

<sup>6</sup>Intrapartum means during birth

## 2.4.4 Cardiotocography (CTG)

Cardiotocography is a method to monitor the fetal heart rate using ultrasound, together with the uterine contractions by measuring the tension of the abdominal wall [Debdas, 2013, p. 1, 14]. CTG is a commonly used method for monitoring: 85% of births in the US get monitored by CTG [Debdas, 2013, p. 1].



**Figure 2.6:** An example of a CTG trace, showing fetal heart rate in the top (in BPM) and the uterine contractions in the bottom. [Byrom, 2012; Fraser and Cooper, 2012]

In a review by [Grivell et al., 2012], no difference in outcome could be found when looking at potential improvements through the use of CTG monitoring. However, when the interpretation of CTG traces was computerized (i.e. made automatic) "the findings looked promising" [Grivell et al., 2012; Pildner von Steinburg et al., 2013]. This suggests that the CTG traces are open for interpretation by the medical staff, and that standardization of their interpretation might help improve the fetal outcome [Pildner von Steinburg et al., 2013].

Monitoring the fetal heart rate using CTG during birth showed a "reduction of perinatal mortality by 50%, but an increase of operative intervention by factor 2.5" [Pildner von Steinburg et al., 2013]. It was found by [Neilson, 2015] that monitoring the baby using ECG and CTG resulted in less surgical assistance with the birth, than with CTG alone. However, there was no difference in the number of caesarean deliveries and little to suggest that babies were in better condition at birth.

Performing some kind of monitoring of the fetal heart rate is thus advantageous, however the effect of using ECG and CTG is rather limited. One reason for this might be "because clinicians have been unable to interpret recordings in an unbiased fashion" [Jenkins, 1989, p. 213]. According to [Jenkins, 1989], more advanced methods might provide better result and improved outcomes are possible by using methods with less room for interpretation than CTG. This serves as a motivator for developing new, more accurate methods for (non-invasive) fetal monitoring. As mentioned, a popular method for fetal monitoring is estimating the fetal heart rate, which is described in the following section.



# 3 | Non-Invasive Fetal Heart Rate Estimation

As it could be seen from the end of the previous chapter, especially Section 2.4.1, a lot of research has been done on the topic of fetal ECG monitoring. The purpose of this chapter is to describe the methods which have previously been used for doing non-invasive fetal heart rate monitoring. This is relevant to investigate even though this thesis is on the topic of morphologically accurate fetal ECG and not FHR estimation. The reason for this is that both the extraction methods and datasets are similar, and observations made about the current extraction methods can therefore also be used when dealing with morphology studies.

In the first part of this chapter, the datasets available for this topic are described. Then, a description of the PhysioNet/Computing in Cardiology Challenge 2013 follows, which is an important milestone in the topic of non-invasive fetal ECG from recent years. Finally, the actual methods used in the literature are described, together with their pros and cons.

## 3.1 Databases of Fetal ECG Recordings

An important part of developing algorithms is having data available for testing the performance of the algorithm. Two overall types of data can be used: Measured (clinical) data and simulated data, which is also true in the topic of fetal ECG. Below is a brief overview of key databases for fetal ECG monitoring presented.

### 3.1.1 Clinical data

Testing the developed algorithm on measured data is of great importance, as the algorithm will eventually be applied to real measurements. This type of data thus gives the most realistic performance measure. In order to really use the measurements, having the ground truth (by means of separate fetal recordings or annotated data) is crucial, which is often challenging.

#### **Abdominal and Direct Fetal Electrocardiogram Database (ADFECGDB)**

The ADFECGDB [Jezewski et al., 2012; Goldberger et al., 2000] contains four channels of maternal abdominal ECG recordings from 5 women. In addition, each recording also contains an ECG measured directly from the fetal scalp (scalp electrode). This allows for directly comparing the extracted fetal ECG from the abdomen recordings with the ground truth fetal ECG from the scalp. The recordings are sampled at a sample rate of 1 kHz, with a 16 bit resolution.

#### **Non-Invasive Fetal Electrocardiogram Database (NIFECGDB)**

The NIFECGDB [Physionet, 2016; Goldberger et al., 2000] contains 55 multichannel recordings from a single woman between 21 to 40 weeks of pregnancy. Each recording contains 4 abdominal signals and two thoracic<sup>1</sup> signals. The recordings are sampled at a sample rate of 1 kHz, with a 16 bit resolution.

---

<sup>1</sup>Recorded at the thorax, i.e. in the chest region.

The advantage of this database is the fact that it contains thoracic recordings. These are placed far away from the fetus, and should therefore not contain any fetal ECG, or only fetal ECG of very small amplitude. This will allow testing methods such as adaptive filters, which require knowledge of the original signal.

### 3.1.2 Simulated data

By using simulated data, one can control exactly how the signals are generated, which is often a big advantage. This also means that the maternal and fetal ECG are known separately, and comparing the output of the algorithm to a ground truth is thus possible. This makes it possible to better evaluate the performance of the algorithm.

A simulation framework called FECGSYN has been presented in [Behar et al., 2014b], which makes it possible to simulate mixtures of fetal and maternal ECG. A wide variety of simulation parameters can be set for both the mother and fetus, such as heart rates, respiration and the heart position. Various noise types can also be included such as baseline wander and electrode movement. The simulator models the hearts as punctual dipoles, which can be moved and rotated.

#### Fetal ECG Synthetic Database (FECGSYNDB)

Developed by [Andreotti et al., 2016; Goldberger et al., 2000], the FECGSYNDB uses the simulation framework from [Behar et al., 2014b], to generate a total of 1750 synthetic signals, totaling 145.8 hours of multichannel data, with 1.1 million fetal peaks. Each simulation has a duration of 5 minutes, and is sampled at 250 Hz with a 16-bit resolution.

Ten different pregnancies were simulated when generating the dataset, and for each simulated pregnant subject, seven different physiological events were simulated, see Table 3.1.

Case	Description
Baseline	Abdominal mixture (no noise or events)
0	Baseline (no events) + noise
1	Fetal movement + noise
2	MHR /FHR acceleration / decelerations + noise
3	Uterine contraction + noise
4	Out-of-place (ectopic) beats (for both fetus and mother) + noise
5	Additional NI-FECG (twin pregnancy) + noise

**Table 3.1:** The seven different simulation settings provided by the FECGSYNDB.

For each pregnancy and each case, 5 different levels of additive noise were applied, and 5 repetitions were made for each combination, yielding the total of 1750 recordings. Three different types of noise can be added [Behar et al., 2014b, p. 1542]: Motion artifact, electrode motion (EM) and baseline wander (BW). In the database, only motion artifact noise is added, which has been generated by fitting a 12'th order auto-regressive (AR) model to a recording of the motion artifact noise. This is a commonly used method to accurately model physiologic and measurement noise, as is done by e.g. [Sæderup et al., 2018]. The poles of the model are allowed to wander a bit, in order to incorporate some variability between recordings.

As can be seen from Table 3.1, it is possible to test the ECG algorithms on very different types of data, varying from a simple mixture of maternal and fetal ECG in ideal conditions, to extreme conditions such as twin pregnancies. This makes it possible to stress-test the algorithm on known signals, making it possible to test its performance in both ideal and non-ideal conditions.

### 3.1.3 PhysioNet/Computing in cardiology challenge 2013

The PhysioNet/Computing in Cardiology Challenge 2013 (or just *the Challenge*) dealt with the topic of noninvasive fetal ECG (NI-FECG). The key questions of the Challenge were: "1) Can accurate FHR measurements be performed using a set of non-invasive abdominal ECG electrodes? and 2) Can an accurate fetal QT measure be performed in an automated way using the extracted signal?" [Clifford et al., 2014, p. 1525]. Two key problems existed in the topic of NI-FECG, namely a lack of data with a ground truth and expert annotations, and a lack of a methodology for assessing the algorithms, both of which the Challenge sought to overcome.

#### Dataset construction

The data used in the Challenge consists of 447 one-minute recordings sampled at 1 kHz, each consisting of 4 abdominal channels. Some of recordings in the database originate from the databases described above, while others have been donated to the Challenge. This means both measured and simulated data are part of the dataset. The data has been annotated using either reference annotations from the dataset, or by crowdsourcing, thus obtaining properly annotated data. The 447 recordings are split up into three sets: [Clifford et al., 2014, p. 1526]

- Set A: 75 recordings. Both records and expert annotations publicly available.
- Set B: 100 recordings: Only the recordings are publicly available.
- Set C: 272 recordings: Both recordings and annotations are private.

The Challenge consisted of five *events*, each with their own purpose and with either public or private test sets as summarized in Table 3.2.

Event	E1	E2	E3	E4	E5
Fetal Heart Rate	x			x	
RR interval		x			x
QT interval			x		
Hidden test set (set C)?	x	x	x		

**Table 3.2:** Overview of the Challenge events, showing their purpose and if they were scored/tested on a publicly available test set. [Clifford et al., 2014, p. 1526-1527]

The PhysioNet 2013 challenge is thus a combination of the above databases, and is a good benchmark database, as it is possible to compare algorithm performance to those that participated in the Challenge. A wide variety of algorithms were submitted to the Challenge, with the purpose of extracting the fetal heart rate, with some of the submitted algorithms being available as open-source code, so they can be used as a benchmark. An overview of the general approach of these methods is given in the following.

## 3.2 Overview of Current Methods

Based on the techniques presented in the Challenge, the authors have summarized that the submitted fetal HR estimation algorithms in general consisted of 5 steps, as shown in Figure 3.1 [Clifford et al., 2014, p. 1528].



**Figure 3.1:** The five steps used in the current methods for FHR estimation.

These will be described in further details in the following sections, followed by a description of the single-channel methods that can also be applied to obtain NI-FECG.

### 3.2.1 Pre-processing

The purpose of this step is to remove noise, baseline wander and noise from power lines (i.e. 50 Hz noise). Some have created augmented channels (as is done with the normal 12-lead ECG, see Appendix A) by subtracting some of the four abdominal channels from others.

From [Clifford et al., 2014, p. 1521] it was concluded that generally, the fetal ECG extraction methods employed later are highly dependent on the pre-processing methods.

To remove the baseline wander, [Lipponen and Tarvainen, 2013] used a 6th order Butterworth high-pass filter (cut-off at 2 Hz). [Varanini et al., 2013] have used a more advanced method, using multiple median filters, and a 50 Hz notch filter, including 2nd and 3rd harmonics.

From [Behar et al., 2014d] it was found that selection of a high cut-off frequency for the high-pass filtering, such as 10 Hz, led to an improved result in detection, possibly due to the removal of the large maternal P and T wave components. A low-pass filter with a cutoff of 99 Hz was applied by [Behar et al., 2014d, p. 15].

Some have removed power line noise by removing the 50 Hz Fourier coefficient, while other have applied a 50 Hz notch filter [Clifford et al., 2014]. Others have made a filter automatically removing either the 50 Hz or 60 Hz component, as some of the recordings in the Challenge dataset are from Europe while others are from the US [Behar et al., 2014d, p. 5].

The signal is typically also normalized - this is usually done by simply scaling the amplitude to the range  $[-1; 1]$ . An interesting normalization procedure is presented by [Behar et al., 2014d, p. 5], who use the the hyperbolic tangent in order to keep the signal in a controlled range. This should make the FQRS detection more accurate by shrinking the high amplitude artifacts in the ECG signal range, as the hyperbolic tangent "squeezes" the signal into  $[-1; 1]$ .

### 3.2.2 Estimation of maternal component

As the biggest problem in detecting the fetal ECG is the presence of the maternal ECG which may have 20 dB larger magnitude than the fetal ECG, see Section 2.2.1, removing the maternal component must be handled. This is typically done by first estimating the maternal part, which can be done using either subspace methods or template methods.

#### Subspace methods

One of the most popular methods from the Challenge for extracting the fetal ECG, were subspace methods. In particular Independent Component Analysis (ICA), Principal Component Analysis (PCA) and Singular Value Decomposition (SVD) were used, which are methods for doing blind source separation (BSS)<sup>2</sup>. The overall idea of BSS is to decompose the signals (i.e. the four channels) into a series of subspaces in such a way that the components of the mixture are put in separate subspaces - ideally, one would like to get the maternal ECG into one subspace, the fetal ECG into another, and the noise into a third.

One might argue that BSS is not a subspace method but a statistical method as one often use statistical measures when doing BSS. But as the sources are separated from each other, they can be interpreted as being in different subspaces. Purely subspace methods such as PCA are often applied as a preprocessing step for BSS, and they are therefore collectively described as subspace methods in this thesis.

The overall difference between PCA and ICA is that PCA uses second-order statistics (i.e. the covariance matrix) and attempts to transform the data into a set of uncorrelated components by using the eigenvalue decomposition (EVD), thus forming an orthogonal basis for the covariance. ICA on the other hand attempts to transform the signals into *independent* components - a much stronger requirement than the signals being uncorrelated - which requires the use of higher-order statistics such as kurtosis [Hyvärinen et al., 2004, p. 7-9]. The SVD is simply a variant of PCA, as we know the SVD and EVD are linked in such a way that the singular values  $\sigma_k$  of the data matrix  $\mathbf{X}$  are the square roots of the eigenvalues  $\lambda_i$  of the (covariance) matrix  $\mathbf{X}^T\mathbf{X}$  [Petersen and Pedersen, 2012, p. 32].

Having separated the mixed signal into subspaces, it is then possible to perform e.g. QRS detection on the maternal ECG subspace alone.

A problem with these subspace decompositions is that the algorithms using them may make the following implicit assumptions [Clifford et al., 2014, p. 1529-1530]:

- The number of signal sources are fixed, discrete, and less than or equal to the number of recorded channels
- The subspace representation is stationary
- The sources are uncorrelated
- The maternal signal has a high signal-to-noise ratio and spans one or more of the dominant spaces

---

<sup>2</sup>Blind source separation is the task of separating a mixture of signals into the original sources with only weak a-priori information about them such as independence [Comon and Jutten, 2010, p. xxi]

This is problematic as the fetal signal, interference and noise may not be linearly separable [Sameni, 2008, p. 10]. Results from [Behar et al., 2014d, p. 17] do however show a rather good performance in practice when using them, see Section 3.3.2.

Another interesting subspace method is Periodic Component Analysis ( $\pi$ CA), which exploits the periodicity that exists in the ECG signal. The idea is to extract "the most periodic" linear components of the observed mixture, which gives better performance than classical source separation algorithms such as ICA and PCA. In  $\pi$ CA, the components are ranked based on their synchronization with the R-peaks, whereas the order of components in regular ICA is random [Sameni et al., 2008]. This is interesting in ECG applications, as the components mostly synchronized with the R-peaks, must be the components making up the maternal ECG.

### Template methods

Another typically used method is templates subtraction (TS). The idea is to make a template of the maternal ECG (MECG) waveform, by averaging the maternal QRS across time and/or channels. This requires detection of the maternal ECG, which is typically done using Pan-Tompkins algorithm [Pan and Tompkins, 1985]. The templates are typically modified over time, in order to adapt to the non-stationary nature of ECG [Behar et al., 2014c]. In order to use templates, one must assume the maternal and fetal components are uncorrelated, and that the morphology of the heart beats are either constant or slowly changing. One can also use (Extended) Kalman filtering designed to track the MECG, which can then be subtracted from the recording, leaving the FECG. This makes it similar to a TS approach [Behar et al., 2014d, p. 13].

The subspace and template methods can also be combined, by e.g. using ICA to decompose the signal to get a "maternal ECG"-channel, from which the maternal QRS complexes (and thus the heart rate) is found, after which templates of the maternal signal can be made [Clifford et al., 2014, p. 1521].

### 3.2.3 Removal of maternal component

After having found the maternal component of the mixture, it must be removed in order to be left with the fetal ECG. For the subspace methods (PCA, ICA), this is done simply by setting the components of non-fetal subspaces to zero [Clifford et al., 2014, p. 1529]. This does however require that one determines which of the components belong to the maternal ECG, which can be challenging.

For the template methods, the QRS template is subtracted from the recording, at the places of the maternal QRS. One may do this adaptively, such that the changes of the ECG are tracked, using e.g. Kalman filtering [Clifford et al., 2014, p. 1529].

### 3.2.4 Estimation of FHR and RR time series

The fetal QRS complexes are typically estimated by either making a simple R-peak amplitude thresholding, or looking at the RS slope. Another option is to use modified versions the QRS estimation algorithms used for adult ECG. The methods can also be combined using a voting system [Clifford et al., 2014, p. 1529]. Others have treated detection as an optimization problem, in which FECG morphology and beat-to-beat interval consistency were considered in the cost function [Clifford et al., 2014, p. 1530].

### 3.2.5 Post-processing

The post-processing step deals with correcting the estimated FHR estimates, by using the a-priori knowledge about the FHR. One such method is to remove beats with an interval outside the range of 70 to 300 ms, as these are not physiologically realistic heart rates [Clifford et al., 2014, p. 1530]. Moving average filters can also be used in order to smoothen the FHR, again due to the physiological constraints to the heart rate variability. However, smoothing only gave a 1% performance improvement [Behar et al., 2014d, p. 15].

### 3.2.6 Single channel methods

As opposed to the subspace methods used in the Challenge mentioned above, the methods presented in this section are using temporal filtering techniques. These methods do not need the recording of multiple channels as (most of) the subspace methods do, but can instead work using just a single recorded abdominal channel. They do however also require a reference signal, such as a chest signal from the mother.

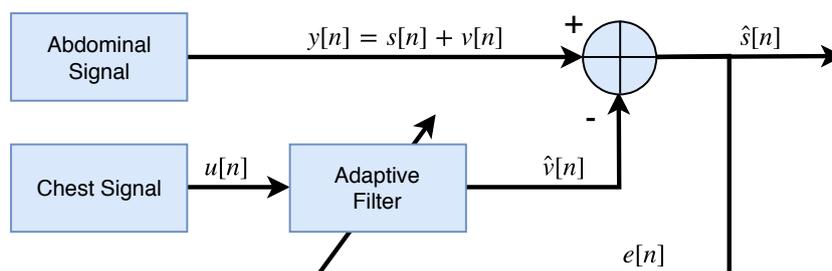
The channel model considered for an adaptive filter is of the additive type:

$$y[n] = s[n] + v[n] \quad (3.1)$$

where:

- $y[n]$  is the observed signal
- $s[n]$  is the signal of interest
- $v[n]$  is the undesired signal (noise and interference)

Here, the adaptive filter aims at learning a model with input  $u[n]$  and output  $\hat{v}[n]$ , such that the mean squared error between  $y[n]$  and the estimate of the undesired signal  $\hat{v}[n]$  is minimized, as illustrated in Figure 3.2. In the case of fetal ECG, the input  $u[n]$  is the chest signal,  $y[n]$  is the abdominal signal and  $\hat{v}[n]$  is the estimated noise + maternal ECG.



**Figure 3.2:** Adaptive noise cancelling with a single reference input  $u(n)$ .  $\hat{v}(n)$  is estimated noise,  $e(n)$  is estimation error and  $\hat{s}(n)$  is the estimate of the fetal ECG signal  $s(n)$ . Figure adapted from [Behar et al., 2014c, fig. 2]

Different classical methods exist for doing this, such as the least mean square (LMS) filter and recursive least square (RLS), which will not be described in further details here.

A more recent method for adaptive filtering is by using recurrent neural networks (RNNs), which are

neural networks capable of non-linear modeling of dynamical systems [Behar et al., 2014c, p. 1343]. In the comparison of these methods made by [Behar et al., 2014c], a variant of RNNs called Echo State Networks (ESN)<sup>3</sup> is used, as this outperformed classical RNNs in many tasks [Lukoševičius and Jaeger, 2009].

The results they obtained are mentioned in the following section, together with a comparison between the other methods for extraction of FECG.

### 3.3 Comparison of Current Methods

The purpose of this section is to compare the above-mentioned methods from both an empirical and theoretical perspective. In order to compare the algorithms, the measures by which the comparison will be done must be established, which is done in the following section.

#### 3.3.1 Evaluation methods

In the Challenge, three factors were of interest in order to evaluate the performance of the algorithms: The beat-to-beat fetal heart rate, the time interval between two RR-peaks and the average QT interval per recording.

When calculating the classification error of heart beats, a *matching window* is normally used. This is done in order to define whether or not an estimated R-peak is close enough to the annotated one, i.e. if it has been correctly or incorrectly classified. In normal adult ECG this window has a length of 150 ms [AAMI, 2012], but to account for the higher FHR, a matching window of 50 ms can be used instead [Behar et al., 2014d, p. 4].

As pointed out by [Behar et al., 2014d, p. 15-16], an inherent problem with the Challenge scoring being based on RMS values (for the RR peaks) is that it encourages estimating the FHR closer to the average FHR, while small variations in the HR, which are often the most interesting, are not rewarded in the scoring. The problem is thus that neither of the scores take the absolute location of the FQRS into account: An algorithm will thus get a good score if it just estimates the average FHR, even though it hasn't estimated the location of a single FQRS correctly, see Table 3.4.

In order to solve this, the two measures, *sensitivity* (Se) and *positive prediction value* (PPV), are introduced as suggested in the guidelines by [AAMI, 2012]. They are defined as [Behar et al., 2014c, p. 1345]:

$$Se = \frac{TP}{TP + FN} \quad PPV = \frac{TP}{TP + FP} \quad (3.2)$$

where:

TP	are true positives
FN	are false negatives
FP	are false positives

These provide ways of detecting if the fetal heart beat have been correctly classified.

---

<sup>3</sup>ESNs remove the need for back-propagation, as it uses a static, randomly generated RNN called the reservoir. Input signals are applied driving nonlinear activations from each neuron, from which the output signal is generated as a trainable linear combination of these neuron responses. [Lukoševičius and Jaeger, 2009, p. 128]

A single measure of the performance of the algorithm is the  $F_1$ -score, which calculates the harmonic mean between the PPV and Se [Behar et al., 2014c, p. 1346]:

$$F_1 = 2 \cdot \frac{PPV \cdot Se}{PPV + Se} = \frac{2 \cdot TP}{2 \cdot TP + FN + FP}, \quad (3.3)$$

from which it can be seen that FN and FP are equally penalizing the measure, which is desired.

Another measure as introduced by [Andreotti et al., 2016, p. 634] is the *mean absolute error* (MAE), which is defined as:

$$MAE = \frac{1}{TP} \sum_{i=1}^{TP} |d_i - \hat{d}_i| \quad (3.4)$$

where:

MAE	is the mean absolute error	[ms]
TP	is the number of true positive QRS detections	
$d_i$	is the time of the reference annotation	[ms]
$\hat{d}_i$	is the time of the detected annotation	[ms]

The advantage of the MAE is that it is able to describe the distance between estimated and true annotations, i.e. quantify any jitter in the estimated annotations, instead of a simple classification measure.

### 3.3.2 Overall performance

With the performance measures described in the previous section, the results can now be compared. This is done by first doing an overall comparison, and then looking into their quantifiable performance measures using the aforementioned scores.

In Table 3.3, a comparison of the most popular methods is listed. The table is sorted into single- and multichannel methods. Single channel methods include those mentioned above, e.g. adaptive filtering and Bayesian filtering such as the Kalman filter. Multichannel methods are the subspace methods mentioned above, i.e. PCA, ICA,  $\pi$ CA and a method not mentioned here, called deflation<sup>4</sup>.

---

<sup>4</sup>Deflation is an algorithm which iteratively performs subspace decomposition, then noise reduction (in the subspaces), and then performs recomposition to get back to the original domain [Sameni, 2008, p. 92].

	Method	Overall Performance	SNR improvement	Computational Cost	Realtime*	Implementation Complexity	Operator interaction	Applications/Comments
Single-Channel	FIR/IIR Filtering	Low	Low	Low	Yes	Simple	Not required	Simple frequency domain denoising
	Wiener Filtering	Medium	Medium	Medium	Possible	Medium	Not required	Denoising using morphologic or spectral priors; realtime implementation is only possible for the causal Wiener filter
	Wavelet Denoising	Medium	Medium	Low	Yes	Medium	Not required	More robust denoising for signal-noise mixtures having different scales
	Adaptive Filtering	Medium	Medium	Median	Yes	Medium	Sometimes required	Low quality maternal ECG cancellation or fetal ECG enhancement for fetal HRV analysis or average morphologic studies; extendable to multichannels
	Nonlinear Filtering	Medium	Medium/High	High	No	Complex	Required	Maternal ECG cancellation or fetal ECG enhancement in single channel recordings or degenerate multichannel mixtures
	Bayesian Filtering	High	High	Median	Yes	Complex	Sometimes required	ECG denoising, maternal ECG removal or fetal ECG enhancement having the R-peaks of the signal and with rather consistent morphologic shape; extendable to multichannels
Multichannel**	PCA/SVD/ Factor Analysis	Low	Low	Low	No	Simple	Not required	Dimension reduction for high-dimensional data and noise removal
	ICA	Medium	Medium	Medium/High	No	Medium/Complex	Required	Blind or semi-blind decomposition of multichannel maternal-fetal mixtures without considerable <i>a priori</i> information and in relatively high SNRs SNR improvement not guaranteed
	$\pi$ CA	Medium	High	Low	No	Simple	Not required	Decomposition of multichannel ECG mixtures having the R-peaks of the signal; dimension reduction with minimal ECG information loss
	Deflation	High	High	Medium/High	No	Complex	Sometimes required	Decomposition of (possibly) degenerate multichannel mixtures of desired and undesired subspaces having <i>a priori</i> information about the signal-noise structures

\* Here we have not considered the possibility of block delayed realtime analysis.

\*\* The indicated performances and complexities of multichannel methods are generally higher than single-channel methods; the qualitative descriptions should only be compared within each group separately.

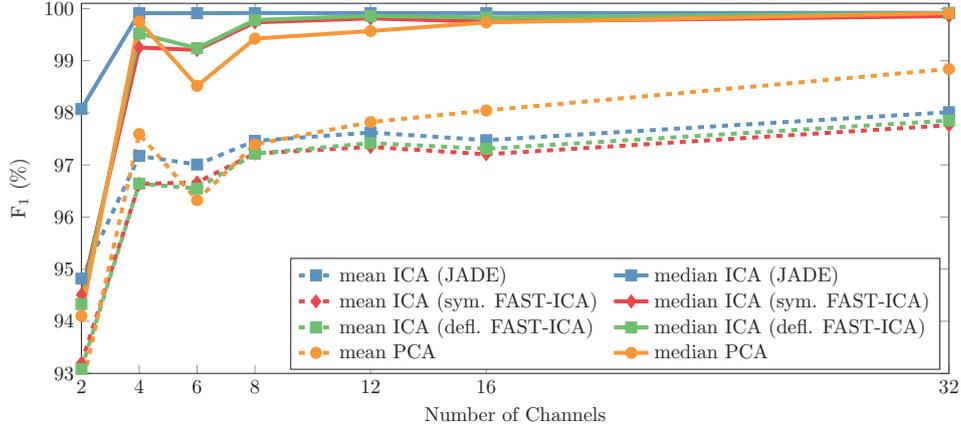
**Table 3.3:** Comparison of the different methods and their applications. Source: [Sameni, 2008, Table 10.1]

What can be concluded from Table 3.3 is that the Bayesian filtering approach gives quite good performance compared to e.g. adaptive filters. Looking at the subspace methods, PCA performs poorly, and ICA slightly better.  $\pi$ CA performs rather well, as does deflation.

With this overview of their performances, a more quantifiable comparison can be made, by evaluating the performance measures mentioned above for each algorithm.

### 3.3.3 Empirical comparison

In the work by [Andreotti et al., 2016], it was investigated how the  $F_1$  score changed as a function of the number of abdominal channels. They found that as soon as at least 8 channels were used, performance increase was negligible, but when using as few as four channels, very good  $F_1$  scores were obtained, as can be seen in Figure 3.3. For this reason, 8 channels were used throughout the tests by [Andreotti et al., 2016].



**Figure 3.3:** BSS performance with respect to the number of abdominal channels, using PCA and ICA. ICA’s performance was assessed using two different algorithmic implementations (JADE and FAST-ICA). Figure taken from [Andreotti et al., 2016, fig. 3]

### Results on Challenge dataset

In the paper by [Behar et al., 2014d], many of the multichannel methods were implemented and compared using the Challenge dataset. The results of a selection of the algorithms tested can be seen in Table 3.4.

	Se (%)	PPV (%)	F1 (%)	F1 <sup>†</sup> (%)
TS	81.8	81.7	81.6	81.2
TS <sub>PCA</sub>	88.1	84.5	86.1	83.6
ICA	69.1	60.0	63.7	61.7
PCA	57.4	47.9	51.6	52.6
TS <sub>PCA</sub> -ICA	93.8	92.2	93.0	92.4
ICA-TS <sub>PCA</sub>	90.1	88.8	89.2	89.4
ICA-TS <sub>PCA</sub> -ICA	92.6	92.2	92.4	91.1
FUSE	95.6	94.3	95.0	94.2
FUSE-SMOOTH	95.9	96.9	96.9	95.2

**Table 3.4:** Performance of the different algorithms on the *training* set, A. HRE: Score for the heart rate Challenge event, RRE: score for the RR Challenge event, N/A: not applicable. F1-score given for  $f_c = 10$  Hz, and  $f_c = 2Hz$  indicated by the <sup>†</sup> symbol. Lower HRE and RRE scores are better. Results from [Behar et al., 2014d, table 3]

Template subtraction (TS) and TS applied on the PCA-components (TS<sub>PCA</sub>) were tested along with ICA and PCA. Combinations of them are also implemented. TS-C simply adapts the template to each beat using a scalar gain and was used during the Challenge by [Podziemski and Gieraltowski, 2013] and [Behar et al., 2014d]. The second, TS-PCA, stacks MEGG cycles, selects some of the principal components after which a back-propagation step takes place on a beat-to-beat basis, thus producing MEGG estimates every cycle [Andreotti et al., 2016, p. 632]. From Table 3.4 it can be concluded that ICA and PCA applied directly on the abdominal channels gives bad performance, while ICA combined with TS performs rather well, with small differences between the order of which they are applied. The best performance was obtained using the authors own *FUSE algorithm*, which simply picks the one of the

other subspace methods which gives the "best"<sup>5</sup> FQRS series. FUSE SMOOTH is the FUSE algorithm onto which a smoothing block is applied, as described in Section 3.2.5. From Table 3.4 it is clear how a higher cutoff frequency of 10 Hz of the baseline removal filter gives a better  $F_1$ -score than a 2 Hz cutoff.

### Results on FECGSYNDB dataset

Using the FECGSYNDB, [Andreotti et al., 2016] compared various non-invasive fetal heart rate extraction methods in different scenarios/cases, see Table 3.1, by looking at the  $F_1$ -score and mean average error (MAE) between true and estimated QRS locations. This comparison is very similar to that of the Challenge, with many of the same algorithms used, except simulated data is used instead of clinical. A table summarizing the results can be seen in Table 3.5, where AM is short for adaptive methods such as RLS and LSM filters, as described in Section 3.2.6.

(a) $F_1$ (%)								
Method	Baseline	Case 0	Case 1	Case 2	Case 3	Case 4	Case 5	Overall
<b>BSS</b> <sub>ica</sub>	100.0 (0.1)	<b>100.0 (0.1)</b>	99.7 (3.9)	<b>100.0 (0.1)</b>	89.4 (23.2)	87.1 (6.6)	<b>99.9 (1.9)</b>	<b>99.9 (6.4)</b>
<b>BSS</b> <sub>pca</sub>	100.0 (0.3)	99.9 (1.1)	99.3 (4.8)	99.8 (1.4)	<b>94.5 (7.2)</b>	86.4 (7.7)	97.2 (11.3)	98.7 (7.9)
<b>TS</b> <sub>c</sub>	100.0 (0.2)	97.7 (13.3)	99.0 (8.5)	98.5 (9.6)	77.4 (11.1)	87.8 (11.0)	94.1 (23.6)	95.0 (18.8)
<b>TS</b> <sub>pca</sub>	<b>100.0 (0.0)</b>	98.3 (11.8)	99.2 (7.6)	98.7 (9.3)	77.5 (10.2)	<b>88.4 (10.8)</b>	94.9 (23.2)	96.0 (18.5)
<b>TS</b> <sub>ckf</sub>	100.0 (0.2)	98.1 (9.9)	98.8 (8.2)	98.2 (9.6)	77.8 (10.0)	83.4 (7.3)	94.5 (24.8)	94.5 (19.8)
<b>AM</b> <sub>lms</sub>	100.0 (0.4)	99.1 (6.9)	99.5 (4.9)	99.5 (3.7)	80.2 (10.6)	87.1 (9.8)	96.0 (19.2)	97.1 (15.0)
<b>AM</b> <sub>rls</sub>	99.9 (0.4)	99.2 (6.0)	99.6 (4.2)	99.6 (3.4)	80.6 (11.1)	87.6 (9.4)	96.3 (17.4)	97.3 (14.6)
<b>AM</b> <sub>esn</sub>	99.7 (1.1)	99.6 (2.9)	<b>99.7 (2.7)</b>	99.6 (2.0)	82.5 (8.8)	87.4 (8.3)	97.0 (14.6)	97.9 (12.5)

(b) MAE (ms)								
Method	Baseline	Case 0	Case 1	Case 2	Case 3	Case 4	Case 5	Overall
<b>BSS</b> <sub>ica</sub>	4.0 (1.1)	4.0 (1.1)	4.2 (1.9)	4.0 (0.1)	5.2 (2.4)	5.2 (3.2)	4.0 (0.8)	4.1 (1.9)
<b>BSS</b> <sub>pca</sub>	4.0 (1.8)	4.2 (1.6)	4.7 (2.0)	4.0 (1.0)	4.6 (2.2)	5.2 (1.4)	4.3 (1.9)	4.4 (1.7)
<b>TS</b> <sub>c</sub>	4.0 (1.8)	4.2 (2.0)	4.1 (0.6)	4.1 (0.7)	4.7 (1.1)	5.6 (0.9)	4.3 (1.8)	4.4 (1.5)
<b>TS</b> <sub>pca</sub>	4.0 (2.0)	4.2 (2.1)	4.2 (0.6)	4.1 (0.8)	4.7 (1.1)	5.6 (0.8)	4.3 (1.8)	4.4 (1.6)
<b>TS</b> <sub>ckf</sub>	4.5 (6.3)	3.8 (4.4)	<b>3.1 (0.9)</b>	<b>3.1 (1.1)</b>	<b>3.9 (1.6)</b>	<b>4.8 (1.4)</b>	<b>3.5 (3.7)</b>	<b>3.8 (2.7)</b>
<b>AM</b> <sub>lms</sub>	3.9 (1.8)	4.0 (1.8)	3.9 (0.7)	3.9 (0.6)	5.0 (1.5)	5.3 (1.1)	4.1 (1.5)	4.2 (1.5)
<b>AM</b> <sub>rls</sub>	3.9 (2.2)	4.0 (1.8)	3.9 (0.8)	3.9 (0.8)	4.9 (1.4)	5.3 (1.2)	4.2 (1.6)	4.2 (1.6)
<b>AM</b> <sub>esn</sub>	<b>3.8 (1.9)</b>	<b>3.8 (1.5)</b>	3.8 (0.6)	3.7 (0.6)	4.6 (1.1)	5.2 (1.1)	3.9 (1.1)	4.0 (1.5)

**Table 3.5:**  $F_1$ -score and MAE for FHR extraction methods on synthetic signals in various cases. Results are shown as median (interquartile range), and best performance in each case highlighted in bold. Table from [Andreotti et al., 2016, table 3]

### 3.3.4 Theoretical differences

Having looked at the statistical comparison of the algorithms, a more heuristic and theoretical comparison can be made.

Starting with the single channel methods, it was concluded that the ESN performed marginally better than the LMS and RLS filters in most of the cases. A problem with the adaptive filters is that they require a reference signal, of which it is assumed that the noise present in the abdominal recording is also present in the reference. This may not be entirely true as e.g. fetal movement will not be seen in the chest signal, however the main noise source, the MECG, is present in both reference and abdominal recording. [Behar et al., 2014c, p. 1350]

As compared to the BSS methods requiring "many" channels (8-16) [Sameni, 2008, p. 68], the adaptive

<sup>5</sup>The authors selected the output with the smoothest detected FQRS time series as the best output [Behar et al., 2014d, p. 19]

filtering methods only require one reference and one abdominal channel. This is both a strength and a weakness, as it requires less computation having fewer channels, although it is also impractical to need reference signals. An advantage of the adaptive filtering methods is that they do not require info about the maternal QRS peaks, which template subtraction and Kalman filtering does [Behar et al., 2014c, p. 1350-1351].

An important difference between the ESN (and neural network approaches in general) is that they allow for non-linear maps between chest and abdominal signals, whereas the adaptive filter only allows for linear maps. The relatively good performance of the adaptive filters could indicate that the mapping is mostly linear, meaning a big, computationally heavy neural network might not be needed for this problem. [Behar et al., 2014c, p. 1351]

Looking at the multichannel source separation methods, one should attempt to exploit the prior information of the signals, as "semi-blind source separation methods based on a priori information, are more effective than totally 'blind' methods" [Sameni, 2008, p. 115]. This would mean that e.g.  $\pi$ CA should perform better than normal ICA, as it utilizes the periodicity of the ECG signal. Finding a comparison between  $\pi$ CA and other methods has however not been possible.

Comparing ICA and PCA, ICA gave over 10% better performance than PCA [Behar et al., 2014d, p. 19]. This might be because ICA assumes the signals to be independent, while PCA is that it assumes the MEKG and FEKG to be in orthogonal subspaces - however there is no reason to make this assumption [Behar et al., 2014d, p. 14]. Both methods used assumed linear mixing, however using nonlinear ICA or Kernel PCA<sup>6</sup> could be attempted, in order to handle nonlinear mixings [Behar et al., 2014d, p. 19], to see if a better performance could be achieved. Applying PCA/ICA to the preprocessed abdominal recordings gave poor results, which may be due to having only 4 channels, as opposed to the 8-16 recommended channels. For this reason, ICA was performed in conjunction with template subtraction, as TS reduces the dimensionality of the problem, making the components more separable. [Behar et al., 2014d, p. 19]

A key concluding point is that in addition to the selected source separation algorithm, the choice of FQRS detector, channel selection and smoothing algorithm made an important difference. This is "likely to be the main element discriminating between the top 10-15 Challenge participants" [Behar et al., 2014d, p. 19].

Finally, it should be noted that "it is not reasonable (nor possible) to present a one-for-all universal filtering solution" [Sameni, 2008, p. 115], meaning different algorithms should be applied for different purposes, as they all have their own pros and cons.

---

<sup>6</sup>Kernel PCA works by first transforming the problem into a higher-dimensional space using a non-linear map, in which the original data might be easier to linearly decompose using PCA



# 4 | Non-Invasive Fetal ECG Morphology Estimation

As can be seen from section 3.2, a wide variety of methods for extraction of fetal heart rate have been developed, with many of them showing very good performance. Looking at the morphology of the fetal ECG could also be of interest, as this is a diagnostics tool used on adult ECGs.

## 4.1 The Challenges of Fetal ECG Morphology

The topic of FECC morphology has not been studied a lot, and "robust and accurate fetal morphology extraction systems have not been properly evaluated" according to [Sameni and Clifford, 2010, p. 16]. The reason for this may be that there is a range of challenges related to extracting fetal morphology [Sameni and Clifford, 2010, p. 14-15]:

- Low amplitude fetal ECG in the abdomen recordings
- Possible movements of the fetus
- High interference of maternal ECG

Apart from the technical challenges, the fetal ECG morphology might not have been considered due to a "limited clinical knowledge on how fetal cardiac function and development map to changes in the FECC" [Behar et al., 2014c, p. 1341] and an uncertainty in the clinical significance of fetal ECG morphology [Sameni and Clifford, 2010, p. 7].

In relation to this, it could be interesting to study the correlation between the QT length and the fetal heart rate, which has not yet been studied [Behar et al., 2016, p. 1394].

Another problem is that in the data provided for the Challenge, no arrhythmias, ST elevations, QT prolongations, or contractions were present [Behar et al., 2014d, p. 20]. This makes it hard to test whether an algorithm is able to spot abnormal ECG morphologies. However, by using simulated data as mentioned in Section 3.1.2 these abnormalities can be generated, although annotated clinical data is of course preferable.

### 4.1.1 Considerations of extracting fetal ECG

When extracting the fetal ECG, one must consider the method for doing it, in order to obtain a correct/representative ECG. Two of the problems to consider are those of temporal averaging and distortion.

#### **Averaging**

It can be discussed whether it is advantageous or not to average the heart beats for morphology studies: By doing running averaging, a more accurate fetal QT annotation can be obtained [Behar et al., 2016, p. 1399]. In fact, "the distributions for scalp and abdominal annotations only matched (i.e. were not significantly different) when using the averaged cycles of the NI-FECC and fetal scalp electrode" [Behar et al.,

2016, p. 1397]. However, for some pathologic cases such as premature ventricular contractions, where the ECG/MCG morphology are not consistent, there can be considerable morphologic variations from beat-to-beat [Sameni, 2008, p. 116]. This implies that using averaging is not helpful in the diagnostics of such diseases, also making  $\pi$ CA less effective due to its periodicity assumptions.

### Distortion

Another key issue is how to prevent distortion of the waveform when extracting the FECCG. As the Challenge participants were primarily interested in obtaining the FQRS to be used for heart rate estimation: Getting an undistorted FECCG has therefore not been a priority, and one can therefore not assume the resulting morphology to be accurate.

The morphology might have been distorted already in the preprocessing, as the high cutoff frequency (10 Hz) used for baseline removal should not be used for FECCG morphology analysis, as the waveform will be highly distorted [Behar et al., 2014d, p. 18]. One way to solve this is to use the high cutoff to get a better FHR estimate and then use these estimated R-peaks in a second run with a lower cutoff frequency to get a less distorted FECCG [Behar et al., 2014d, p. 18].

However, in the study by [Behar et al., 2016, p. 1399] where the MindChild [Mindchild Medical, 2018] fetal monitoring system was used for obtaining NI-FECCG, the authors looked at the RMSE of the fetal QT interval estimate, and found it comparable with the RMSE of adult QT interval estimates (17.9ms vs 17.07 ms) [Behar et al., 2016, p. 1399-1400]. The authors concluded that there was "not any significant reason to believe the signal processing of our FECCG would lead to significant distortions". This extraction process is not described in the paper, although the authors claim to have shown that the low frequency components of the FECCG are not distorted by the extraction process, as described in [Clifford et al., 2011].

In this study, the ST elevation of the fetus was considered, by comparing results from a scalp electrode with those estimated from abdominal recordings using a Kalman filter as described in [Sameni et al., 2007b]. It was found that the fetal HR and ST change from the abdomen was "highly accurate and on average clinically indistinguishable from FHR and ST change calculated using fetal scalp electrode data". The RMS error between the ST change calculated by both methods averaged over all processed segments was 3.2%, showing how a well-functioning estimator can provide good results.

As pointed out in [Andreotti et al., 2016, p. 634], a direct comparison of the morphology between a scalp electrode and that from abdominal leads is not representative, as the scalp electrode sees a different projection of the cardiac electric activity than the abdominal leads, meaning the ECG waveform will look different.

## 4.2 Overview of Current Methods

The study of FECG morphology has not yet been explored greatly, and few methods are currently in use. An overview of the relatively newly-developed non-invasive methods is given in this section.

### 4.2.1 Commercial methods

The most well-known device for FECG morphology studies currently in use is the STAN device [Medical, 2015]. STAN combines traditional CTG monitoring with ST-segment analysis of the fetal ECG, obtained from a scalp electrode. The STAN method is thus invasive, calling for new, non-invasive methods for looking at FECG morphology. STAN works by finding the ratio between the amplitude of the fetal T-wave and the fetal QRS-complex (T/QRS ratio), and the presence and absence of a biphasic ST segment [J, 2012, p. 133], as this is a proxy (i.e. approximate) measure of the ST segment.

### 4.2.2 Methods used in Physionet challenge

As part of the Challenge, QT intervals were to be extracted in one of the events, although only one official submission was made to this event. This was made by [Podziemski and Gieraltowski, 2013], who made a relative simple QT interval estimator: The Q position was found as the first maximum before the R wave and the T position was defined to be at the minimal value between 100 ms and 400 ms from the position of the Q wave [Podziemski and Gieraltowski, 2013, p. 336]. The other submission, although unofficial, was made by [Behar et al., 2013], who estimated the QT interval by fitting Gaussians onto the ECG, as proposed by [Clifford and Villarroel, 2006]. Note that in neither papers, accuracies of the QT estimates were mentioned - the only comparison is thus their abstract test scores on the hidden test set C.

### 4.2.3 Framework for stress-testing ECG extraction algorithms

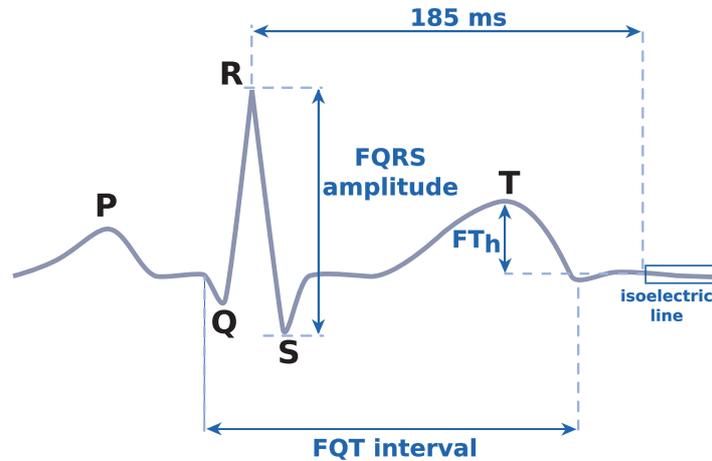
Using the FECGSYNDB data, see Section 3.1.2, [Andreotti et al., 2016] did a comparison of various ECG extraction algorithm, with an emphasis on their ability to estimate the QT interval and the T/QRS ratio (used as an ST-elevation proxy in the STAN device), resulting in the following performance measures, which are illustrated in Figure 4.1:

$$e_{\text{FQT}} = |\text{FQT}_{\text{est}} - \text{FQT}_{\text{ref}}| \quad (4.1)$$

$$e_{\text{FTQRS}} = \left| \left| \frac{\text{FT}_{\text{h, est}}}{\text{FQRS}_{\text{est}}} \right| - \left| \frac{\text{FT}_{\text{ref}}}{\text{FQRS}_{\text{ref}}} \right| \right| \quad (4.2)$$

where:

$e_{\text{FQT}}$	is the fetal QT-interval error	[ms]
$\text{FQT}_{\text{est}},$ $\text{FQT}_{\text{ref}}$	is the estimated and reference fetal QT-interval	[ms]
$e_{\text{FTQRS}}$	is the T/QRS-ratio error	
$\text{FQRS}_{\text{est}},$ $\text{FQRS}_{\text{ref}}$	is the estimated and reference fetal T/QRS-ratio	[ms]



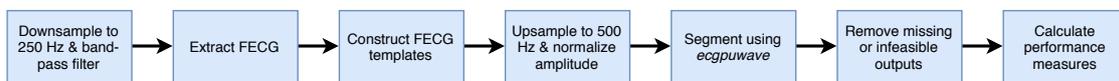
**Figure 4.1:** A template beat showing the fetal T-wave amplitude  $FT_h$ , FQRS amplitude and the isoelectric line, starting 185 ms after the R-peak. Figure taken from [Andreotti et al., 2016, fig. 2]

Looking at these performance measures, it is of interest to know what the average fetal QT-interval and T/QRS ratio is. However, it has interestingly enough not been possible to find any literature on what the average fetal QT interval is, but the T/QRS ratio was found to be less than 0.3 in lamb fetuses [Allan et al., 2001, p. 104]

In order to simplify the comparison, only two cases were considered by [Andreotti et al., 2016], namely the baseline case (no noise) and case 0 with additive noise at different SNRs. The authors also decided to use the reference fetal QRS annotations in order to make the morphology study independent from the accuracy of the FQRS estimation methods [Andreotti et al., 2016, p. 627].

### Algorithm steps

The algorithm used by [Andreotti et al., 2016, p. 635-636] is available online from [Behar et al., 2017], and is summarized in Figure 4.2, and described in further details in the following.



**Figure 4.2:** The algorithms steps used for evaluating the morphology performance of FECG extraction methods.

1. Bandpass-filter recording using a 13th-order<sup>1</sup> LPF with a cutoff at 100 Hz and an eight-order HPF with a cutoff at 0.5 Hz (both are zero-phase Butterworth filters with 20 dB attenuation at the stop-band). These choices are based on the recommendations from [Kligfield et al., 2007], although a LPF cutoff of at least 150 Hz is recommended by [Kligfield et al., 2007]. Implicit downsampling to 250 Hz is assumed, as the FECGSYNDB data is sampled at this rate.
2. Extract the fetal ECG using one of the extraction methods from Section 3.2

<sup>1</sup>When running the code, a 13th order filter is applied, even though the paper states that a 7th order filter is used.

3. Fetal ECG templates are constructed on a minute basis using the QRS reference annotations. As each recording is 5 minutes long, 5 templates are made for each channel of the recording. The segment is discarded if a template can not be built for either test or reference signal. Templates are generated using the following steps [Oster et al., 2015]:
  - (a) Wrap each ECG cycle to the interval  $[-\pi; \pi]$ , with the R peak located at  $-\pi/3$ , as to allow the T-wave to be fully included.
  - (b) Templates are resampled such that each template is 250 samples long.
  - (c) Beats are clustered by comparing each beat to the mean heart cycle of each pre-existing cluster, using cross correlation. If the correlation coefficient is above a certain threshold, the beat is added to the dominant cluster. If it does not match any of the clusters, it is added to a new cluster. This is repeated while decreasing the correlation threshold (starting at 0.9 and decreasing in steps of 0.1 to 0.6) until all beats are clustered, or until the correlation threshold falls below 0.6.
  - (d) The cluster with the highest number of members is selected.
  - (e) Beats within this cluster are averaged, resulting in a template.
4. Upsample to 500 Hz and normalize amplitude to 2 mV as this is the input expected by the segmentation algorithm, *ecgpwave*<sup>2</sup>.
5. Segment each template using *ecgpwave*, thus finding the Q-onset, T-offset and T-peak.
6. Remove segmentation outputs that are either missing (i.e. template could not be segmented) or are physiologically impossible ( $FQT < 100$  ms or  $FQT > 500$  ms)
7. Calculate performance measures  $e_{FQT}$ ,  $e_{FTQRS}$  from segmentation outputs

With the ECG extraction algorithm explained, the results obtained from it can be described.

---

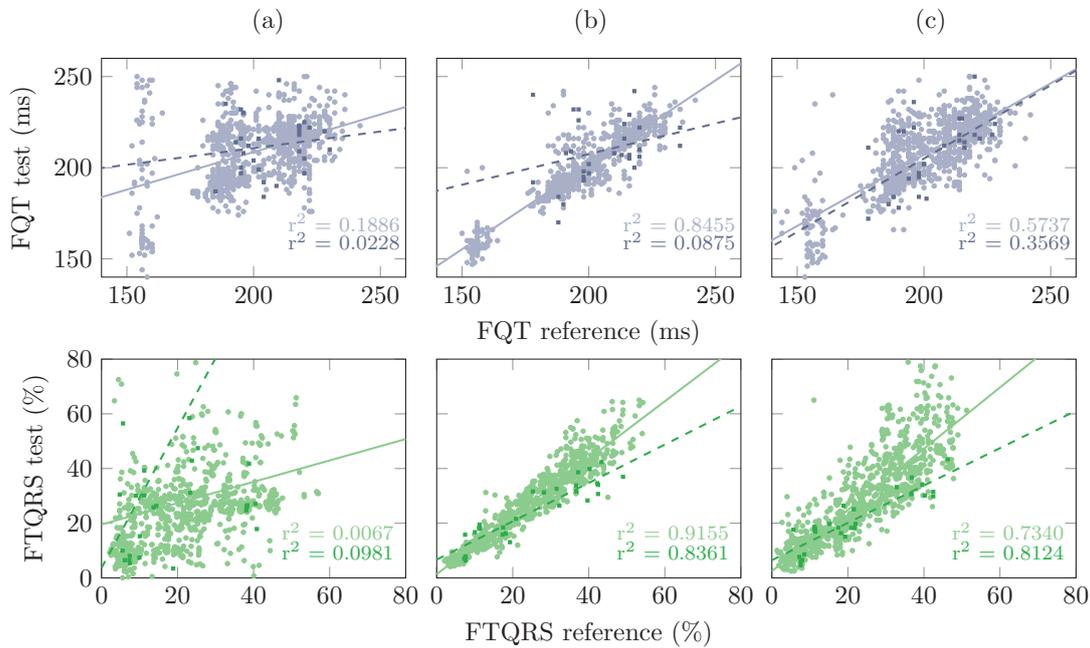
<sup>2</sup>Based on the algorithm by [Jane et al., 1997], *ecgpwave* is freely available in the *WFDB* toolbox [Goldberger et al., 2000].

### 4.3 Comparison of Current Methods

In the paper by [Andreotti et al., 2016], a comparison between the current methods for extracting fetal ECG morphology has been conducted. In the paper, only two cases were considered for morphology studies: The baseline case with no noise, and case 0 see Table 3.1, i.e. with different noise levels (12, 9, 6, 3 and 0 dB SNR).

#### 4.3.1 FQT and FTQRS estimation accuracy

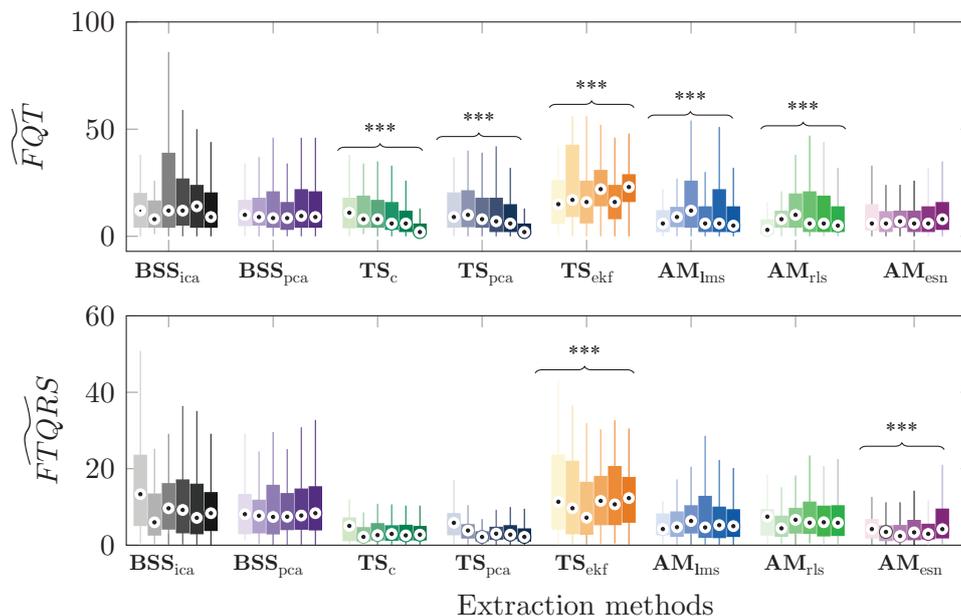
In Figure 4.3, plots of the estimated and true FQT and FTQRS can be seen, using three different extraction methods, namely independent component analysis (ICA), template subtraction (TS) and echo state network (ESN).



**Figure 4.3:** Measured/estimated FQT interval (top) and FTQRS (bottom) vs and FQT/FTQRS reference from original FECCG signal. Extraction methods with highest coefficient of determination ( $r^2$ ) for each category are shown: (a) BSS-ICA, (b) TS and (c) AM-ESN. Results are shown for the baseline (lighter colors, solid lines) and case 0 with SNR 0 = dB (darker colors, dashed lines). Figure from [Andreotti et al., 2016, fig. 7]

From Figure 4.3, it is clear that there is a great uncertainty in these estimates, as the  $r^2$  value is very low ( $< 0.2$ ) for the BSS-ICA case, and never grows above 0.92 for any of the methods. It should however be noted that this performance is also dependent on the segmentation of the ECG, and not just on the extraction.

A comparison of the accuracy of the FQT and FTQRS estimates can be seen in Figure 4.4, where all the tested methods are compared at different SNRs as described above. From the figure, it is clear that BSS techniques do not provide good FQT or FTQRS measures. Meanwhile, TS and AM produced pretty good FTQRS estimates, even in the presence of noise, but their FQT estimates are less robust to noise [Andreotti et al., 2016, p. 643]. The reason for the poor performance using BSS might be because source domain methods are adaptive, meaning the content of the the output channels change over time, thus also changing the morphology [Andreotti et al., 2016, p. 643]. From these results, it is not recommended

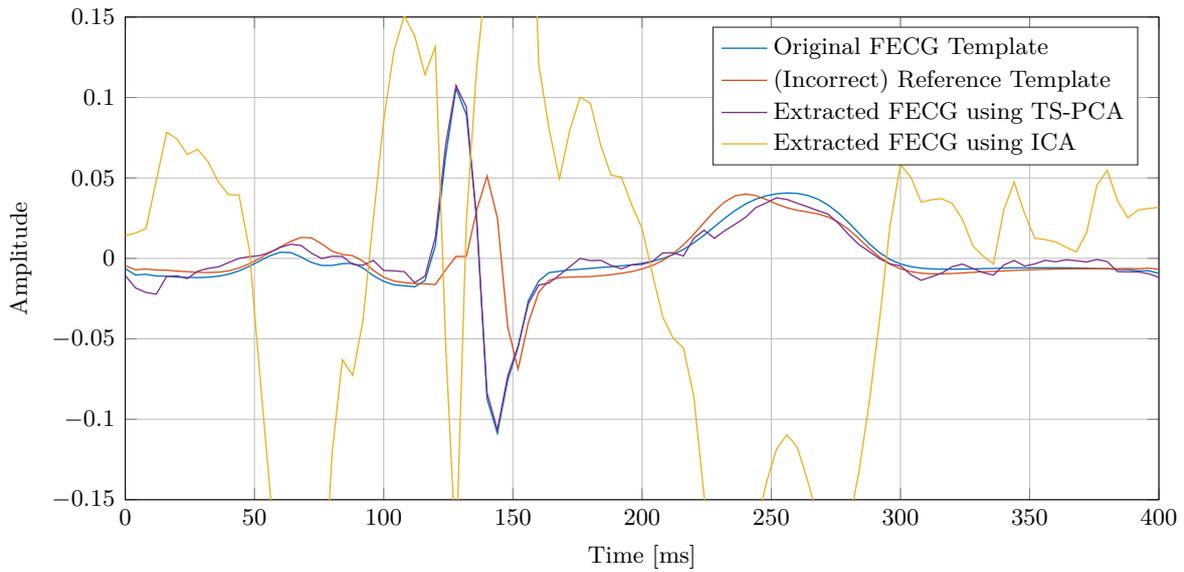


**Figure 4.4:** Accuracy of FQT and FTQRS estimates for the different extraction methods with different SNRs (increasing from left to right, i.e. baseline case furthest to the right). A Kruskal-Wallis test was performed to show statistically significant differences across SNR levels ( $p < 0.001$ ), indicated by \*\*\*. Figure from [Andreotti et al., 2016, fig. 8]

by [Andreotti et al., 2016, p. 643] to use BSS methods for morphology analysis, and it was thus concluded by the authors that less adaptive methods (such as TS) provide fewer distortions in the output FECG estimate.

### 4.3.2 Template comparison

These distortions are also illustrated by comparing the templates obtained from the original FECG, compared to those extracted from the abdominal mixture, see Figure 4.5. In the FECGSYNDB code, 32 abdominal channels are generated, of which 4 or 8 of them are selected as the "recorded" channels. However, in the implementation of the FECGSYNDB code, the channel used for generating the reference template might not be one of these "recorded" channels, hence it is not expected to get a perfect match between between extracted and original template, as can also be seen in Figure 4.5. It is clear that while template subtraction methods perform rather well, the subspace methods such as ICA perform very poorly.



**Figure 4.5:** A comparison of the templates extracted when doing morphology studies, as compared to a template made from the original FECG.

Another interesting finding not clear from the figures is the fact that the percentage of excluded beats increased with the noise level, ranging from 8 to 78% for template subtraction (TS) methods and 19 to 48% for adaptive methods (AM). However, it was relatively constant for blind source separation (BSS) methods (14–20%) [Andreotti et al., 2016, p. 639].

The authors conclude the following from the results of the comparison: "Considerable efforts need to be made in improving currently available techniques, so that clinical relevant information can be obtained from the FECG morphology. The expressive number of excluded beats were either due to problems in the template generation, segmentation or due to the applied methods themselves, therefore further studies should focus on improving these individual steps" [Andreotti et al., 2016, p. 643].

Based on this discussion, it is clear that it would be of great interest to explore methods for accurately extracting fetal morphology. Considerations regarding this is described in the following chapter.

# 5 | Extraction Methods Considerations

The purpose of this chapter is to get a brief overview of potential extraction methods of maternal and fetal ECGs. While the previous two chapters have mostly focused on the extraction algorithms' performance on e.g. the Challenge or the FECGSYN dataset, this chapter will instead focus more on general methods that can be relevant to apply to the problem, but which have either not been developed, or have not yet been applied to fetal ECG extraction. A few potential methods will be investigated throughout the chapter, after which a single method is chosen for full implementation.

## 5.1 Problems with Existing Methods

The first part in identifying which methods to use, is to first study which problems can be identified with the currently used methods, as described in Section 3.2 and 4.3, and why better methods must be identified.

### 5.1.1 Template Subtraction (TS)

Template subtraction (TS) models the maternal ECG, and generates a template ECG waveform, which is subtracted from the recorded ECG waveform. By subtracting the maternal ECG, only the fetal ECG should (ideally) be left. The performance of the TS method depends on the accuracy of the template - the more it resembles the actual maternal beat, the less distorted the residual FECG will be. One of the places where template subtraction falls short, is when the maternal and fetal QRS complexes overlap, as the MECG model must here be very accurate as to not distort the fetal QRS.

To ensure this, many maternal heart beats are rejected when constructing the template, in order to only use those that have a "typical" characteristic. This means TS is not very adaptive (meaning it does not handle it does not handle beat-to-beat variations well), as any large variations of the waveform are ignored when constructing the template. This also requires the averaging of many beats, in order to build a satisfyingly "average" template.

### 5.1.2 Extended Kalman Filter (EKF)

The classical Kalman Filter (KF) uses a dynamic (state-space) model of the maternal ECG, to predict what the waveform will look like in the next sample. By comparing this predicted value with the observed sample, an updated estimate of the true value is generated, by using both the model and the observation. This is done recursively, i.e. on a sample-by-sample basis, where noise covariance matrices are also updated for each sample. An extended Kalman filter (EKF) behaves just like a KF, but while a KF assumes a linear model, the EKF can handle non-linear models. The non-linear model is adapted to the KF by linearizing the model in the operating point, determined by the current sample. The dynamic model used in EKF in the Challenge [Behar et al., 2014d], is a model proposed by [McSharry et al., 2003], which is commonly used for modeling ECGs. As seen in [Sameni et al., 2007a, fig. 7], the model is able to accurately model the ECG, which might be a reason for its popularity. The reason for this accuracy is that the model can be seen as a sum of Gaussians (a Gaussian mixture) [Sameni et al., 2007a, p. 3]: Due

to the *universal approximation* property of Gaussian mixtures, any signal can be modeled to arbitrary precision, assuming a sufficient number of Gaussians are used [Ben-Arie and Rao, 1995].

The KF and EKF assumes the noise to be white, in which case it is optimal in a minimum mean-square (MMSE) sense [Matisko and Havlena, 2012, p. 1523], however the (E)KF does also work in other noise conditions, although it might be suboptimal.

The measurement residual or prediction error is called *Innovation* and is a white, zero-mean Gaussian process if the assumptions posed by the KF are fulfilled, and can thus be used as an optimality check [Matisko and Havlena, 2012, p. 1524]. These assumptions are however *not* fulfilled in the case of fetal ECG: When the maternal ECG has been subtracted, the fetal ECG remain which is of course not a white Gaussian process. The basic assumptions of the (E)KF are therefore not fulfilled, which will reduce performance.

### 5.1.3 Independent Component Analysis (ICA)

Independent Component Analysis (ICA) is a method to perform blind source separation (BSS). The standard formulation of the ICA problem is that  $K$  independent, non-Gaussian, unobserved sources  $\mathbf{s}(t) = [s_1(t), s_2(t), \dots, s_K(t)]$  are linearly combined by the (unknown) mixing matrix  $\mathbf{A} \in \mathbb{R}^{M \times K}$ , to form  $M$  observed mixed signals  $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_M(t)]$  [Hesse and James, 2006]:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) \tag{5.1}$$

ICA aims at finding the  $K \times M$  de-mixing matrix  $\mathbf{B}$ , which can be interpreted as the inverse of the mixing matrix  $\mathbf{A}$  if  $\mathbf{A}$  is non-singular, thus making it possible to estimate the source signals  $\mathbf{s}(t)$ :

$$\hat{\mathbf{s}}(t) = \mathbf{B}\mathbf{x}(t) \tag{5.2}$$

The actual ICA algorithm estimating  $\mathbf{B}$  is not described in this section. Normally, linear mixing is assumed in ICA, although non-linear variants exists, where  $\mathbf{A}$  is replaced by a function  $f : \mathbb{R}^K \rightarrow \mathbb{R}^M$ , such that  $\mathbf{x}(t) = f(\mathbf{s}(t))$ . However, infinitely many solutions exist for  $f$  in the unconstrained case of nonlinear ICA [Hyvärinen et al., 2004, p. 317], hence linear ICA is preferred. The problem with this linearity assumption is that the maternal and fetal ECGs are not simple linear combinations: As one of the key purposes of ECG is to observe the heart from different directions as described in Appendix A, the ECG observed at different leads are not simple linear transformations of a single source signal.

While the classical 12-lead ECG observes projections of the cardiac electric activity which are somewhat arbitrarily chosen (from a mathematical but not clinical) viewpoint, a possible interpretation of the different independent components extracted from multichannel ECG is that they can be considered as projections of the heart activity (in a  $K$ -dimensional hyperspace) onto the directions of 'significant' importance [Sameni, 2008, p. 61].

As ICA has no physical knowledge about the signals, there is no guarantee that the extracted components bear any physical meaning. For instance, the maternal QRS complex may be split up into a positive and negative part, and any other "bumps" of the signal might be considered separate components [Sameni, 2008, fig. 6.13], as these are considered (mathematically) 'significant' due to their independence. Picking the number of ICs is therefore not a straight-forward procedure, especially as it also depends on the number of available channels, which determines if the problem is under- or overdetermined. Usually, one assumes  $\mathbf{A}$  to be square, i.e.  $M = K$ , meaning there are as many mixtures as sources.

## 5.2 Potential Extraction Methods

The purpose of this section is to introduce some extraction methods which may help reduce the problems posed in the previous sections. Note that this is only a brief introduction done in order to be able to compare the methods, and choose a method to implement, which will then be described in further details and using a greater mathematical foundation.

### 5.2.1 Constrained Independent Component Analysis (cICA)

While classical ICA does not pose any constraints to the  $\mathbf{A}$ -matrix, it is possible to introduce constraints, reflecting any prior knowledge about  $\mathbf{A}$  or the source signals  $\mathbf{s}(t)$ .

#### Signal propagation constraint

One such prior is the assumption that the signal amplitude decreases proportional to the inverse square of the distance from the source. This means the value of the matrix element  $\mathbf{A}_{mk}$ , is dependent on two parameters, namely the distance between the  $m$ 'th detector and the  $k$ 'th source and the amplitude  $a_k$  of the source signal [Knuth, 2002, p. 286]:

$$\mathbf{A}_{mk} = \frac{a_k}{4\pi|r_{mk}|^2} \quad (5.3)$$

where:

$A_{mk}$	is the $m, k$ 'th element of the mixing matrix'
$a_k$	is the amplitude of the $k$ 'th source
$r_{mk}$	is the position vector between the $k$ 'th source and the $m$ 'th detector

By exploiting this prior knowledge as a constraint, it might be possible to obtain ICs which better represent the actual signals, and not just arbitrarily (in a physical sense) chosen signal components. This does however not solve the problem with the mixing matrix assuming linear mixing of components, and there is therefore still no guarantee that physically meaningful ICs are returned.

#### Temporal constraint

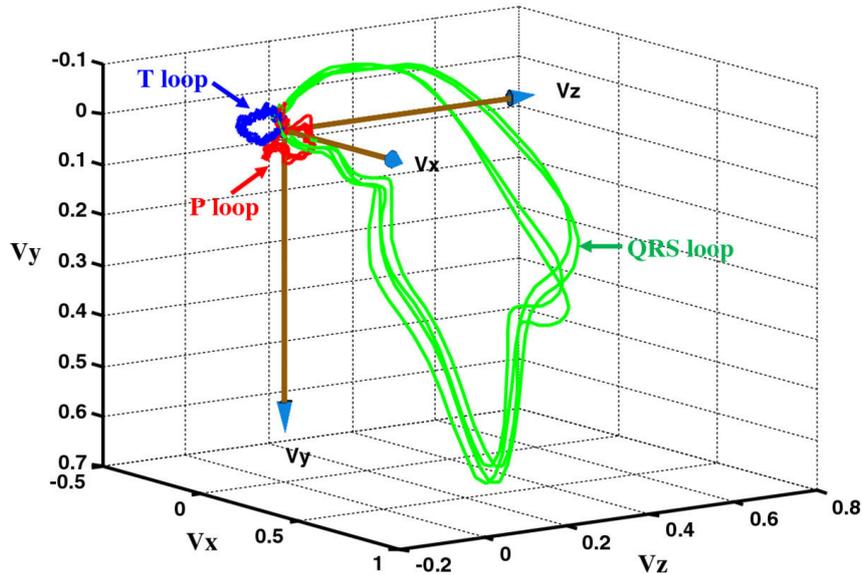
Another way to constrain ICA is by utilizing prior information on the source signal  $\mathbf{s}(t)$ . One such prior is that the extracted signals should be "close" to some reference signal. One such measure of closeness is for instance correlation, meaning the constrained ICA problem aims at finding the sources that are statistically independent and are correlated with some reference signals. Another measure could be the MSE between estimated source and reference. [Lu and Rajapakse, 2006]

This way, one can "help" the ICA finding the sources that have a structure and periodicity similar to the fetal and maternal ECG.

The idea of temporally constrained ICA has not yet been applied to ECG signals, but has been applied to electroencephalographic (EEG) signals, showing good results [James and Gibson, 2003].

## 5.2.2 Vectorcardiogram Independent Component Analysis (VCA-ICA)

This method is based on the idea of the vectorcardiogram, VCG. The VCG is an alternative to the ECG, but does in fact represent the same information. The VCG has 3 leads - one in each of the (orthogonal) x- y- and z-axes of the body, thus giving a 3D vector of the cardiac activity, see Figure 5.1.



**Figure 5.1:** A vectorcardiogram with the different parts of the electric cardiac activity highlighted. Source: [Yang et al., 2012]

As suggested by [Sameni et al., 2007a, p. 3], each of the x-, y-, z- directions can be modeled as sum of Gaussians. In this model, all model parameters for each axis are separate, meaning the axes are independent.

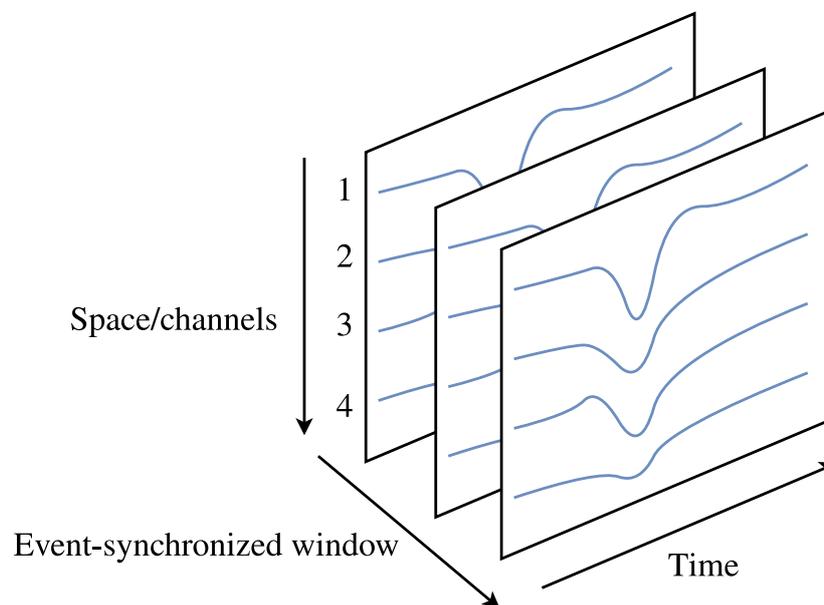
Assuming this is an accurate model of the heart, the independence assumptions made by ICA are now fulfilled, assuming the ICs to be found are the three orthogonal VCG axes. From the VCG, the different ECG leads can be derived using a simple projection matrix [Edenbrandt and Pahlm, 1988], which means all components are linearly separable and the ICA assumption of linear mixing is thus also true. It should therefore be possible to apply ICA and obtain the VCG leads. The ICA method applied can be the constrained ICA mentioned above, in order to incorporate more prior knowledge.

However, it might still be necessary to apply an EKF to filter the estimated VCG channels. After filtering, the cleaned VCG can be back-projected to obtain the original (but filtered and separated) ECG leads.

### 5.2.3 Tensor decomposition

Tensor<sup>1</sup> decomposition is a method to perform blind source separation where prior knowledge about the fact that ECGs are quasi-periodic is exploited. The key idea of tensor decomposition "consists in constructing a tensor by synchronizing on the symbol rate of a certain source, and decomposing the tensor using the Canonical Polyadic (CP) decomposition to extract the characteristics of the source" [Niknazar et al., 2014, p. 53].

This means that a tensor is formed, where all ECGs are segmented and aligned across the third dimension of the tensor, with the other two dimensions being time and space, i.e. the individual channels, see Figure 5.2.



**Figure 5.2:** Illustration of a tensor with event-synchronized windows. Figure based on [Niknazar et al., 2014, fig. 1]

Two problems arise when using tensor decomposition: The first is, that it requires accurate determination of both maternal and fetal R-peaks. Finding maternal is rather easy, using e.g. the Pan-Tompkins algorithm, however fetal R-peaks require having a somewhat functioning FECG extraction algorithm already. The second problem is that the temporal structure of the sources are ignored, meaning all time-variations between periods (heart beats) are ignored. If one wishes to recover these dynamics, a Kalman filter can be applied on the separated channels, as the mixing matrix is estimated when using tensor decomposition.

Note that it is not necessary to derive a dynamic model for the Kalman filter applied - one can simply use the temporal structure estimated when performing the decomposition.

<sup>1</sup>A tensor is a multidimensional array of numbers. While numbers are combined to form a vector, which can be combined to form a matrix, matrices can be combined to form tensors. Note that a tensor is general term for any tensor having a so-called tensor rank higher than two, where the tensor rank describes the number of independent dimensions of the underlying space. For instance, a vector has a tensor rank of 1 and a matrix a tensor rank of 2.

## 5.2.4 Bayesian filtering

Bayesian filters allow for either filtering, predicting or smoothing some observed data to obtain an estimate of an underlying, hidden, set of states/variables. This is done by using a statistical model of the system, from which Bayesian inference can be performed. Many different types of Bayesian filtering exist, but in this case, two variants of the Kalman filter are of interest, together with the particle filter.

### Parallel extended Kalman filter

As described in section 5.1, the EKF approach assumes the residual as white, which is not a true assumption since this residual includes the fetal ECG, as only the maternal ECG is modeled with the EKF. One could also model the fetal ECG using an EKF and use a deflation/sequential approach, where the maternal ECG is first removed, after which the fetal ECG can be filtered using a secondary EKF. Using this approach, the problem with the fetal ECG being assumed white is however still present. To solve this, [Niknazar et al., 2013] proposed a so-called parallel-EKF (par-EKF), as opposed to the so-called sequential approach (seq-EKF) just mentioned. The par-EKF models the *sum* of maternal and fetal ECG in the Kalman Filter, which means the residual should just be noise, which can be assumed white, meaning the par-EKF should perform better than the seq-EKF.

By using the par-EKF approach, the case where the fetal and maternal QRS occur simultaneously can be handled better than the seq-EKF, as shown in [Niknazar et al., 2013, fig. 5]. Due to the overlap, the seq-EKF sees the fetal QRS as part of the maternal, meaning the remaining fetal ECG is distorted. This is not the case (or at least to a much lesser extent) the case of par-EKF, as both signals are modeled. Par-EKF has also shown good performance at the extraction of fetal ECG in the case of multiple fetuses, i.e. twins, triplets etc. The method has however not been applied to any of the datasets mentioned above, meaning its performance compared to the other, more common, methods is not well-known.

### Multichannel extended Kalman filter

The multichannel EKF builds on an observation made by [Niknazar et al., 2013], namely that both par- and seq-EKF are only applied to a single ECG recording: Multichannel methods, where the EKF is applied to multiple ECG channels has not been attempted, and could therefore prove an interesting extraction method.

A multichannel EKF would however require a multidimensional model of the heart vector, in order to make a model, which is able to describe the signals received at different sensors as projections of this 3D heart vector.

A dynamic 3D dipole model as proposed by [Sameni et al., 2007a] could be used for this purpose, which is based upon the model by [McSharry et al., 2003]. A description of the propagation from heart to sensor must also be made part of the model.

### Particle filter

A particle filter (PF) is a method to perform statistical inference, i.e. solving the filtering problem consisting of estimating a set of hidden states, based on a set of observations. This is the same problem solved by the (Extended) Kalman Filter, but while the KF assumes linear models in white noise, the PF can handle any function and any kind of noise distribution [Arulampalam et al., 2002].

The PF is a so-called sequential Monte Carlo method, meaning the estimate of the posterior distribution of interest (of the hidden states) is updated sequentially, as new samples of the (observed) prior distribution are observed. A PF is thus using Bayes theorem to perform statistical inference, but uses sample-distributions instead of closed-form pdfs [Arulampalam et al., 2002].

The ability to handle non-linearities would remove the need to linearize the dynamic model used for the EKF, and the fetal ECG can be kept as a part of the residual, as the PF does not assume the noise as white.

Only a single example of the PF being applied to ECGs has been found [Lin et al., 2011], however no applications to fetal ECG has been found. In the paper by [Lin et al., 2011], a noisy ECG was to be filtered using EKF and PF, and in the worst-case scenario with input SNR of - 5 dB, the PF showed a 10 dB greater SNR improvement of the output SNR than the EKF.

Being a Monte Carlo method, the computational cost is higher than e.g. the EKF, as particles must be sampled, which can be computationally heavy, depending on the number of particles needed - 500-3000 were used by [Lin et al., 2011]. This shows that the PF typically performs better than an EKF, but at a higher computational cost.

### 5.2.5 Beamforming methods

Beamforming is a signal processing technique often used in wireless communications such as cellular communication and radar systems. Beamforming methods can be interpreted as a filtering technique, but instead of filtering in the frequency domain, filtering is performed in the spatial domain. This often includes estimating the direction of arrival (DoA) of the signal of interest, after which a (spatial) filter can be designed, such that the interfering signals can be attenuated.

In order to perform beamforming, multiple sensors, i.e. a sensor array, is needed, as the beamforming techniques exploit the (phase) delay between array elements. In the case of fetal ECG, the signal of interest is the fetal ECG, and the maternal ECG is the interfering signal. By performing beamforming, the fact that the maternal and fetal heart have different locations in the mother is exploited, such that the maternal ECG is attenuated.

Beamforming methods has not yet been applied to fetal ECG signals, and will therefore be interesting to explore in greater details. This is done in Appendix B.

### 5.2.6 Overview

Having described some possible methods to improve the performance of the extraction, they can now be summarized.

The previously used methods can be grouped based on the type of information they exploit:

- Multichannel methods, e.g. ICA and beamforming
- Source model methods, e.g. Bayesian filtering

A basic but intuitive assumption for the separation problem at hand is that the more information is exploited, the better an estimate of the FECCG can be obtained. For this reason, it is of interest to use methods that can both use all the channels available, *and* use the prior information on the source signals (e.g. a source model).

It is therefore of interest that the chosen method uses all available channels, and utilizes prior information, such as a propagation model or a source model.

Another point to consider is that it is of interest to use novel approaches that has not yet been applied to the topic of non-invasive fetal ECG. Based on these criteria, a direct comparison of the mentioned methods is possible, as is shown in Table 5.1.

Extraction method vs. properties	Constained ICA	VCG-ICA + EKF	Tensor decomposition	Bayesian Filtering	Beamforming Methods
Multichannel	x	x	x	x (multichannel EKF)	x
Propagation model	x	x		x (multichannel EKF)	x
Source model	x	x	x (periodicity)	x	
Novel to NI-FECCG	x	x		x (multichannel EKF)	x

**Table 5.1:** An overview of the properties of the extraction methods considered throughout this section.

From Table 5.1, it is clear that e.g. tensor decomposition does not exploit much information - only the periodicity of the ECG is exploited. Of the Bayesian filters, only the multichannel EKF uses multichannel information and a propagation model. The multichannel EKF requires estimates of the propagation between sources and receivers, which can be done using a-priori knowledge on the propagation, or as part of the EKF updates. However, estimating the propagation matrix is somewhat equivalent to estimating the (constrained) mixing matrix as presented under the VCG-ICA-method. This means that if the EKF is applied to the channels extracted from the VCG-ICA method, (or the propagation matrix in the EKF is found using ICA) the multichannel EKF and VCG-ICA methods actually become one. This is the reason why these methods have the same properties as shown in Table 5.1. Bayesian filtering techniques have already been applied to NI-FECCG, although the multichannel EKF has not yet been developed. The other methods are also novel approaches, except for the tensor decomposition.

Generally, the more fields checked off in Table 5.1, the more desirable the method is, as more information should yield better estimates. For this reasons, the methods to be investigated further is the constrained ICA and the VCG-ICA + EKF/ multichannel EKF. The EKF used can also be replaced by a particle filter, however this will increase the computational cost.

From this analysis, three methods are investigated further: VCG-ICA + EKF, beamforming and constrained ICA. Beamforming is investigated in Appendix B, while VCG-ICA + EKF and constrained ICA are described in the following sections.

### 5.3 VCG-ICA + EKF

The purpose of this section is to describe the vector-cardiogram (VCG) independent component analysis (ICA) method, which can be combined with the extended Kalman filter (EKF).

The intuition behind this approach comes from the way ECG signals are simulated, as described in [Behar et al., 2014b], which is explained in the following.

#### 5.3.1 Generating artificial ECG signals

The simulation framework presented in [Behar et al., 2014b] models the vector-cardiogram as a sum of  $N = 7$  Gaussians for each of the  $x, y, z$ -axes. This is done for both the maternal and fetal VCGs, indexed by  $k \in m, f$ .

The dynamic model for the VCG signal is given as [Behar et al., 2014b]<sup>2</sup>:

$$VCG_p : \begin{cases} \dot{x} = -\sum_{i=1}^N \frac{a_i^x \omega}{(b_i^x)^2} \Delta\theta_i^x \exp\left(-\frac{(\Delta\theta_i^x)^2}{2(b_i^x)^2}\right), \\ \dot{y} = -\sum_{i=1}^N \frac{a_i^y \omega}{(b_i^y)^2} \Delta\theta_i^y \exp\left(-\frac{(\Delta\theta_i^y)^2}{2(b_i^y)^2}\right), \\ \dot{z} = -\sum_{i=1}^N \frac{a_i^z \omega}{(b_i^z)^2} \Delta\theta_i^z \exp\left(-\frac{(\Delta\theta_i^z)^2}{2(b_i^z)^2}\right), \\ \Delta\theta_i^j = (\theta - \xi_i^j) \bmod 2\pi, \quad j \in \{x, y, z\} \\ \dot{\theta} = \omega \\ \omega = 2\pi f \end{cases}, \quad p \in m, f \quad (5.4)$$

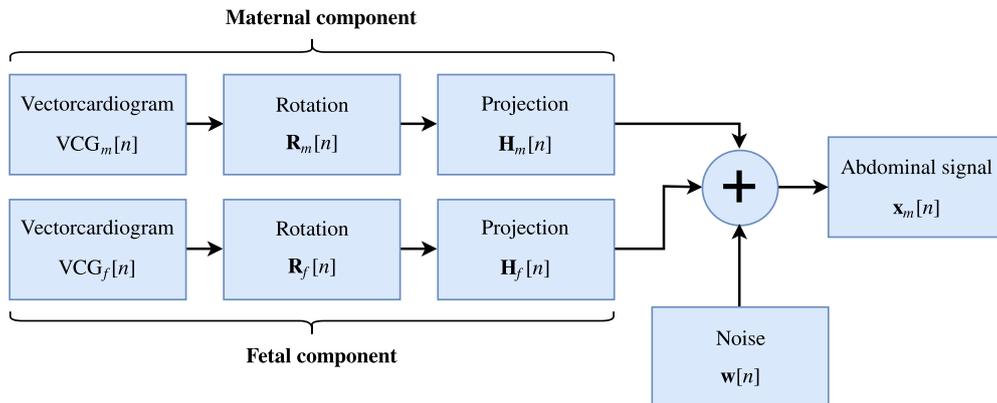
where:

$a_i^j$	is the peak amplitude of the $i$ 'th kernel of the $j$ 'th axis	
$b_i^j$	is the width of the $i$ 'th kernel of the $j$ 'th axis'	
$\xi_i^j$	is the center location of the $i$ 'th kernel of the $j$ 'th axis	[s]
$\omega$	is the velocity of the VCG model	[rad/s]
$f$	is the beat-to-beat heart rate	[Hz]
$k$	denotes either fetal or maternal VCG	

Using this model, 3D vector-cardiograms of both the maternal and fetal hearts are generated. The VCGs are hereafter rotated in 3D using a (time-varying) rotation matrix  $\mathbf{R}_p[n]$ , after which they are projected onto the abdomen, using a projection matrix  $\mathbf{H}_p[n]$ . Finally, noise is added to the abdominal signal, such that the whole simulation chain can be written as can be seen in Equation 5.5, which is also shown in Figure 5.3.

$$\mathbf{x}[n] = \sum_{p \in m, f} \mathbf{H}_p[n] \cdot \mathbf{R}_p[n] \cdot VCG_p[n] + w[n] \quad (5.5)$$

<sup>2</sup>In the expression for  $\dot{y}$ , in [Behar et al., 2014b], the term  $(\Delta x \theta_i^y)$  is assumed to contain a typo, and should actually read  $(\Delta \theta_i^y)$ . The reason for this is that there is no apparent reason why the y-direction should be dependent on the x-direction, especially as the paper itself states the VCG leads are (theoretically) uncorrelated [Behar et al., 2014b, p. 1536].



**Figure 5.3:** The processing chain for generating artificial MECG + FECG signals, as described by [Behar et al., 2014b, p. 1537].

By combining the maternal and fetal components into one, the notation can be simplified. In this case, the source signal is modeled as a 6 dimensional time series,  $\mathbf{s}[n] \in \mathbb{R}^{6 \times 1}$ , for  $n = 0, 1, \dots, N - 1$ :

$$\mathbf{s}[n] = \begin{bmatrix} \text{VCG}_m[n] & \text{VCG}_f[n] \end{bmatrix}^T \quad (5.6)$$

$$= \begin{bmatrix} x_m[n] & y_m[n] & z_m[n] & x_f[n] & y_f[n] & z_f[n] \end{bmatrix}^T \quad (5.7)$$

The rotation and projection matrices are also combined in the following way, where the time index  $n$  is omitted for simplicity.

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}_m & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_f \end{bmatrix}, \quad \mathbf{H} = \begin{bmatrix} \mathbf{H}_m \\ \mathbf{H}_f \end{bmatrix} \quad (5.8)$$

Hereby, the channel model becomes:

$$\mathbf{x} = \mathbf{H} \cdot \mathbf{R} \cdot \mathbf{s} + \mathbf{w} \quad (5.9)$$

### 5.3.2 Separation of MECG and FECG as a BSS problem

The approach presented in this section assumes that the problem at hand can be handled as a blind source separation problem. However, it is not immediately clear why this is the case, which is therefore briefly described in the following.

- MECG and FECG are independent sources, and the signals are therefore also statistically independent. This is described in Section 2.2.1 and is supported by [Jafari and Chambers, 2005, p. 393].
- The propagation of signals from the heart to the abdomen can be considered linear as well as instantaneous [Jafari and Chambers, 2005, p. 393], as the propagation delays are negligible, which is investigated further in Appendix B.
- Looking at Equation 5.4, the parameters describing the ECG cycles in each of the  $x, y, z$ -axes of  $\mathbf{s}$  are independent variables, meaning the axes are mutually independent. This is a necessary condition for ICA, which, as the name implies, assumes components as being independent.

### 5.3.3 The VCG-ICA algorithm

The overall idea of VCG-ICA is to perform operations in Equation 5.9 in the reverse order.

By defining,  $\mathbf{A} = \mathbf{H} \cdot \mathbf{R}$ , the problem can now be identified as a BSS problem:

$$\mathbf{x} = \mathbf{A}\mathbf{s} + \mathbf{w} \quad (5.10)$$

Where one is to estimate  $\mathbf{B} \approx \mathbf{A}^{-1}$  and  $\mathbf{s}$  given  $\mathbf{x}$ , and where the noise vector  $\mathbf{w}$  can be omitted if the noise is assumed white Gaussian, for reasons explained in Section 6.2. This is done by applying independent component analysis.

When  $\mathbf{s}$  has been estimated, an (extended) Kalman Filter can be applied, in order to remove any noise artifacts. The Kalman filter would use a model based on the Gaussian mixture model of the ECGs. In case one is not interested in the VCGs but the original measured leads, the filtered ECGs can be reprojected onto the abdomen using the  $\mathbf{B}$ -matrix, from which  $\mathbf{A}$  can be estimated. Note that as  $\mathbf{H}$  and  $\mathbf{s}$  is simply the maternal and fetal projection matrix/VCGs stacked, it is possible to discard the maternal part, and reproject only the fetal ECG.

The VCG-ICA algorithm is briefly summarized in the following:

1. Estimate  $\mathbf{B} \approx \mathbf{A}^{-1}$  using ICA.
2. Estimate VCG:  $\hat{\mathbf{V}}\mathbf{C}\mathbf{G} = \mathbf{B} \cdot \mathbf{x}$ .
3. Filter VCGs using extended Kalman filter.
4. Reproject fetal VCG onto abdomen:  $\hat{\mathbf{x}}_f = \hat{\mathbf{A}}_f \cdot \hat{\mathbf{V}}\mathbf{C}\mathbf{G}_f, \quad \hat{\mathbf{A}} = \hat{\mathbf{B}}_f^{-1}$ .

### 5.3.4 Problems with VCG-ICA

A series of difficulties are related to the presented approach, which are briefly stated here.

The noise term  $\mathbf{w}$  is for ICA assumed Gaussian, as ICA is not affected by additive Gaussian noise, see Section 6.2.3. This is not an entirely true assumption, as the noise modelled in [Behar et al., 2014b] is not white, but instead represents muscle movement. However, preprocessing should remove much of the noise, as to make its contribution smaller. This is also why an EKF can be applied to the estimated VCGs, in order to remove any noise.

The main problem with VCG-ICA is the following: The mixing matrix  $\mathbf{A} \in \mathbb{R}^{6 \times 6}$  is typically assumed square in ICA, as is also the case here. The reason for this is that the system becomes underdetermined if  $\mathbf{A}$  does not have full rank. This means some of its rows are free variables, which has the implication that the estimated source signals provided by those rows are by no means guaranteed to be the desired VCG components. It is therefore crucial that the matrix has full rank, meaning at least 6 sensors are needed. This can be obtained when using simulated signals, but as most of the signals in the databases are recorded using 4 channels, this requirement cannot be fulfilled. Using VCG-ICA on recorded signals is therefore not expected to result in the desired VCG signals. In other words, one cannot expect ICA to estimate 6 VCG channels from only 4 recorded channels, unless a lot of prior knowledge is applied, such as  $\mathbf{A}$  being sparse. This has also been demonstrated by [Mohammad-Djafari, 2001] where a Bayesian

approach to BSS is taken, but it is still clear that three sources cannot be successfully estimated using only two sensors [Mohammad-Djafari, 2001, fig. 7].

To summarize: VCG-ICA can be seen as an "optimal" method in the theoretical case, as it simply applies the inverse operations as those used for generating artificial ECG recordings. However, its performance on real signals is expected to be less optimal, as 6 abdominal channels is rarely available. This is also the case for the signals available in Physionet databases (e.g. the Challenge dataset or the ADFECGDB set), meaning it will not be possible to compare VCG-ICA with other methods on these widely used datasets. Finally, there is no guarantee that the extracted components are indeed the VCG components. For these reasons, the VCG-ICA method is not an ideal choice as an extraction method.

## 5.4 Constrained ICA

Constrained ICA is based on the idea of independent component analysis, i.e. finding the components of a received signal which have the properties of being as independent as possible.

The problem ICA aims at solving is for convenience stated again (ignoring noise):

$$\mathbf{x}[n] = \mathbf{A}\mathbf{s}[n] \quad (5.11)$$

where:

$\mathbf{x}[n]$  are the received, mixed signal,  $\mathbf{x} \in \mathbb{R}^{M \times N}$   
 $\mathbf{A}$  is the mixing matrix,  $\mathbf{A} \in \mathbb{R}^{M \times K}$   
 $\mathbf{s}[n]$  are the (independent) source signals,  $\mathbf{s} \in \mathbb{R}^{K \times N}$

Here, ICA aims at finding  $\mathbf{B} = \mathbf{A}^{-1}$  (assuming  $\mathbf{A}$  is invertible), such that the estimated sources  $\hat{\mathbf{s}}[n] = \mathbf{B}\mathbf{x}[n]$  are as independent as possible. ICA can be formulated as an optimization problem (where the time index  $n$  is omitted, and sequences are treated as random variables):

$$\hat{\mathbf{s}} = \arg \max_{\mathbf{B}} f(\mathbf{B}\mathbf{x}) \quad (5.12)$$

where:

$f(x)$  is some independency measure

Constrained ICA builds upon this optimization problem, but adds constraints to the problem. One such constraint could be a temporal constraint, stating that the extracted signals must be similar to some pre-defined reference signals  $r[n]$  [Lu and Rajapakse, 2006].

$$\begin{aligned} \hat{\mathbf{s}} &= \arg \max_{\mathbf{B}} f(\mathbf{B}\mathbf{x}) \\ \text{s.t.} \quad & g(\mathbf{B}\mathbf{x}, \mathbf{r}) \leq 0 \end{aligned} \quad (5.13)$$

where:

$f(x)$  is some independency measure  
 $g(x, y)$  is some similarity measure  
 $\mathbf{r}$  is a reference signal

Another possible constraint is the one stated in Equation 5.3, where prior knowledge on the mixing matrix elements is exploited. However, this constraint cannot easily be included, as it requires formulating prior pdfs for the amplitudes of the sources,  $a_k$ , and the distances between sources and sensors,  $r_{mk}$  [Knuth, 2002]. These are hard to form, and cannot easily be combined with other constraints, as they are based on a Bayesian approach forming a prior probability of the mixing matrix elements, whereas the temporal constraint is incorporated as a constraint function to the optimization problem, and not as a statistical prior. For this reason, the square mixing constraint is not considered in the following.

#### 5.4.1 Constrained ICA in non-invasive fetal ECG estimation

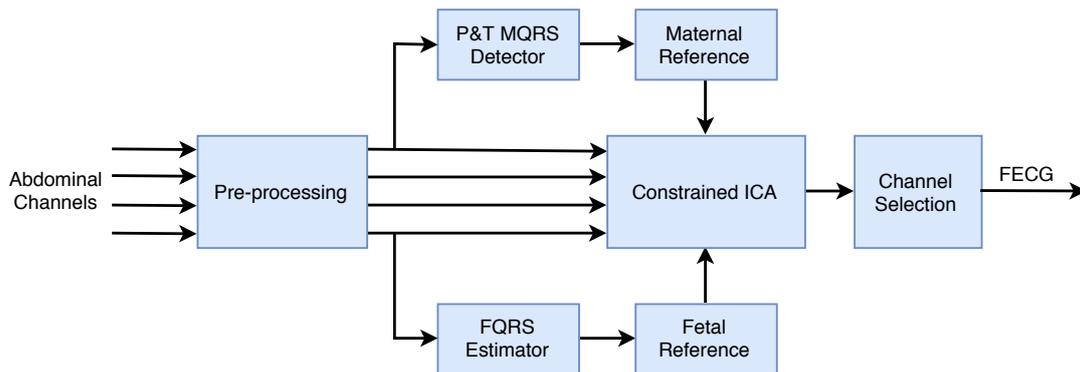
The idea of temporally constrained ICA (cICA) was proposed by [Lu and Rajapakse, 2006], but has been shown to be useful for EEG and MEG signals in the paper by [James and Gibson, 2003].

Only a few examples of applying cICA to fetal ECGs can be found, for instance the paper by [Lee et al., 2005], who used square signals at MQRS locations as a reference signal. However, the method was only tested on a small dataset (with 4 channels), and results are therefore hard to compare to other methods, although they look promising.

Another example is the paper by [Lin et al., 2007], who used 8 abdominal recordings of 10 s duration. Here, a seemingly good performance is also obtained using cICA, but it is not directly comparable to other methods.

Ideally, the more prior knowledge is included, the better the expected performance of cICA. It would therefore be preferable to have both maternal and fetal reference signals as constraints to ICA. After performing ICA, some kind of channel selection should be applied, as to select the best fetal channel. Finally, some post-filtering may be included to smoothen the extracted ECG, e.g. an extended Kalman smoother.

An overview of this constrained ICA processing chain can be seen in Figure 5.4.



**Figure 5.4:** A schematic drawing of constrained ICA applied to non-invasive fetal ECG extraction.

In Figure 5.4, the maternal QRS annotations are found using the Pan-Tompkins algorithm [Pan and Tompkins, 1985], which, in a modified version, might also be able to estimate the fetal QRS complexes, although other methods such as adaptive thresholding can also be used. Reference signals can be generated in a variety of different ways - from very simple square signals at the QRS locations, to more physiologically correct ways, by building up templates of the actual ECG signal and inserting them at the QRS locations. The fetal reference signal could be generated from e.g. a scalp electrode from other recordings (assuming the ECG waveform is similar), or from the ECG model presented in Section 5.3.1.

The steps of the cICA algorithm can therefore be summarized as follows:

1. Apply Pan-Tompkins to find maternal QRS locations
2. Generate maternal ECG reference signal using QRS locations
3. Estimate fetal QRS locations using e.g. modified Pan-Tompkins
4. Generate fetal ECG reference signal using e.g. scalp electrode or Gaussian ECG model
5. Perform constrained ICA (cICA)
6. Identify optimal fetal channel (channel selection)
7. Optional: Apply EKF to extracted fetal channel
8. Generate templates
9. Segment templates
10. Estimate performance measures (FT/QRS and FQT)

Here, the last three steps are included for being able to measure the performance of the cICA algorithm.

### 5.4.2 Problems with constrained ICA

The main problem with cICA, is the choice and generation of reference signals. Generating maternal references is rather easy as the maternal signal is very strong compared to the fetal ECG. However, generating fetal references either requires having pre-determined FQRS annotations, or some algorithm able to extract the FQRS location accurately. This algorithm could be an ICA algorithm, and it just seems contradictory to apply ICA before being able to apply cICA. However, as is clear from Table 3.4, the FQRSs can be estimated rather accurately using pre-existing extraction methods. As the purpose of this thesis is to extract morphologically accurate FECG and not do accurate FQRS estimation, it would be acceptable to assume the FQRS locations as known. This is also the assumption made by [Andreotti et al., 2016], and it would therefore still be a fair comparison to the methods tested on this dataset.

As cICA seems to show promising results [Lin et al., 2007] and the problems related to using cICA appear surmountable, cICA is chosen as the method of extraction. In the following section, the pre-processing performed on the abdominal recordings is described, as this must be applied before performing cICA.

## 5.5 Preprocessing

Most signal processing algorithms include some kind of preprocessing, intended to increase the quality of the signal, thus preparing the signal for the main part of the algorithm, such that it will perform better than if no preprocessing was applied.

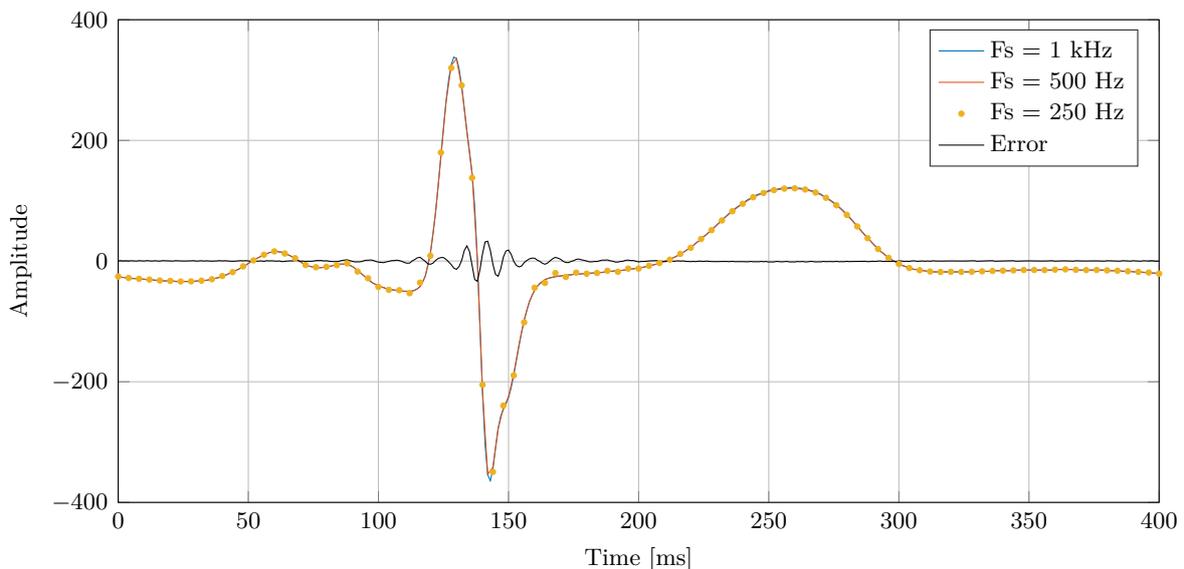
The purpose of preprocessing as described in Section 3.2, is to 1) remove high-frequency noise, 2) remove power line (50 Hz) noise and 3) remove baseline wander from e.g. muscle movement. Each of these three steps are described in the following, and can be summarized in Figure 5.5. Two types of data are considered when choosing the preprocessing methods: The simulated data from FECGSYNDB and the measured channels from ADFECGDB, as it is important that the preprocessing fulfills its purpose in both synthetic and measured data.



**Figure 5.5:** A block diagram of the preprocessing chain.

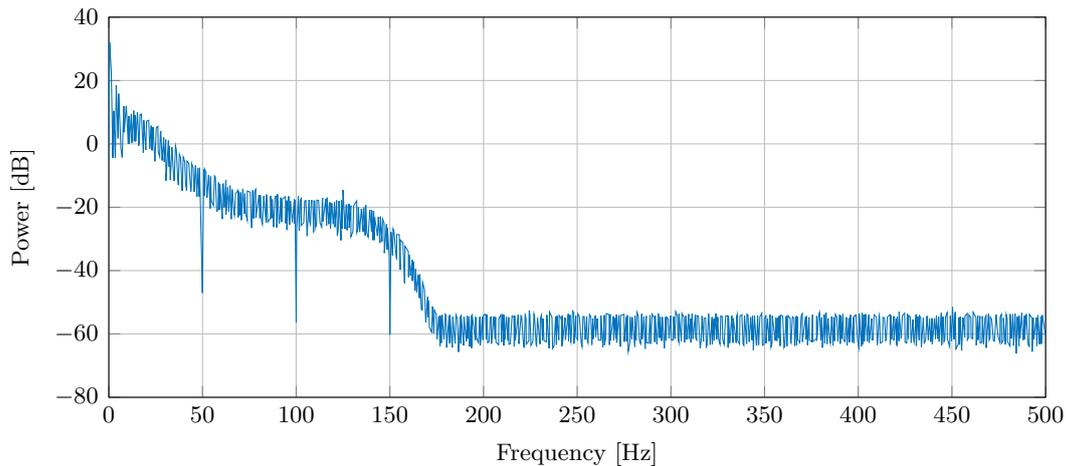
### 5.5.1 Downsampling

The first thing to consider regarding preprocessing, is which sample rate to use. Both the recorded and synthetic data is sampled at  $f_s = 1$  kHz, however in the benchmark algorithm, the templates are downsampled to 250 Hz, see Section 4.2.3. In order to investigate the effects of this downsampling, a single fetal ECG waveform is downsampled, using a downsampling factor of  $M = 1, 2, 4$ . The results of this can be seen in Figure 5.6.

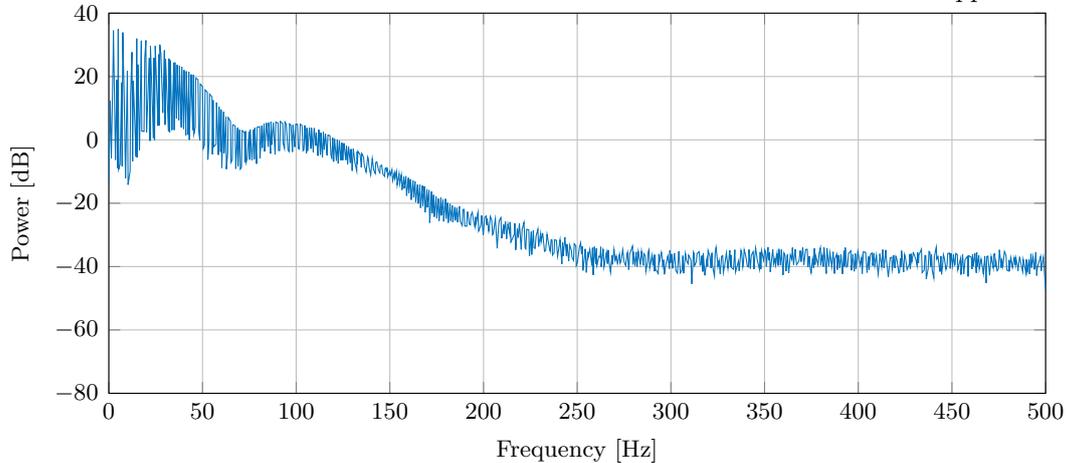


**Figure 5.6:** A single simulated ECG waveform sampled at different rates (i.e. using downsampling factors  $M = 1, 2, 4$ ).

From Figure 5.6, it is clear that no significant difference between the different sample rates, except from a slightly smaller R-peak, meaning downsampling to 250 Hz gives errors that are deemed acceptable with respect to the fetal waveform. This claim is also supported by looking at the power spectra of the recorded and artificial signals, as shown in Figure 5.7. Here, the power spectrum is estimated using Welch's method, and the input signal is a full-length abdominal recording in the ADFECDB case and a full-length fetal ECG signal in the synthetic case. Note that the ADFECDB signals available from Physionet has already been preprocessed, as a 50 Hz notch filter and 150 Hz LPF has been applied as is also clear from Figure 5.7a.



(a) Power spectrum of a recorded ECG signal. Data is first abdominal channel from the ADFECDB-dataset. Note that a 50 Hz notch filter and a 150 Hz LPF has been applied.



(b) Power spectrum of a simulated fetal ECG signal using the simulator from [Behar et al., 2017].

**Figure 5.7:** Power spectral density (PSD) estimates of recorded and simulated ECG signals using the Welch's method.

By down-sampling to 250 Hz, one would remove all frequency content above the Nyquist frequency, i.e. 125 Hz. As a 150 Hz LPF has been applied to the recorded signals from ADFECDB, one must assume the information present above this frequency is not critical to keep.

Looking at the spectrum of the simulated signal in Figure 5.7b, it is clear that only noise is present above

250 Hz, and an exponential decrease of the signal power from about 125 to 250 Hz is clear, which again emphasizes that the frequency content about 125 Hz is negligible.

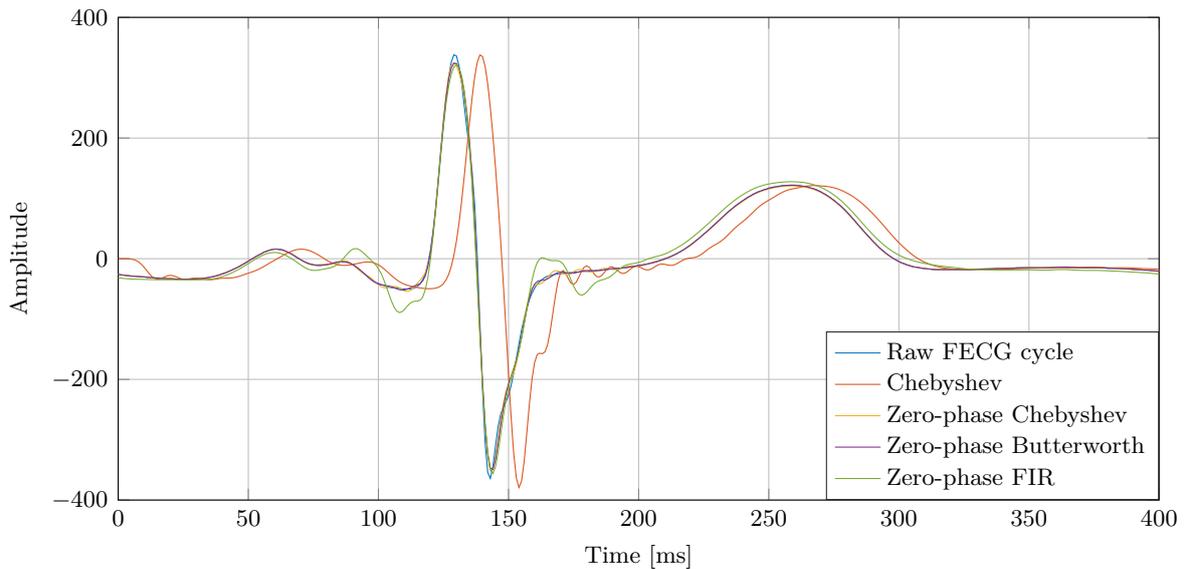
Based on this discussion, downsampling to  $f_s = 250$  Hz is therefore deemed acceptable, without critical loss of information.

### 5.5.2 Low-pass filtering

The purpose of the low-pass filter is to remove any high-frequency noise component which are not part of the ECG and not needed for morphology studies.

As a downsampling to 250 Hz is performed, a low pass (anti-aliasing) filter must be applied, such that the frequency content above the Nyquist rate is negligible.

This is done implicitly in Matlab's `decimate`-function, where an 8th order Chebyshev Type I lowpass anti-aliasing filter with cutoff frequency at 100 Hz is applied before downsampling. A 10'th order 100 Hz Butterworth was applied by [Behar et al., 2014d, p. 5, p. 15], showing that a 100 Hz cutoff is common in the litterature. Another commonly used filter method is a linear phase FIR filter which, like the other filters can be implemented as either non-causal zero-phase filters or as causal filters. These filters are applied onto a single FECG waveform, in order to investigate how they perform, see Figure 5.8.



**Figure 5.8:** A single fetal ECG cycle with different low pass filters applied. For all filters,  $f_c = 100$  Hz.

As can be seen from Figure 5.8, the Chebyshev filter introduces an undesired group delay, while the FIR filter distorts the waveform. The zero-phase Chebyshev and Butterworth filters both perform well and have almost identical frequency responses, but due to its zero passband ripple, the zero-phase Butterworth filter is chosen as the anti-aliasing filter applied before downsampling.

### 5.5.3 50 Hz notch filter

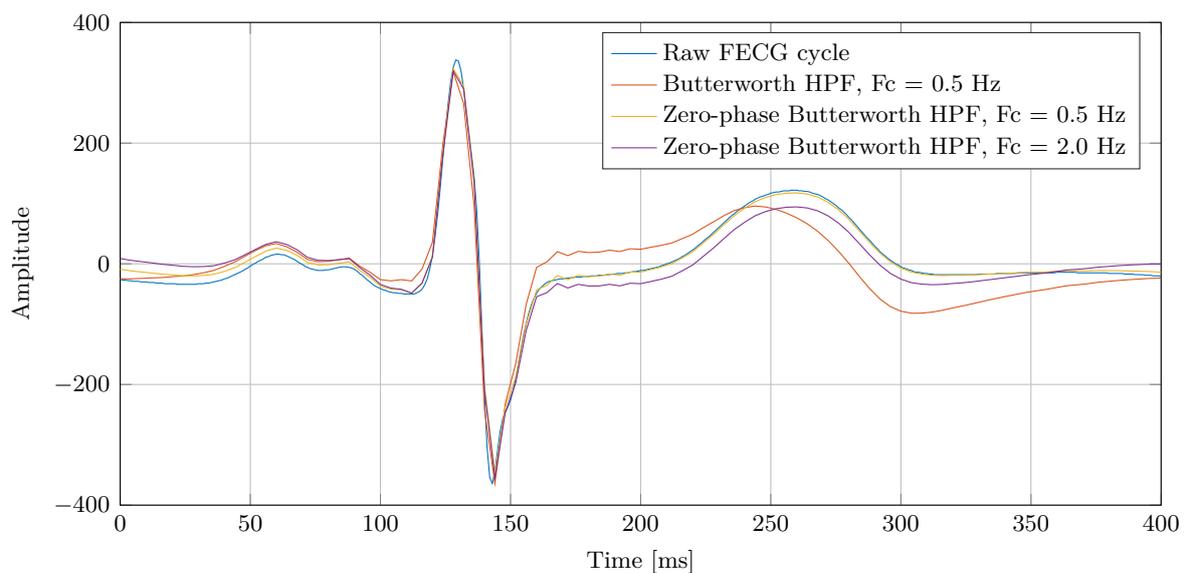
A 50 Hz comb filter has already been applied to the signals from ADFECDB, and no 50 Hz noise is present on the simulated signals. A 50 Hz notch filter is therefore only necessary in case other recordings are used. The filter can be implemented in different ways, as described in Section 3.2. The 50 Hz filter is however not described in further details here, as it is not needed for the testing methods used.

### 5.5.4 Baseline wander removal (HPF)

The baseline wander removal filter is a high pass filter which serves the purpose of removing any low frequency components arising from e.g. movement of the subject or the electrodes.

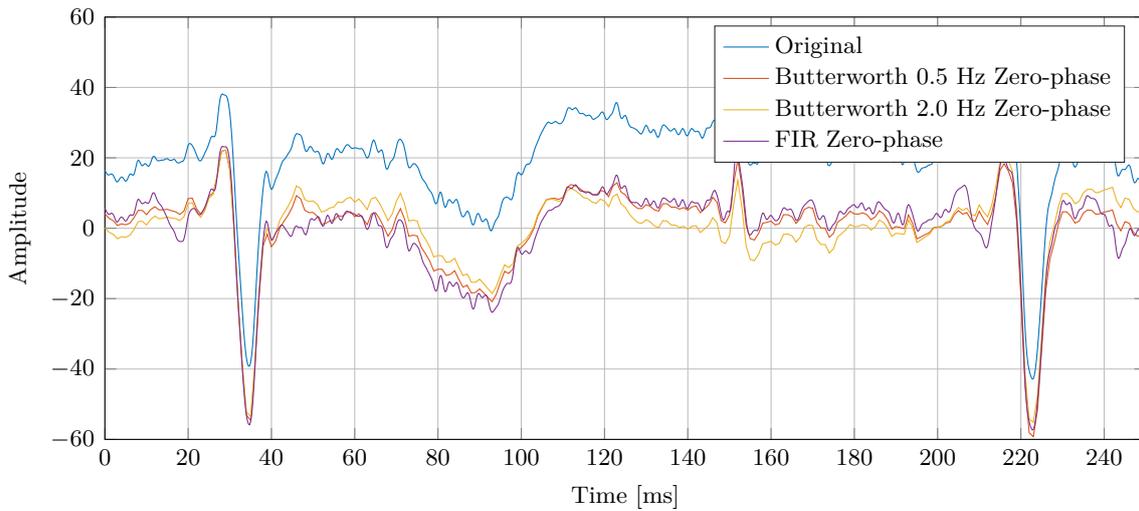
The cutoff frequency of the baseline filter must be very low as to not distort the ECG: [Lipponen and Tarvainen, 2013] used a 6th order Butterworth high-pass filter (cut-off at 2 Hz), while [Andreotti et al., 2016] suggest an 8th order Butterworth filter with a cutoff frequency of 0.5 Hz as to prevent undesired distortion of the FECG.

These two filter types are implemented and applied to a single simulated fetal ECG signal using both normal and zero-phase implementation, see Figure 5.9. From the figure, it is clear that as with the LPF from Figure 5.8, the zero-phase implementation of the HPF is crucial as to not distort the waveform. However, it is also clear that the difference between the 0.5 Hz and 2 Hz cutoff frequencies is rather small.



**Figure 5.9:** A single simulated fetal ECG cycle, with different high pass filters (HPF) applied.

In order to see how well the filters can remove baseline wander, they are applied to a segment of a ADFECDB recording, as seen in Figure 5.10. Also applied is a traditional linear-phase FIR filter. From Figure 5.10, it can be concluded three filters remove the baseline in a satisfying manner, although the FIR filter distort the waveform (see e.g. near 200 ms on the figure).

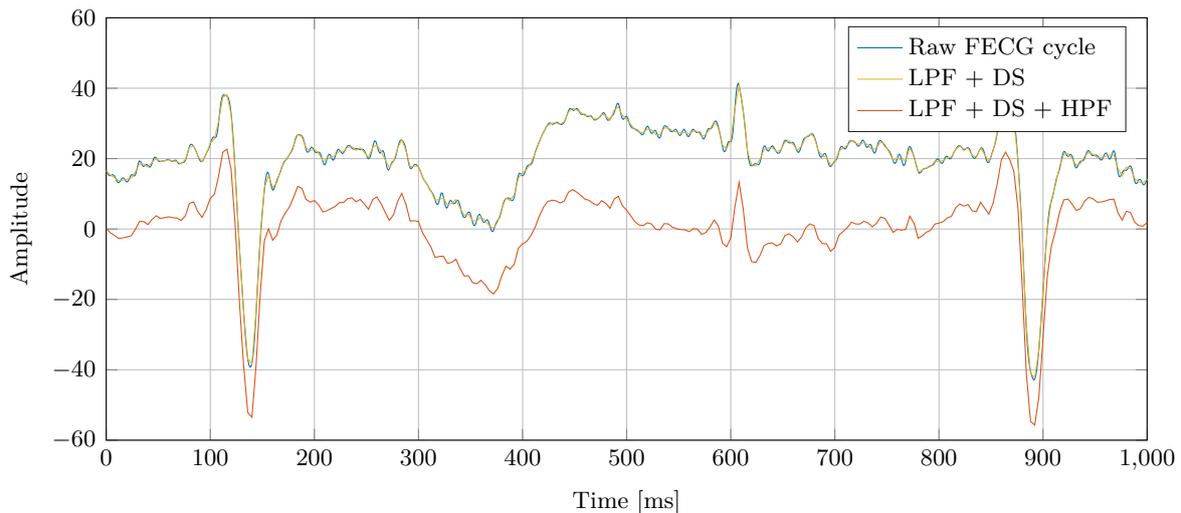


**Figure 5.10:** A 250 ms segment of an abdominal recording from ADFECDB before and after applying high pass filtering/baseline removal.

Based on this investigation and the argument from [Andreotti et al., 2016] regarding keeping the distortion of the waveform to a minimum, the 8th order filter with a 0.5 Hz cutoff is chosen. Another reason to choose this filter is that it allows for a fair comparison with the benchmark methods, as the same HPF is applied.

### 5.5.5 Full preprocessing chain

The LPF, downsampling and HPF can be applied in succession, forming the full preprocessing chain. Figure 5.11 shows an abdominal signal before and after preprocessing.



**Figure 5.11:** A 250 ms segment of an abdominal recording from ADFECDB before and after applying the preprocessing.



# 6 | Independent Component Analysis

The purpose of this chapter is to introduce the concept of independent component analysis (ICA). First, some of the preliminary theory needed is presented, and hereafter some of the most popular ICA algorithms (fastICA, InfoMax) are introduced. An example of ICA being applied to a random signal is also presented.

## 6.1 General ICA Concepts

Before introducing ICA, some preliminary theory must be described. This theory is in the field of probability theory, information theory and optimization. In this section, the theory which a reader is not expected to know about is presented, namely higher-order statistics and information theory.

### 6.1.1 First and second-order moments

Statistics is a measure of properties of the random variable, and are calculated by applying a function to the samples of that variable. The power of which the function raises the sample to, is called the order of the statistics. The most commonly used statistics are first and second order statistics, which are mentioned here as to have introduced the notation and concepts for when the higher-order statistics presented later.

Assume a random vector  $\mathbf{x} \in \mathbb{R}^{N \times 1}$ , with a probability density function (pdf) given as  $p_{\mathbf{x}}(\mathbf{x})$ . The first moment is the mean of  $\mathbf{x}$  and can be found as:

$$\mathbf{m}_{\mathbf{x}} = E[\mathbf{x}] = \int_{-\infty}^{\infty} \mathbf{x} p_{\mathbf{x}}(\mathbf{x}) \, d\mathbf{x} \quad (6.1)$$

The second moment of  $\mathbf{x}$  is the correlation between elements of  $\mathbf{x}$ ,  $x_i, x_j$ , and forms the  $N \times N$  correlation matrix  $\mathbf{R}$ , with elements  $r_{ij}$ :

$$\mathbf{R}_{\mathbf{x}} = E[\mathbf{x}\mathbf{x}^T] \quad (6.2)$$

$$r_{ij} = E[x_i x_j] = \int_{-\infty}^{\infty} x_i x_j p_{\mathbf{x}}(\mathbf{x}) \, d\mathbf{x} \quad (6.3)$$

Typically, one uses the central moments, which are found just as the usual moments, except the means of the random variables are subtracted before calculating the expectation. The second order central moment is the covariance, and is found as:

$$\mathbf{C}_{\mathbf{x}} = E[(\mathbf{x} - \mathbf{m}_{\mathbf{x}})(\mathbf{x} - \mathbf{m}_{\mathbf{x}})^T] \quad (6.4)$$

$$c_{ij} = E[(x_i - m_i)(x_j - m_j)^T] \quad (6.5)$$

The elements of  $\mathbf{x}$  are said to be uncorrelated, if  $c_{ij} = 0$ ,  $\forall i, j$ ,  $i \neq j$ , i.e.  $\mathbf{C}_{\mathbf{x}}$  is a diagonal matrix.  $\mathbf{x}$  is a white random vector if  $\mathbf{C}_{\mathbf{x}} = \mathbf{I}$ .

Two random vectors  $\mathbf{x}$  and  $\mathbf{y}$  are uncorrelated if their cross-covariance matrix  $\mathbf{C}_{\mathbf{xy}}$  is a zero matrix:

$$\mathbf{C}_{\mathbf{xy}} = E[(\mathbf{x} - \mathbf{m}_{\mathbf{x}})(\mathbf{y} - \mathbf{m}_{\mathbf{y}})^T] = \mathbf{0} \quad (6.6)$$

### 6.1.2 Higher order moments

For moments above two, one usually refers to them as higher order moments. However, moments higher than four are rarely used, and will therefore not be described in the following.

The  $j$ 'th moment  $\alpha_j$  of  $x$  is defined as [Hyvärinen et al., 2004, p. 41]:

$$\alpha_j = E[x^j] = \int_{-\infty}^{\infty} x^j p_x(x) dx, \quad j = 1, 2, \dots \quad (6.7)$$

and the  $j$ 'th central moment is defined as

$$\mu_j = E[(x - \alpha_1)^j] = \int_{-\infty}^{\infty} (x - m_x)^j p_x(x) dx, \quad j = 1, 2, \dots \quad (6.8)$$

The third central moment is called *skewness* and can be used as a measure of asymmetry of the pdf.

The fourth order central moment measures the tailed-ness of the distribution, and is defined as

$$\mu_4 = E[(x - m_x)^4] \quad (6.9)$$

One often uses a normalized fourth order moment, where the fourth order moment is divided by the second moment squared. This statistical measure is called *kurtosis*:

$$\text{Kurt}[x] = \frac{E[(x - m_x)^4]}{E[(x - m_x)^2]^2} \quad (6.10)$$

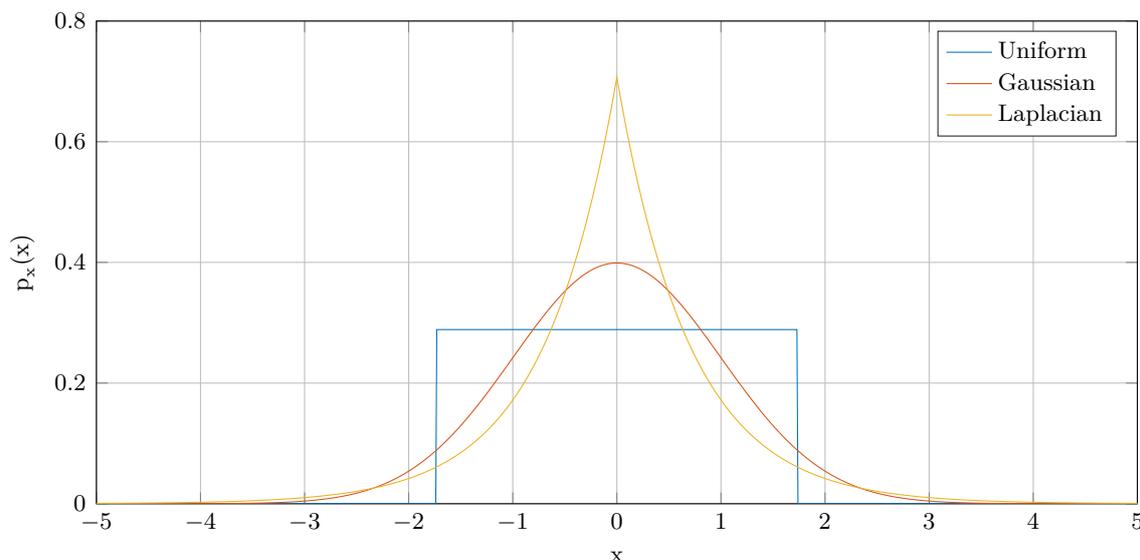
Typically, one uses the statistics called *excess kurtosis*, which is defined as [Hyvärinen et al., 2004, p. 38]

$$\bar{\kappa}(x) = \frac{E[(x - m_x)^4]}{E[(x - m_x)^2]^2} - 3 \quad (6.11)$$

An important feature of excess kurtosis is that it can be used as a simple statistical measure of non-Gaussianity. It can be shown that the excess kurtosis of a Gaussian variable is equal to 0 [Hyvärinen et al., 2004, p. 38]. This is one reason why excess kurtosis is commonly used over kurtosis or the fourth central moment, as it is "normalized" so it is zero for Gaussian variables. As the fourth moment measures tailed-ness, a pdf can be categorized based on its tailedness, relative to a Gaussian [Hyvärinen et al., 2004, p. 38-39]:

- Gaussian/Mesokurtic ( $\bar{\kappa}(x) = 0$ ): : Has the same tailedness as Gaussian.
- Sub-Gaussian/platykurtic ( $\bar{\kappa}(x) < 0$ ): Has less tails than a Gaussian. Sub-Gaussians tends to be flatter than Gaussians or multimodal. A uniform pdf is subGaussian.
- Super-Gaussian/leptokurtic ( $\bar{\kappa}(x) > 0$ ): Has longer tails and sharper peak than a Gaussian. A Laplacian distribution is super-Gaussian.

Examples of sub- and super-Gaussian pdfs can be seen in Figure 6.1. Note how the super-Gaussian has longer tails and a higher spike than the Gaussian, while the opposite is true for the sub-Gaussian.



**Figure 6.1:** Three different probability densities. One is sub-Gaussian (the uniform pdf), one is Gaussian, and one is super-Gaussian (the laplacian pdf). All three have zero mean and unit variance.

The reason why excess kurtosis is often used, is because it is zero for Gaussian variables, and has some other practical properties. These properties come from the fact that excess kurtosis is in fact a (normalized) fourth order cumulant, which will be explained in the following section.

### 6.1.3 Cumulants

The moment-generating function  $\psi(\omega)$  is defined as the continuous Fourier transform of the pdf  $p_x(x)$  [Hyvärinen et al., 2004, p. 41]:

$$\psi(\omega) = \int_{-\infty}^{\infty} \left( \sum_{k=0}^{\infty} \frac{x^k (j\omega)^k}{k!} \right) p_x(x) dx = \sum_{k=0}^{\infty} E[x^k] \frac{(j\omega)^k}{k!} \quad (6.12)$$

where the expression has been expanded by use of the Taylor series. By taking the natural log of  $\psi(\omega)$ , the cumulant generating function  $\phi(\omega)$  is obtained [Hyvärinen et al., 2004, p. 41]:

$$\phi(\omega) = \ln(\psi(\omega)) = \ln(E[\exp(j\omega x)]) \quad (6.13)$$

By writing the Taylor expansion of Equation 6.13, the  $k$ 'th cumulant  $\kappa_k$  becomes part of the expression:

$$\phi(\omega) = \sum_{k=0}^n \kappa_k \frac{(j\omega)^k}{k!} \quad (6.14)$$

From which the  $k$ 'th cumulant is found as the derivative:

$$\kappa_k = (-j)^k \left. \frac{d^k \phi(\omega)}{d\omega^k} \right|_{\omega=0} \quad (6.15)$$

For a zero-mean random variable  $x$ , the first four cumulants are given as:

$$\kappa_1 = 0, \quad \kappa_2 = E[x^2], \quad \kappa_3 = E[x^3], \quad \kappa_4 = E[x^4] - 3E[x^2]^2 \quad (6.16)$$

As can be seen, the cumulants are very similar to the moments, but are slightly different for higher-order statistics. The cumulants are preferred over the moments, as the cumulants have the following properties [Hyvärinen et al., 2004, p. 42]:

- For statistically independent random vectors of the same dimension,  $\mathbf{x}$  and  $\mathbf{y}$ , the cumulants of their sum  $\mathbf{z} = \mathbf{x} + \mathbf{y}$  is equal to the sum of their cumulants.
- If a random process  $\mathbf{x}$  is a multivariate Gaussian, all its cumulants of order three and higher are all zero.

By dividing  $\kappa_4$  with  $E[x^2]^2$  (in the zero-mean case), the expression for excess kurtosis,  $\bar{\kappa}$ , is obtained as mentioned in Equation 6.11. In the following, we will simply refer to excess kurtosis as kurtosis, as "raw" kurtosis is rarely used, and this is the notation used by [Hyvärinen et al., 2004].

### 6.1.4 Independency

Two random variables  $x, y$  are independent, if and only if their joint pdf is the product of their marginal pdfs [Hyvärinen et al., 2004, p. 27]:

$$p_{x,y}(x, y) = p_x(x)p_y(y) \quad (6.17)$$

One can also formulate this using expectations, and any two absolutely integratable<sup>1</sup> functions  $g(x)$  and  $h(y)$  [Hyvärinen et al., 2004, p. 27]:

$$E[g(x)h(y)] = E[g(x)]E(h(y)) \quad (6.18)$$

Note that this is a much stronger criterion than uncorrelatedness, which only states that the second order moments are factorizable:

$$r_{xy} = E[xy] = E[x]E[y] \quad (6.19)$$

For two variables to be independent, all their moments must be factorizable, as one can choose  $g(x) = x^n$ ,  $n = 1, 2, \dots$  and likewise for  $h(y)$ , and Equation 6.18 must hold for all of them.

With the fundamentals of these statistical measures explained, the topic of principal component analysis can be described, which makes use of second-order statistics.

### 6.1.5 Principal Component Analysis (PCA) and whitening

Principal component analysis (PCA) as a commonly used method for feature extraction and data compression. In PCA, the principal components of a dataset  $\mathbf{x}$  are extracted by using second-order statistics: The aim is to find a transformation  $\mathbf{y} = \mathbf{W}\mathbf{x}$  of the data such that the transformed data becomes uncorrelated, i.e.  $E[\mathbf{y}\mathbf{y}^T] = \mathbf{D}$ , where  $\mathbf{D}$  is a diagonal matrix [Hyvärinen et al., 2004, p. 126].

By doing this, the variances, i.e. the  $d_i$ 's are also maximized. This means the first principal component (i.e. the first dimension of  $\mathbf{y}$ ) corresponds to the biggest variance, the next dimension has the largest variance in a direction orthogonal to the first one, etc. [Hyvärinen et al., 2004, p. 126]. Different methods exist for performing PCA, although a commonly used method is through eigenvalue decomposition.

---

<sup>1</sup>Absolute integratable means the integral of the absolute value over the whole domain is finite.

Building upon the PCA idea, one can also perform whitening. Here the idea is to find a transformation of the data  $\mathbf{z} = \mathbf{V}\mathbf{x}$  such that

$$E[\mathbf{z}\mathbf{z}^T] = \mathbf{I} \quad (6.20)$$

i.e. that the elements of  $\mathbf{z}$  become uncorrelated and have unit variance. In the following, a whitening matrix  $\mathbf{V}$  will be presented.

### Whitening matrix

The first step in performing whitening is to center the data by removing the mean:

$$\mathbf{x} \leftarrow \mathbf{x} - E[\mathbf{x}] \quad (6.21)$$

The covariance matrix of  $\mathbf{x}$  is then

$$\mathbf{C}_\mathbf{x} = E[\mathbf{x}\mathbf{x}^T], \quad \mathbf{C}_\mathbf{x} \in \mathbb{R}^{N \times N} \quad (6.22)$$

The eigenvalue decomposition of  $\mathbf{C}_\mathbf{x}$  is:

$$\mathbf{C}_\mathbf{x} = \mathbf{E}\mathbf{D}\mathbf{E}^T, \quad (6.23)$$

where:

- $\mathbf{E}$  is a square  $N \times N$  matrix of normalized eigenvectors  $\mathbf{E} = [\mathbf{e}_1 \ \mathbf{e}_2 \ \dots \ \mathbf{e}_N]$
- $\mathbf{D}$  is an  $N \times N$  diagonal matrix of eigenvalues  $d_1, d_2, \dots, d_N$

As  $\mathbf{C}_\mathbf{x}$  is a real, symmetric matrix (in fact, it is positive semidefinite), the eigenvalue decomposition exists, and the eigenvalues  $d_i$  are real. The eigenvectors can be chosen as orthogonal, and for this reason,  $\mathbf{E}$  is an orthogonal matrix, i.e.  $\mathbf{E}^T = \mathbf{E}^{-1}$ .

Given this decomposition, a linear whitening transform is given by [Hyvärinen et al., 2004, p. 140]:

$$\mathbf{V} = \mathbf{D}^{-1/2}\mathbf{E}^T \quad (6.24)$$

This can easily be shown as being a whitening transform:

$$E[\mathbf{z}\mathbf{z}^T] = E[\mathbf{V}\mathbf{x}(\mathbf{V}\mathbf{x})^T] \quad (6.25)$$

$$= \mathbf{V}E[\mathbf{x}\mathbf{x}^T]\mathbf{V}^T \quad (6.26)$$

$$= \mathbf{V}\mathbf{C}_\mathbf{x}\mathbf{V}^T \quad (6.27)$$

$$= \mathbf{D}^{-1/2}\mathbf{E}^T \mathbf{E}\mathbf{D}\mathbf{E}^T \mathbf{E}\mathbf{D}^{-1/2} \quad (6.28)$$

$$= \mathbf{D}^{-1/2}\mathbf{D}\mathbf{D}^{-1/2} \quad (6.29)$$

$$= \mathbf{I} \quad (6.30)$$

where we have used the fact that  $\mathbf{E}^T\mathbf{E} = \mathbf{I}$ , as  $\mathbf{E}$  is orthogonal.

### 6.1.6 Information theory

In this section, the basic concepts of information theory are introduced. The basic quantity in information theory is entropy, and other variations of it will also be presented. These will prove useful when deriving the ICA algorithms.

## Entropy

The entropy  $H$  of a discrete-value random variable  $X$  is defined as [Hyvärinen et al., 2004, p. 105]

$$H(X) = - \sum_i P(X = x_i) \log(P(X = x_i)) \quad (6.31)$$

where  $P(X = x_i)$  is the pmf of  $X$ . One can use different logarithms for calculating entropy - if the base 2 logarithm is used, the entropy is measured in bits. The logarithm only affects the output unit, so the choice is not important - in the following the natural log is used.

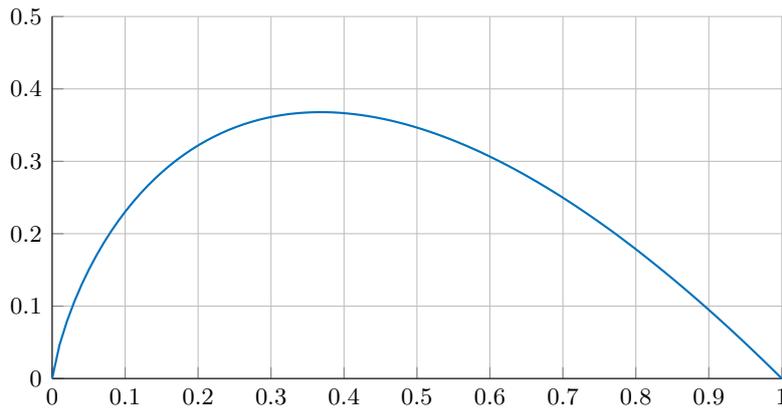
By defining a function  $f(p)$  [Hyvärinen et al., 2004, p. 106]:

$$f(p) = -p \ln(p), \quad \text{for } 0 \leq p \leq 1 \quad (6.32)$$

the entropy can be found as

$$H(X) = \sum_i f(P(X = x_i)) \quad (6.33)$$

A plot of  $f(p)$  can be seen in Figure 6.2, from which it is clear that the entropy is low when the probabilities  $P(X = x_i)$  are close to 0 or 1, and large in between.



**Figure 6.2:** The function  $f(p)$  plotted in the interval  $[0, 1]$ .

The entropy somehow describes the "randomness" of the random variable: By having  $P(X = x_i)$  small for all  $x_i$  except one (as the probabilities must sum to one), there is little randomness in  $X$ , as it almost always takes the same value. However, when probabilities are of equal size, it is not easy to predict which value  $X$  will take, and its entropy is therefore high.

The concept of entropy can be extended to continuous variables, in which case it is called differential entropy, being defined as [Hyvärinen et al., 2004, p. 108]:

$$H(x) = - \int_i p_x(x) \log(p_x(x)) dx = \int f(p_x(x)) dx \quad (6.34)$$

An interesting property of Gaussians is that a standard normal Gaussian variable has the largest entropy among all random variables of unit variance [Hyvärinen et al., 2004, p. 112].

### Mutual information

Mutual information is a measure of how much members of a set of random variables have in common with the other members of the set. The mutual information  $I$  between  $N$  random variables  $x_1, x_2, \dots, x_N$  is defined as [Hyvärinen et al., 2004, p. 110]

$$I(x_1, x_2, \dots, x_N) = \sum_{i=1}^N H(x_i) - H(\mathbf{x}), \quad (6.35)$$

where:

$\mathbf{x}$  is the vector containing all the  $x_i$ 's  
 $H(p)$  is the entropy of  $p$

Mutual information can also be defined using Kullback-Leibler (KL) divergence, which is a way to compare two pdfs,  $p_1$  and  $p_2$  [Hyvärinen et al., 2004, p. 110]:

$$\delta(p_1, p_2) = \int p_1(\mathbf{x}) \log \left( \frac{p_1(\mathbf{x})}{p_2(\mathbf{x})} \right) d\mathbf{x} \quad (6.36)$$

KL divergence is a popular performance measure in deep learning methods, as one typically estimates probabilities when doing classification, and KL divergence is a good method for comparing probabilities.

### Negentropy

The fact that a Gaussian variable has the largest entropy of all pdfs (assuming they have the same variance) makes entropy a useful measure of non-Gaussianity. By using the differential entropy defined above, one can define negentropy  $J(\mathbf{x})$  as [Hyvärinen et al., 2004, p. 112]:

$$J(\mathbf{x}) = H(\mathbf{x}_{\text{Gauss}}) - H(\mathbf{x}) \quad (6.37)$$

where  $\mathbf{x}_{\text{Gauss}}$  is a Gaussian random variable with the same covariance matrix  $\Sigma$  as  $\mathbf{x}$ . Negentropy has the useful properties of being non-negative, and being 0 if and only if  $\mathbf{x}$  is Gaussian.

Similarly, negentropy is invariant to invertible linear transformations, i.e. that for a non-singular matrix  $\mathbf{M}$ ,  $J(\mathbf{M}\mathbf{x}) = J(\mathbf{x})$  [Hyvärinen et al., 2004, p. 113].

### 6.1.7 Optimization preliminaries

In this section, the concepts of update rules and Lagrangians are briefly explained.

#### First- and second order update rules

The first-order Taylor-expansion  $J_T(\mathbf{x}_n + \Delta\mathbf{x})$  of the cost function  $J(\mathbf{x})$  around a point  $\mathbf{x}_n$  is:

$$J_T(\mathbf{x}_n + \Delta\mathbf{x}) \approx J(\mathbf{x}_n) + \nabla J(\mathbf{x}_n) \cdot \Delta\mathbf{x} \quad (6.38)$$

In optimization, we are looking for stationary points, i.e. points where the derivative of the function is equal to zero. Using the first-order Taylor expansion, we simply get that the optimum are the points where the gradient of the cost function  $J(\mathbf{x})$  is equal to zero:

$$\mathbf{0} = \frac{d}{d\mathbf{x}} J_T(\mathbf{x}_n) = \nabla J(\mathbf{x}) \quad (6.39)$$

This fact leads to the classical first-order minimization approach, known as gradient descent. Gradient descent exploits the fact that we will converge towards to minimum of the cost function by moving in the direction of the negative gradient, i.e.

$$\Delta \mathbf{x} \propto -\nabla J(\mathbf{x}) \quad (6.40)$$

This is also called the steepest-descent method.

One may also use a second-order Taylor expansion, which includes the Hessian  $\mathbf{H}(\mathbf{x}_n)$ :

$$J_{T2}(\mathbf{x}_n + \Delta \mathbf{x}) \approx J(\mathbf{x}_n) + \nabla J(\mathbf{x}_n) \cdot \Delta \mathbf{x} + \Delta \mathbf{x} \cdot \mathbf{H}(\mathbf{x}_n) \cdot \Delta \mathbf{x}^T \quad (6.41)$$

Taking the derivative of this equation also leads to the stationary points (optima):

$$\mathbf{0} = \frac{d}{d\mathbf{x}} J_{T2}(\mathbf{x}_n) = \nabla J(\mathbf{x}_n) + \mathbf{H}(\mathbf{x}_n) \cdot \Delta \mathbf{x}^T \quad (6.42)$$

Rearranging and assuming  $\mathbf{H}$  to be non-singular leads to the following second-order update rule, which is also known as Newton's method:

$$\Delta \mathbf{x} \propto -\mathbf{H}^{-1}(\mathbf{x}) \cdot \nabla J(\mathbf{x}) \quad (6.43)$$

By using a second-order approximation, the curvature of the optimization landscape is taken into consideration, which leads to faster convergence.

The proportionality factor used in these update rules is called the step-size learning rate,  $\alpha$ . It is an important hyper-parameter which must be tuned based on the application.

### The Lagrangian of an optimization problem

Considering a general constrained optimization problem

$$\begin{aligned} \max_{\mathbf{x}} \quad & J(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^N \\ \text{s.t.} \quad & g(\mathbf{x}) = \mathbf{0}, \end{aligned} \quad (6.44)$$

The Lagrangian of this problem is formed by introducing a Lagrange multiplier  $\lambda$  and combining the cost function and the constraints in the following way:

$$\mathcal{L}(\mathbf{x}) = J(\mathbf{x}) - \lambda g(\mathbf{x}) \quad (6.45)$$

The Karush-Kuhn-Tucker (KKT) conditions state that a local minima of a constrained optimization problem must satisfy that the gradient of the Lagrangian is zero:

$$\mathbf{0} = \nabla \mathcal{L}(\mathbf{x}) \quad (6.46)$$

Having laid the foundations for ICA, the working principle of ICA can now be explained, as is done in the following section.

## 6.2 ICA by Maximization of Non-Gaussianity

The purpose of this section is to describe the overall working principle of independent component analysis (ICA), namely to maximize the non-Gaussianity of the estimated components.

This basic idea of ICA builds on the central limit theorem, which is a well-known theorem in probability theory. It states that the sum of independent, identically distributed random variables will approach a Gaussian distribution as the number of samples goes to infinity [Hyvärinen et al., 2004, p. 166]. Loosely speaking, it means that the sum of two independent random variables will usually be more Gaussian than the original two variables.

Looking at the traditional ICA problem, but with the source and received signals treated as random vectors  $\mathbf{s}$  and  $\mathbf{x}$  instead of time series, the mixing model becomes:

$$\mathbf{x} = \mathbf{A}\mathbf{s} \quad (6.47)$$

where:

$\mathbf{x}$             are the received, mixed signal,  $\mathbf{x} \in \mathbb{R}^{M \times 1}$   
 $\mathbf{A}$             is the mixing matrix,  $\mathbf{A} \in \mathbb{R}^{M \times K}$   
 $\mathbf{s}$             are the (independent) source signals,  $\mathbf{s} \in \mathbb{R}^{K \times 1}$

From which we want to estimate the demixing matrix  $\mathbf{B}$ , such that we obtain the estimated source signal  $\hat{\mathbf{s}}$ , which should, ideally, be equal to  $\mathbf{s}$ . The de-mixing matrix fulfills the following:

$$\hat{\mathbf{s}} = \mathbf{B}\mathbf{x} \quad (6.48)$$

As the  $s_i$ 's are independent random variables and  $x_i$  is a linear combination (i.e. a scaled sum of the sources  $s_i$ ), the central limit theorem states that the distribution of  $\mathbf{x}$  is more Gaussian than that of  $\mathbf{s}$ . The underlying idea of ICA is built on the following intuition: By using the central limit theorem in reverse, one would expect the independent sources to be those sources that are the least Gaussian.

An argument for why this must hold is that if there are two sources which are less Gaussian than the original ones, the observed (mixed) variables would also be less Gaussian.

This can also be seen by looking at a single estimated source signal  $\hat{s}_i$  and defining  $\mathbf{q} = \mathbf{b}^T \mathbf{A}$ , where  $\mathbf{b}^T$  is a single row of  $\mathbf{B}$ . We can now write  $\hat{s}_i$  as:

$$\hat{s}_i = \mathbf{b}^T \mathbf{x} = \mathbf{b}^T \mathbf{A}\mathbf{s} = \mathbf{q}^T \mathbf{s} \quad (6.49)$$

From the central limit theorem,  $\hat{s}_i$  is usually more Gaussian than any of the  $s_i$ 's, and will become least Gaussian when it is indeed one of the original sources,  $s_i$ . Choosing  $\mathbf{q}^T$  to have more than one non-zero element, means  $\hat{s}_i$  will be a linear combination (weighted sum) of more than one of the original  $s_i$ 's, which will make it more Gaussian, due to the central limit theorem. This is however only strictly true if the original sources are identically distributed, as is assumed here, as required by the central limit theorem.

In the following, a slightly more in-depth reasoning behind this logic is presented. Although not a proof, it should convince the reader that by maximizing (the absolute value of) kurtosis, and thus non-Gaussianity, the independent sources are obtained.

### 6.2.1 Kurtosis as an independency measure

Looking at a simple 2D-case, the ICA channel model becomes:

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} \quad (6.50)$$

Using the additive property of kurtosis from Section 6.1.1, the kurtosis of the first mixed signal  $x_1$  can be written as

$$\text{Kurt}[x_1] = \text{Kurt}[a_{11}s_1 + a_{12}s_2] \quad (6.51)$$

$$= \text{Kurt}[a_{11}s_1] + \text{Kurt}[a_{12}s_2] \quad (6.52)$$

$$= a_{11}^4 \text{Kurt}[s_1] + a_{12}^4 \text{Kurt}[s_2] \quad (6.53)$$

As kurtosis has the property that for any scalar  $\beta$ ,  $\text{Kurt}[\beta x] = \beta^4 \cdot \text{Kurt}[x]$ . We now observe  $x_1$ , and can therefore estimate the (sample) kurtosis  $\text{Kurt}[x_1]$ . We now have two cases:

**Case 1:** The sources  $s_1, s_2$  are the most non-Gaussian sources which can be extracted from  $x_1$ . This means that for any pair of hypothesized sources  $s'_1, s'_2$ , the following must hold:

$$|\text{Kurt}[s_1]| > |\text{Kurt}[s'_1]| \wedge |\text{Kurt}[s_2]| > |\text{Kurt}[s'_2]|, \quad \forall s'_1, s'_2$$

The reason for why the absolute value is used is because a bigger absolute value means less Gaussian, since kurtosis can be negative, and a kurtosis of 0 implies Gaussianity. In this case, ICA works as desired, as the independent, estimated sources are also the sources with the smallest kurtosis among all possible sources.

**Case 2:** The sources  $s_1, s_2$  are *not* the most non-Gaussian sources. That means there exists some  $s'_1, s'_2$  where at least one of them has a smaller (absolute) kurtosis than the true sources, i.e.  $|\text{Kurt}[s'_1]| \geq |\text{Kurt}[s_1]|, |\text{Kurt}[s'_2]| \geq |\text{Kurt}[s_2]|$ .

Having these new variables  $s'_1, s'_2$  implies we would obtain a new sample kurtosis of the variable  $x'_1$ :

$$\text{Kurt}[x'_1] = a_{11}^4 \text{Kurt}[s'_1] + a_{12}^4 \text{Kurt}[s'_2] \quad (6.54)$$

But as the true mixing matrix  $\mathbf{A}$  is constant, the coefficients  $a_{11}, a_{12}$  are also constant. This implies that there is an observed variable  $x'_1$  such that

$$|\text{Kurt}[x'_1]| \geq |\text{Kurt}[x_1]|$$

However, as we have observed  $x_1$  and want to find the sources of it, we know that  $x_1 = x'_1$ , and thus also that  $\text{Kurt}[x'_1] = \text{Kurt}[x_1]$ .

This in turn implies equality for the estimated sources, i.e. that  $\text{Kurt}[s'_1] = \text{Kurt}[s_1]$  and  $\text{Kurt}[s'_2] = \text{Kurt}[s_2]$ .

The only choice for  $s'_1 \neq s_1, s'_2 \neq s_2$  would therefore be some distributions  $s'_1, s'_2$  having exactly the same kurtosis as  $s_1, s_2$ , and which still satisfy the relation  $\mathbf{s}' = \mathbf{B}\mathbf{x}$  for some  $\mathbf{B}$ . This is highly unlikely, as that would require the two distributions to have the same first, second and fourth order moments, but the distributions would still have to be different. With this intuition in mind, it is clear how kurtosis can be used a measure of independency, by applying the central limit theorem in reverse.

## 6.2.2 Mutual information and non-Gaussianity

An interesting connection exists between mutual information and negentropy: For zero mean, uncorrelated, unit variance random variables  $y_i$ , the following relation holds [Hyvärinen et al., 2004, p. 223]:

$$I(y_1, y_2, \dots, y_N) = k - \sum_i J(y_i) \quad (6.55)$$

where:

$I(y_i)$	is the mutual information between $y_i$ 's
$k$	is some constant
$J(y_i)$	is the negentropy of $y_i$

This shows that "ICA estimation by minimization of mutual information is equivalent to maximizing the sum of non-Gaussianities of the estimates of the independent components" [Hyvärinen et al., 2004, p. 223].

A small difference is however that by using negentropy, the estimates of the ICs are forced to be uncorrelated, which is not necessary when using mutual information [Hyvärinen et al., 2004, p. 223]. Also, using negentropy allows a deflationary approach, where ICs are extracted one-by-one, while the mutual information approach requires estimation of all ICs simultaneously [Hyvärinen et al., 2004, p. 223]. This thus shows that ICA can be interpreted from an information theoretical or from a stochastic perspective, but that they fortunately lead to equivalent optimization problems.

### 6.2.3 Limitations of ICA

Having described the overall concept behind ICA, some of the limitations of ICA must now be mentioned.

#### Gaussian sources

As ICA works by maximizing non-Gaussianity, one might ask how ICA performs if the sources have a Gaussian distribution. Unfortunately, ICA does not work on Gaussian sources. This can easily be seen by considering two (zero-mean, unit variance) independent Gaussian sources,  $s_1, s_2$ , with a joint pdf given by [Hyvärinen et al., 2004, p. 161]:

$$p(s_1, s_2) = \frac{1}{2\pi} \exp\left(-\frac{s_1^2 + s_2^2}{2}\right) = \frac{1}{2\pi} \exp\left(-\frac{\|\mathbf{s}\|^2}{2}\right) \quad (6.56)$$

For simplicity, we now assume the mixing matrix  $\mathbf{A}$  to be orthogonal, implying that  $\mathbf{A}^{-1} = \mathbf{A}^T$ . This assumption is true if the data is pre-whitened, and all that remains to be estimated is a rotation matrix, as will be described in greater detail in Section 6.3.1. One can then show the joint density of the mixtures  $x_1, x_2$  to be given as [Hyvärinen et al., 2004, p. 162]:

$$p(x_1, x_2) = \frac{1}{2\pi} \exp\left(-\frac{\|\mathbf{A}^T \mathbf{x}\|^2}{2}\right) |\det \mathbf{A}^T| \quad (6.57)$$

Due to the orthogonality of  $\mathbf{A}$ , we have that  $\|\mathbf{A}^T \mathbf{x}\|^2 = \|\mathbf{x}\|^2$ , and  $|\det \mathbf{A}^T| = 1$ . This yields

$$p(x_1, x_2) = \frac{1}{2\pi} \exp\left(-\frac{\|\mathbf{x}\|^2}{2}\right) \quad (6.58)$$

From this, it is clear that  $\mathbf{A}$  has not changed the pdf, as it is not a part of the it. This means it is not possible to perform inference about  $\mathbf{A}$  from the mixtures. This comes from the fact that for Gaussian variables, uncorrelatedness *does* imply independence, as all moments higher than 2 are zero for Gaussians.

Performing ICA on Gaussians does therefore not do more than performing whitening - the final rotation matrix (the orthogonal  $\mathbf{A}$ -matrix from above) cannot be estimated.

If there are any Gaussian sources in the mixture, it is possible to estimate all non-Gaussian sources, but all the Gaussian sources cannot be separated. This can in fact prove useful, as many noise sources are assumed Gaussian. The fact that ICA can extract non-Gaussian ICs in the presence of Gaussian ICs means that ICA can actually be useful as a method of de-noising, as all Gaussian noise sources will be separated into one single Gaussian source.

### Variations of independent components

In ICA, the variances/amplitudes of the independent components cannot be estimated. This is because any scalar  $\alpha_i$  multiplied onto the sources can be cancelled by an inverse scalar multiplied onto the mixing matrix  $\mathbf{A}$ :

$$\mathbf{x} = \sum_{i=1}^N (\alpha_i^{-1} \mathbf{a}_i) (s_i \alpha_i) \quad (6.59)$$

For this reason, it is possible to constrain the rows  $\mathbf{B}$  to have unit norm, as to simplify the problem.

### Sign ambiguity

Note that as  $\alpha_i$  might be negative, ICA cannot determine the sign of the components either. This means the estimated sources might be inverted. Typically, this is not a problem.

### Order of independent components

As both  $\mathbf{s}$  and  $\mathbf{A}$  are unknown, one can freely choose the order of the components of the corresponding rows of  $\mathbf{s}$  and columns of  $\mathbf{A}$  and still get the same  $\mathbf{x}$ . In some cases, the ordering is not needed, but if it is, some prior knowledge must be applied to perform ordering of the estimated sources.

### Square mixing

Another limitation posed by ICA is that square mixing is often assumed, meaning that we have an equal amount of sources and sensors, meaning  $\mathbf{A}$  is square. This can be circumvented, but it simplifies things greatly.

## 6.3 FastICA

FastICA is a commonly used method for performing independent component analysis, and was proposed by [Hyvarinen, 1999].

As described in the previous section, mutual information can be used as a measure of non-Gaussianity. But, as described in Section 6.2.2, minimizing mutual information is equivalent to maximizing negentropy, under the constraint of the each of the estimated sources  $\hat{\mathbf{s}}_i = \mathbf{b}_i^T \mathbf{x}$  must be uncorrelated, i.e. that  $E[\hat{\mathbf{s}}_j \hat{\mathbf{s}}_k^T] = \delta_{jk}$ .

Note that here, the sources are estimated one by one, i.e. the de-mixing matrix  $\mathbf{B}$  is found by stacking the individual de-mixing vectors,  $\mathbf{b}_i$ , i.e.  $\mathbf{B} = [\mathbf{b}_1 \mathbf{b}_2 \dots \mathbf{b}_N]^T$ .

By using negentropy as the Gaussianity-measure, the optimization problem which FastICA tries to solve is [Hyvarinen, 1999, p. 627]:

$$\begin{aligned} \tilde{\mathbf{b}}_i &= \arg \max_{\mathbf{b}_i} J(\mathbf{b}_i^T \mathbf{x}), \quad i = 1, 2, \dots, N \\ \text{s.t.} \quad & E[(\mathbf{b}_k^T \mathbf{x})(\mathbf{b}_j^T \mathbf{x})^T] = \delta_{jk} \end{aligned} \quad (6.60)$$

where:

- $\mathbf{b}_i$  is the  $i$ 'th row of the de-mixing matrix  $\mathbf{B}$
- $J(\mathbf{y})$  is the negentropy of  $\mathbf{y}$
- $\mathbf{x}$  is the observed mixture
- $\delta_{jk}$  is the Kronecker delta function

Before describing how to solve this problem, we first consider the case of which the data is whitened, as this turns out to simplify the problem.

### 6.3.1 Whitening as preprocessing

By applying a whitening matrix  $\mathbf{V}$  to the mixed signals, the mixing matrix  $\mathbf{A}$  is transformed into another,  $\tilde{\mathbf{A}}$  [Hyvärinen et al., 2004, p. 160]:

$$\mathbf{z} = \mathbf{V}\mathbf{x} = \mathbf{V}\mathbf{A}\mathbf{s} = \tilde{\mathbf{A}}\mathbf{s} \quad (6.61)$$

We now apply an orthogonal transformation  $\mathbf{U}$  to  $\mathbf{z}$ :

$$\mathbf{y} = \mathbf{U}\mathbf{z} \quad (6.62)$$

From the orthogonality of  $\mathbf{U}$ , we have that

$$E[\mathbf{y}\mathbf{y}^T] = E[\mathbf{U}\mathbf{z}\mathbf{z}^T\mathbf{U}^T] = \mathbf{U}E[\mathbf{z}\mathbf{z}^T]\mathbf{U}^T = \mathbf{U}\mathbf{I}\mathbf{U}^T = \mathbf{I}, \quad (6.63)$$

implying that  $\mathbf{y}$  is white as well. This means that by applying whitening alone, we cannot see if  $\mathbf{z}$  or  $\mathbf{y}$  are the independent components. This means that PCA gives the independent components up to an orthogonal transformation [Hyvärinen et al., 2004, p. 160]. However, applying whitening before ICA as a preprocessing step is highly useful: The new mixing matrix  $\tilde{\mathbf{A}} = \mathbf{V}\mathbf{A}$  is orthogonal:

$$E[\mathbf{z}\mathbf{z}^T] = E[\tilde{\mathbf{A}}\mathbf{s}\mathbf{s}^T\tilde{\mathbf{A}}^T] = \tilde{\mathbf{A}}E[\mathbf{s}\mathbf{s}^T]\tilde{\mathbf{A}}^T = \tilde{\mathbf{A}}\tilde{\mathbf{A}}^T = \mathbf{I}, \quad (6.64)$$

as we know that both the sources  $\mathbf{s}$  and the whitened variables  $\mathbf{z}$  are uncorrelated.

This means that after whitening, instead of estimating the arbitrary mixing matrix  $\mathbf{A}$ , we can limit the search space to estimating an orthogonal matrix  $\tilde{\mathbf{A}}$ , corresponding to a rotation or reflection.

### 6.3.2 FastICA on whitened data

By using whitened input data  $\mathbf{z} = \mathbf{V}\mathbf{x}$ , the ICA problem is simplified.

We now introduce a new de-mixing matrix  $\mathbf{W}$ , which operates on the whitened data  $\mathbf{z}$ . This de-mixing matrix satisfies the following relation:

$$\hat{\mathbf{s}} = \mathbf{B}\mathbf{x} = \mathbf{W}\mathbf{z} = \mathbf{W}\mathbf{V}\mathbf{x} \quad (6.65)$$

from which it is clear that  $\mathbf{B} = \mathbf{W}\mathbf{V}$ . As described in Section 6.3.1, by first applying whitening  $\mathbf{V}$ , the de-mixing matrix  $\mathbf{W}$  is an orthogonal matrix.

Looking at the uncorrelatedness-constraint from Equation 6.60, we see the same:

$$\mathbf{I} = E[(\mathbf{B}\mathbf{x})(\mathbf{B}\mathbf{x})^T] \quad (6.66)$$

$$= E[(\mathbf{W}\mathbf{z})(\mathbf{W}\mathbf{z})^T] \quad (6.67)$$

$$= \mathbf{W}E[\mathbf{z}\mathbf{z}^T]\mathbf{W}^T \quad (6.68)$$

$$= \mathbf{W}\mathbf{I}\mathbf{W}^T \quad (6.69)$$

$$= \mathbf{W}\mathbf{W}^T \quad (6.70)$$

which we know holds, as for an orthogonal matrix,  $\mathbf{M}^{-1} = \mathbf{M}^T$ . Thus, constraining  $\mathbf{W}$  to be orthogonal will ensure that the correlation constraint is satisfied. This means the optimization problem can be formulated as:

$$\begin{aligned} \tilde{\mathbf{w}}_i &= \arg \min_{\mathbf{w}_i} J(\mathbf{w}_i^T \mathbf{z}), \quad i = 1, 2, \dots, N \\ \text{s.t. } \mathbf{z} &= \mathbf{V}\mathbf{x} \\ \mathbf{w}_k \mathbf{w}_j^T &= \delta_{jk}, \quad \forall k, j \end{aligned} \quad (6.71)$$

where:

$\mathbf{w}_i$  is the  $i$ 'th row of the de-mixing matrix  $\mathbf{W}$

$\mathbf{V}$  is a whitening matrix

Having established which problem ICA aims to solve, we can now attempt to solve it.

### 6.3.3 Approximating negentropy

The negentropy is in some sense an optimal estimator of non-Gaussianity with regards to statistical performance [Hyvärinen et al., 2004, p. 183], but also has some practical problems. In order to calculate negentropy, sample estimates of the pdf are required, which is very difficult to obtain in an accurate way. For this reason, it is of interest to find methods to approximate negentropy.

One can show that negentropy can be approximated using higher-order cumulants [Hyvärinen et al., 2004, p. 115]:

$$J(y) \approx \frac{1}{12} E[y^3]^2 + \frac{1}{48} \text{Kurt}[y]^2 \quad (6.72)$$

where  $y$  is a random variable with zero mean and unit variance.

Notice how this reduces to maximizing the absolute value of the kurtosis, as the first term (the skewness) can often be ignored as we in many cases deal with symmetric densities, and maximizing the squared kurtosis is similar to maximizing the absolute value [Hyvärinen et al., 2004, p. 183].

This cumulant-based approach can be generalized to use general non-quadratic functions or "non-polynomial moments" [Hyvärinen et al., 2004, p. 183]. These methods also turn out to be more accurate than using cumulants [Hyvarinen, 1999]. Negentropy can more generally be approximated as [Hyvarinen, 1999, p. 627]

$$J(y_i) \approx c (E[F(y_i)] - E[F(v)])^2 \quad (6.73)$$

where:

- $c$  is an (irrelevant) constant
- $F(y_i)$  is some non-quadratic function
- $v$  is a standard normal variable

If one for instance chooses  $F(y) = y^4$ , one obtains the estimate of the kurtosis, since the fourth order moment contains an  $E[y^4]$ -term [Hyvarinen, 1999, p. 627]. Other possible contrast functions  $F(y)$  are:

$$F_1(y) = a_1^{-1} \log \cosh a_1 y, \quad (6.74)$$

$$F_2(y) = -\exp(-y^2/2), \quad (6.75)$$

$$F_3(y) = y^4 \quad (6.76)$$

where  $1 \leq a_1 \leq 2$  is some suitable constant, often just chosen to be one for simplicity.

It is now possible to derive the fastICA algorithm, which is done in the following section.

### 6.3.4 Derivation of FastICA

We now look at the ICA optimization problem in Equation 6.71, which we for now simplify by only considering a single  $\mathbf{w}$ , i.e. a single independent component is extracted. Doing this, the orthogonality constraint reduces to  $\mathbf{w}$  having unit norm:

$$\begin{aligned} \tilde{\mathbf{w}} = \arg \min_{\mathbf{w}} J(\mathbf{w}^T \mathbf{z}) \dots, N \\ \text{s.t. } \|\mathbf{w}\|_2^2 = \mathbf{w}^T \mathbf{w} = 1 \end{aligned} \quad (6.77)$$

The Lagrangian can now be formed:

$$\mathcal{L}(\mathbf{w}) = J(\mathbf{w}^T \mathbf{z}) - \lambda(\mathbf{w}^T \mathbf{w} - 1) \quad (6.78)$$

$$= (E[F(\mathbf{w}^T \mathbf{z})] - E[F(v)])^2 - \lambda(\mathbf{w}^T \mathbf{w} - 1) \quad (6.79)$$

$$= E[F(\mathbf{w}^T \mathbf{z})]^2 + E[F(v)]^2 - 2E[F(\mathbf{w}^T \mathbf{z})]E[F(v)] - \lambda(\mathbf{w}^T \mathbf{w} - 1) \quad (6.80)$$

Taking the gradient (w.r.t.  $\mathbf{x}$ ) of the Lagrangian gives [Hyvarinen, 1999, p. 630]:

$$0 = \nabla \mathcal{L}(\mathbf{w}) = E[\mathbf{z}f(\mathbf{w}^T \mathbf{z})] - \lambda \mathbf{w}, \quad (6.81)$$

Where the expectation  $E[F(v)]$  has been removed as it is a constant w.r.t  $\mathbf{w}$  since  $v$  is i.i.d, and other constants have been ignored as they do not change the optimizer. Also note the function  $f(y) = F'(y)$  has been introduced. For this relation to be zero, the Lagrange multiplier can be found by re-arranging the expression [Hyvarinen, 1999, p. 630]:

$$\lambda = E[\mathbf{w}_0^T \mathbf{z} f(\mathbf{w}_0^T \mathbf{z})] \quad (6.82)$$

where:

- $\mathbf{w}_0$  is the value of  $\mathbf{w}$  at the optimum.

Denoting the gradient of the Lagrangian as  $L$ , the Hessian is found by taking the gradient of  $L$ , such that an outer product is formed (the Hessian is an  $N \times N$  matrix) [Hyvarinen, 1999, p. 630]:

$$H(\mathbf{x}) = \nabla(L)^T = \nabla(\nabla\mathcal{L}(\mathbf{w}))^T = E[\mathbf{z}\mathbf{z}^T f'(\mathbf{w}_0^T \mathbf{z})] - \lambda \mathbf{I} \quad (6.83)$$

To simplify the inversion of  $H(\mathbf{x})$ , the first term of Equation 6.83 is approximated. Since  $\mathbf{z}$  is whitened, i.e.  $E[\mathbf{z}\mathbf{z}^T] = \mathbf{I}$ , a reasonable approximation is [Hyvarinen, 1999, p. 630]:

$$E[\mathbf{z}\mathbf{z}^T f'(\mathbf{w}_0^T \mathbf{z})] \approx E[\mathbf{z}\mathbf{z}^T]E[f'(\mathbf{w}_0^T \mathbf{z})] = E[f'(\mathbf{w}^T \mathbf{z})]\mathbf{I} \quad (6.84)$$

Which means the Hessian is now a diagonal matrix, which can easily be inverted. By approximating  $\lambda$  using the current value of  $\mathbf{w}$  instead of  $\mathbf{w}_0$ , the Newton step becomes [Hyvarinen, 1999, p. 630]:

$$\mathbf{w} \leftarrow \mathbf{w} - \frac{E[\mathbf{z}f(\mathbf{w}^T \mathbf{z})] - \lambda \mathbf{w}}{E[f'(\mathbf{w}^T \mathbf{z})] - \lambda} \quad (6.85)$$

Multiplying both sides by  $E[f'(\mathbf{w}^T \mathbf{z})] - \lambda$  and simplifying yields the following FastICA update rule [Hyvarinen, 1999, p. 630]:

$$\mathbf{w} \leftarrow E[\mathbf{z}f(\mathbf{w}^T \mathbf{z})] - E[f'(\mathbf{w}^T \mathbf{z})]\mathbf{w} \quad (6.86)$$

$$\mathbf{w} \leftarrow \mathbf{w}/\|\mathbf{w}\| \quad (6.87)$$

Where the normalization step is introduced to constrain  $\mathbf{w}$  to have unit norm.

For the contrast functions in Equation 6.74, the functions  $f(y)$ ,  $f'(y)$  become:

$$f_1(y) = \tanh(a_1 y) \quad f'_1(y) = a_1(1 - \tanh^2(a_1 y)) \quad (6.88)$$

$$f_2(y) = y \exp(-y^2/2) \quad f'_2(y) = (1 - y^2) \exp(-y^2/2) \quad (6.89)$$

$$f_3(y) = y^3 \quad f'_3(y) = 3y^2 \quad (6.90)$$

FastICA is summarized in Algorithm 1:

---

**Algorithm 1** FastICA

---

- 1: Subtract mean from data:  $\mathbf{x} \leftarrow \mathbf{x} - E[\mathbf{x}]$
  - 2: Apply whitening:  $\mathbf{z} = \mathbf{V}\mathbf{x}$
  - 3: Choose an initial (e.g. random)  $\mathbf{w}$  of unit norm
  - 4:  $\mathbf{w} \leftarrow E[\mathbf{z}f(\mathbf{w}^T \mathbf{z})] - E[f'(\mathbf{w}^T \mathbf{z})]\mathbf{w}$
  - 5:  $\mathbf{w} \leftarrow \mathbf{w}/\|\mathbf{w}\|$
  - 6: If not converged, go back to step 4
- 

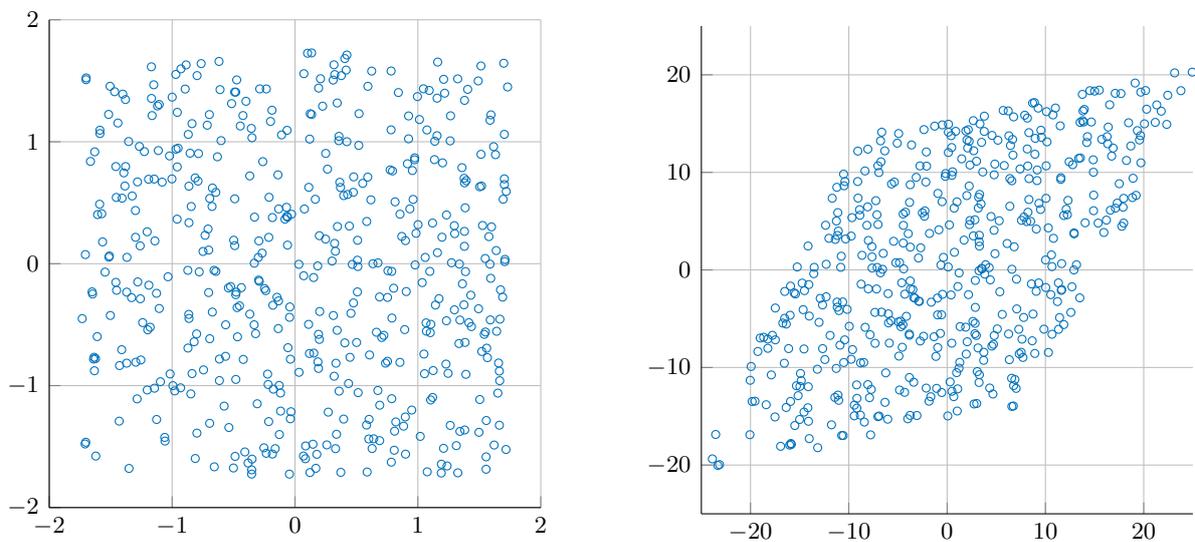
With FastICA explained, we can now present a small example of FastICA being applied to a simple case of two independent signals being mixed.

## 6.4 An Example of ICA

The purpose of this section is to show an example of how ICA works on a very simple case of only two sources, inspired by the example used by [Hyvärinen et al., 2004]. The sources  $\mathbf{s}$  in this example consists of two independent uniform random variables  $\mathbf{s} = [s_1 \ s_2]^T$ , both having zero mean and unit variance. 500 realizations of each source is used throughout this example. The mixing matrix to be used in this example is defined as:

$$\mathbf{A} = \begin{bmatrix} 5 & 10 \\ 10 & 2 \end{bmatrix}$$

A scatter plot of the sources  $\mathbf{s}$  and the mixed signals  $\mathbf{x} = \mathbf{A}\mathbf{s}$  can be seen in Figure 6.3.

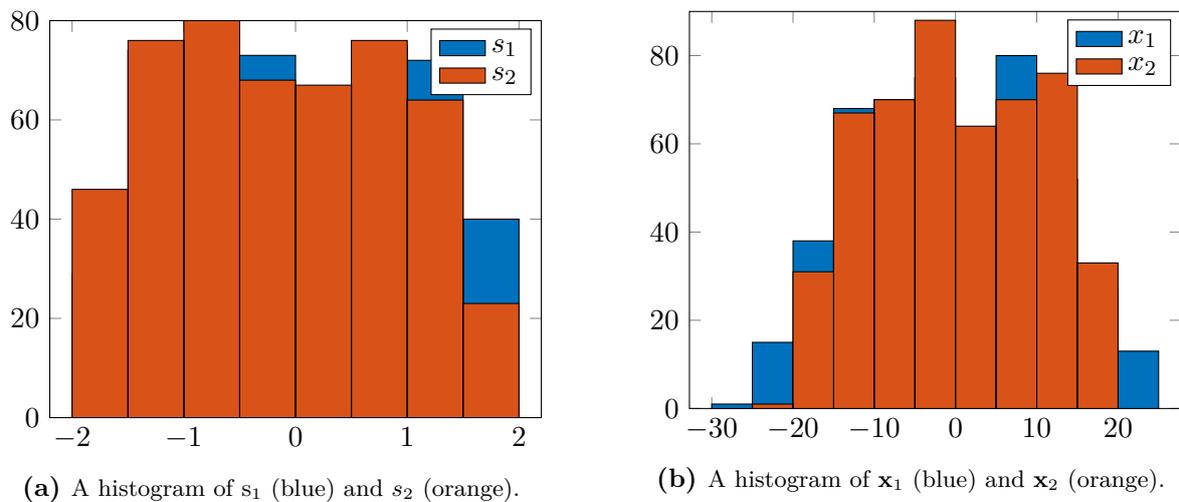


(a) 500 realizations of a 2-dimensional uniform distribution with zero mean and unit variance. This is the source signal  $\mathbf{s} \in \mathbb{R}^{2 \times 500}$

(b) The mixed signal  $\mathbf{x} \in \mathbb{R}^{2 \times 500}$ , i.e.  $\mathbf{s}$  after multiplication with the mixing matrix  $\mathbf{A}$ .

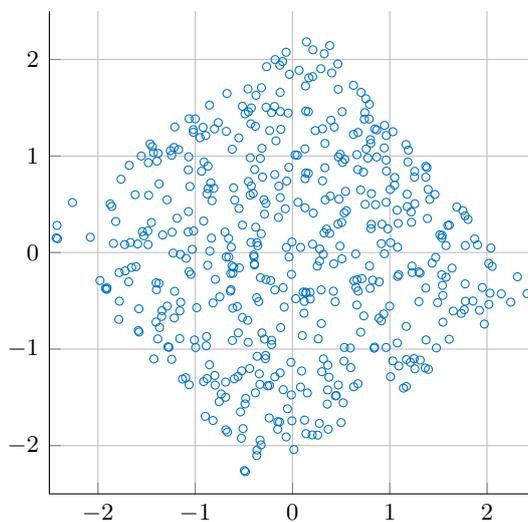
**Figure 6.3:** A 2D uniform source signal  $\mathbf{s}$  and the mixed signals  $\mathbf{x} = \mathbf{A}\mathbf{s}$ .

As ICA looks at the Gaussianity of the sources and mixtures, it is interesting to look at a histogram of  $\mathbf{s}$  and  $\mathbf{x}$ , to get an idea of their distributions. This is done in Figure 6.4, where it is clear that the shape of the histogram of  $\mathbf{s}$  approaches the shape of the uniform pdf, as  $\mathbf{s}$  is a multivariate uniform distribution. It is also clear that  $\mathbf{x}$  is much more Gaussian in its shape than the uniform distributions of  $\mathbf{s}$ , due to the central limit theorem.



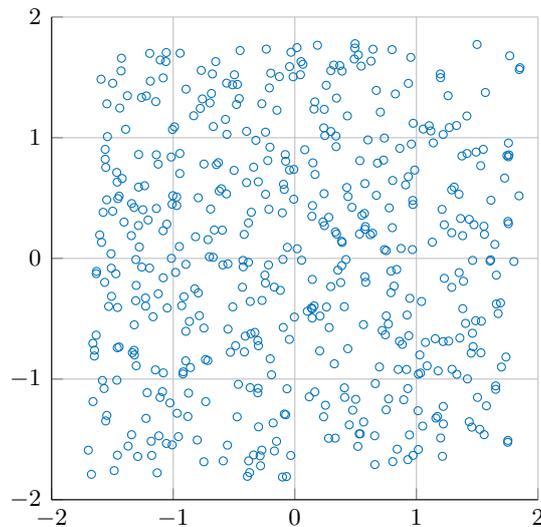
**Figure 6.4:** Histograms of sources  $\mathbf{s}$  and the mixed signals  $\mathbf{x} = \mathbf{A}\mathbf{s}$ .

The first step in most ICA algorithms is to apply whitening/PCA, as this corresponds to making the sources uncorrelated. The result of doing this can be seen in Figure 6.5. This removes the scaling caused by  $\mathbf{A}$ , such that the rhombic shape of the mixture seen in Figure 6.3b is removed, and the original square shape from Figure 6.3a is obtained. But whitening cannot estimate the rotation of the pdf, hence ICA must be applied.



**Figure 6.5:** The mixed signal after performing whitening/PCA. Notice how the rhombic shape of  $\mathbf{x}$  has been removed, but that the overall shape is still rotated relative to  $\mathbf{s}$ .

In order to remove the rotation still present in Figure 6.5, ICA must be applied. In this case, FastICA is applied, as described in Section 6.3. After applying ICA, it is clear that the rotation of the mixture is removed, such that the estimated sources  $\hat{\mathbf{s}}$  in Figure 6.6 are very similar to the original distribution in Figure 6.3a.



**Figure 6.6:** The estimated sources  $\hat{\mathbf{s}}$  after performing fastICA. Notice how the shape and rotation of the original  $\mathbf{s}$  has been estimated.

An interesting thing to notice by comparing Figure 6.6 and Figure 6.3a is how hard it is to see if  $\hat{\mathbf{s}}$  should be rotated by 90, 180 or 270 degrees, or if it should be flipped, either horizontally or vertically, as either will still result in the same square shape of the pdf. These ambiguities correspond exactly to the ambiguities present in ICA: Rotating and flipping the scatter plot  $\hat{\mathbf{s}}$  corresponds to swapping  $\hat{s}_1$  and  $\hat{s}_2$ , and/or changing their signs, and all are equally good estimates of the original uniform distributions.

From this small example it is clear how ICA works, why ICA is better than whitening/PCA, as well as the limitations of ICA.

## 6.5 Infomax

Infomax is another classical algorithm for doing ICA, which is also known as the Bell-Sejnowski algorithm [Bell and Sejnowski, 1995]. Infomax gets its name from the fact that it can be seen as "maximizing the output entropy, or information flow, of a neural network with nonlinear outputs" [Hyvärinen et al., 2004, p. 211].

Infomax can be seen as an optimal estimator of  $\mathbf{B}$ , in a maximum likelihood (ML) sense, assuming the densities  $p_i$  of the original sources  $s_i$  are known. In the following, infomax will be derived using an ML approach. However, as it turns out, the ML approach and information flow approach lead to the same result, and both methods are thus equally valid [Hyvärinen et al., 2004, p. 213].

Infomax assumes the standard mixing and de-mixing equations as with the other ICA algorithms:

$$\mathbf{x} = \mathbf{A}\mathbf{s} \tag{6.91}$$

$$\hat{\mathbf{s}} = \mathbf{B}\mathbf{x} \tag{6.92}$$

Where we are interested in estimating the de-mixing matrix  $\mathbf{B}$ .

Assuming a linear transformation from one random variable  $\mathbf{s} \in \mathbb{R}^K$  to another  $\mathbf{x} = \mathbf{A}\mathbf{s}$ ,  $\mathbf{x} \in \mathbb{R}^M$ , then the densities of  $\mathbf{s}$  and  $\mathbf{x}$  share the following relation [Hyvärinen et al., 2004, p. 36]:

$$p_{\mathbf{x}}(\mathbf{x}) = \frac{1}{|\det(\mathbf{A})|} p_{\mathbf{s}}(\mathbf{s}) \quad (6.93)$$

As we want to find  $\mathbf{B}$  such that  $\mathbf{B}\mathbf{A} = \mathbf{I}$ , we must have that  $1 = \det(\mathbf{B}\mathbf{A}) = \det(\mathbf{A})\det(\mathbf{B})$ . This implies that the expression can be re-written as:

$$p_{\mathbf{x}}(\mathbf{x}) = |\det(\mathbf{B})| p_{\mathbf{s}}(\mathbf{s}) \quad (6.94)$$

$$= |\det(\mathbf{B})| \prod_{k=1}^K p_k(s_k) \quad (6.95)$$

Where we have used the fact that the sources  $s_k$  are mutually independent, meaning their joint pdf is the product of their marginal pdfs.

We can now use the fact that  $\hat{s}_k = \mathbf{b}_k^T \mathbf{x}$  and substitute the true  $s_k$  with this estimate, such that we obtain an expression for the likelihood of  $\mathbf{x}$ , which is a function of  $\mathbf{b}_k$  and  $\mathbf{x}$ :

$$p_{\mathbf{x}}(\mathbf{x}|\mathbf{B}) = |\det(\mathbf{B})| \prod_{k=1}^K p_k(\mathbf{b}_k^T \mathbf{x}) \quad (6.96)$$

We can now form the likelihood-function  $l(\mathbf{x}|\mathbf{B})$ , by taking the product over  $N$  observations of the likelihood  $p_{\mathbf{x}}(\mathbf{x}|\mathbf{B})$ . This means the random variables  $\mathbf{x}$  is now replaced by an observation matrix,  $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N]$ ,  $\mathbf{X} \in \mathbb{R}^{K \times N}$ .

$$l(\mathbf{X}|\mathbf{B}) = \prod_{n=1}^N \prod_{k=1}^K p_k(\mathbf{b}_k^T \mathbf{x}_n) |\det(\mathbf{B})| \quad (6.97)$$

Taking the logarithm of the likelihood function gives us the log-likelihood:

$$\mathcal{L}(\mathbf{X}|\mathbf{B}) = \log l(\mathbf{X}|\mathbf{B}) = \sum_{n=1}^N \sum_{k=1}^K \log(p_k(\mathbf{b}_k^T \mathbf{x}_n)) + N \log(|\det(\mathbf{B})|) \quad (6.98)$$

Dividing both sides by the number of observations  $N$  means the sum over  $N$  on the right hand side can be expressed as an expectation, and we can thus re-introduce  $\mathbf{x}$  as a random variable:

$$\frac{1}{N} \mathcal{L}(\mathbf{x}|\mathbf{B}) = E \left[ \sum_{k=1}^K \log(p_k(\mathbf{b}_k^T \mathbf{x})) \right] + \log(|\det(\mathbf{B})|) \quad (6.99)$$

In order to find the ML estimate of  $\mathbf{B}$ , we must take the derivative of  $\mathcal{L}(\mathbf{B})$  w.r.t.  $\mathbf{B}$  and set it equal to zero, as to find the maximum:

$$0 = \frac{1}{N} \frac{\partial \log(\mathcal{L})}{\partial \mathbf{B}} = (\mathbf{B}^{-1})^T + E[g(\mathbf{B}\mathbf{x})\mathbf{x}^T] \quad (6.100)$$

We have here used the fact that [Bell and Sejnowski, 1995, p. 31]

$$\frac{\partial}{\partial \mathbf{B}} \log |\det(\mathbf{B})| = \mathbf{B}^{-T} \quad (6.101)$$

and introduced the function  $g_k$  which is defined as:

$$g_k(s_k) = \frac{\partial}{\partial \mathbf{B}} \log(p_k(s_k)) = (\log(p_k))' = \frac{p'_k(s_k)}{p_k(s_k)} \quad (6.102)$$

From Equation 6.100 follows the update-rule for infomax, also known as the Bell-Sejnowski-algorithm [Bell and Sejnowski, 1995]:

$$\Delta \mathbf{B} \propto (\mathbf{B}^{-1})^T + E[g(\mathbf{B}\mathbf{x})\mathbf{x}^T] \quad (6.103)$$

In order to avoid inverting  $\mathbf{B}$ , the whole expression is multiplied by  $\mathbf{B}^T \mathbf{B}$  - this rescales the gradient and removes the need for an inversion - this is referred to as the "natural" gradient descent algorithm [Amari et al., 1996, p. 761]

$$\Delta \mathbf{B} \propto \mathbf{B} + E[g(\mathbf{B}\mathbf{x})\mathbf{x}^T \mathbf{B}^T] \mathbf{B} \quad (6.104)$$

$$\propto (\mathbf{I} + E[g(\mathbf{B}\mathbf{x})\hat{\mathbf{s}}]) \mathbf{B}, \quad (6.105)$$

and is also known as the infomax update rule.

As shown in the derivations, by using  $\Delta \mathbf{B}$  as the update step of an iterative solver, one can obtain the ML estimate of  $\mathbf{B}$ , assuming that  $g_k(s_k) = (\log(p_k))' = p'_k(s_k)/p_k(s_k)$ , where  $p_k$  is the pdf of the  $k$ 'th source  $s_k$ . However, the pdfs of the sources is often unknown in practice. Fortunately, one can show that the algorithm is stable and that estimation is possible, if one simply chooses  $g_k(s_k)$  such that it has the correct sign, based on whether  $p_k$  is sub- or super-Gaussian [Hyvärinen et al., 2004, p. 206-207]. This is because the non-polynomial function  $g_k(s_k)$  measures the shape of the density, much like kurtosis - in fact, by using  $g_k = -s_k^3$  we would actually obtain kurtosis [Hyvärinen et al., 2004, p. 207]. This is also why the choice of  $g_k$  coincides very much with the functions measuring non-Gaussianity as mentioned in Equation 6.88.

This means that for super-Gaussian sources, one can simply choose

$$g_k^+(s_k) = -2 \tanh(s_k) \quad (6.106)$$

which is the nonlinearity used in the Bell-Sejnowski-algorithm [Bell and Sejnowski, 1995; Lee et al., 1999].

The infomax algorithm can now be summarized, see Algorithm 2. Here, a step-size  $\alpha$  has been introduced and the de-mixing matrix  $\mathbf{W}$  is used which operates on the whitened data  $\mathbf{z}$ :

---

#### Algorithm 2 Infomax

---

- 1: Subtract mean from data:  $\mathbf{x} \leftarrow \mathbf{x} - E[\mathbf{x}]$
  - 2: Apply whitening:  $\mathbf{z} = \mathbf{V}\mathbf{x}$
  - 3: Choose an initial (e.g. random)  $\mathbf{W}$  where each row is of unit norm
  - 4:  $\hat{\mathbf{s}} = \mathbf{W}\mathbf{z}$  {Estimate sources}
  - 5:  $\mathbf{W} \leftarrow \alpha (\mathbf{I} + E[g(\mathbf{W}\mathbf{z})\hat{\mathbf{s}}]) \mathbf{W}$
  - 6:  $\mathbf{w}_i \leftarrow \mathbf{w}_i / \|\mathbf{w}_i\|, \quad \forall i$  {Normalize rows}
  - 7: If not converged, go back to step 4
  - 8:  $\mathbf{B} = \mathbf{W}\mathbf{V}$
- 

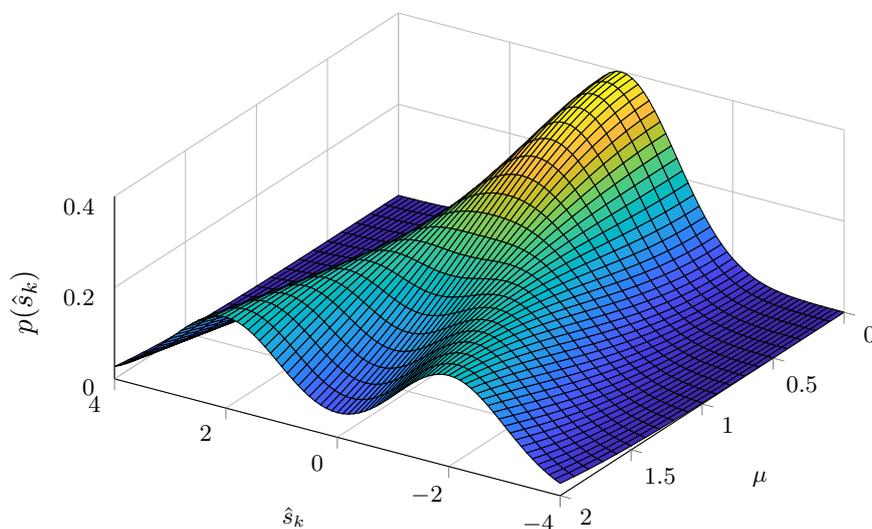
Having derived the infomax algorithm, we can now look into a slight modification of it, called Extended infomax.

### 6.5.1 Extended Infomax

While infomax provides a general framework for estimating  $\mathbf{B}$  when the sub- and super-Gaussianity of the sources are known, extended infomax aims at making "a simple update rule with a fixed nonlinearity that can separate sources with a variety of distributions" [Lee et al., 1999, p. 422]. This is done by first constructing a symmetric pdf which is strictly sub-Gaussian [Lee et al., 1999, p. 422]:

$$p_{\hat{s}_k}(\hat{s}_k) = \frac{1}{2} \left( \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\hat{s}_k - \mu)^2}{2\sigma^2}\right) + \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\hat{s}_k + \mu)^2}{2\sigma^2}\right) \right). \quad (6.107)$$

The pdf  $p_{\hat{s}_k}(\hat{s}_k)$  is thus a sum of two Gaussians with means of  $\pm\mu$ . For  $\mu = 0$  the distribution is Gaussian, and for bigger values of  $\mu$ ,  $p_{\hat{s}_k}(\hat{s}_k)$  becomes more bimodal (i.e. more sub-Gaussian). This is especially clear from looking at a plot of  $p_{\hat{s}_k}(\hat{s}_k)$  for different values of  $\mu$ , which can be seen in Figure 6.7.



**Figure 6.7:** The sub-Gaussian density used for deriving the extended infomax learning rule, for different values of  $\mu$ .

We now want to find an expression for  $g_k(\hat{s}_k)$  for the pdf in Equation 6.107 [Lee et al., 1999, p. 422]:

$$g_k(\hat{s}_k) = \frac{p'_k(\hat{s}_k)}{p_k(\hat{s}_k)} \quad (6.108)$$

$$= -\frac{\hat{s}_k}{\sigma^2} + a \left( \frac{\exp(a\hat{s}_k) - \exp(-a\hat{s}_k)}{\exp(a\hat{s}_k) + \exp(-a\hat{s}_k)} \right) \quad (6.109)$$

where

$$a = \frac{\mu}{\sigma^2}.$$

Using the definition of the hyperbolic tangent, we can simplify this expression:

$$g_k(\hat{s}_k) = -\frac{\hat{s}_k}{\sigma^2} + \frac{\mu}{\sigma^2} \tanh\left(\frac{\mu}{\sigma^2} \hat{s}_k\right), \quad (6.110)$$

which can be further simplified by setting  $\mu = \sigma^2 = 1$  and generalizing the expression to the vector case of  $\hat{\mathbf{s}}$ :

$$g(\hat{\mathbf{s}}) = -\hat{\mathbf{s}} + \tanh(\hat{\mathbf{s}}) \quad (6.111)$$

Inserting this into Equation 6.105 and using the fact that  $\hat{\mathbf{s}} = \mathbf{B}\mathbf{x}$  yields [Lee et al., 1999, p. 423]:

$$\Delta \mathbf{B} \propto (\mathbf{I} + g(\mathbf{B}\mathbf{x})\hat{\mathbf{s}}^T) \mathbf{B} \quad (6.112)$$

$$\propto (\mathbf{I} + (\tanh(\hat{\mathbf{s}}) - \hat{\mathbf{s}})\hat{\mathbf{s}}^T) \mathbf{B} \quad (6.113)$$

$$\propto (\mathbf{I} + \tanh(\hat{\mathbf{s}})\hat{\mathbf{s}}^T - \hat{\mathbf{s}}\hat{\mathbf{s}}^T) \mathbf{B}, \quad (6.114)$$

where the expectation from Equation 6.105 has been omitted to allow for online processing (i.e. using a stochastic gradient approach), although sample means can also be used.

One can do a similar proof for a super-Gaussian source, which results in the following update rule [Lee et al., 1999, p. 423]:

$$\Delta \mathbf{B} \propto (\mathbf{I} - \tanh(\hat{\mathbf{s}})\hat{\mathbf{s}}^T - \hat{\mathbf{s}}\hat{\mathbf{s}}^T) \mathbf{B} \quad (6.115)$$

Combining Equation 6.114 and 6.115, the extended infomax update rule then becomes:

$$\Delta \mathbf{B} \propto (\mathbf{I} - \mathbf{K} \tanh(\hat{\mathbf{s}})\hat{\mathbf{s}}^T - \hat{\mathbf{s}}\hat{\mathbf{s}}^T) \mathbf{B} \quad (6.116)$$

where  $k_i$  is defined as:

$$k_i = \begin{cases} 1 & \text{if } \hat{s}_i \text{ is super-Gaussian} \\ -1 & \text{if } \hat{s}_i \text{ is sub-Gaussian} \end{cases} \quad (6.117)$$

## 6.5.2 Estimating sub- and super-Gaussianity

A crucial part of extended infomax is knowing the sign of the kurtosis, i.e. if the signal is sub- or super-Gaussian. The sign  $k_i$  of the kurtosis of the  $i$ 'th estimated source  $\hat{s}_i$  can be found as [Lee et al., 1999, p. 425]:

$$k_i = \text{sign}(E[\text{sech}^2(\hat{s}_i^2)]E[\hat{s}_i] - E[\tanh(\hat{s}_i)]\hat{s}_i) \quad (6.118)$$

The extended infomax algorithm can now be summarized, see Algorithm 3, where a step-size  $\alpha$  has also been introduced and the matrix  $\mathbf{W}$  is used instead of  $\mathbf{B}$  as the input data is pre-whitened:

---

### Algorithm 3 Extended Infomax

---

- 1: Subtract mean from data:  $\mathbf{x} \leftarrow \mathbf{x} - E[\mathbf{x}]$
  - 2: Apply whitening:  $\mathbf{z} = \mathbf{V}\mathbf{x}$
  - 3: Choose an initial (e.g. random)  $\mathbf{W}$  where each row is of unit norm
  - 4:  $\hat{\mathbf{s}} = \mathbf{W}\mathbf{z}$  {Estimate sources}
  - 5: **for**  $i = 0; i \leq N; i = i + 1$  **do**
  - 6:    $k_i = \text{sign}(E[\text{sech}^2(\hat{s}_i^2)]E[\hat{s}_i] - E[\tanh(\hat{s}_i)]\hat{s}_i)$  {Estimate sub/super-Gaussianity}
  - 7: **end for**
  - 8:  $\mathbf{W} \leftarrow \alpha (\mathbf{I} - \mathbf{K} \tanh(\hat{\mathbf{s}})\hat{\mathbf{s}}^T - \hat{\mathbf{s}}\hat{\mathbf{s}}^T) \mathbf{W}$
  - 9:  $\mathbf{w}_i \leftarrow \mathbf{w}_i / \|\mathbf{w}_i\|, \quad \forall i$  {Normalize rows}
  - 10: If not converged, go back to step 4
  - 11:  $\mathbf{B} = \mathbf{W}\mathbf{V}$
- 

With the classical ICA methods explained, the constrained ICA algorithm can now be derived, which is done in the following section.



## 7 | ICA with Reference

ICA with reference (also referred to as constrained ICA, cICA) is a method for performing independent component analysis, but where some prior information about the sources to be estimated is exploited, as proposed by [Lu and Rajapakse, 2006]. The purpose of this is to "help" the ICA algorithm to extract the desired ICs, as there may be many ways to solve the ICA problem in addition to the true/desired one. In this chapter, cICA is derived, and different methods for creating reference signals are presented.

### 7.1 Deriving the cICA Algorithm

As with the other ICA methods, we want to estimate the de-mixing matrix  $\mathbf{B} = [\mathbf{b}_1^T, \mathbf{b}_2^T, \dots, \mathbf{b}_N^T]^T$  which satisfies that

$$\hat{\mathbf{s}} = \mathbf{B}\mathbf{x} \quad (7.1)$$

ICA with reference builds upon the optimization problem which FastICA tries to solve, as presented in Equation 6.60, which is re-stated in the following:

$$\begin{aligned} \tilde{\mathbf{b}}_i = \arg \min_{\mathbf{b}_i} & -J(\mathbf{b}_i^T \mathbf{x}), \quad i = 1, 2, \dots, N \\ \text{s.t.} & h(\mathbf{B}\mathbf{x}) = \mathbf{0} \end{aligned} \quad (7.2)$$

where:

$\mathbf{b}_i$	is the $i$ 'th row of the de-mixing matrix $\mathbf{B}$
$J(\mathbf{y})$	is the negentropy of $\mathbf{y}$
$\mathbf{x}$	is the observed mixture
$h(\mathbf{y})$	is an equality constraint

Here, an equality constraint function  $h(\mathbf{y})$  is introduced, which constrains the estimated sources as being uncorrelated and having unit norm, i.e. the orthogonality assumption posed in Section 6.2.3:

$$h(\mathbf{y}) = (E[\mathbf{y}\mathbf{y}^T] - \mathbf{I})^2 = \mathbf{0}, \quad (7.3)$$

where the square has been included to follow the notation used in [Lu and Rajapakse, 2006].

What constrained ICA does, is to introduce a similarity/closeness-measure  $g(\hat{\mathbf{s}}, \mathbf{r})$ , between estimated sources  $\hat{\mathbf{s}}$  and some reference signals  $\mathbf{r}$ . The closeness-measure is included in the optimization problem as an inequality constraint, meaning that the estimated sources must be similar to the reference signals:

$$\begin{aligned} \tilde{\mathbf{b}}_i = \arg \min_{\mathbf{b}_i} & -J(\mathbf{b}_i^T \mathbf{x}), \quad i = 1, 2, \dots, N \\ \text{s.t.} & g(\hat{\mathbf{s}}, \mathbf{r}) \leq \mathbf{0} \\ & h(\hat{\mathbf{s}}) = \mathbf{0} \end{aligned} \quad (7.4)$$

In order to solve this constrained ICA problem, a fixed-point<sup>1</sup> algorithm is derived in the following, using the gradient descent-approach. The first step in doing this, is to convert the inequality constraint to an

<sup>1</sup>A fixed-point algorithm is a method for computing fixed-points (e.g. maxima and minima) of a cost function  $J(x)$ , by iteratively evaluating a function  $x_{n+1} = f(x_n)$ , such that the function outputs  $x_n$  converge to the fixed point.

equality constraint, by introducing slack-variables  $\mathbf{v} = [v_1, v_2, \dots, v_N]$ , which satisfies that:

$$\hat{g}_i(\hat{s}_i, r_i) = g_i(\hat{s}_i, r_i) + v_i^2 = 0 \quad (7.5)$$

from which we can write up the augmented Lagrangian [Lu and Rajapakse, 2006, p. 2247]:

$$\begin{aligned} \mathcal{L}(\mathbf{B}, \boldsymbol{\mu}, \boldsymbol{\lambda}, \mathbf{v}) = & -J(\hat{\mathbf{s}}) + \boldsymbol{\mu}^T \hat{g}(\hat{\mathbf{s}}) + \frac{1}{2} \gamma \|\hat{g}(\hat{\mathbf{s}})\|_2^2 \\ & + \boldsymbol{\lambda}^T h(\hat{\mathbf{s}}) + \frac{1}{2} \gamma \|h(\hat{\mathbf{s}})\|_2^2, \end{aligned} \quad (7.6)$$

where the quadratic regularization term  $\frac{1}{2} \gamma \|\cdot\|_2^2$  ensures local convexity by punishing large values of  $\mathbf{B}$ , and  $\gamma > 0$  is a penalty parameter.

We are here interested in minimizing this Lagrangian w.r.t. all the variables, that is  $\mathbf{B}, \boldsymbol{\mu}, \boldsymbol{\lambda}, \mathbf{v}$ . The minimization w.r.t.  $\mathbf{v}$  can be done for fixed  $\mathbf{B}$ :

$$\begin{aligned} \tilde{v}_i = \arg \min_{v_i} \quad & \mathcal{L}(\mathbf{B}, \boldsymbol{\mu}, \boldsymbol{\lambda}, \mathbf{v}), \quad \forall i = 1, 2, \dots, N \\ = \arg \min_{v_i^2 \geq 0} \quad & \mu_i(g(\hat{s}_i) + v_i) + \frac{1}{2} \gamma \|g(\hat{s}_i) + v_i\|_2^2, \end{aligned} \quad (7.7)$$

where all terms in  $\mathcal{L}(\mathbf{B}, \boldsymbol{\mu}, \boldsymbol{\lambda}, \mathbf{v})$  not involving  $\mathbf{v}$  have been ignored, as they do not change the maximizer.

At the minimum,  $\nabla_{v_i} \mathcal{L} = 0$  and  $v_i^2 \geq 0$ . Setting  $\nabla_{v_i} \mathcal{L} = 0$  and solving for  $v_i^2$  gives

$$\tilde{v}_i^2 = \max \left( 0, - \left( \frac{\mu_i}{\gamma} + g_i(y_i) \right) \right), \quad (7.8)$$

where the second term comes from direct isolation of  $v_i^2$  and the first term comes from the fact that  $v_i^2$  must be non-negative. Substituting  $\tilde{v}_i^2$  into Equation 7.6 gives the following Lagrangian (where the slack variable  $\mathbf{v}$  has been removed) [Lu and Rajapakse, 2006, p. 2247]:

$$\mathcal{L}(\mathbf{B}, \boldsymbol{\mu}, \boldsymbol{\lambda}) = -J(\hat{\mathbf{s}}) + \mathcal{G}(\hat{\mathbf{s}}, \boldsymbol{\mu}) + \mathcal{H}(\hat{\mathbf{s}}, \boldsymbol{\lambda}), \quad (7.9)$$

Where the function  $\mathcal{G}_i(\hat{s}_i, \mu_i)$  is defined as

$$\mathcal{G}_i(\hat{s}_i, \mu_i) = \frac{1}{2\gamma} (\max(0, \mu_i + \gamma g(\hat{s}_i))^2 - \mu_i^2) \quad (7.10)$$

$$= \begin{cases} \mu_i g(\hat{s}_i) + \frac{1}{2\gamma} \|g(\hat{s}_i)\|_2^2 & \text{if } \mu_i + \gamma g_i(\hat{s}_i) \geq 0 \\ -\mu_i^2 / 2\gamma & \text{otherwise} \end{cases} \quad (7.11)$$

and  $\mathcal{H}(\hat{\mathbf{s}}, \boldsymbol{\lambda})$  is:

$$\mathcal{H}(\hat{\mathbf{s}}, \boldsymbol{\lambda}) = \boldsymbol{\lambda}^T h(\hat{\mathbf{s}}) + \frac{1}{2} \gamma \|h(\hat{\mathbf{s}})\|_2^2 \quad (7.12)$$

### 7.1.1 Update rule for B-matrix

Having formed an augmented Lagrangian with the slack-variables accounted for, we can now take the gradient w.r.t.  $\mathbf{B}$  [Lu and Rajapakse, 2006, p. 2248]:

$$\nabla_{\mathbf{B}} \mathcal{L}(\mathbf{B}, \boldsymbol{\mu}, \boldsymbol{\lambda}) = E[(-\nabla_{\hat{\mathbf{s}}} J(\hat{\mathbf{s}})) \mathbf{x}^T] + E[(\nabla_{\hat{\mathbf{s}}} \mathcal{G}(\hat{\mathbf{s}}, \boldsymbol{\mu})) \mathbf{x}^T] + (2\nabla_{\Sigma_{\hat{\mathbf{s}}}} \mathcal{H}(\hat{\mathbf{s}}, \boldsymbol{\lambda})) E[\hat{\mathbf{s}} \mathbf{x}^T], \quad (7.13)$$

where the expectations originate from the definitions of  $J(\hat{\mathbf{s}})$ ,  $g(\hat{\mathbf{s}}, \mathbf{r})$  and  $h(\hat{\mathbf{s}})$ . Calculating the gradients yield the following [Lu and Rajapakse, 2006, p. 2248]:

$$-\nabla_{\hat{\mathbf{s}}_i} J(\hat{\mathbf{s}}_i) = -\hat{\rho} \cdot f'_{\hat{\mathbf{s}}_i}(\hat{\mathbf{s}}_i) \quad (7.14)$$

$$= 2(E[f_i(\hat{\mathbf{s}}_i)] - E[f_i(v)]) \cdot f'_{\hat{\mathbf{s}}_i}(\hat{\mathbf{s}}_i) \quad (7.15)$$

$$\nabla_{\hat{\mathbf{s}}_i} \mathcal{G}(\hat{\mathbf{s}}_i, r_i) = \mu_i \cdot g'_{\hat{\mathbf{s}}_i}(\hat{\mathbf{s}}_i) \quad (7.16)$$

$$\nabla_{\Sigma_{\hat{\mathbf{s}}_i, \hat{\mathbf{s}}_j}} \mathcal{H}(\hat{\mathbf{s}}_i) = 2\lambda_{ij} \cdot (E[(\hat{\mathbf{s}}_i)(\hat{\mathbf{s}}_j)] - \delta_{ij}) \quad (7.17)$$

where:

$\delta_{ij}$	is the Kronecker delta function
$f'_{\hat{\mathbf{s}}_i}(\hat{\mathbf{s}}_i)$	is the derivative of the contrast function, see Equation 6.88
$g'_{\hat{\mathbf{s}}_i}(\hat{\mathbf{s}}_i)$	is the derivative of the constraint function $g_i(\hat{\mathbf{s}}_i)$

The actual function rules for  $f, f', g$  and  $g'$  will be described in Section 7.1.3 and 7.1.4, where it is investigated which contrast function and closeness measure to use.

The update rule used by [Lu and Rajapakse, 2006] is the well-known *method of multipliers* [Boyd et al., 2011, p. 11]:

$$\mathbf{B}_{n+1} = \arg \min_{\mathbf{B}} \mathcal{L}(\mathbf{B}, \boldsymbol{\mu}_n, \boldsymbol{\lambda}_n) \quad (7.18)$$

$$\boldsymbol{\mu}_{n+1} = \boldsymbol{\mu}_n + \arg \min_{\boldsymbol{\mu}} \mathcal{L}(\mathbf{B}_{n+1}, \boldsymbol{\mu}, \boldsymbol{\lambda}_n) \quad (7.19)$$

$$\boldsymbol{\lambda}_{n+1} = \boldsymbol{\lambda}_n + \arg \min_{\boldsymbol{\lambda}} \mathcal{L}(\mathbf{B}_{n+1}, \boldsymbol{\mu}_n, \boldsymbol{\lambda}) \quad (7.20)$$

By first looking at the update rule for  $\mathbf{B}$ , we can use the steepest descent algorithm:

$$\mathbf{B}_{n+1} = \mathbf{B}_n - \alpha \cdot \Delta \mathbf{B} \quad (7.21)$$

where:

$\alpha$	is the step-size/learning rate
----------	--------------------------------

The update step  $\Delta \mathbf{B}$  is simply found as:

$$\Delta \mathbf{B} = \nabla_{\mathbf{B}} \mathcal{L}(\mathbf{B}, \boldsymbol{\mu}, \boldsymbol{\lambda}) \quad (7.22)$$

Having described the update rule for  $\mathbf{B}$  by means of the gradient of the Lagrangian, the Lagrange multipliers must also be updated. This is done in the following.

## 7.1.2 Update rule for Lagrange multipliers

Building upon the method of multipliers-rule, we here aim to find an update rule for the two Lagrange multipliers  $\boldsymbol{\mu}, \boldsymbol{\lambda}$ .

This is done by taking the derivative of  $\mathcal{L}(\mathbf{B}, \boldsymbol{\mu}, \boldsymbol{\lambda})$  w.r.t.  $\boldsymbol{\mu}$  and  $\boldsymbol{\lambda}$ .

For  $\boldsymbol{\mu}$ , all terms but  $\mathcal{G}(\hat{\mathbf{s}}, \mathbf{r})$  are constants w.r.t.  $\boldsymbol{\mu}$ , and we thus need to only consider this term. The update rule for  $v\boldsymbol{\mu}$  thus becomes:

$$\Delta\boldsymbol{\mu} = \nabla_{\boldsymbol{\mu}}\mathcal{L}(\mathbf{B}, \boldsymbol{\mu}, \boldsymbol{\lambda}) \quad (7.23)$$

$$= \nabla_{\boldsymbol{\mu}}\mathcal{G}(\hat{\mathbf{s}}, \boldsymbol{\mu}) \quad (7.24)$$

$$= \nabla_{\boldsymbol{\mu}} \left( \frac{1}{2\gamma} (\max(0, \boldsymbol{\mu} + \gamma g(\hat{\mathbf{s}}))^2 - \boldsymbol{\mu}^2) \right) \quad (7.25)$$

$$= \max(-\boldsymbol{\mu}, \gamma g(\hat{\mathbf{s}})), \quad (7.26)$$

where  $\max$  is the element-wise maximum function, and where the end-result has been multiplied by  $\gamma$  as is also done by [Lu and Rajapakse, 2006].

For  $\boldsymbol{\lambda}$ , we only need to consider  $\mathcal{H}(\hat{\mathbf{s}})$ , as all other terms of  $\mathcal{L}$  are constants w.r.t  $\boldsymbol{\lambda}$ . The update rule for  $\boldsymbol{\lambda}$  thus becomes:

$$\Delta\boldsymbol{\lambda} = \nabla_{\boldsymbol{\lambda}}\mathcal{L}(\mathbf{B}, \boldsymbol{\mu}, \boldsymbol{\lambda}) \quad (7.27)$$

$$= \nabla_{\boldsymbol{\lambda}}\mathcal{H}(\hat{\mathbf{s}}, \boldsymbol{\lambda}) \quad (7.28)$$

$$= \nabla_{\boldsymbol{\lambda}} \left( \boldsymbol{\lambda}^T h(\hat{\mathbf{s}}) + \frac{1}{2}\gamma \|h(\hat{\mathbf{s}})\|_2^2 \right) \quad (7.29)$$

$$= \gamma \cdot h(\hat{\mathbf{s}}), \quad (7.30)$$

where the end-result has been multiplied by  $\gamma$  as is also done by [Lu and Rajapakse, 2006].

We have now specified the update rules for both  $\mathbf{B}, \boldsymbol{\mu}$  and  $\boldsymbol{\lambda}$ . We can now describe how to choose the contrast function  $f(\hat{\mathbf{s}})$ .

### 7.1.3 Choice of contrast function

As with FastICA, the functions used to approximate the negentropy depend on the sign of the kurtosis of the (estimated) sources. The following super- and sub-Gaussian contrast functions are proposed by [Lu and Rajapakse, 2006, p. 2249]:

$$f_{\text{sup}}(y) = \frac{1}{a} \log \cosh(ay) - \frac{a}{2}y^2 \quad (7.31)$$

$$f_{\text{sub}}(y) = \frac{b}{4}y^4 \quad (7.32)$$

where:

$a, b$  are some positive constants

And their derivatives are found as

$$f'_{\text{sup}}(y) = \tanh(ay) - ay \quad (7.33)$$

$$f'_{\text{sub}}(y) = by^3 \quad (7.34)$$

Which are, interestingly enough, the same contrast functions proposed for FastICA (see Equation 6.74 and 6.88) and extended infomax (see Equation 6.111).

For determining which contrast function to use, the sign of the Kurtosis can be estimated using the same equation as in Extended Infomax, see Equation 6.118.

### 7.1.4 Choice of closeness measure

The final step in deriving the constrained ICA algorithm is to look into which closeness-measure to use.

The measure is a function of the type

$$g_i(y_i, r_i) = \xi_i(y_i, r_i) - \epsilon_i \leq 0 \quad (7.35)$$

where:

- $\xi(y_i, r_i)$  is a closeness measure between  $y_i$  and  $r_i$
- $\epsilon_i$  is a threshold parameter

Two obvious candidates of closeness measures are

- Mean-squared error (MSE):  $\xi_i(y_i, r_i) = E[(y_i - r_i)^2]$
- (Squared) correlation:  $\xi_i(y_i, r_i) = E[y_i r_i]^2$

In this case, the correlation measure is chosen: Partly as correlation is more robust towards (white) noise (as it is uncorrelated with all signals), and partly because the correlation measure is used by [Lu and Rajapakse, 2006]. By squaring the correlation, the constraint becomes non-negative, and simultaneously makes it invariant to the sign of the estimated sources. This is an advantage, as the sign of the sources may not be known.

However, the correlation is inverted by [Lu and Rajapakse, 2006], such that

$$\xi_i(y_i, r_i) = 1/(E[y_i r_i]^2) \quad (7.36)$$

This is partly to satisfy that  $g(y_i, r_i) \leq 0$  (and not  $\geq 0$ ), but the inversion also ensures that (too) small correlations are punished hard, due to the reciprocal. Note also that the correlation is normalized, such that the autocorrelations of  $y_i$  and  $r_i$  equal one:

$$R(y_i, r_i) = \frac{E[y_i r_i]}{\sqrt{E[y_i^2] \cdot E[r_i^2]}} \quad (7.37)$$

The correlation constraint then becomes:

$$g(y_i, r_i) = \frac{1}{R(y_i, r_i)^2} - \frac{1}{\epsilon_i^2} \quad (7.38)$$

And the derivative of  $g(y_i, r_i)$  is (ignoring normalizations) [Lu and Rajapakse, 2006, p. 2249]:

$$g'(y_i, r_i) = -2 \frac{E[r_i]}{(E[y_i r_i])^3} \quad (7.39)$$

As the threshold parameter  $\epsilon_i$  is also inverted and squared, we can interpret  $0 \leq \epsilon_i \leq 1$  as being the desired (minimum) correlation between  $y_i$  and  $r_i$ .

Choosing  $\epsilon_i$  is however not straight-forward. If one for instance chooses  $\epsilon_i = 0.9$ , this will most likely not result in a feasible solution. The reason is that the initial guess for an optimization problem should be feasible, and should thus fulfill that  $g(y_i, r_i) = \frac{1}{R(y_i, r_i)^2} - \frac{1}{0.9^2} \leq 0$ . In order to ensure the initial guess is feasible, one should start the algorithm by setting  $\epsilon_i$  very low (e.g. 0), and then slowly increase  $\epsilon_i$  for each iteration, towards the desired threshold. This will ensure the estimated sources always remain feasible.

### 7.1.5 Momentum

In steepest descent, the following update rule of the decision variable  $\boldsymbol{\theta}$  (i.e. the variable we are optimizing over) is used:

$$\Delta\boldsymbol{\theta}_n = \alpha \cdot \nabla J(\boldsymbol{\theta}_n) \quad (7.40)$$

$$\boldsymbol{\theta}_{n+1} = \boldsymbol{\theta}_n - \Delta\boldsymbol{\theta}_n \quad (7.41)$$

where:

$\boldsymbol{\theta}$	is the decision variable
$\alpha$	is the step size/learning rate
$\nabla J(\boldsymbol{\theta})$	is the gradient of the cost function $J(\boldsymbol{\theta})$

In steepest descent, a step is taken in the direction of the negative gradient, after which the gradient is evaluated again, and a new step is taken. There is therefore no "memory" in this approach, as the next update step only takes the current gradient into account.

The intuition behind using momentum is to imagine a ball rolling on a field or a hill (the cost function landscape), down towards the minimum. As there might be many local minima (bumps on the hill), it might sometimes be necessary to go over small hill, in order to be able to continue further down towards the global minimum.

The name momentum comes from the analogy of the momentum this ball rolling on the hill has: Once it gets rolling (down hill) in a certain direction, it will mostly continue rolling in that direction, and not change direction all the time as with steepest descent - despite any small bumps in the way [Qian, 1999, p. 146]. This means the algorithm will converge faster, and is less likely to get stuck in a local minimum.

The update rule for momentum becomes [Qian, 1999, p. 145]:

$$\Delta\boldsymbol{\theta}_n = \eta \cdot \Delta\boldsymbol{\theta}_{n-1} + \alpha \cdot \nabla J(\boldsymbol{\theta}_n) \quad (7.42)$$

$$\boldsymbol{\theta}_{n+1} = \boldsymbol{\theta}_n - \Delta\boldsymbol{\theta}_n \quad (7.43)$$

where:

$\eta$	is the momentum term
--------	----------------------

Typically, one chooses  $\eta = 0.9$  or a similar value [Qian, 1999, p. 149], but depends on how many variables the gradient is taken over, i.e. the sample variance of the data for the current application. Varieties of momentum exists, such as Nesterov accelerated momentum, but here, the basic version is used.

## 7.2 Summary of cICA Algorithm

The cICA has now been derived, and the various functions to be used have also been described. A few modifications are however made between what is described above, and the final algorithm:

It is decided to pre-whiten the input,  $\mathbf{z} = \mathbf{V}\mathbf{x}$  as proposed by [Lin et al., 2007]. even though [Lu and Rajapakse, 2006] state that it is not necessary. Applying pre-whitening does not change the separation performance, but makes convergence twice as fast [Lin et al., 2007, p. 1274].

Pre-whitening has the implication that the de-mixing matrix  $\mathbf{W}$  is estimated instead of  $\mathbf{B}$ , where  $\mathbf{W}$  satisfies that  $\hat{\mathbf{s}} = \mathbf{W}\mathbf{z}$ .

It also means the equality constraint  $\mathcal{H}(\hat{\mathbf{s}})$  can be omitted, and instead replaced by a normalization of

the rows of  $\mathbf{W}$ . This also removes the need for calculating the gradient of  $\mathcal{H}(\hat{\mathbf{s}})$ , thus making calculations faster.

A while-loop in the algorithm ensures that it does not run more than  $\text{MAX\_ITER} = 5000$  iterations. It is however needed to check for each iteration if the algorithm has converged - this is the case if:

$$\|W_n - W_{n-2}\|_\infty < \epsilon_W \quad (7.44)$$

where:

$W_n$	is the de-mixing matrix at the $n$ 'th iteration
$\epsilon_W$	is the convergence threshold

A secondary criterion for convergence is of course also that  $g(\hat{\mathbf{s}}) \leq \mathbf{0}$ , i.e. the algorithm converged to a *feasible* solution. It is thus possible to write up the full cICA algorithm, as summarized in Algorithm 4.

---

**Algorithm 4** cICA
 

---

```

1:  $\mathbf{x} \leftarrow \mathbf{x} - E[\mathbf{x}]$  {Subtract mean from data}
2:  $\mathbf{r} \leftarrow (\mathbf{r} - E[\mathbf{r}])/\sigma_{\mathbf{r}}$  {Subtract mean from references and scale to unit variance}
3:  $\mathbf{z} = \mathbf{V}\mathbf{x}$  {Apply whitening}
4: Choose an initial (e.g. random)  $\mathbf{w}$  of unit norm
5: while  $n < \text{MAX\_ITER}$  do
6:    $\hat{\mathbf{s}} = \mathbf{W}_n\mathbf{z}$ 
7:    $\Delta\boldsymbol{\mu} = \max(-\boldsymbol{\mu}, \gamma g(\hat{\mathbf{s}}))$ 
8:    $\boldsymbol{\mu} \leftarrow \boldsymbol{\mu} + \Delta\boldsymbol{\mu}$  {Update Lagrange multiplier}
9:   Increase  $\epsilon_i$  towards desired  $\epsilon_i$ 
10:  for  $i = 0; i \leq N; i = i + 1$  do
11:     $k_i = \text{sign}(E[\text{sech}^2(\hat{s}_i^2)]E[\hat{s}_i] - E[\tanh(\hat{s}_i)]\hat{s}_i)$  {Estimate sub/super-Gaussianity}
12:    if  $k_i > 0$  then
13:       $-\nabla_{\hat{s}_i} J = 2(E[f_{i,\text{sup}}(\hat{s}_i)] - E[f_{i,\text{sup}}(v)]) \cdot f'_{i,\text{sup}}\hat{s}_i(\hat{s}_i)$  {Super-Gaussian}
14:    else
15:       $-\nabla_{\hat{s}_i} J = 2(E[f_{i,\text{sub}}(\hat{s}_i)] - E[f_{i,\text{sub}}(v)]) \cdot f'_{i,\text{sub}}\hat{s}_i(\hat{s}_i)$  {Sub-Gaussian}
16:    end if
17:  end for
18:   $\nabla_{\hat{\mathbf{s}}}\mathcal{G}(\hat{\mathbf{s}}, \boldsymbol{\mu}) = \boldsymbol{\mu} \cdot g'_{\hat{\mathbf{s}}}(\hat{\mathbf{s}})$  {Constraint-gradient}
19:   $\nabla_{\mathbf{W}}\mathcal{L}(\mathbf{B}, \boldsymbol{\mu}) = -\nabla_{\hat{\mathbf{s}}}J(\mathbf{B}) + \nabla_{\hat{\mathbf{s}}}\mathcal{G}(\hat{\mathbf{s}}, \boldsymbol{\mu})$  {Full gradient of augmented Lagrangian}
20:   $\Delta\mathbf{W} \leftarrow \eta\Delta\mathbf{W} + \alpha\nabla_{\mathbf{W}}\mathcal{L}(\mathbf{B}, \boldsymbol{\mu})$  {Step-direction}
21:   $\mathbf{W}_n = \mathbf{W}_{n-1} - \Delta\mathbf{W}$  {Momentum update}
22:   $\mathbf{w}_i \leftarrow \mathbf{w}_i/\|\mathbf{w}_i\|_2, \quad \forall i = 1, 2, \dots, N$  {Normalize the rows of  $\mathbf{W}$ }
23:   $n = n + 1$ 
24:  if  $\|W_n - W_{n-2}\|_\infty < \epsilon_W$  and  $g(\hat{\mathbf{s}}) \leq \mathbf{0}$  then
25:    break {Converged at a feasible solution}
26:  end if
27: end while
28:  $\hat{\mathbf{s}} = \mathbf{W}\mathbf{x}$ 
    
```

---

With the constrained ICA algorithm described, the next step is to consider which reference signals to use, as this is of great importance for how the algorithm will perform. This is therefore described in the following section.

## 7.3 Generating Reference Signals

An important part of the constrained ICA algorithm which has not yet been considered, is choosing the fetal and maternal reference signals. This will therefore be described in the following.

Three different reference types are described in this section:

- Square and biphasic signals
- Maternal ECG template
- Fetal ECG template from scalp electrode data

In the original cICA paper, pulse signals are used as references, which have the same periodicity as the signals(s) of interest [Lu and Rajapakse, 2006, p. 2250]. In the case of non-invasive fetal ECG, the periodicity of the signals are given by the maternal and fetal heart rates. In order to estimate these, MQRS and FQRS locations must be estimated.

Estimating FQRS locations is not a straight-forward procedure: This is obvious from the fact that it was a part of the Challenge, see Section 3.1.3. As described in Section 3.3.2, different algorithms exist, which are able to estimate FQRS locations rather well. Implementing and tuning these algorithms is however beyond the scope of this project, and the FQRS locations will therefore be assumed known, as is also described in section 5.4.

A method for estimating maternal QRS locations is however needed. The de-facto standard algorithm for QRS estimation is the Pan-Tompkins algorithm [Pan and Tompkins, 1985], which is therefore also used here. With the maternal and fetal QRS locations known, or at least accurately estimated, the reference signals can be generated.

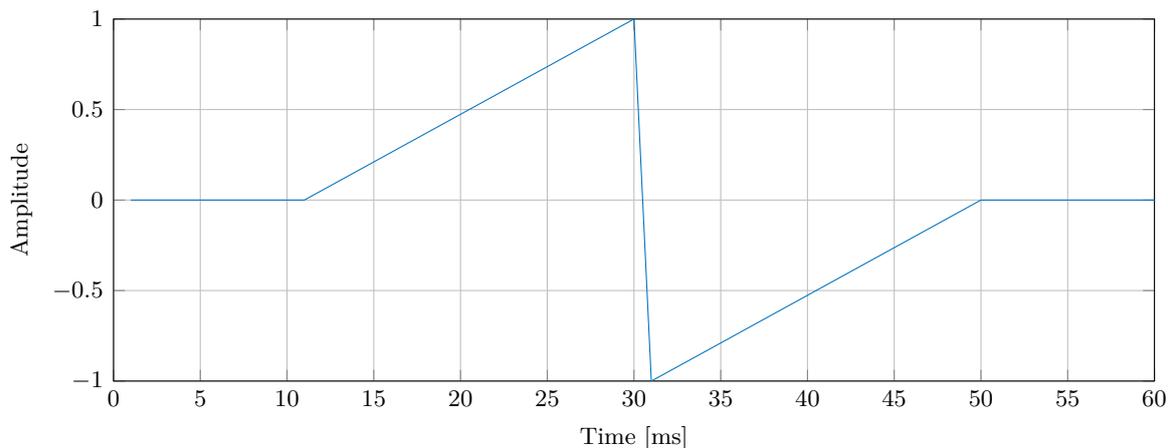
Examples of reference signals generated based on a recorded abdominal signal can be seen in Figure 7.3.

### 7.3.1 Square & biphasic signals

The simplest reference signal is a series of unit-amplitude pulses, located at the QRS locations. The only parameter to choose here is the pulse-width, as it should be similar to the QRS complex width. Experiments show a width of 16 ms for both the maternal and fetal references gives good results, in agreement with the (maternal) QRS duration ranging from 6-100 ms, see section 2.3.

The fetal QRS has a typical duration of 20-50 ms, see Section 2.3. Having a slightly narrower reference signal thus allows for some uncertainty in the exact location of where the QRS is located.

Another option is to use a biphasic signal, which is more representative of the biphasic nature of the QRS complex (depending on which ECG lead is used), see Figure 7.1.



**Figure 7.1:** A biphasic reference signal, centered around 30 ms.

### Including a T-wave

In order to make the (fetal) reference signal more accurate, a T-wave can be included, which is modeled as having a Gaussian shape. From Figure 4.3, the mean QT interval is approx. 200 ms, which should also be the case for the reference signal.

It is decided to make the T-wave have an amplitude of 0.25, based on the fact that the T/QRS ratio is less than 0.3, see Section 4.2.3. The duration of the T-wave must also be determined: As no literature is available on the subject, the duration is found by visual inspection of scalp electrode data and simulated FECGs, from which a duration of 100 ms is chosen. An example of a biphasic signal with T-wave can be seen in Figure 7.3c.

### 7.3.2 Maternal ECG template

Generating MECG reference signals from templates, is based on the template subtraction (TS) algorithm, used for estimating FECG, see Section 3.3.3.

#### Unweighted template

The (unweighted) template generation method uses the QRS annotations of the MECG to extract MECG cycles. These cycles are aligned and stacked, after which a template is generated as the mean of the stacked cycles.

In order to account for the morphological change of the ECG over time, the templates are continuously updated for each beat, i.e. the template becomes a running median of the last  $K$  MECG cycles. The code implemented this algorithm is taken from [Behar et al., 2014d].

#### Weighted template

The weighted template subtraction method builds upon the paper by [Martens et al., 2007]. Here, a template is generated separately for the P-wave, QRS and T-wave (all templates are updated over time) in an approach similar to the unweighted template. However, each sub-template is then scaled/weighted

individually [Behar et al., 2014d, p. 1569]. Assuming a template  $\mathbf{T}$ :

$$\mathbf{T} = \begin{bmatrix} \mathbf{t}_P & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{t}_{QRS} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{t}_T \end{bmatrix}, \quad (7.45)$$

we want to find the scaling vector  $\mathbf{a} \in \mathbb{R}^3$  which minimizes the MSE between the template and the maternal ECG cycle  $\mathbf{m}$ :

$$\|\mathbf{a}\mathbf{T} - \mathbf{m}\|_2^2 \quad (7.46)$$

Where we know the solution is provided by the pseudo-inverse of  $\mathbf{T}$ :

$$\mathbf{a} = (\mathbf{T}^T\mathbf{T})^{-1}\mathbf{T}^T\mathbf{m} \quad (7.47)$$

### Template from PCA

As described in Section 3.3.3, TS-PCA stacks MECG cycles, then selects some of the principal components after which a back-propagation step takes place on a beat-to-beat basis, thus producing MECG estimates every cycle [Andreotti et al., 2016, p. 632]. This approach also updates the templates over time. The implementation of TS-PCA is taken from [Behar et al., 2014d].

### 7.3.3 Fetal ECG template from scalp electrode

The template methods mentioned above require that clear ECG cycles available, and are therefore not applicable for generating FECG templates.

For this reason, another approach must be taken: As the scalp electrode is the golden standard in FECG morphology analysis, one could generate a template using the scalp data.

This is done using the data from the ADFECGDB data set, where the FQRS notations have been annotated. Using these annotations, all FECG cycles for all 5 subject are extracted and stacked. A cycle is taken as being from 200 ms before the QRS location, to 300 ms after, i.e. the cycles have a length  $L = 500$  samples as  $f_s = 1$  kHz.

From this stack of cycles, templates are generated using three approaches:

- Taking the mean of all cycles
- Taking the median of all cycles
- Taking the first principal component (PC) of all cycles, and then taking the mean of these simplified cycles.

The PCA approximation is found by first performing an EVD of the sample covariance matrix of the stacked cycles,  $\mathbf{X} \in \mathbb{R}^{N \times L}$ :

$$\mathbf{U}\mathbf{\Lambda}\mathbf{U}^T = \frac{1}{N-1} \sum_{i=1}^N (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T \quad (7.48)$$

where:

- $\mathbf{U}$  is the matrix of (sorted) normalized eigenvectors corresponding to  $\mathbf{\Lambda}$
- $\mathbf{\Lambda}$  is a diagonal matrix of sorted eigenvalues .
- $\mathbf{x}_i$  is the  $i$ 'th row of the  $N \times L$  observation matrix  $\mathbf{X}$

and where  $\bar{\mathbf{x}}$  is the mean observation, i.e. the mean of all rows of  $\mathbf{X}$ :

$$\bar{\mathbf{x}} = \frac{1}{N} \sum_{i=1}^N \mathbf{x}_i \quad (7.49)$$

The  $D$ 'th order PCA approximation of the  $n$ 'th ECG cycle can be found as [Bishop, 2006, p. 566]:

$$\tilde{\mathbf{x}}_n = \bar{\mathbf{x}} + \sum_{i=1}^D (\mathbf{x}_n^T \mathbf{u}_i - \bar{\mathbf{x}}^T \mathbf{u}_i) \mathbf{u}_i \quad (7.50)$$

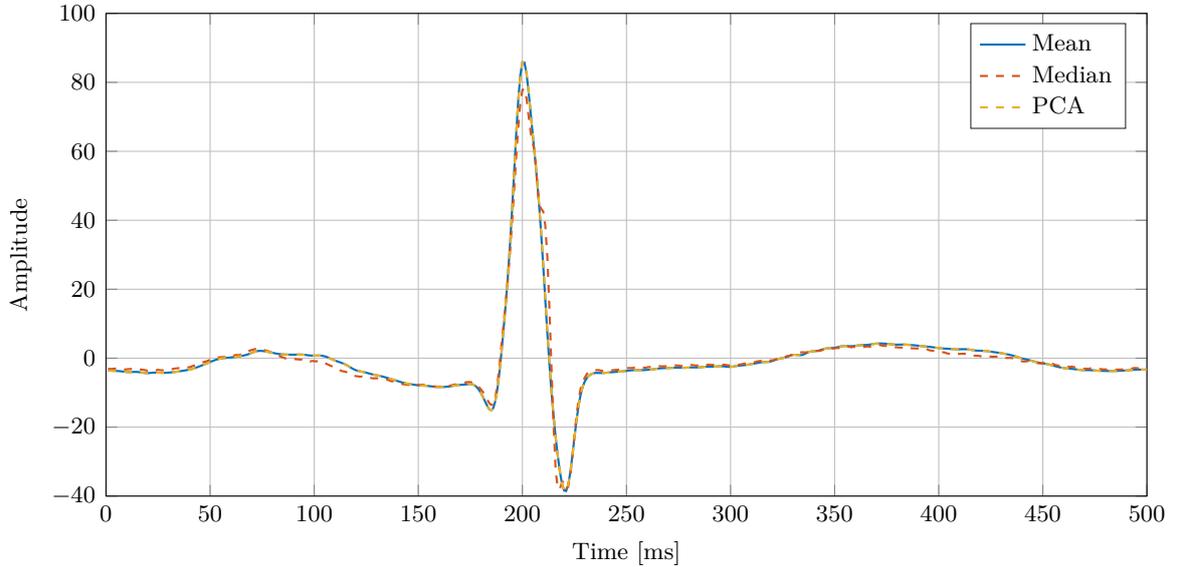
where:

- $\mathbf{u}_i$  is the  $i$ 'th eigenvector

The template  $\tilde{\mathbf{x}}$  is then found as the average of the approximated cycles:

$$\tilde{\mathbf{x}} = \frac{1}{N} \sum_{i=1}^N \tilde{\mathbf{x}}_n \quad (7.51)$$

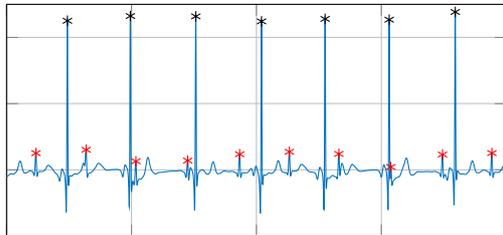
We here choose  $D = 1$ , as this provides satisfactory results. The resulting FECC templates can be seen in Figure 7.2.



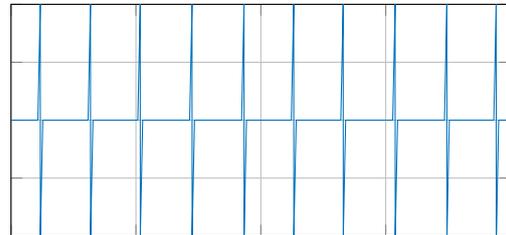
**Figure 7.2:** The generated scalp electrode template, using either the mean of all FECC cycles, the median of all cycles or the first principal component.

As there is no significant difference between the three template types, the PCA approach is selected.

Having described different methods for generating reference signals, both maternal and fetal, the different methods can now be compared to a recorded abdominal signal. This is done in Figure 7.3.



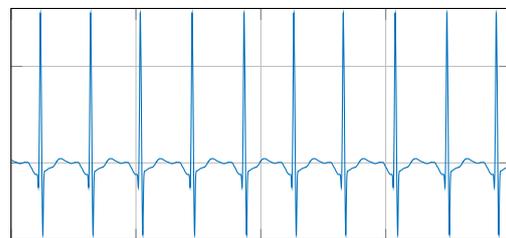
(a) Recorded signal. Black is MQRS, red is FQRS.



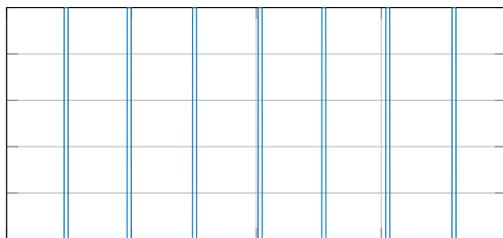
(b) Biphasic fetal reference.



(c) Biphasic + T-wave fetal reference.



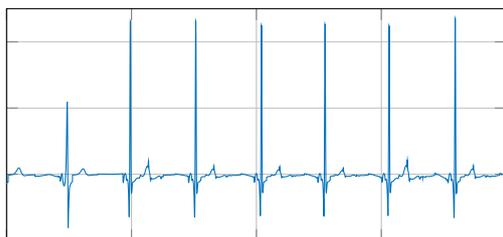
(d) Fetal scalp electrode reference.



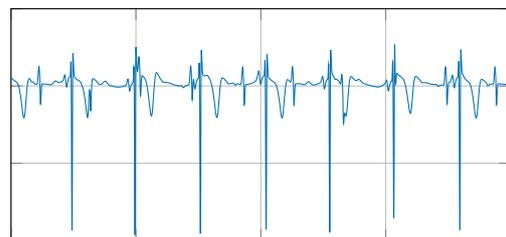
(e) Maternal square wave reference.



(f) Maternal unweighted template reference.



(g) Maternal weighted template reference.



(h) Maternal PCA template reference.

**Figure 7.3:** Original signal (a) and corresponding fetal references (c-d) and maternal references (e-h). Duration of all signals is 5 sec, and amplitudes are not included for visualization purposes

### 7.3.4 Using references as an initial guess for $\mathbf{B}$

Previously, a random initialization of the de-mixing matrix  $\mathbf{B}$  has been used. However, other initial guesses have been proposed in the literature, two of which are presented here:

#### **B-initialization using pseudo-inverse**

The de-mixing matrix  $\mathbf{B}$  satisfies that:

$$\hat{\mathbf{S}} = \mathbf{B}\mathbf{X} \quad (7.52)$$

where:

$\hat{\mathbf{S}}$	is the $K \times N$ matrix of estimated sources
$\mathbf{X}$	is the $M \times N$ observation matrix
$\mathbf{B}$	is the $K \times M$ de-mixing matrix

Assuming we know the true signals  $\hat{\mathbf{S}}$ , we may use these as reference signals, i.e.  $\mathbf{R} = \hat{\mathbf{S}}$ . Doing this allows us to find the de-mixing matrix by using the Moore-Penrose (pseudo)-inverse of the observation matrix  $\mathbf{X}^\dagger$  [Sun and Shang, 2010, p. 1014]:

$$\mathbf{B} = \hat{\mathbf{S}}\mathbf{X}^\dagger \quad (7.53)$$

As we know the pseudo-inverse is the solution to the least-squares problem

$$\min_{\mathbf{B}} \left\| \hat{\mathbf{S}} - \mathbf{B}\mathbf{X} \right\|_2^2 \quad (7.54)$$

Having shown that the pseudo-inverse can be used for finding the true de-mixing matrix  $\mathbf{B}$ , we can use the same approach for finding an initial guess  $\mathbf{B}_0$  using the reference signals  $\mathbf{r}$  which are not necessarily the true signals [Sun and Shang, 2010, p. 1014]:

$$\mathbf{B}_0 = \arg \min_{\mathbf{B}} \left\| \mathbf{r} - \mathbf{B}\mathbf{X} \right\|_2^2 = \mathbf{r}\mathbf{X}^\dagger \quad (7.55)$$

#### **B-initialization using correlation**

Another approach for finding an initial guess for  $\mathbf{B}$  is to choose a  $\mathbf{B}$  which maximizes the correlation between sources and references [Mi and Gui, 2010, 158]. Looking at the initial guess of a single row of  $\mathbf{B}$ ,  $\mathbf{b}_0^T$ , we are interested in solving [Mi and Gui, 2010, 158]:

$$\begin{aligned} \mathbf{b}_0 &= \arg \min_{\mathbf{b}} E[\mathbf{b}^T \mathbf{x} \cdot \mathbf{r}] \\ \text{s.t.} \quad & \|\mathbf{b}\|_2^2 = \mathbf{b}^T \mathbf{b} = 1 \end{aligned} \quad (7.56)$$

We write up the Lagrangian as:

$$\mathcal{L}(\mathbf{b}, \lambda) = E[\mathbf{b}^T \mathbf{x} \cdot \mathbf{r}] + \lambda(\mathbf{b}^T \mathbf{b} - 1) \quad (7.57)$$

Solving this, we know that at the minimum

$$0 = \frac{\partial}{\partial \mathbf{b}} \mathcal{L}(\mathbf{b}, \lambda) = E[\mathbf{x} \cdot \mathbf{r}] - 2\lambda \mathbf{b} \quad (7.58)$$

Solving for  $\mathbf{b}$  yields:

$$\mathbf{b} = \frac{E[\mathbf{x} \cdot \mathbf{r}]}{2\lambda} \quad (7.59)$$

As the Lagrange multiplier should ensure that the constraint is satisfied, i.e. that  $\mathbf{b}$  has unit norm, we can simply write this as [Mi and Gui, 2010, 158]:

$$\mathbf{b} = \frac{E[\mathbf{x} \cdot \mathbf{r}]}{\|E[\mathbf{x} \cdot \mathbf{r}]\|_2} \quad (7.60)$$

Tests show that the pseudo-inverse provides better results than the correlation-approach, and this initialization is therefore used in the following. Having presented two methods for initialization of  $\mathbf{B}$ , all parts of the constrained ICA algorithm has now been described, and the algorithm can be implemented. In the following section, the test results of the cICA algorithm are presented and compared to other extraction algorithms.

# 8 | Test and Performance Results

In this chapter, the test results of the developed algorithm are presented. First, the test-measures are presented, followed by an overview over which tests are to be performed. Finally, the actual results are shown - both results on non-ECG signals, on simulated ECGs and on real ECGs.

## 8.1 Test Measures

Two types of performance measures are used: ICA performance measures and morphology performance measures (e.g. QT-interval and T/QRS ratio).

In order to determine how well the sources have been separated, i.e. how well ICA works from a purely mathematical perspective, some performance measures must be defined, which can compare the estimated sources with the true ones. Two such performance measures are used in this case: The performance index (PI) and the signal to noise ratio (SNR). Note that in addition to these, the morphological performance measures described previously, i.e. the QT interval and T/QRS ratio are also used, however they are more contextual performance measures, and not directly describing how well the ICA algorithm performs. For this reason, the PI and SNR measures are needed. These are introduced in the following, after which the morphology performance measures are described as well.

### 8.1.1 Performance Index (PI)

The performance index measures how well the de-mixing matrix has been found. In the following, square mixing is assumed, i.e.  $\mathbf{A} \in \mathbb{R}^{M \times K}$ ,  $M = K$ , but the concepts can be generalized for  $M \neq K$ .

Ideally, one would like  $\mathbf{B} = \mathbf{A}^{-1}$ , (or, for singular  $\mathbf{A}$ ,  $\mathbf{B}\mathbf{A} = \mathbf{I}$ ) which would imply perfect separation. However, due to ambiguity in ICA, the extracted sources in ICA may be scaled, inverted and permuted relative to the original sources.

This can be described by introducing  $\mathbf{B}_0$ , which is an ICA-equivalent to  $\mathbf{B}$ , meaning they are, from an ICA perspective, equally good solutions [Nordhausen et al., 2011, p. 486]:

$$\mathbf{B}_0 = (\mathbf{QDL})\mathbf{B} \quad (8.1)$$

where:

- Q** is an  $K \times K$  permutation matrix, obtained by permuting the rows of an identity matrix
- D** is an  $K \times K$  scaling matrix, (diagonal matrix with positive diagonal elements)
- L** is an  $K \times K$  sign-change matrix, (diagonal elements are  $\pm 1$ )

Assuming there exists a  $\mathbf{B} \in \mathbb{R}^{M \times K}$  such that  $\mathbf{B}\mathbf{A} = \mathbf{I}$ , this does *not* imply that  $\mathbf{B}_0\mathbf{A} = \mathbf{I}$ . It is therefore not possible to simply consider how close  $\mathbf{B}_0\mathbf{A}$  is to identity, by using some  $p$ -norm as a measure of similarity:

$$PI = |\mathbf{B}_0\mathbf{A} - \mathbf{I}|_p \quad (8.2)$$

Instead, a permutation matrix  $\mathbf{P} = \mathbf{B}\mathbf{A}$ , where the  $i, j$ 'th element of  $\mathbf{P}$  is denoted as  $p_{ij}$ , is introduced. The performance index of the de-mixing matrix  $\mathbf{B}$  can be found as [Amari et al., 1996, p. 762]:

$$rPI_i = \sum_{j=1}^n \frac{|p_{ij}|}{\max_k |p_{ik}|} - 1 \quad (8.3)$$

$$cPI_j = \sum_{i=1}^n \frac{|p_{ij}|}{\max_k |p_{kj}|} - 1 \quad (8.4)$$

$$PI = \sum_{i=1}^n rPI_i + \sum_{j=1}^n cPI_j \quad (8.5)$$

The term  $rPI_i$  can be interpreted as a row-wise normalization of  $\mathbf{P}$ , and describes the error of the separation between the output  $\hat{\mathbf{s}}_i$  and the sources. The term  $cPI_j$  is a column-wise normalization of  $\mathbf{P}$ , and measures the degree of which the desired IC  $\mathbf{s}_j$  appears multiple times at the output [Lu and Rajapakse, 2006, p. 2250]. The PI is non-negative, and is zero when the ICs are perfectly separated. By using this PI, it is invariant to permutations and sign matrices  $\mathbf{Q}, \mathbf{L}$ , but not *necessarily* invariant to scalings,  $\mathbf{D}$  [Nordhausen et al., 2011, p. 487]. It is nevertheless still a very popular performance index used throughout the literature.

### 8.1.2 Signal to noise ratio (SNR)

The signal to noise ratio is a typically used performance measure, but is here slightly adapted to the topic of ICA.

#### SNR for ICA

The SNR is for ICA applications defined as the ratio between the variance of the  $k$ 'th source signal and the mean-squared error (MSE) between the true and estimated  $k$ 'th source [Lu and Rajapakse, 2006, p. 2250]. This means the SNR in dB for the  $k$ 'th source is defined as [Lu and Rajapakse, 2006, p. 2250]:

$$\text{SNR}_k = 10 \log_{10} \left( \frac{\sigma_{\mathbf{s}_k}^2}{\frac{1}{N} \sum_{n=0}^{N-1} (\mathbf{s}_k[n] - \hat{\mathbf{s}}_k[n])^2} \right) \quad (8.6)$$

where:

$\sigma_{\mathbf{s}_k}^2$  is the variance of  $\mathbf{s}_k$   
 $N$  is the number of observations

Although being widely used in the literature, there are a couple problems with the SNR as a performance measures, namely it's lack of invariance to the permutation, scaling and sign-changing matrices  $\mathbf{Q}, \mathbf{D}, \mathbf{L}$ : By changing the sign of  $\hat{\mathbf{s}}$ , the MSE becomes very big. This can be seen by assuming ideal (but inverted) de-mixing, such that  $\hat{\mathbf{s}}[n] = -\mathbf{s}[n]$ . This yields an MSE of  $\frac{1}{N} \sum_{n=0}^{N-1} (2\mathbf{s}_k[n])^2 = 4 \cdot \sum_{n=0}^{N-1} (\mathbf{s}_k[n])^2$ .

By scaling the estimated source, the MSE will also grow, and by permuting the estimated sources, incorrect waveforms are compared when calculating the SNR. This means a very low SNR will result, as the signals are assumed independent, and a wrong source will therefore appear noise-like.

In order to overcome these problems, the SNR must be calculated in a different way.

### Adapted SNR calculation

The first step in making the SNR calculation a useful measure, is to normalize the signals to unit variance. By doing this, the SNR will not depend on scaling of this signals, but purely on how well-separated they are.

In order to overcome the problem with permutations, the SNR is calculated for all possible combinations of sources and estimated sources. This also includes negated versions of the sources, as to overcome the problem with sign-changes. The calculations done in order to find the "optimal" SNR (i.e. the estimated source maximizing the SNR for each true source) are shown in the following.

First, the source vectors are repeated  $M$  times, creating a vector  $\mathbf{s}'$ :

$$\mathbf{s}' = \mathbf{s}[n] \otimes \mathbf{1}_{M \times 1} = \begin{bmatrix} \underbrace{[\mathbf{s}_1[n]^T \quad \mathbf{s}_1[n]^T \quad \cdots \quad \mathbf{s}_1[n]^T]^T}_{M \text{ times}} \\ \underbrace{[\mathbf{s}_2[n]^T \quad \mathbf{s}_2[n]^T \quad \cdots \quad \mathbf{s}_2[n]^T]^T}_{M \text{ times}} \\ \vdots \\ \underbrace{[\mathbf{s}_K[n]^T \quad \mathbf{s}_K[n]^T \quad \cdots \quad \mathbf{s}_K[n]^T]^T}_{M \text{ times}} \end{bmatrix} \quad (8.7)$$

This vector is then concatenated with a negated version of it self, forming a big matrix  $\mathbf{S} \in \mathbb{R}^{2KM \times N}$ :

$$\mathbf{S} = [(\mathbf{s}')^T \quad -(\mathbf{s}')^T]^T \quad (8.8)$$

The estimated sources are repeated  $2K$  times, obtaining a matrix  $\hat{\mathbf{S}} \in \mathbb{R}^{2KM \times N}$ :

$$\hat{\mathbf{S}} = \left[ \underbrace{[\hat{\mathbf{s}}[n]^T \quad \hat{\mathbf{s}}[n]^T \quad \cdots \quad \hat{\mathbf{s}}[n]^T]_{2K \text{ times}}^T \right]^T \quad (8.9)$$

It is now possible to calculate the mean squared error for each element of these matrices, thus obtaining the MSE between each pair of sources and estimated sources, as well as the negated sources.

$$\text{MSE}_i = \sum_{j=1}^N (\mathbf{S}_{ij} - \hat{\mathbf{S}}_{ij})^2, \quad \forall i = 1, 2, \dots, 2KM \quad (8.10)$$

Note that due to the formulation of MSE, it is not important if it's  $\mathbf{s}$  or  $\hat{\mathbf{s}}$  that is negated - the results are the same.

Having the MSE, the SNR can now be calculated as usual, see Equation 8.11. Note that the variance  $\sigma_i$  is not explicitly stated in the formula, as we normalized the signals such that  $\sigma_i = 1$ , as this reduces the problem of scale invariance.

$$\text{SNR}_i = -10 \log_{10}(\text{MSE}_i) \quad (8.11)$$

We now have a long vector  $\mathbf{SNR} \in \mathbb{R}^{2KM \times 1}$ , which is reshaped to the following matrix,  $\mathbf{SNR} \in \mathbb{R}^{2M \times K}$ :

$$\mathbf{SNR} = \begin{bmatrix} \text{SNR}_{\hat{s}_1, \mathbf{s}_1} & \cdots & \text{SNR}_{\hat{s}_1, \mathbf{s}_k} & \cdots & \text{SNR}_{\hat{s}_1, \mathbf{s}_K} \\ \vdots & \ddots & \vdots & \vdots & \\ \text{SNR}_{\hat{s}_m, \mathbf{s}_1} & \cdots & \text{SNR}_{\hat{s}_m, \mathbf{s}_k} & \cdots & \text{SNR}_{\hat{s}_m, \mathbf{s}_K} \\ \vdots & \ddots & \vdots & \vdots & \\ \text{SNR}_{\hat{s}_M, \mathbf{s}_1} & \cdots & \text{SNR}_{\hat{s}_M, \mathbf{s}_k} & \cdots & \text{SNR}_{\hat{s}_M, \mathbf{s}_K} \\ \text{SNR}_{-\hat{s}_1, \mathbf{s}_1} & \cdots & \text{SNR}_{-\hat{s}_1, \mathbf{s}_k} & \cdots & \text{SNR}_{-\hat{s}_1, \mathbf{s}_K} \\ \vdots & \ddots & \vdots & \vdots & \\ \text{SNR}_{-\hat{s}_m, \mathbf{s}_1} & \cdots & \text{SNR}_{-\hat{s}_m, \mathbf{s}_k} & \cdots & \text{SNR}_{-\hat{s}_m, \mathbf{s}_K} \\ \vdots & \ddots & \vdots & \vdots & \\ \text{SNR}_{-\hat{s}_M, \mathbf{s}_1} & \cdots & \text{SNR}_{-\hat{s}_M, \mathbf{s}_k} & \cdots & \text{SNR}_{-\hat{s}_M, \mathbf{s}_K} \end{bmatrix} \quad (8.12)$$

Finally, the optimal SNR and sources can be found. This is done by finding the estimated source (both positive and negative) that maximizes the SNR for each true source,  $k$ . The argument (modulo  $M$ ) gives which estimated source is the optimal, denoted  $\tilde{i}_k$ .

$$\tilde{\text{SNR}}_k = \max_i \text{SNR}_{i,k} \quad (8.13)$$

$$\tilde{i}_k = \left( \arg \max_i \text{SNR}_{i,k} \right) \bmod M \quad (8.14)$$

where:

$\tilde{\text{SNR}}_k$  is optimal SNR of the  $k$ 'th source  
 $\tilde{i}_k$  is the optimal estimated source for the  $k$ 'th source

If the argument of the optimal source is greater than  $M + 1$ , the source is in the lower half of the  $\mathbf{SNR}$ -matrix, meaning the sign is changed. The optimal sign can therefore be found by subtracting the argument of the optimization from  $M + 1$  and taking the sign, i.e.:

$$\gamma_i = \text{sign} \left( (M + 1) - \left( \arg \max_i \text{SNR}_{i,k} \right) \right) \quad (8.15)$$

where:

$\gamma_i$  is the sign of optimal source signal

By doing this, meaningful SNRs can be found, even though the estimated sources may be permuted, scaled and inverted.

### 8.1.3 Channel selection

Despite having made different performance measures, it is still of interest to determine which channel is the "best" (by some criterion which is to be determined) so that channel can be used for segmentation for calculating QT-intervals etc. There is no interest in calculating this for the maternal channels and "poor" fetal channels, so an optimal fetal channel must be selected.

There are multiple ways of doing this: In the paper by [Behar et al., 2014d], a smoothing indicator is used, which looks at the heart rate variability and a beat comparison between estimated FQRS and MQRS. Others, such as the Challenge benchmarking, use signal quality measures, as proposed by [Andreotti et al., 2017]. Here, a signal quality index (SQI) is computed based on different categories of measures, such as time (e.g. variance, skewness, kurtosis), frequency (e.g. power of FQRS), detection-based measures (e.g. percentage of correctly detected beats) and FECG-specific measures (e.g. spectral properties of the FECG). Doing these calculations are however rather complicated, and a simpler approach is desired.

A straight-forward idea is to simply pick the channel with the highest SNR, however this will most likely lead to maternal channel being chosen, as the maternal SNR will inherently be higher than that of fetal channels. Another option is to utilize the fact that the maternal and fetal QRS locations are assumed know (or accurately estimated). One could therefore compare the extracted channel with these reference, using e.g. correlation:

$$\tilde{m} = \arg \min_m \rho(\hat{\mathbf{s}}_m, \mathbf{MQRS}) - \rho(\hat{\mathbf{s}}_m, \mathbf{FQRS}), \quad m = 1, 2, \dots, M \quad (8.16)$$

where:

$\tilde{m}$	is the optimal channel
$\hat{\mathbf{s}}_m$	is the $m$ 'th estimated channel
$\rho(x, y)$	is the (normalized) correlation between $x$ and $y$
MQRS	is a reference signal with peaks at maternal QRS locations
FQRS	is a reference signal with peaks at fetal QRS locations

By minimizing the positive correlation to MQRS and the negative correlation to the FQRS, we are finding whichever channel is least correlated to the maternal signals and most correlated to the fetal signals.

However, it is assumed that a better estimate can be obtained by using mutual information instead of correlation. This is because correlation only uses second-order statistics, whereas mutual information (MI) considers the whole pdf, i.e. all moments. The use of mutual information for channel selection is used by [Vrins et al., 2004], who use it as a pre-processing technique, i.e. choosing which channels to use for ICA. Here, we use it as a post-processing method, and the algorithm proposed in [Vrins et al., 2004] is therefore also modified.

The developed channel selection algorithm is instead inspired by [Sameni et al., 2006b], who use an approach similar to the one in Equation 8.16, where the MI between signal and mother is minimized, while that between signal and fetus is maximized. It is however done as two separate stages in [Sameni et al., 2006b], as opposed to the joint optimization presented in this work:

$$\tilde{m} = \arg \min_m I(\hat{\mathbf{s}}_m, \mathbf{MQRS}) - I(\hat{\mathbf{s}}_m, \mathbf{FQRS}), \quad m = 1, 2, \dots, M \quad (8.17)$$

where:

$I(x, y)$	is the mutual information between $x$ and $y$
-----------	---

The mutual information between two random variables  $X, Y$  is defined as [Sameni et al., 2006b, eq. 1]:

$$I(X, Y) = \int \int p_{X,Y}(x, y) \cdot \log \left( \frac{p_{X,Y}(x, y)}{p_X(x)p_Y(y)} \right) dx dy \quad (8.18)$$

With the test measures defined and a channel selection algorithm presented, it is now possible to test the developed ICA algorithm.

### 8.1.4 Morphological test measures

In order to compare the quality of the extracted ECGs from a more clinical perspective, some performance measures must be introduced. In Equation 4.1, the QT interval and T/QRS ratio were introduced as performance measures, and these will also be used in this comparison. A problem is however that these measures require the use of a segmentation algorithm, meaning it is both the performance of the extraction- and segmentation algorithm that is tested. This is problematic, as the segmentation algorithm often does not return the relevant features (T-peak, T-end), meaning a high number of templates are rejected [Andreotti et al., 2016, p. 645].

For this reason, a direct comparison of the true and estimated templates is suggested: By calculating the root mean squared error (RMSE) between the two templates, a similarity measure of the two is obtained. A desired property of this measure is that by minimizing the RMSE between the true and estimated templates, the errors in QT and T/QRS must also be minimized, as these implicitly depend on the difference between the waveforms. However, as mentioned in [Behar et al., 2014a, p. 56], the RMSE measure is "weighted towards large amplitude features (like the QRS complex) but can provides little insight into subtle but clinically significant changes". This means that despite the RSME being an objective measure not relying on correct segmentation of the ECG, it may not accurately capture the small morphological differences that are of interest.

A table of the three morphology performance measures used can be seen in Table 8.1.

Name and Definition	Description	Unit
$e_{\text{FQT}} =  \text{FQT}_{\text{est}} - \text{FQT}_{\text{ref}} $	Absolute error of the QT interval estimate	ms
$e_{\text{FTQRS}} = \left  \left  \frac{\text{FT}_{\text{h, est}}}{\text{FQRS}_{\text{est}}} \right  - \left  \frac{\text{FT}_{\text{ref}}}{\text{FQRS}_{\text{ref}}} \right  \right $	Difference between estimated and true T/QRS ratio	.
$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{n=0}^{N-1} (x_{\text{est}}[n] - x_{\text{ref}}[n])^2}$	RMS error between true and extracted FEKG template	ms

**Table 8.1:** The performance measures used to compare the different extraction methods.

## 8.2 Test Plan

The purpose of this section is to give an overview of which tests are to be performed and why. Two overall test cases are performed, each with two sub-tests:

1. Tests of the ICA algorithm using the ICA test measures described in section 8.1.
  - A) ICA test on non-ECG signals
  - B) ICA test on ECG signals

2. Tests of the constrained ICA algorithm on ECG mixtures using the ECG morphology test measures described in Section 8.1.4
  - A) Morphology test on FECGSYNDB
  - B) Morphology test on ADFECGDB

One could also test the algorithm's ability to estimate fetal heart rate. But as the algorithm assumes the fetal QRS locations as known, it does not make sense to see how well it can estimate these locations. Also, as the purpose of this thesis is to provide morphologically more accurate fetal ECG, it is not needed to evaluate the ability to estimate fetal heart rate.

An overview of the tests to be performed can be seen in Figure 8.1 and will be each be described in further details in the following sections.

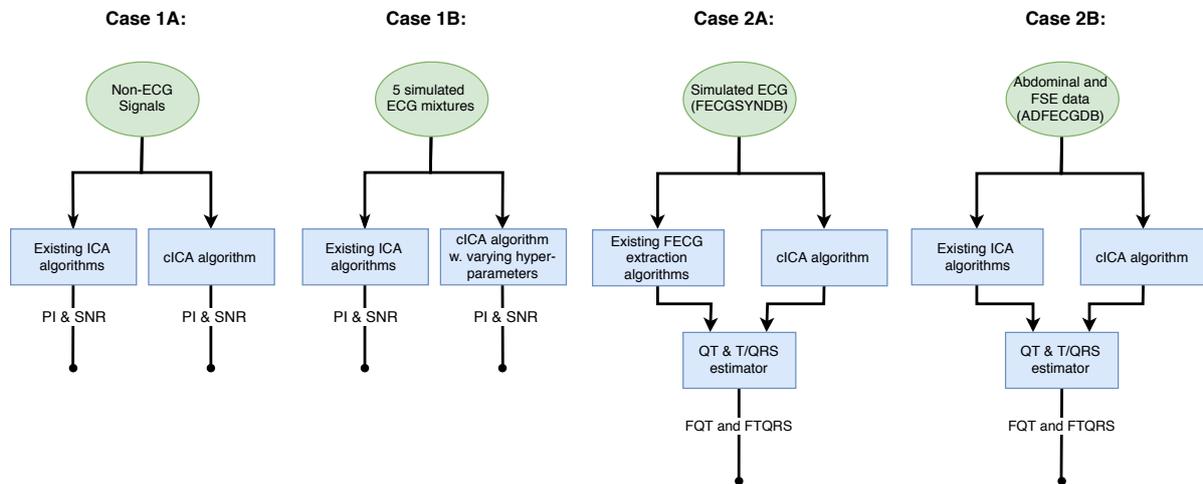


Figure 8.1: A visualization of the tests to be conducted.

### 8.2.1 Test case 1: ICA test measures

There are two sub-tests for this ICA test:

- A) ICA test on non-ECG signals
- B) ICA test on ECG signals

The reason for the test on non-ECG signals is that before performing tests on ECG signals, it is of interest to test the cICA algorithm on well-known, "simple" signals, and compare its performance with other ICA algorithms. Hereafter, the algorithms can be applied to artificial ECGs with known mixing matrices. The performance measures used for this test case are the performance index (PI) and the signal to noise ratio (SNR) between true and estimated sources, as described in section 8.1.

The four algorithms tested are:

1. FastICA
2. InfoMax
3. Extended InfoMax
4. cICA: Temporally constrained ICA (ICA with reference)

### **Test 1A): ICA test on non-ECG signals**

For test 1A), 5 artificial source signals are generated:

- Sinusoid
- Triangle
- Amplitude-modulated sine
- Pulse-signals of varying width
- Uniform noise

These signals are scaled in different ways, from a peak amplitude of 2 to an amplitude of 0.005, in order to make the separation harder, and to emulate the amplitude ranges seen in the ECG mixtures between fetal and maternal ECGs. A random mixing matrix  $\mathbf{A}$  is generated, after which the ICA algorithms are applied. The test includes a no-noise case, and a case with white Gaussian noise added, at an SNR of 12 dB. For the cICA case, the reference signals used are the true signals, in order to test the algorithms in an ideal case before moving on to more realistic cases, which is done in test 1B.

### **Test 1B): ICA test on ECG signals**

In this test, the FECSYN-toolbox, see Section 3.1.2, is used to generate 5 artificial ECG mixtures, i.e. from 5 different subjects, with six sensors used to record the abdominal mixture. The signals are generated in two cases: Without noise and with a signal (maternal ECG) to noise ratio (SNR) of 6 dB. In this test, the original signals and the mixing matrix are well-known, and the PI and SNR can therefore be evaluated to allow comparison of the results.

This test features two sub-parts:

1. Comparison between ICA algorithms
2. Hyper-parameter search for cICA

#### **1): Comparison between ICA algorithms**

In this test, the four ICA algorithms mentioned in Section 8.2.1 are tested with and without a reference signal. For simplicity and to give an ideal foundation for the algorithms, the reference signals used are the true ECG signals.

In the cases of FastICA and InfoMax, the reference is used as initialization for the estimated de-mixing matrix  $\mathbf{B}$ . For the constrained ICA case, the reference is used as the reference signal for the algorithm.

The step-size used for the cICA algorithm is 0.1, and 0.0001 for the InfoMax methods. No momentum is used.

The performance measures used are the performance index (PI) of the permutation matrix  $\mathbf{P} = \mathbf{BA}$  and the signal to noise ratio (SNR) between true and estimated sources.

## 2): Hyper-parameter search for cICA

The purpose of this test is to investigate which hyper-parameters deliver the best performance of the constrained ICA method.

The parameters to be determined are:

- Step-size/learning rate,  $\alpha$
- Correlation-threshold,  $\epsilon$
- Momentum-factor,  $\gamma$
- Maternal reference type
- Fetal reference type

A grid search of different values for these parameters is done on all five subjects, for all combinations of maternal and fetal reference types. Tests are performed in noise-free and in noisy (6 dB SNR) cases. Another parameter investigated is the execution time.

The results of this test are of interest for when the cICA algorithm must be applied to other ECG signals, as is done in test case 2.

### 8.2.2 Test case 2: ECG morphology test measures

Having tested the ICA algorithms in test case 1, the cICA algorithm can now be tested using the contextual performance measures as described in Section 8.1.4, namely the QT-interval, T/QRS ratio and RMSE. This test is of great importance as it is these results that are of interest for the end user.

The tests are conducted on two data sets:

- A) The FECGSYNDB data set (synthetic data)
- B) The ADFECGDB data set (measured data)

By applying the algorithm to both synthetic and measured data, it can be seen how well the algorithm performs in both cases.

#### Case 2A): FECGSYNDB

The purpose of this test is to compare the developed algorithm with the currently existing methods. The basis for this comparison, i.e. the benchmarking algorithms, are those available in [Behar et al.,

2017]. This test case is identical to the testing framework presented in [Andreotti et al., 2016], as described in Section 4.2.3. The code has been modified slightly in order to make it executable, however no performance-altering modifications to the algorithms themselves have been made, except that only four sensors are used instead of the eight sensors used by [Andreotti et al., 2016]. This is done in order to make the results more realistic, as most FECG databases only have four sensors available. The algorithms are thus not re-implemented, but just re-run on a similar dataset.

Data is generated artificially as part of the code, however the data is generated in the same way as that in the FECGSYNDB, see Section 3.1.2, which was used as the code for the results presented in section 4.3. This means the results are comparable, as the data used is just different realizations with the same parameters, meaning they are statistically identical.

### **Case 2B): ADFECGDB**

The purpose of this test is to see how the algorithm performs on real, measured data from the ADFECGDB data set. Being real data, there is no knowledge of the original ECG source signals or the mixing matrix. There is however scalp electrode recordings, which can be used as a reference for the morphology, as the scalp electrode is the golden standard in fetal ECG morphology. Since the scalp electrode sees a different projection of the cardiac potentials than the abdominal recordings, the morphology might be different, but a scalp electrode is the best one can do on real recordings. The data is also annotated, so the fetal QRS locations are known in advance. An RMSE measure between extracted FECG the scalp electrode ECG is also included, to compare the accuracy of the estimated ECG. However, the RMSE is not expected to be zero for an ideal extraction method, as the cardiac electric potentials at the scalp and abdominal are not the same.

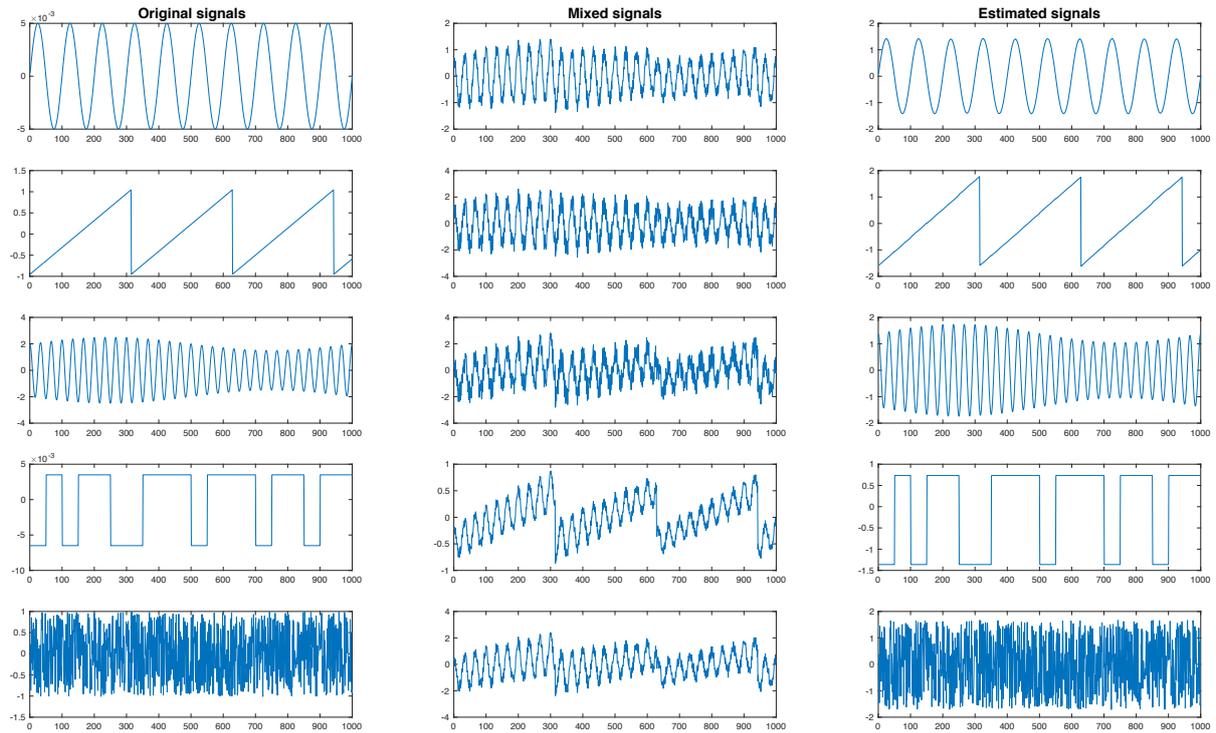
Note that it is not directly possible to test the template-subtraction methods on the real signals as they require an (accurate) maternal reference signal which is not trivial to generate. Only subspace methods (ICA, PCA) are therefore tested on these signals.

Having described the test cases that are performed, the results can now be presented, as is done in the following section.

## **8.3 Test Results 1A - ICA on Non-ECG Signals**

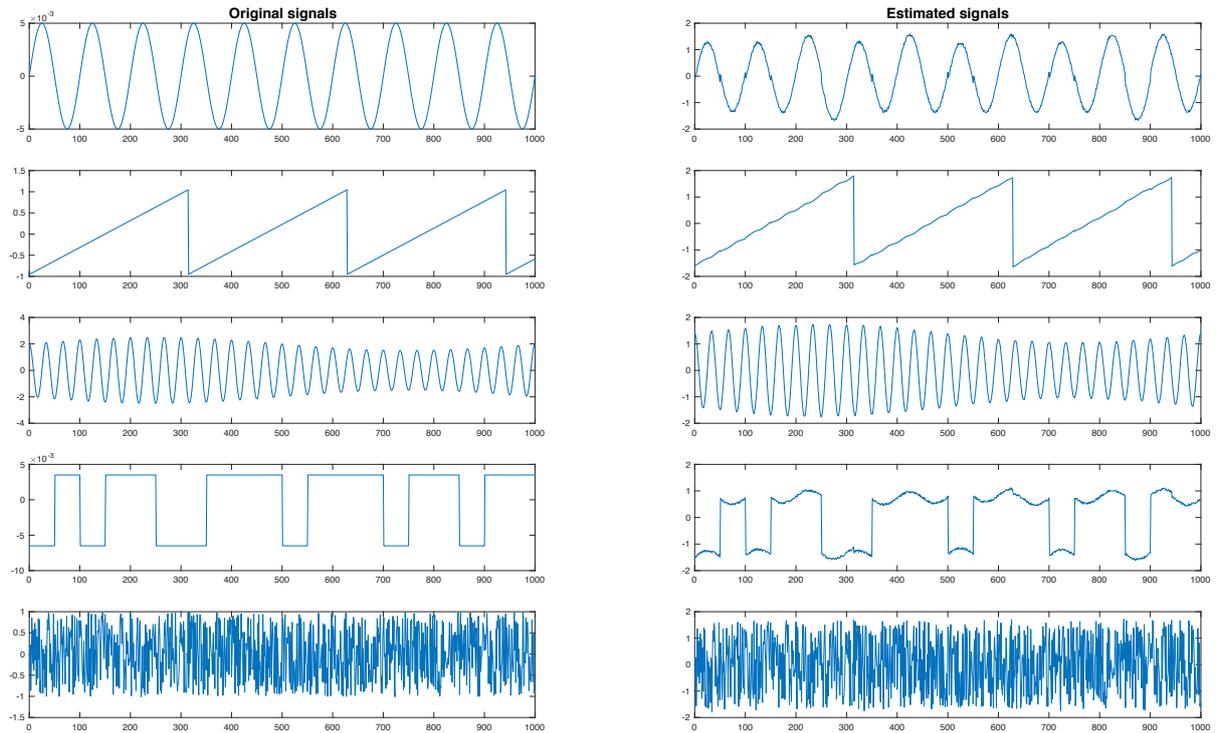
In this section, the test results for the ICA test on non-ECG signals are presented.

Five signals are generated as described in Section 8.2.1, and are mixed using a random mixing matrix. In Figure 8.2, the sources and mixed signals can be seen, together with the estimated sources. The sources are here estimated using constrained ICA with true signals as references, and with a correlation threshold of 0.1. From the figure, it is clear that cICA performs quite good when good reference signals are available.



**Figure 8.2:** The original, mixed and estimated source signals for the non-ECG signals. The ICA method used is constrained ICA with true references.

If a different ICA algorithm is used, the estimated sources might not be extracted as well, as is clear from Figure 8.3 where FastICA is used for extraction.



**Figure 8.3:** The original and estimated source signals for the non-ECG signals. The ICA method used is FastICA with true references.

A comparison of the PI and SNRs of the different extraction algorithms can be seen in Table 8.2. The algorithms are run both with and without reference signals used for  $\mathbf{B}$ -matrix initialization. Different correlation thresholds are set for the cICA method, which also has its  $\mathbf{B}$ -matrix initialization set either randomly or using the references.

ICA method	No noise		12 dB SNR	
	PI	Mean SNR	PI	Mean SNR
FastICA	0.68	25.8	2.17	1.36
FastICA w. ref	0.68	25.8	2.17	1.36
InfoMax	2.40	1.81	2.98	-0.67
InfoMax w. ref.	2.96	0.66	4.48	-0.97
Ext. InfoMax	1.07	17.2	3.36	1.15
Ext. InfoMax w. ref	0.63	25.3	2.63	1.43
cICA	2.89	0.98	4.74	-1.20
cICA w. ref, $\epsilon = 0.1$	0.37	73.2	2.49	2.00
cICA w. ref, $\epsilon = 0.9$	0.37	73.2	2.49	2.00
cICA w. ref, rand. init, $\epsilon = 0.1$	2.89	0.98	4.02	-1.01
cICA w. ref, rand. init, $\epsilon = 0.9$	2.89	0.98	4.02	-1.01

**Table 8.2:** A comparison between ICA methods on non-ECG signals. Results are found by taking the mean over 50 random starting values of the de-mixing matrix  $\mathbf{B}$ . No noise has been to the observations.

In Appendix D, the SNRs can be seen for all 5 test signals, in both noisy and noise-free settings.

From Table 8.2, it is clear that FastICA performs best of the classical ICA algorithms. Also, cICA performs well when the true signals are used as reference and the reference is used as an initial guess for the de-mixing matrix. If a random initialization is used, the results of cICA are poor.

## 8.4 Test Results 1B - ICA on ECG signals

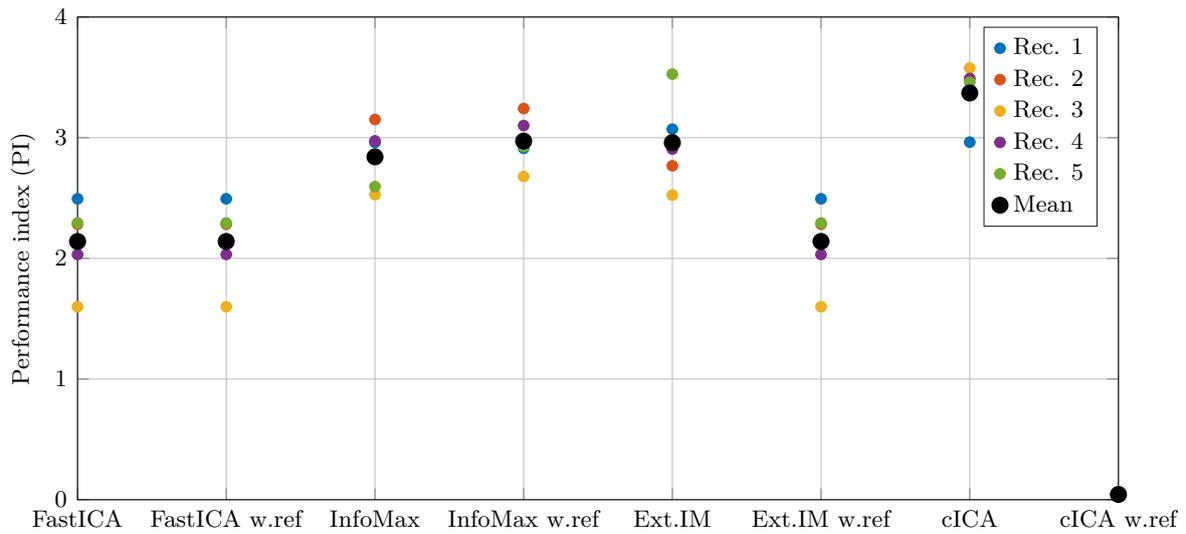
The purpose of this section is to compare various ICA methods on a small set of artificial ECG signals. There are two purposes of this test: The first is to compare different ICA methods, and see how the proposed constrained ICA method compares to other ICA algorithms. The other is to do a hyper-parameter search for constrained ICA, in order to find out which parameters deliver the best performance.

### 8.4.1 Comparison between ICA methods

The purpose of this test is to compare different ICA methods to see which perform better, as described in Section 8.2.1.

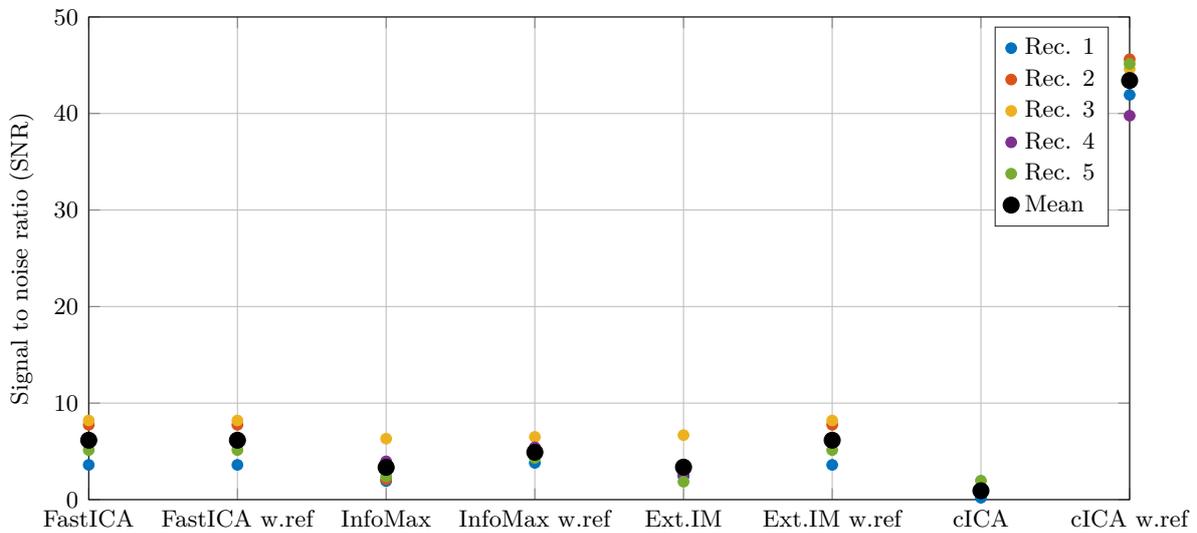
#### No noise

In this section, the results for the no-noise test case are presented. Results are thus based on averages of 5 artificially generated ECG mixtures, with no additional noise. The raw results for this test are presented in tables in Appendix E, but are summarized in the following figures. In Figure 8.4, the performance index for the ICA algorithms tested can be seen. It is clear that cICA with reference has superior performance over the other methods.



**Figure 8.4:** A comparison of the performance index (PI) for different ICA algorithms on 5 different simulated ECG recordings. Lower numbers are better.

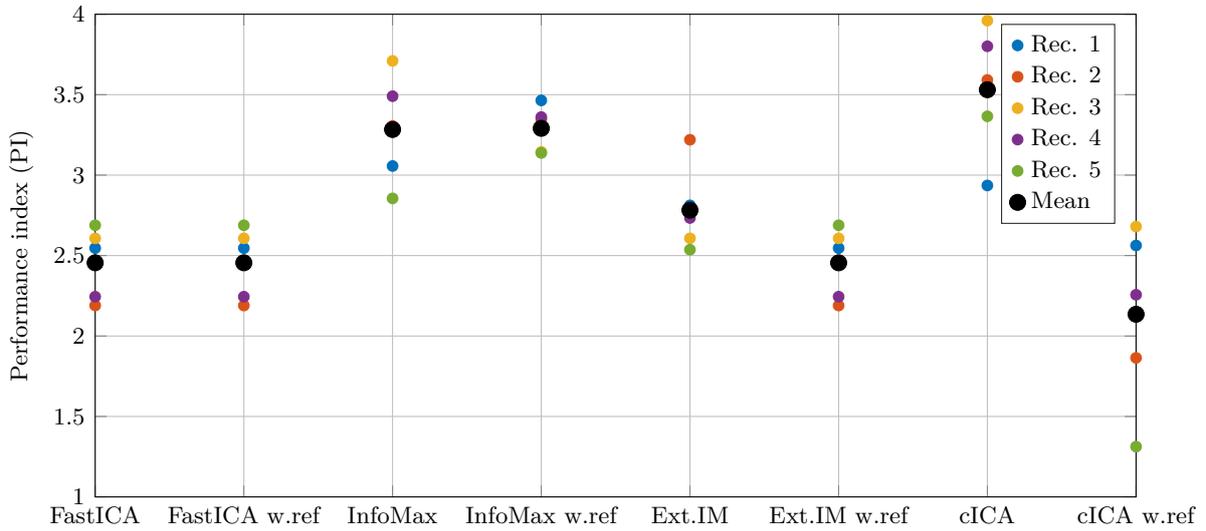
In Figure 8.5, the SNR for the ICA algorithms tested can be seen. Here, it is clear that the SNR for cICA with reference is much higher than for any of the other methods, showing that cICA outperforms the other methods.



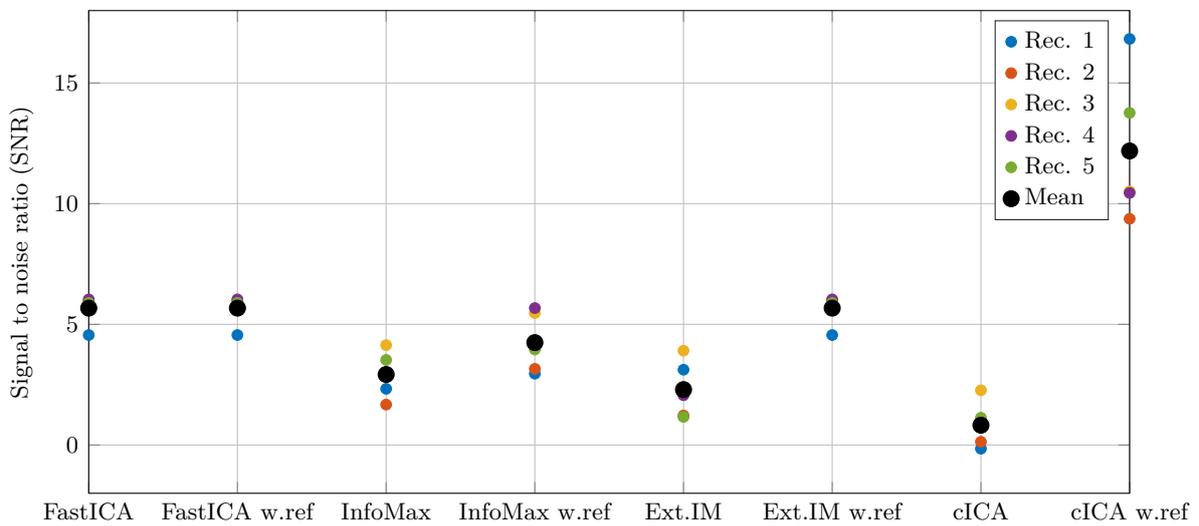
**Figure 8.5:** A comparison of the signal to noise ratio (SNR) of extracted ECG channels, for different ICA algorithms on 5 different simulated ECG recordings with an SNR of 6 dB (fetus to noise). Higher numbers are better.

## 6 dB SNR

In this section, noise has been added to the artificial signals which the tests are performed on, such that the signal to noise ratio between maternal signal and the noise floor is 6 dB. The performance index for the different ICA algorithms is shown in Figure 8.6, while the SNR between true and estimated sources is shown in Figure 8.7.



**Figure 8.6:** A comparison of the performance index (PI) for different ICA algorithms on 5 different simulated ECG recordings. Lower numbers are better.



**Figure 8.7:** A comparison of the signal to noise ratio (SNR) of extracted ECG channels, for different ICA algorithms on 5 different simulated ECG recordings with an SNR of 6 dB (fetus to noise). Higher numbers are better.

Another interesting comparison to make, is the execution time between algorithms. This can be seen in Table 8.3.

ICA method	No noise	6 dB SNR
FastICA	0.18	0.13
FastICA w. ref	0.16	0.14
InfoMax	4.06	4.00
InfoMax w. ref.	7.78	8.18
Ext. InfoMax	22.6	22.0
Ext. InfoMax w. ref	1.26	1.63
cICA	0.09	0.08
cICA w. ref	0.09	0.08

**Table 8.3:** Execution times (in seconds) for different ICA methods. Shown values are averages over 5 subjects in no noise.

A general observation is that most ICA methods converged within 5000 iterations, which was defined as the maximum number of allowed iterations. However, Infomax w. reference and extended Infomax without reference, did not manage to converge within 5000 iterations for any of the five subjects. This explains their relatively long execution time.

Finally, looking at the feasibility of the algorithms, it is clear that all the algorithms converge in all cases, except for subject no. 2 for some unknown reason.

With the ICA algorithms compared, a more in-depth testing of constrained ICA will be performed.

### 8.4.2 Constrained ICA hyper-parameter search

In this section, a comparison of constrained ICA with different reference signals will be performed. A comparison of other hyper-parameters such as the step-size (learning rate) and the correlation threshold are also investigated.

Six different types of maternal references are tested:

- No reference
- True signal
- Square signal from maternal QRS locations
- Weighted maternal ECG template
- Unweighted maternal ECG template
- Maternal template from principal component analysis (PCA)

Similarly, six different types of fetal references are tested:

- No reference
- True signal
- Square signal from fetal QRS locations
- Biphasic signal from fetal QRS locations
- Biphasic signal from fetal QRS locations and a T-wave
- Template generated from scalp electrode

Just as with the different ICA methods, tests are carried out on 5 artificial ECG mixtures, either with no noise applied or with a (maternal) signal to noise ratio of 6 dB. For these tests, the following test parameters are used:

- Step-size:  $\alpha = 0.1$
- Maternal correlation threshold  $\epsilon_m = 0.3$
- Fetal correlation threshold  $\epsilon_f = 0.1$
- Momentum:  $\eta = 0$

### Reference types - no noise

In Table 8.4, the performance index for the different fetal and maternal references can be seen. Results are obtained by taking the mean over all five test subjects.

Mat. vs fet. ref. type	No reference	True signal	Square from QRS	Biphasic from QRS	Biphasic + T-wave	Template from scalp
No reference	3.37	3.66	3.66	3.74	3.74	3.66
True signal	4.16	0.04	1.55	1.52	1.50	1.36
Square from QRS	4.03	2.03	3.65	3.63	3.56	3.46
Weighted template	4.14	1.59	3.22	3.21	3.25	2.93
Unweighted template	4.15	1.54	3.17	3.16	3.12	2.87
Template from PCA	3.79	1.57	2.70	2.68	2.62	2.54

**Table 8.4:** A comparison of performance index (PI) of ICA, for different types of maternal and fetal references in no noise conditions.

In Table 8.5, the signal to noise ratio between true and estimated sources for the different fetal and maternal references can be seen. Results are obtained by taking the mean over all five test subjects, and by taking the mean SNR over all six ECG channels to be estimated.

Mat. vs fet. ref. type	No reference	True signal	Square from QRS	Biphasic from QRS	Biphasic + T-wave	Template from scalp
No reference	0.92	1.74	1.74	1.47	1.46	1.74
True signal	2.84	43.42	23.07	21.65	21.61	24.16
Square from QRS	2.07	24.60	4.25	2.83	2.79	5.34
Weighted template	3.19	26.21	5.86	4.44	4.40	6.95
Unweighted template	2.88	26.33	5.99	4.56	4.52	7.07
Template from PCA	4.80	22.43	7.17	6.88	6.84	7.97

**Table 8.5:** A comparison of the signal to noise ratio (SNR) between sources for different types of maternal and fetal references in no noise conditions.

### Reference types - 6 dB SNR

The different reference types are now compared in more realistic conditions, i.e. when the maternal SNR is 6 dB. The performance indices can be seen in Table 8.6, while the average SNR between true and estimated sources can be seen in Table 8.7.

Mat. vs fet. ref. type	No reference	True signal	Square from QRS	Biphasic from QRS	Biphasic + T-wave	Template from scalp
No reference	3.53	4.20	4.20	4.25	4.09	4.20
True signal	4.53	2.14	2.88	3.08	3.12	2.66
Square from QRS	4.53	2.89	3.85	4.01	4.05	3.74
Weighted template	4.49	2.49	3.33	3.57	3.63	2.96
Unweighted template	4.49	2.55	3.46	3.62	3.69	3.02
Template from PCA	3.79	2.38	2.84	3.23	3.18	2.50

**Table 8.6:** A comparison of performance index (PI) of ICA, for different types of maternal and fetal references, with recordings having 6 dB (maternal) signal to noise ratio.

Mat. vs fet. ref. type	No reference	True signal	Square from QRS	Biphasic from QRS	Biphasic + T-wave	Template from scalp
No reference	0.82	1.55	1.55	0.44	0.84	1.55
True signal	2.05	12.18	7.49	5.74	5.69	8.08
Square from QRS	2.05	8.95	4.25	2.48	2.43	4.85
Weighted template	2.52	9.88	5.19	3.42	3.37	5.78
Unweighted template	2.61	9.95	5.25	3.49	3.44	5.85
Template from PCA	3.75	8.64	5.11	4.02	4.06	5.51

**Table 8.7:** A comparison of the signal to noise ratio (SNR) between sources for different types of maternal and fetal references with signals having a 6 dB (maternal) signal to noise ratio.

### Step size and correlation threshold

The purpose of this test is to compare different step sizes and correlation thresholds, and investigate their influence on the performance of the cICA algorithm. The following test cases are considered:

- Step size:  $\alpha = 0.01, 0.1, 1$
- Fetal and maternal correlation threshold:  $\epsilon = 0.1, 0.3, 0.99$

These are tested on all 5 artificial subjects, with and without noise, for all combinations of 6 maternal and 6 fetal reference types.

In Table 8.8, a summary of these test results are shown. To generate the table, the maternal/fetal reference signal combination giving the best performance is selected for each of the different test cases. Hereafter, mean PI and SNR are calculated over all five subjects. In addition to the overall average SNR, the average SNR for just the three fetal VCG channels is considered. Finally, the average execution time over all five subjects is shown.

	$\alpha$ , no noise			$\alpha$ , 6 dB SNR			$\epsilon$ , no noise			$\epsilon$ , 6 dB SNR		
Value	0.01	0.1	1	0.01	0.1	1	0.1	0.3	0.99	0.1	0.3	0.99
<b>PI</b>	2.55	2.54	2.48	2.45	2.50	2.49	2.51	2.54	2.40	2.44	2.50	2.46
<b>SNR</b>	7.98	7.98	7.92	5.85	5.85	5.85	7.69	7.97	7.67	5.85	5.85	5.80
<b>Fetal SNR</b>	7.15	6.99	6.97	5.20	5.27	5.28	6.90	6.99	7.34	5.00	5.27	4.94
<b>Exec. time [s]</b>	16.8	17.4	16.5	0.20	0.18	0.19	0.18	17.0	57.9	0.11	0.11	36.2

**Table 8.8:** Hyper-parameter comparison, where optimal performance using different step-sizes and varying correlation thresholds can be seen, for artificial signals with and without noise.

## Momentum

For the momentum test, we set  $\alpha = 0.1$  and  $\epsilon = 0.1$ , and test three different momentum values:  $\eta = 0$ , 0.1 and 0.9. The results of this test can be seen in Table 8.9.

	Momentum, no noise			Momentum, 6 dB SNR		
Value	0	0.1	0.9	0	0.1	0.8
<b>PI</b>	2.51	2.51	2.52	2.44	2.44	2.37
<b>SNR</b>	7.69	7.69	7.71	5.85	5.85	5.85
<b>Fetal SNR</b>	6.90	6.90	7.05	5.00	5.00	5.01
<b>Exec. time [s]</b>	0.17	0.18	0.16	0.16	0.16	0.17

**Table 8.9:** Comparison between optimal results when using different momentum-factors, for artificial signals with and without noise.

## Feasibility

There are two ways in which the ICA methods can deliver infeasible results: Either by the constraints not being satisfied or by the algorithm not converging within a predefined number of iterations - in this case 5000.

All reference combinations for all step sizes deliver feasible results, and likewise for the correlation threshold, except from the largest threshold of  $\epsilon = 0.99$ . In this case, it was only the test case where the true references that were used that the algorithm converged.

### 8.4.3 Summary

Based on this comparison, it is possible to select which hyper-parameters to use for further testing, when the morphology studies are to be conducted:

It is clear from the comparison of reference types, that the best performance is obtained by using the

true signals as reference signals. This is of course not possible for real scenarios, and a different reference type combination must be chosen. The reference types which for the no-noise case result in the lowest PI and highest average SNR, is the maternal reference using a PCA template and a fetal reference made from a template from a scalp electrode.

In the 6 dB SNR case it is still the scalp electrode fetal reference that gives better performance, while both weighted, unweighted and PCA templates give good results - however, as the SNR difference is only 1 dB, the difference is not that significant.

It is therefore chosen to use the maternal reference using PCA, and the fetal reference using the scalp electrode template.

For the step-sizes, the performance difference between values is negligible, both for PI and SNR in noisy and noise-free conditions. The various correlation thresholds also result in practically the same results, although one should note that the high threshold of  $\epsilon = 0.99$  resulted in infeasible results. For this reason, a step size in the middle of the testing range is chosen, namely  $\alpha = 0.1$ . For the correlation threshold, a value of  $\epsilon = 0.1$  is chosen as it gives the same PI/SNR performance as the others, but the algorithm converges much faster. Finally, the momentum factor did not affect performance in any significant way, and the momentum-factor will therefore be set equal to zero, such that steepest descent is used for simplicity. The hyper-parameters to be used in the following tests are therefore:

- Step-size  $\alpha = 0.1$
- Correlation threshold  $\epsilon = 0.1$
- Momentum:  $\eta = 0$
- Maternal reference: PCA template
- Fetal reference: Scalp electrode template

Comparing the results from the classical ICA algorithms in Figure 8.4 and Appendix E with those obtained using cICA with the selected hyper-parameters, it is clear that FastICA has a comparable performance to cICA: PI = 2.14 vs 2.51, SNR = 6.16 vs 7.69 dB in the no-noise case.

With the hyper-parameters selected, it is now possible to estimate the morphology of the extracted fetal ECGs, which is done in the following test.

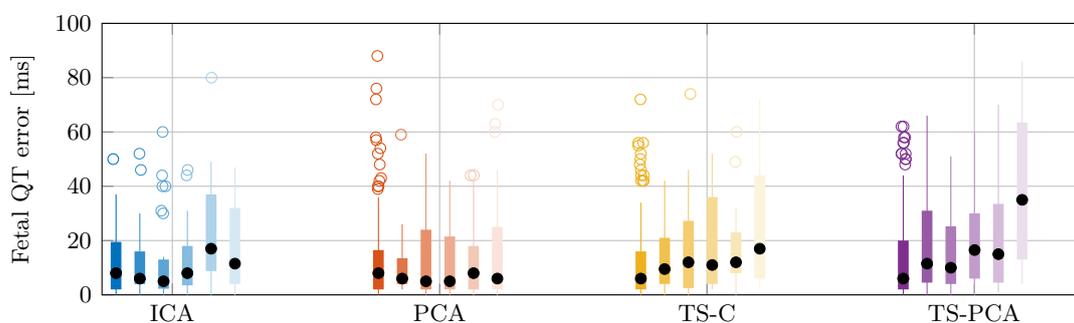
## 8.5 Test Results 2A - Morphology of Artificial Signals

In this section, the test results obtained using the FECGSYNDB dataset are presented. This makes it possible to compare the developed constrained ICA algorithm with the algorithms used by [Andreotti et al., 2016].

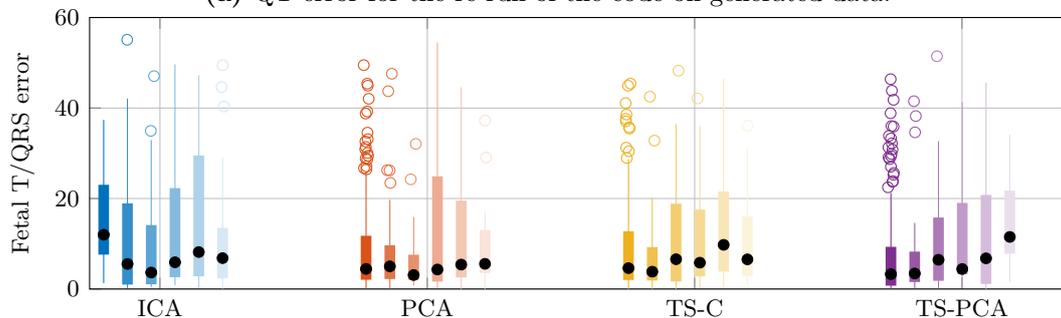
### 8.5.1 Comparison with original results

In this section, the results of the benchmarking algorithms are compared with those presented in the original paper by [Andreotti et al., 2016], which have been shown in section 4.3.

It was decided to run the code using four of the possible extraction methods, namely ICA, PCA, TS-C and TS-PCA, see Section 3.3.3 for more details. The results of this benchmark can be seen in Figure 8.8.



(a) QT error for the re-run of the code on generated data.



(b) T/QRS error for the re-run of the code on generated data.

**Figure 8.8:** Performance measures for different extraction methods at different SNRs. SNR cases are (from left to right): No noise, 12 dB, 9 dB, 6 dB, 3 dB and 0 dB.

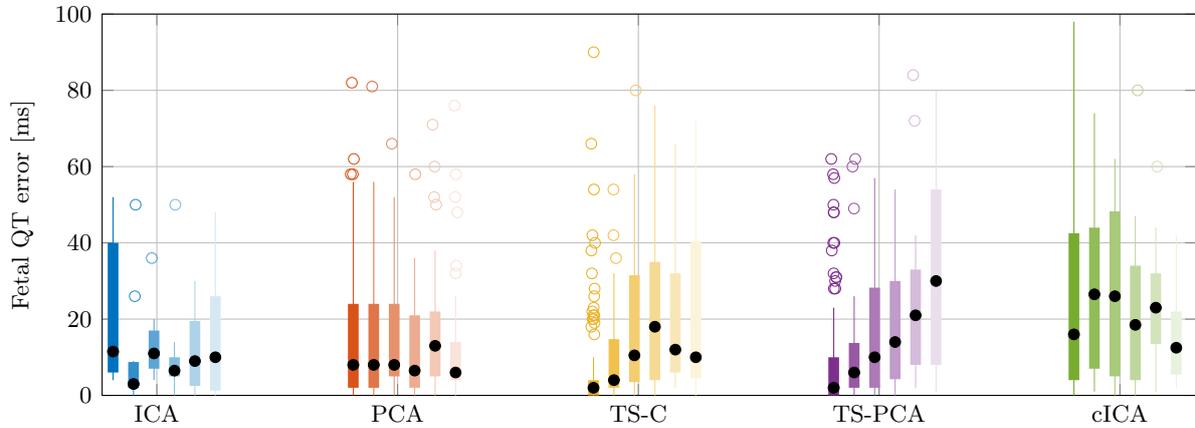
Comparing Figure 8.8 to the original results as shown in Figure 4.4, some differences are seen and in general, the results we obtained are not as good as those in Figure 4.4, which is described further in the discussion, see chapter 9.

### 8.5.2 Comparison with constrained ICA

In this section, the constrained ICA (cICA) algorithm is compared to other extraction methods as mentioned above (ICA, PCA, TS-C and TS-PCA). This is done by three performance measures: QT-interval, T/QRS ratio and RMSE. In Appendix F, tables showing the mean values of the three performance measures for each extraction method for each SNR.

### QT interval

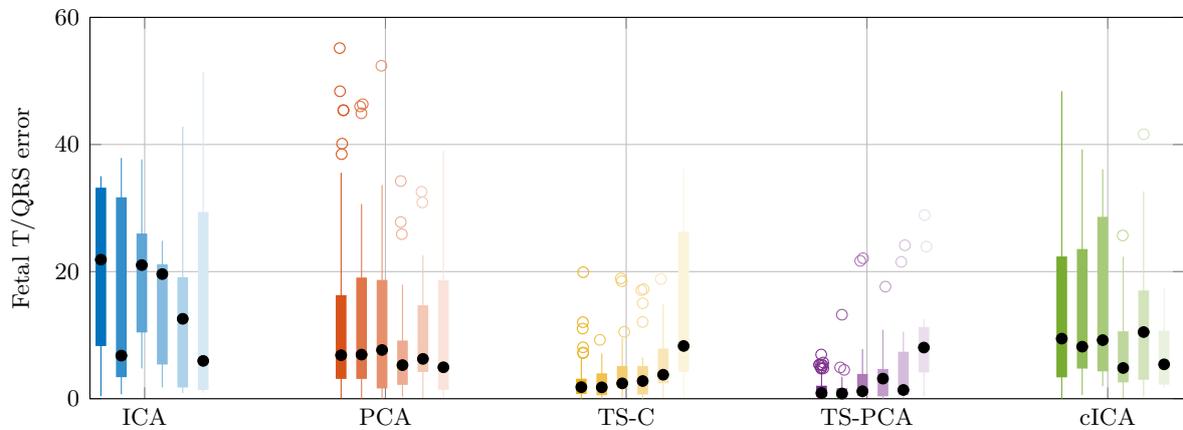
The QT interval error can be seen in Figure 8.9, for different SNR levels.



**Figure 8.9:** QT-interval error error for different extraction methods at different SNR levels. From left to right: No noise, 12 dB, 9 dB, 6 dB, 3 dB and 0 dB.

### T/QRS ratio

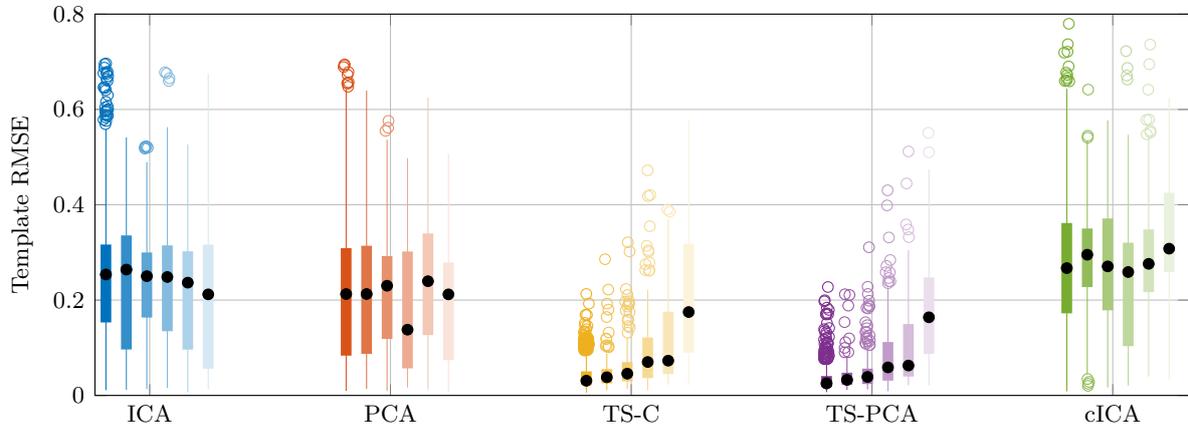
In Figure 8.10, a boxplot of the T/QRS error is shown for different extraction methods.



**Figure 8.10:** T/QRS ratio error for different extraction methods at different SNR levels. From left to right: No noise, 12 dB, 9 dB, 6 dB, 3 dB and 0 dB.

### Root-mean squared error

Finally, the (normalized) RMSE between true and estimated fetal ECG templates are shown, for different extraction algorithms and SNR levels.



**Figure 8.11:** Root-mean square error between true and estimated template for different extraction methods at different SNR levels. From left to right: No noise, 12 dB, 9 dB, 6 dB, 3 dB and 0 dB.

From this test, it is clear that TS-methods deliver the best performance in regards to the T/QRS ratio and the template RMSE. For the QT-ratio, ICA and PCA both perform well. The developed cICA algorithm is slightly better than ICA at estimating T/QRS, but the RMSE and QT-error are higher for cICA than any of the other tested methods.

## 8.6 Test Results 2B - Morphology of Real Signals

In this test, the developed constrained ICA algorithm is applied to abdominal mixtures recorded on real subjects, originating from the ADFECGDB database. There is no "true" unmixed signals, although a scalp electrode recording is used as the ground truth signal in this test. The ecgpuwave-algorithm is used for segmenting the extracted fetal ECG, and estimate the desired morphological performance measures. An RMSE estimate between extracted FECG and scalp electrode is also provided as a performance measure.

The results from this test are summarized in Table 8.10.

Subj. No.	FQT [ms]					FT/QRS					RMSE				
	1	4	7	8	10	1	4	7	8	10	1	4	7	8	10
ICA	100	N/A	N/A	N/A	N/A	216	N/A	N/A	N/A	N/A	0.32	0.38	0.57	0.47	0.57
PCA	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0.37	0.36	0.36	0.26	0.36
cICA	N/A	NA	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0.36	0.46	0.58	0.40	0.36

**Table 8.10:** Morphology test results on the signals from the ADFECGDB. N/A means segmentation did not return valid results.

From this table, it is clear that either the extraction methods do not return useful FEEGs, or the segmentation algorithm is struggling to find valid QT and T/QRS ratios, as it was only possible to segment a single single recording and only by using the classical ICA method.

Looking at the RMSE values in Table 8.10, it appears that PCA performs the best, while cICA is slightly better than classical ICA. However, due to the poor ability to segment the data and the low number of subjects, one should be careful drawing conclusions based on these results.

To compare the estimated QT-interval and T/QRS ratio with those from the scalp electrode, a graphical user interface (GUI) has been developed, which is used for annotating the QT and T/QRS of the scalp ECG. Due to the poor performance of the extraction and segmentation algorithms, the scalp data is not annotated. The GUI is however still presented, and can be seen in Appendix C.

## 9 | Discussion

In this section, the methods used and obtained results are discussed. First, some general comments are presented, after which the reproduction of the results found in literature are discussed. Finally, possible explanations for the poor performance of constrained ICA are mentioned.

### General comments

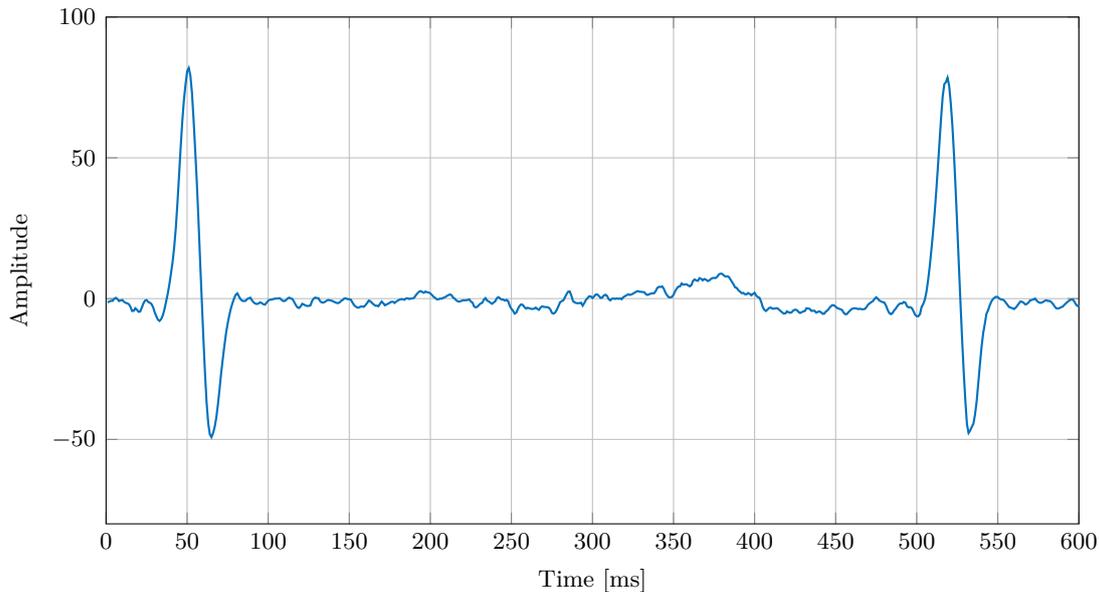
Looking at the results, it is clear that the proposed method does not beat the current methods in the ability to extract morphologically accuracy FECG. It is observed that QT-interval errors appear bigger for cICA than any other method, and the RMSE is also higher. For the T/QRS ratio, cICA performs better than ICA, but poorer than the other methods. Comparing cICA to other ICA methods, FastICA seems to perform very well, and is only beaten by cICA if the true references are used. The TS-methods do however consistently deliver the best morphology results. The reasons why these rather poor results are seen, is explained in the later in this chapter.

The RMSE measure used to compare the accuracy of the extracted ECGs might not be a meaningful performance measure in some cases, as the true signals used for the RMSE comparison are the VCGs. However, the algorithm is not estimating the VCG components, which means the ECGs are not expected to have a similar waveform to the VCGs (although the morphology measure might be similar). Similarly for the real signals where the scalp electrode is used as the true signal, we do not expect the extracted FECG to have the same waveform as the scalp electrode data. This is the reason why the more generalizable measures of QT and T/QRS are more meaningful, as they should be somewhat identical for both VCG, ECG and scalp.

It turns out that the segmentation algorithm rejects a lot of signals, i.e. it fails at providing physiologically possible morphology estimates. This is partly due to poorly extracted MECGs, but as the algorithm also fails at extracting morphology features from the true sources, it shows that the segmentation algorithm could be improved. The high number of rejected signals explain why a lot of N/A's were seen in the test on the true recorded ECGs.

This might also be a reason for why the variance of the results between runs is quite high: If a high number of signals are rejected, a smaller batch size is used for calculating the performance, and the sample variance will be higher. An observation related to this, is the fact that the results from the FECGSYN framework could not be reproduced, as described in the following.

In some cases, it is however obvious why the segmentation algorithm fails: Looking at a median of 10 ECG cycles from the scalp electrode, it is observed that there is no clear T-wave present in (some of) the signals, see Figure 9.1. This obviously makes it near-impossible for the algorithm to estimate the QT-interval and T/QRS ratio, meaning it fails. The reason why no T-wave is seen in some scalp electrode data is not known, and require more knowledge on FECGs than what has been covered in this thesis.



**Figure 9.1:** The median of 10 ECG cycles of the scalp electrode data from recording *r01* from the ADFECGDB dataset. Note the lack of a clear T-wave.

### Reproducing FECGSYN results

As described in Section 8.5.1, the results obtained by running the benchmarking algorithms as shown in Figure 8.8 are not identical to the results presented in Figure 4.4 i.e. the original results from [Andreotti et al., 2016]. The reasons for this are discussed in the following.

Firstly, it is noted that the quantile bars in the boxplot of Figure 8.8 are larger in the re-run than in the original one. The exact reason for this is unknown, but one reason might be the fact that only four abdominal sensors have been used to generate the results in this thesis, as opposed to the eight sensors of the original paper. This is based on Figure 3.3, where a slight performance increase is seen between using 4 vs 8 sensors, and the intuition that by using 8 sensors, we better capture the true nature of the ECGs, by having more projections/more data. Another option is that many runs have been performed in the paper by [Andreotti et al., 2016] in order to obtain a realization yielding a low error, but a more likely explanation is due to the high rejection ratio of the algorithm: As described in Section 4.3, a very high number of templates could not be generated or segmented, yielding rejection ratios as high as 90 % for the re-run. This will obviously affect the performance, as the data not rejected might not be representative for the statistics of all the data.

Secondly, the outliers are omitted in the results from [Andreotti et al., 2016]. This makes it hard to compare the two, as the outliers are only such in a statistical sense - the algorithm itself has not detected them as outliers, and they are therefore valid results - just very poor ones. From the re-run of the tests, the amount of outliers might be as high as 10 % of the total number of samples per case - this is quite high, which again emphasizes that they should not be removed.

Most outliers are present in the no-noise case, which, when excluded, is the category performing the best, as would also be expected. However, when the outliers are not excluded, there is no clear performance difference between the SNRs, or the methods. This is quite different than the results seen in Figure 4.4.

Despite it not being able to reproduce the results obtained by [Andreotti et al., 2016], the re-run is still a valid benchmark tool, as the developed algorithm is tested on the same data (i.e. simulated data using FECGSYN), meaning a direct comparison is fair.

### **Possible reasons for poor performance of cICA**

As mentioned when describing the VCG-ICA approach, we know the underlying electrical activity of the heart has three components, namely the x, y and z-directions - this is true for both the maternal and fetal hearts, meaning the mixture is described by a 6-dimensional source space. Having ICA to take something described in 6 dimensions and reduce it to just four (assuming four sensors are available), one cannot assume that nicely separated, morphologically accurate ECGs will result, as the information in the six dimensions must somehow be incorporated in the projection onto the four-dimensional space. This is suspected to be a key reason why cICA does not work well on this problems, as all tests are done using only four recorded channels.

An interesting observation is that the cICA algorithm has very fast convergence (2-3 iterations) when the references are used to generate the initial guess for the de-mixing matrix by using the pseudo-inverse. The reason for this is that the initial guess is located close to a local minimum, which cICA then converges to. The fact that local minima's are present, is clear from the fact that the solution obtained from a random guess is not the same as that obtained when using the pseudo-inverse, which again is not the same as the true solution. As this is seen for both the ECG- and non-ECG-test, this clearly demonstrates how the highly non-convex nature of the problem. This could be because there are some linear combinations of the sources which appear independent, (i.e. are very non-Gaussian by having a high negentropy), but there is some other linear combination (the true one) which has an even higher negentropy. Another reason for the non-convexity might be that the ECG signals have similar statistics, meaning it is difficult to distinguish them - mixing them in various ways may therefore not alter the cost significantly.

A method which might help prevent converging to a local minima, is to add noise to the references to make them less "accurate", thereby forcing cICA to search for a better solution, which should hopefully lead to a different solution than that found by just using the pseudo-inverse as initial guess.

Another possible solution is to use a grid search of random guesses, thereby trying many different initializations of the de-mixing matrix spanning the entire search-space, thus increasing the chance of finding the global minimum.

Convergence of cICA is tested by evaluating the norm of the difference between the current de-mixing matrix and the one from the second-newest - if it is below a certain threshold, the algorithm has converged. Setting this convergence threshold it lower, cICA is forced to run for longer periods of time, which affects the end result. In the tests, the threshold was set rather high, in order to ensure that the algorithm did converge, as it was observed that lowering value would mean cICA did not converge, especially if it was strongly constrained. A small test of tuning this parameter showed that a lower threshold produced better results on cICA without reference, but not with the reference, as it made it hard for the algorithm to converge within the maximum allowed number of iterations.

From the hyper-parameter search, it was seen that more realistic reference signals lead to better results. This is because a the pseudo-inverse used for the initial guess was then closer to the true inverse of the mixing matrix. The fast convergence also explains why the choice of step-size, momentum and correlation threshold did not change the performance significantly.



# 10 | Conclusion

In this thesis, the topic of non-invasive fetal electrocardiogram (NI-FECG) estimation has been studied. First, the necessary clinical background in terms of the components of the ECG, ECG morphology and the motivation for doing electronic fetal ECG monitoring have been described.

A literature review was conducted, from which the current methods for estimating both fetal heart rate (FHR) and FECG morphology were described in overall terms and compared. The methods currently in use are template subtraction methods, subspace methods (such as principal component analysis (PCA)), source separation methods (such as independent component analysis (ICA)) and adaptive methods such as least-means-square (LMS) filters.

Using the FECGSYN framework, which makes use of the fetal ECG synthetic database (FECGSYNDB), was is possible to compare fetal morphology estimation methods, using the morphological performance measures QT-interval and T/QRS ratio.

Based on tests of the different extraction methods (using the FECGSYN framework), some of the problems with the current methods are identified. Based on this analysis, three potential new extraction methods are presented: Beamforming, vectorcardiogram independent component analysis (VCG-ICA) and constrained ICA. Of these, the constrained ICA method is investigated further as it proves the most useful from both a theoretical and practical perspective.

Classical ICA algorithms aim at finding the (statistically) maximally independent components/sources of an observed mixture. This is done using some independency measure such as mutual information or negentropy which is based on kurtosis, a fourth-order statistics which can be interpreted as a non-Gaussianity measure. Two such algorithms are the widely known algorithms FastICA and Infomax, which have both been derived.

In constrained ICA (cICA)/ICA with reference, the independency of extracted components is still maximized, but a correlation constraint between the estimated sources and a set of reference sources is included in the optimization problem. The cICA algorithm is derived as a solution to this constrained optimization problem. Different reference signals were tested, ranging from simple pulse signals at the QRS locations, to maternal ECG templates using PCA and fetal ECG templates using scalp electrode data.

The performance index (PI) and signal-to-noise-ratio between true and estimated sources were used for testing the constrained ICA algorithm against FastICA and Infomax, on both non-ECG signals and synthetic ECGs. From this test, it was clear that cICA provides excellent results if the true signals were used as references, and if the references were used to generate an initial guess for the de-mixing matrix instead of a random guess. Generally though, FastICA provides an acceptable estimate which is hard to outperform unless very good reference signals are provided to cICA.

A hyper-parameter search was performed, in order to find out which reference signals, step-size, and correlation threshold give the best performance. With the parameters selected, the cICA algorithm was tested using the FECGSYN framework and compared to other methods such as "classical" ICA, PCA and template subtraction.

Results indicate that cICA does not perform better than the other methods when looking at the QT-interval, but does perform slightly better than ICA when estimating the T/QRS ratio. The quality of the test results is however debatable as the algorithm used to segment the estimated FECGs often fails, and the results reported in the literature could not be reproduced. Developing a more robust segmentation algorithm is therefore of great interest.

From these tests, it is observed that the algorithm quickly converges to a local minimum close to the initial guess, which is based on the reference signals. Due to the this non-convex nature of the optimization problem, the performance of cICA comes down to the accuracy of the reference signals.

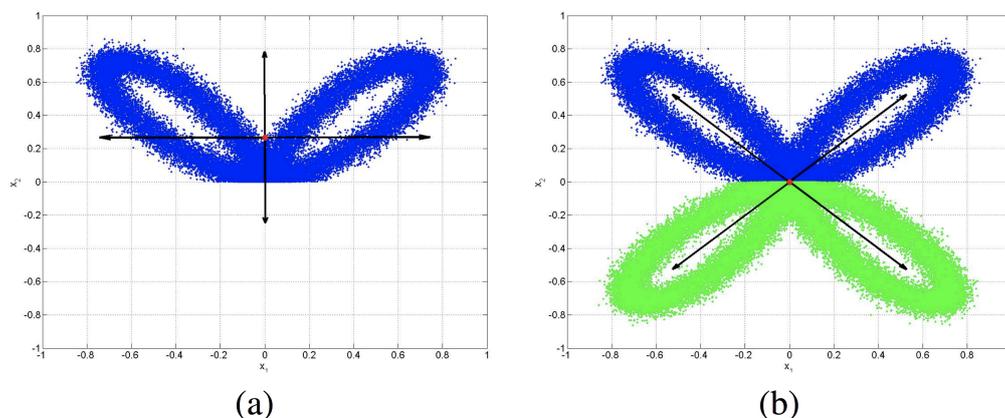
Finally, the extraction methods are tested on real recordings, however none of the performance measures could be estimated for any extraction method, thus proving that the topic of NI-FECG estimation is very difficult in practice.

In conclusion, a constrained ICA algorithm has been derived, which makes use of reference signals to obtain more accurate FECGs. The algorithm does not outperform other FECG extraction algorithms, (neither ICA- nor non-ICA-methods), thus demonstrating that ICA methods, and in particular constrained ICA, is not as good as template subtraction methods for extracting morphologically accurate fetal ECG.

# 11 | Future Work

In the current implementation of cICA, a gradient-descent solver is used, instead of the Newton-like update rule proposed by [Lu and Rajapakse, 2006]. This was done in order to simplify the algorithm. One could include the Hessian to obtain a Newton-like update rule which should converge faster as the curvature of the optimization landscape is also taken into account. But, being a second-order method, Newton's method may converge to a saddle point or maximum when being applied to higher-dimensional systems instead of a minimum. A small test in which the Hessian was included in the cICA update step did show a slightly faster convergence, but no obvious changes to the quality of the solution. One could do more experiments with using Newton's method in order to obtain faster convergence.

An interesting point is brought up by [Sameni et al., 2006a, p. 657]: "The ECG and VCG signals are not symmetric around their mean values, and the isoelectric point of the heart slightly differs from the mean values of the data". What this means in practice is that while ICA assumes the data as symmetric (since the sign cannot be determined and the mean is subtracted), this may not be a correct assumption when dealing with ECG signals, as is demonstrated in Figure 11.1. This means the extracted ICs do not correspond to the expected dimensions of the VCG space. This can be overcome by "mirroring" the signal, i.e. by stacking the observation vector with an inverted version of itself, thus making the observations symmetric around zero as ICA expects. Small tests indicate that this does not change the result significantly, but combined with the other improvements it might give rise to better results.



**Figure 11.1:** The effect of pdf asymmetry on the extracted ICs. (a) asymmetric pdf (b) pdf made by stacking with inverted observations. The mixing vectors extracted by ICA are plotted in each case. Figure taken from [Sameni et al., 2006a, fig. 1]

It is observed that none of the subspace and source separation methods work on real signals. It would be interesting to generate accurate maternal references, so the TS-methods could be applied to real signals, to see if these methods would be able to provide morphologically accurate estimates of real ECG data. Getting accurate ECGs on real data is however still difficult due to unexpected morphology (i.e. missing T-wave) of the scalp ECGs.

An important improvement would be to develop a better segmentation algorithm, to reduce the number of ECG cycles being rejected. This is however not straight forward as the *ecgpwave* segmentation algorithm

currently used in the FECGSYN framework is one of the best open-source algorithms available. It could also be interesting to apply the developed algorithms to eight abdominal channels using the FECGSYN framework, to get results that are directly comparable with those from [Behar et al., 2014d].

Another thing to consider is the ability to do real-time implementations of the extraction algorithm, as this might be preferred in clinical situations, over having a delay of several minutes due to processing. For this reason, simpler algorithms than ICA might be of interest, as they will have a lower execution time.

In conclusion, the results obtained point towards ICA not being the best method for extracting morphologically accurate FECG. For this reason, one could consider using a completely different method - one such could be by using neural networks (NNs), which have proven highly useful in a wide range of applications from speech processing and image recognition to playing chess. The literature on NNs being applied to the problem of NI-FECG is rather limited, and no literature describing a deep neural network (DNN) being applied to the problem has been found. It therefore seems like a very promising method to apply such a network.

A challenge which must be handled when applying a NN on this problem is that the input and output variables are time-series, and not just a series of random samples. This means that the commonly used NN structure of multilayer perceptrons (MLPs) is not suitable, as the number of variables becomes very large: A 500 element input and a single hidden layer with e.g. 1000 units will result in  $500 \cdot 1000 + 1000 = 501000$  parameters.

It also means that as correlations are present over time, due to the periodicity of the ECGs, the NN must be able to handle having large correlations between samples far from each other. This rules out the use of convolutional neural networks (CNN), typically used for images, as these assume local coherence/correlation, exactly the assumption which we need to overcome. Instead, one might use recurrent neural networks (RNNs) or long short-term memory networks (LSTMs), which are designed for modeling sequences such as time-series. One might also be able to reformulate the problem, such that the time-series need not be considered, and an MLP might then be used.

Implementing the proposed improvements may lead to better extractions, but many things also point towards the conclusion that ICA might not be the method to use for extracting morphologically accurate fetal ECG, and that other approaches such as a neural network should be attempted instead.

# Bibliography

- AAMI (2012). *AAMI EC57:2012 Testing and Reporting Performance Results of Cardiac Rhythm and ST Segment Measurement Algorithms*. AAMI.
- Allan, C., Singh, S., and Malcolm, S. (2001). *Fetal Electrocardiography*. Cardiopulmonary Medicine From Imperial College Press. World Scientific Publishing Company.
- Amari, S.-i., Cichocki, A., and Yang, H. H. (1996). A new learning algorithm for blind signal separation. In *Advances in neural information processing systems*, pages 757–763.
- Anatomybodysystem (2017). Electrical Conduction System Of The Heart Diagram. <https://anatomybodysystem.com/wp-content/uploads/2017/11/electrical-conduction-system-of-the-heart-diagram-anatomy-of-the-cardiovascular-system-the-cardiovascular-system.jpg>. Downloaded February 5, 2018.
- Andreotti, F., Behar, J., Zaunseder, S., Oster, J., and Clifford, G. D. (2016). An open-source framework for stress-testing non-invasive foetal ecg extraction algorithms. *Physiological measurement*, 37(5):627.
- Andreotti, F., Gräber, F., Malberg, H., and Zaunseder, S. (2017). Non-invasive fetal ecg signal quality assessment for multichannel heart rate estimation. *IEEE Transactions on Biomedical Engineering*, 64(12):2793–2802.
- Arulampalam, M. S., Maskell, S., Gordon, N., and Clapp, T. (2002). A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking. *IEEE Transactions on signal processing*, 50(2):174–188.
- Atkielski, A. (2007). Sinus Rhythm. <https://upload.wikimedia.org/wikipedia/commons/9/9e/SinusRhythmLabels.svg>. Downloaded February 6, 2018.
- Behar, J., Andreotti, F., Oster, J., and Clifford, G. D. (2014a). A bayesian filtering framework for accurate extracting of the non-invasive fecg morphology. In *Computing in Cardiology Conference (CinC), 2014*, pages 53–56. IEEE.
- Behar, J., Andreotti, F., Oster, J., Zaunseder, S., and Clifford, G. D. (2017). Fecgsyn. <http://fernandoandreotti.github.io/fecgsyn/index.html>. Accessed March 8, 2018.
- Behar, J., Andreotti, F., Zaunseder, S., Li, Q., Oster, J., and D Clifford, G. (2014b). An ecg simulator for generating maternal-foetal activity mixtures on abdominal ecg recordings. 35:1537–1550.
- Behar, J., Johnson, A., Clifford, G. D., and Oster, J. (2014c). A comparison of single channel fetal ecg extraction methods. *Annals of Biomedical Engineering*, 42(6):1340–1353.
- Behar, J., Oster, J., and Clifford, G. D. (2013). Non-invasive fecg extraction from a set of abdominal sensors. In *Computing in Cardiology Conference (CinC), 2013*, pages 297–300. IEEE.
- Behar, J., Oster, J., and Clifford, G. D. (2014d). Combining and benchmarking methods of foetal ecg extraction without maternal or scalp electrode data. *Physiological measurement*, 35(8):1569.
- Behar, J., Zhu, T., Oster, J., Niksch, A., Mah, D. Y., Chun, T., Greenberg, J., Tanner, C., Harrop, J., Sameni, R., et al. (2016). Evaluation of the fetal qt interval using non-invasive fetal ecg technology. *Physiological measurement*, 37(9):1392.

- Bell, A. J. and Sejnowski, T. J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural computation*, 7(6):1129–1159.
- Ben-Arie, J. and Rao, K. R. (1995). Nonorthogonal signal representation by gaussians and gabor functions. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, 42(6):402–413.
- Bishop, C. (2006). *Pattern Recognition and Machine Learning*. Information Science and Statistics. Springer.
- Bobrow, C. S. and Soothill, P. W. (1999). Causes and consequences of fetal acidosis. *Archives of Disease in Childhood - Fetal and Neonatal Edition*, 80(3):F246–F249.
- Bongaarts, J. (2009). Human population growth and the demographic transition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1532):2985–2990.
- Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J., et al. (2011). Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine learning*, 3(1):1–122.
- Byrom, A. (2012). Cardiotocography (CTG). <http://www.practisingmidwife.co.uk/images/a-z/F7.jpg>. Downloaded February 12, 2018.
- Cables and Sensors, LLC (2018). 12-Lead ECG Placement Guide with Illustrations. <https://www.cablesandsensors.com/pages/12-lead-ecg-placement-guide-with-illustrations>. Accessed February 7, 2018.
- Clifford, G., Sameni, R., Ward, M. J., Robinson, J., and Wolfberg, A. J. (2011). Clinically accurate fetal ecg parameters acquired from maternal abdominal sensors. *American journal of obstetrics and gynecology*, 205(1):47.e1–47.e5.
- Clifford, G. and Villarroel, M. (2006). Model-based determination of qt intervals. In *Computers in Cardiology, 2006*, pages 357–360. IEEE.
- Clifford, G. D., Silva, I., Behar, J., and Moody, G. B. (2014). Noninvasive fetal ecg analysis. *Physiological measurement*, 35(8):1521–1536.
- Comon, P. and Jutten, C. (2010). *Handbook of Blind Source Separation: Independent component analysis and applications*. Academic press.
- Connor, S. (2006). Overpopulation 'is main threat to planet'. <http://www.independent.co.uk/environment/overpopulation-is-main-threat-to-planet-521925.html>. Accessed February 13, 2018.
- Curry, T., Dowdey, J., and Murry, R. (1990). *Christensen's Physics of Diagnostic Radiology*. M - Medicine Series. Lea & Febiger.
- Debdas, A. (2013). *Practical Cardiotocography*. Jaypee Brothers, Medical Publishers Pvt. Limited.
- Deonandan, R. (2017). The World Isn't Facing An Overpopulation Problem. [http://www.huffingtonpost.ca/dr-raywat-deonandan/overpopulation\\_b\\_16216754.html](http://www.huffingtonpost.ca/dr-raywat-deonandan/overpopulation_b_16216754.html). Accessed February 13, 2018.

- DiPietro, J. A., Costigan, K. A., and Gurewitsch, E. D. (2003). Fetal response to induced maternal stress. *Early Human Development*, 74(2):125 – 138.
- Edenbrandt, L. and Pahlm, O. (1988). Vectorcardiogram synthesized from a 12-lead ecg: Superiority of the inverse dower matrix. *Journal of Electrocardiology*, 21(4):361 – 367.
- Fraser, D. and Cooper, M. (2012). *Survival Guide to Midwifery*. A Nurse’s Survival Guide. Elsevier Health Sciences.
- Freeman, R., Garite, T., Nageotte, M., and Miller, L. (2012). *Fetal Heart Rate Monitoring*. Wolters Kluwer Health.
- Fuster, V., O’Rourke, R., Walsh, R., and Poole-Wilson, P. (2007). *Hurst’s the Heart, 12th Edition*. McGraw-Hill Education.
- Goldberger, A. L., Amaral, L. A. N., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., Mietus, J. E., Moody, G. B., Peng, C.-K., and Stanley, H. E. (2000). Physiobank, physiotoolkit, and physionet. *Circulation*, 101(23):e215–e220.
- Grivell, R. M., Alfirevic, Z., Gyte, G. M., and Devane, D. (2012). Antenatal cardiotocography for fetal assessment. *Cochrane Database of Systematic Reviews*, (12).
- Hall, J. (2016). *Guyton and Hall Textbook of Medical Physiology*. Guyton Physiology Series. Elsevier.
- Hesse, C. W. and James, C. J. (2006). On semi-blind source separation using spatial constraints with applications in eeg analysis. *IEEE Transactions on Biomedical Engineering*, 53(12):2525–2534.
- Hon, E. H. (1958). The electronic evaluation of the fetal heart rate. *American Journal of Obstetrics & Gynecology*, 75(6):1215–1230.
- Hyvarinen, A. (1999). Fast and robust fixed-point algorithms for independent component analysis. *IEEE transactions on Neural Networks*, 10(3):626–634.
- Hyvärinen, A., Karhunen, J., and Oja, E. (2004). *Independent Component Analysis*. Adaptive and Cognitive Dynamic Systems: Signal Processing, Learning, Communications and Control. Wiley.
- IT’IS Foundation (2016). Tissue properties database v3.1.
- J, A. (2012). The future of fetal monitoring. *Reviews in Obstetrics and Gynecology*, 5(3-4):e132–e136.
- Jafari, M. G. and Chambers, J. A. (2005). Fetal electrocardiogram extraction by sequential source separation in the wavelet domain. *IEEE Transactions on Biomedical Engineering*, 52(3):390–400.
- James, C. J. and Gibson, O. J. (2003). Temporally constrained ica: an application to artifact rejection in electromagnetic brain signal analysis. *IEEE Transactions on Biomedical Engineering*, 50(9):1108–1116.
- Jane, R., Blasi, A., Garcia, J., and Laguna, P. (1997). Evaluation of an automatic threshold based detector of waveform limits in holter ecg with the qt database. In *Computers in Cardiology 1997*, pages 295–298.
- Jenkins, H. M. (1989). Thirty years of electronic intrapartum fetal heart rate monitoring: discussion paper. *Journal of the Royal Society of Medicine*, 82(4):210–214.

- Jezewski, J., Matonia, A., Kupka, T., Roj, D., and Czabanski, R. (2012). Determination of fetal heart rate from abdominal signals: evaluation of beat-to-beat accuracy in relation to the direct fetal electrocardiogram. *Biomedizinische Technik/Biomedical Engineering*, 57(5):383–394.
- Johnson, D. H. and Dudgeon, D. E. (1993). *Array signal processing: concepts and techniques*. PTR Prentice Hall Englewood Cliffs.
- Kligfield, P., Gettes, L. S., Bailey, J. J., Childers, R., Deal, B. J., Hancock, E. W., van Herpen, G., Kors, J. A., Macfarlane, P., Mirvis, D. M., Pahlm, O., Rautaharju, P., and Wagner, G. S. (2007). Recommendations for the standardization and interpretation of the electrocardiogram: Part i: The electrocardiogram and its technology a scientific statement from the american heart association electrocardiography and arrhythmias committee, council on clinical cardiology; the american college of cardiology foundation; and the heart rhythm society endorsed by the international society for computerized electrocardiology. *Journal of the American College of Cardiology*, 49(10):1109 – 1127.
- Knuth, K. H. (2002). A bayesian approach to source separation. *arXiv preprint physics/0205032*.
- Lee, J., Park, K., and Lee, K. (2005). Temporally constrained ica-based foetal eeg separation. *Electronics Letters*, 41(21):1158–1160.
- Lee, T.-W., Girolami, M., and Sejnowski, T. J. (1999). Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources. *Neural computation*, 11(2):417–441.
- Lin, C., Bugallo, M., Mailhes, C., and Tourneret, J.-Y. (2011). Ecg denoising using a dynamical model and a marginalized particle filter. In *Signals, Systems and Computers (ASILOMAR), 2011 Conference Record of the Forty Fifth Asilomar Conference on*, pages 1679–1683. IEEE.
- Lin, Q.-H., Zheng, Y.-R., Yin, F.-L., Liang, H., and Calhoun, V. D. (2007). A fast algorithm for one-unit ica-r. *Information Sciences*, 177(5):1265–1275.
- Lipponen, J. A. and Tarvainen, M. P. (2013). Advanced maternal eeg removal and noise reduction for application of fetal qrs detection. In *Computing in Cardiology Conference (CinC), 2013*, pages 161–164. IEEE.
- Lu, W. and Rajapakse, J. C. (2006). Ica with reference. *Neurocomputing*, 69(16-18):2244–2257.
- Lukoševičius, M. and Jaeger, H. (2009). Reservoir computing approaches to recurrent neural network training. *Computer Science Review*, 3(3):127–149.
- Macones, G. A., Hankins, G. D. V., Spong, C. Y., Hauth, J., and Moore, T. (2008). The 2008 national institute of child health and human development workshop report on electronic fetal monitoring: update on definitions, interpretation, and research guidelines. *Obstet Gynecol*, 112(3):661–666.
- Martens, S. M., Rabotti, C., Mischi, M., and Sluijter, R. J. (2007). A robust fetal eeg detection method for abdominal recordings. *Physiological measurement*, 28(4):373.
- Matisko, P. and Havlena, V. (2012). Optimality tests and adaptive kalman filter. *IFAC Proceedings Volumes*, 45(16):1523–1528.
- McSharry, P. E., Clifford, G. D., Tarassenko, L., and Smith, L. A. (2003). A dynamical model for generating synthetic electrocardiogram signals. *IEEE transactions on biomedical engineering*, 50(3):289–294.

- Medical, N. (2015). The stan method. <http://www.neoventa.com/st-analysis/>. Downloaded March 5, 2018.
- MedlinePlus (2017). Cerebral Palsy. <https://medlineplus.gov/ency/article/000716.htm>. Accessed February 8, 2018.
- Mi, J.-X. and Gui, J. (2010). A method for ica with reference signals. In *Advanced Intelligent Computing Theories and Applications. With Aspects of Artificial Intelligence*, pages 156–162. Springer.
- Miklavčič, D., Pavšelj, N., and Hart, F. X. (2006). Electric properties of tissues. *Wiley encyclopedia of biomedical engineering*.
- Mindchild Medical, I. (2018). MindChild. <http://www.mindchild.com/>. Accessed February 23, 2018.
- Mohammad-Djafari, A. (2001). A bayesian approach to source separation. In *AIP Conference proceedings*, volume 567, pages 221–244. AIP.
- Monk, C., Fifer, W. P., Myers, M. M., Sloan, R. P., Trien, L., and Hurtado, A. (2000). Maternal stress responses and anxiety during pregnancy: Effects on fetal heart rate. *Developmental Psychobiology*, 36(1):67–77.
- Moody, O. (2017). Overpopulation is the biggest threat to mankind, Nobel laureates say. <https://www.thetimes.co.uk/article/overpopulation-is-the-biggest-threat-to-mankind-nobel-laureates-say-zxdnv2bcv>. Accessed February 13, 2018.
- Neilson, J. P. (2015). Fetal electrocardiogram (ecg) for fetal monitoring during labour. *Cochrane Database of Systematic Reviews*, (12).
- Niknazar, M., Becker, H., Rivet, B., Jutten, C., and Comon, P. (2014). Blind source separation of underdetermined mixtures of event-related sources. *Signal Processing*, 101:52–64.
- Niknazar, M., Rivet, B., and Jutten, C. (2013). Fetal ecg extraction by extended state kalman filtering based on single-channel recordings. *IEEE Transactions on Biomedical Engineering*, 60(5):1345–1352.
- Nordhausen, K., Ollila, E., and Oja, H. (2011). On the performance indices of ica and blind source separation. In *Signal Processing Advances in Wireless Communications (SPAWC), 2011 IEEE 12th International Workshop on*, pages 486–490. IEEE.
- OPENPediatrics Staff (2015). Electrocardiogram waveform: ST segment changes. <https://www.openpediatrics.org/sites/default/files/assets/Electrocardiogram-Waveform-ST-Segment-Changes-.jpg>. Downloaded June 12, 2018.
- Oster, J., Behar, J., Sayadi, O., Nemati, S., Johnson, A. E. W., and Clifford, G. D. (2015). Semisupervised ecg ventricular beat classification with novelty detection based on switching kalman filters. *IEEE Transactions on Biomedical Engineering*, 62(9):2125–2134.
- Oudijk, M. A., Kwee, A., Visser, G. H., Blad, S., Meijboom, E. J., and Rosén, K. G. (2004). The effects of intrapartum hypoxia on the fetal qt interval. *BJOG: An International Journal of Obstetrics & Gynaecology*, 111(7):656–660.
- Pan, J. and Tompkins, W. J. (1985). A real-time qrs detection algorithm. *IEEE transactions on biomedical engineering*, (3):230–236.

- Pardi, G., Brambati, B., Dubini, S., Luchetti, D., Polvani, F., and Candiani, G. (1971). Analysis of the fetal electrocardiogram by the group averaging technique: Preliminary report1. In *Perinatal Medicine*, pages 75–84. Karger Publishers.
- Pathirana, J., Muñoz, F. M., Abbing-Karahagopian, V., Bhat, N., Harris, T., Kapoor, A., Keene, D. L., Mangili, A., Padula, M. A., Pande, S. L., Pool, V., Pourmalek, F., Varricchio, F., Kochhar, S., Cutland, C. L., and Group, T. B. C. N. D. W. (2016). Neonatal death: Case definition & guidelines for data collection, analysis, and presentation of immunization safety data. *Vaccine*, 34(49):6027–6037.
- Petersen, K. B. and Pedersen, M. S. (2012). The matrix cookbook. [http://www2.imm.dtu.dk/pubdb/views/edoc\\_download.php/3274/pdf/imm3274.pdf](http://www2.imm.dtu.dk/pubdb/views/edoc_download.php/3274/pdf/imm3274.pdf). Downloaded March 5, 2018.
- Physionet (2016). Non-Invasive Fetal Electrocardiogram Database. <https://physionet.org/physiobank/database/nifecgdb/>. Accessed February 15, 2018.
- Physionet (2017). Noninvasive Fetal ECG: the PhysioNet/Computing in Cardiology Challenge 2013. <https://www.physionet.org/challenge/2013>. Accessed February 7, 2018.
- Pildner von Steinburg, S., Boulesteix, A.-L., Lederer, C., Grunow, S., Schiermeier, S., Hatzmann, W., Schneider, K.-T. M., and Daumer, M. (2013). What is the "normal" fetal heart rate? *PeerJ*, 1:e82.
- Podziemski, P. and Gieraltowski, J. (2013). Fetal heart rate discovery: algorithm for detection of fetal heart rate from noisy, noninvasive fetal ecg recordings. In *Computing in Cardiology Conference (CinC), 2013*, pages 333–336. IEEE.
- Qian, N. (1999). On the momentum term in gradient descent learning algorithms. *Neural networks*, 12(1):145–151.
- Roser, M. (2018). Child Mortality. <https://ourworldindata.org/child-mortality>. Accessed February 7, 2018.
- Roser, M. and Ortiz-Ospina, E. (2018). World Population Growth. <https://ourworldindata.org/world-population-growth>. Accessed February 13, 2018.
- Sachs, J. and WHO (2001). *Macroeconomics and Health: Investing in Health for Economic Development: Report of the Commission on Macroeconomics and Health*. Health Reference Center Academic. OMS.
- Sæderup, R. G., Hoang, P., Winther, S., Bøttcher, M., Struijk, J., Schmidt, S., and Østergaard, J. (2018). Estimation of the second heart sound split using windowed sinusoidal models. *Biomedical Signal Processing and Control*, 44:229–236.
- Sameni, R. (2008). *Extraction of fetal cardiac signals from an array of maternal abdominal recordings*. PhD thesis, Institut National Polytechnique de Grenoble-INPG; Sharif University of Technology (SUT).
- Sameni, R. and Clifford, G. D. (2010). A review of fetal ecg signal processing; issues and promising directions. *The open pacing, electrophysiology & therapy journal*, 3:4–20.
- Sameni, R., Clifford, G. D., Jutten, C., and Shamsollahi, M. B. (2007a). Multichannel ecg and noise modeling: Application to maternal and fetal ecg signals. *EURASIP Journal on Advances in Signal Processing*, 2007(1):043407.

- Sameni, R., Jutten, C., and Shamsollahi, M. B. (2006a). What ica provides for ecg processing: Application to noninvasive fetal ecg extraction. In *Signal Processing and Information Technology, 2006 IEEE International Symposium on*, pages 656–661. IEEE.
- Sameni, R., Jutten, C., and Shamsollahi, M. B. (2008). Multichannel electrocardiogram decomposition using periodic component analysis. *IEEE Transactions on Biomedical Engineering*, 55(8):1935–1940.
- Sameni, R., Shamsollahi, M. B., Jutten, C., and Clifford, G. D. (2007b). A nonlinear bayesian filtering framework for ecg denoising. *IEEE Transactions on Biomedical Engineering*, 54(12):2172–2185.
- Sameni, R., Vrins, F., Parmentier, F., Herail, C., Vigneron, V., Verleysen, M., Jutten, C., and Shamsollahi, M. (2006b). Electrode selection for noninvasive fetal electrocardiogram extraction using mutual information criteria. In *AIP Conference Proceedings*, volume 872, pages 97–104. AIP.
- Samuel, J. and Franklin, C. (2008). *Hypoxemia and Hypoxia*, pages 391–394. Springer New York, New York, NY.
- Sepulveda W, S. N. (2013). Fetal Circulation. <https://basicsofpediatricanesthesia.files.wordpress.com/2013/07/1-fetal-circulation.jpg>. Accessed February 9, 2018.
- Stinstra, J. (2001). The reliability of the fetal magnetocardiogram.
- Sun, Z.-L. and Shang, L. (2010). An improved constrained ica with reference based unmixing matrix initialization. *Neurocomputing*, 73(4-6):1013–1017.
- Thaler, M. (2007). *The Only EKG Book You’ll Ever Need*. Board Review Series. Lippincott Williams & Wilkins.
- UN Inter-agence Group for Child Mortality Estimation (2017). *Levels and Trends in Child Mortality: Report 2017*. United Nations Children’s Fund.
- Van Trees, H. (2004). *Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory*. Detection, Estimation, and Modulation Theory. Wiley.
- VanPutte, C., Regan, J., and Russo, A. (2016). *Seeley’s essentials of anatomy and physiology*. McGraw-Hill Education, 9th edition.
- Varanini, M., Tartarisco, G., Billeci, L., Macerata, A., Pioggia, G., and Balocchi, R. (2013). A multi-step approach for non-invasive fetal ecg analysis. In *Computing in Cardiology Conference (CinC), 2013*, pages 281–284. IEEE.
- Vrins, F., Jutten, C., and Verleysen, M. (2004). Sensor array and electrode selection for non-invasive fetal electrocardiogram extraction by independent component analysis. In *International Conference on Independent Component Analysis and Signal Separation*, pages 1017–1024. Springer.
- Yang, H., Bukkapatnam, S. T., and Komanduri, R. (2012). Spatiotemporal representation of cardiac vectorcardiogram (vcg) signals. *Biomedical engineering online*, 11(1):16.



# Appendices

## A Electrocardiographic Leads

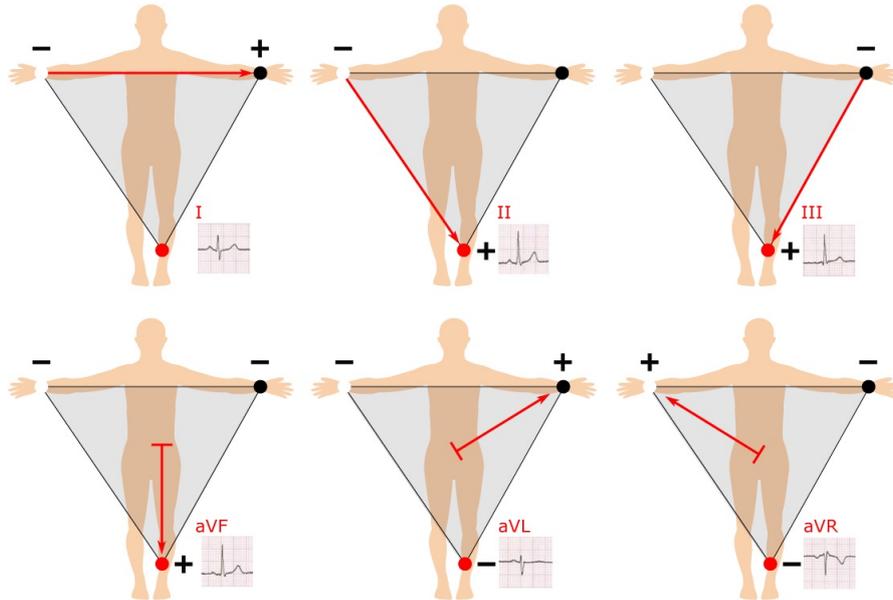
A normal ECG consists of 12 *leads*, which can be split up into two categories: The 6 limb lead (also called extremity leads), and the 6 chest leads (also called precordial leads) [Hall, 2016, p. 134-136]. These two types will be described in the following.

### Limb leads

The extremity leads be split up into two types: Three standard leads and three *augmented* leads. All six limb leads are obtained by placing three electrodes on the body: On the right arm (RA), the left arm (LA) and left leg (LL). These three electrodes form an equilateral triangle called Einthoven's triangle, as can be seen in Figure A.1. The idea of this triangle is that the distances between the cardiac electric vectors and extremities is great enough to be assumed infinity, which means the heart can be represented by a single electric dipole [Fuster et al., 2007, p. 295].

The standard leads are also called bipolar limb leads, as they measure the potential difference between two of the electrodes as described in Table A.1. The augmented leads are formed by connecting two of the electrodes through a resistance (thus taking their average) to the negative terminal, forming a common ground and using the remaining electrode as the positive, see Figure A.1 and Table A.1 [Hall, 2016, p. 136].

Each lead measures the electrical activity from a certain angle, called the angle of orientation, and can most easily be found looking at Einthoven's triangle in Figure A.1. These angles are also shown in Table A.1.



**Figure A.1:** Einthoven's triangle and the six limb leads [Cables and Sensors, LLC, 2018].

An interesting relation can be derived from Kirchoff's voltage law, which states that the sum of voltages around a closed network is zero. This means that the sum of the potentials recorded by lead I and II equal the potential recorded by lead III. This is also known as Einthoven's law [Hall, 2016, p. 135].

Electrode choice/ Lead number	Negative	Positive	Angle
Lead I	RA	LA	$0^\circ$
Lead II	RA	LL	$60^\circ$
Lead III	LA	LL	$120^\circ$
Lead aVF	$0.5 (RA+LA)$	LL	$90^\circ$
Lead aVL	$0.5 (RA+LL)$	LA	$-30^\circ$
Lead aVR	$0.5 (LA+LL)$	RA	$-150^\circ$

**Table A.1:** The relationship between limb leads and electrodes, as well as the angle of orientation. [Thaler, 2007, p. 40]

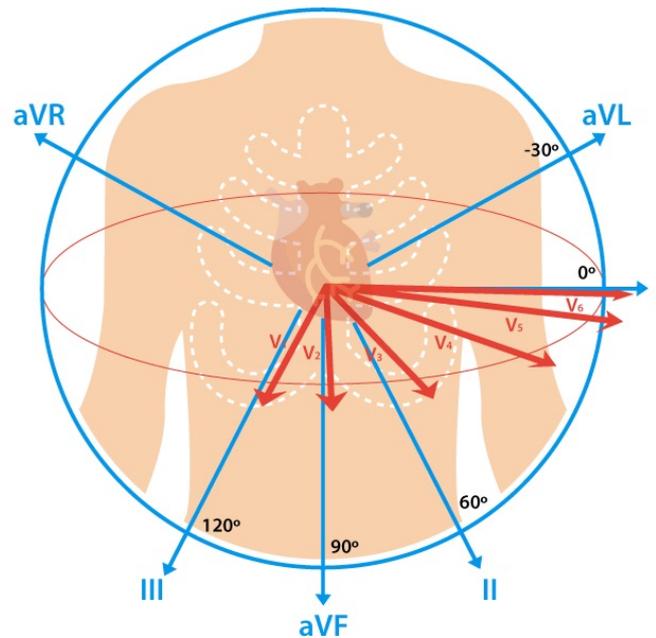
### Chest leads

Chest leads are often called *unipolar leads*, as they consist of a single electrode for each lead. This is of course not true, as one always must measure a voltage relative to some ground potential. This ground potential is in ECG terminology called Wilson's Central Terminal (WCT), and is formed by taking the average of the LA, RA and LL electrodes, by connecting them together through  $5000\Omega$  resistors. The sum potentials of these electrodes is approximately<sup>1</sup> equal to zero throughout the heart cycle with respect

<sup>1</sup>As the LA, RA and LL are not exactly equidistant from each other and the heart, the potential of the WCT is approx.  $0.3 \text{ mV}$  on average [Fuster et al., 2007, p. 296]

to any point on the body surface, making it a good grounding point [Fuster et al., 2007, p. 295].

The six chest leads thus measure the electric potential between the electrodes and the WCT. The six electrodes are placed in the front and left part of the chest, meaning they measure in the horizontal plane of the chest. The measurement directions obtained by chest leads is shown in Figure A.2, together with the directions of the six limb leads.



**Figure A.2:** The different measurement angles obtained by a 12-lead ECG. Notice how the extremity leads measure in the vertical plane, while the chest leads measure in the horizontal plane. [Cables and Sensors, LLC, 2018].

## 12-lead ECG

Combining these 12 leads provides a wide range of information about the electric activity of the heart, with each lead giving information in a angle of orientation. For instance, "leads II, III, and AVF are called the inferior leads because they most effectively view the inferior surface of the heart." [Thaler, 2007, p. 42].

The orientation of the leads determine how the ECG looks: For instance, the P-wave comes from the depolarization from the right atrium to the SA node, i.e. moving in a downwards direction to the left. As lead II measures in this direction, the P-wave will on lead II be positive. However, the aVR lead measures in the opposite direction, meaning the P-wave will be negative on this lead. [Thaler, 2007, p. 47].

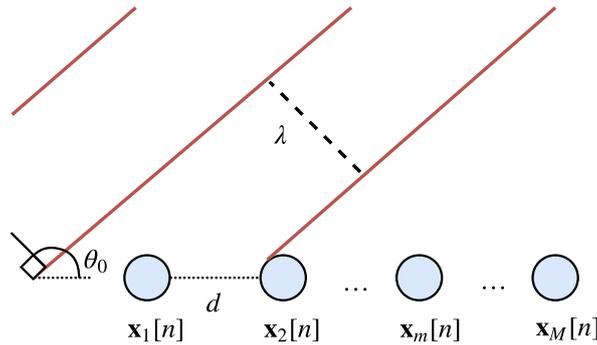
The electrical axis is defined as a vector originating at the WCT, and shows the direction of the electrical activation process of the heart projected onto the vertical plane given by the limb leads [Fuster et al., 2007, p. 295]. The normal electrical axis of the ventricles is  $59^\circ$  [Hall, 2016, p. 145], and for this reason one sometimes chooses to find the rhythm (heart rate) by looking only at lead II, as lead II will give a very strong QRS complex, making it easy to find RR-interval.

## B Beamforming Methods

Beamforming is a method used to perform spatial filtering, by forming a directional sensor from multiple sensors placed in an array configuration. By spatial domain is meant that instead of sampling over time, samples are spread over different locations in space [Van Trees, 2004, p. 17]. As part of this, it is often necessary to estimate the direction of arrival (DoA) of the signal. This can be done by doing spatial spectral analysis, which can be interpreted as the spatial equivalent of applying the Fourier transform to a temporal signal.

This means that when the DoA has been found, a spatial filter can be applied though beamforming, such that signals not coming from the desired DoA, i.e. noise sources, are attenuated, just as a FIR filter is used to remove undesired frequency components of a temporal signal. A beamformer is therefore an option for attenuating the maternal ECG, by exploiting the different DoAs of the maternal and fetal ECG signals. The purpose of this section is to explore the opportunities and limitations related to applying a beamformer to FECG extraction.

In Figure B.3,  $M$  sensors are placed in a Uniform Linear Array (ULA) separated by a distance  $d$ , onto which an wave with wavelength  $\lambda$  is coming from the DoA  $\theta_0$ . At every time instance  $n$ , each sensor produces an output  $\mathbf{x}_m[n]$ .



**Figure B.3:** A wave with wavelength  $\lambda$  hitting an  $M$ -element uniform linear array at a direction of arrival (DoA) of  $\theta_0$ .

### B.1 Relation Between Spatial and Temporal Sampling

The similarity between spatial spectral analysis and the discrete-time Fourier Transform (DTFT) is obvious when comparing their definitions - note however that this direct comparison is only valid for ULA's [Van Trees, 2004, p. 38]:

$$X(\theta, n) = \sum_{m=-\infty}^{\infty} x_m[n] \cdot \exp(2\pi j \cdot d \cdot m \cdot (\xi \cos(\theta))) \quad (\text{B.1})$$

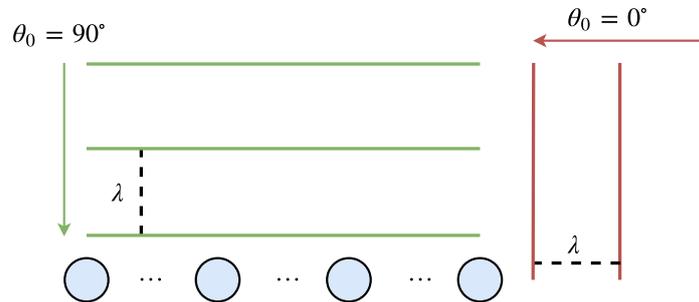
$$X(f) = \sum_{n=-\infty}^{\infty} x[n] \cdot \exp(2\pi j \cdot T_s \cdot n \cdot f) \quad (\text{B.2})$$

As is clear from Equation B.1 and B.2 is that both are transforms on the same form, but with different summing indices and variable names in the exponential. In Table B.2, a direct comparison of these variables in the temporal and spatial domain can be seen.

	Frequency	Period	Number of samples	Sampling rate
<b>Spatial</b>	$\xi \cdot \cos(\theta_k)$	$\lambda$	$M$	$d$
<b>Temporal</b>	$f$	$T$	$N$	$T_s$

**Table B.2:** An overview of the similarities between temporal and spatial domains.

While the Fourier domain describes the frequency content of the signal for a temporal signal, the Fourier domain for a spatial signal describes the DoA, ranging from 0 to  $2\pi$  (or  $-\pi$  to  $\pi$ ), relative to the broadside of the array, as can be seen in Figure B.4.



**Figure B.4:** Two planar waves with the same wavelength  $\lambda$  hitting a sensor array. The red wave represents a spatial frequency of  $\xi = \lambda^{-1}$ , while the green wave represents a spatial frequency of 0, as the wavefront hits all sensors simultaneously.

The intuition behind the term spatial frequency is, just as with temporal frequency, the frequency/phase difference between consecutive samples: The closer the wave is to coming from an angle of 0 or  $\pi$  radians (relative to the array axis), the bigger the phase difference between each sample, and thus the higher the frequency. Conversely, when the wavefront is normal to the array, the phase difference between samples is 0, corresponding to DC. This is the reason why the spatial frequency in Equation B.1 and Table B.2 contains a cosine-term, as this encodes the DoA information into the expression. In order to avoid spatial aliasing, the sensors must have a spacing of  $d < \lambda/2$ .

## B.2 Propagation Speed and Sampling Resolution

As is clear from the concept of spatial frequency, beamforming relies on using the phase information present in the received signal and deriving the direction of arrival (DoA) from this phase difference between array elements. It is therefore critical that there is a significant phase difference between array elements.

To investigate the size of this problem, the maximum phase difference between two array elements is calculated. In order to do this, the velocity of the electromagnetic wave of the ECG in the human body must be found. At low frequencies ( $< 100$  kHz), it is primarily the extracellular fluid which conducts the electric signal through the body. For this calculation, this fluid is assumed being water, having a relative permittivity of  $\epsilon_r = 85$  [IT'IS Foundation, 2016] and a relative permeability of  $\mu_r \approx 1$ . The frequency of the signal is also needed - in this case we will assume a best-case frequency of 100 Hz, which is the cutoff frequency of the low-pass filter described in Section 5.5.2. Using these values, the wavelength of

the electric signal in the human body can be found as:

$$\lambda = \frac{v}{f} = \frac{c}{f\sqrt{\epsilon_r\mu_r}} = \frac{3 \cdot 10^8}{100 \cdot \sqrt{85 \cdot 1}} = 325 \cdot 10^3 \text{ m} \quad (\text{B.3})$$

The wavenumber  $k$  is defined as [Van Trees, 2004, p. 30]:

$$k = \frac{d\theta}{dx} = \frac{2\pi}{\lambda} \quad (\text{B.4})$$

where:

$\theta$	is the phase	[rad]
$x$	is the position	[m]
$\lambda$	is the wavelength	[m]

Using this definition and substituting the differential with simple differences, we can find the phase difference between two points on the wave, separated by  $\Delta x$ . In this case, we are interested in the phase difference between two electrodes on the maternal stomach, so we set  $\Delta x = 0.2$  m, as this is a realistic yet optimistic electrode distance:

$$\Delta\theta = \frac{2\pi\Delta x}{\lambda} = \frac{2\pi \cdot 0.2}{325 \cdot 10^3} = 3.86 \cdot 10^{-6} \text{ rad} \quad (\text{B.5})$$

Converting this number to decibel gives -108 dB, meaning that the SNR must be at least 108 dB in order for the phase difference not to be overshadowed by the noise. Also, by converting the inverse of this number to base 2 gives the number of bits needed to represent this number, i.e. the bit-resolution needed to sample the signal with the precision:

$$N_{\text{bits}} = \lceil \log_2((\Delta\theta)^{-1}) \rceil = 18 \quad (\text{B.6})$$

where:

$N_{\text{bits}}$	is the necessary bit-resolution
-------------------	---------------------------------

Thus, an ADC with a bit resolution of at least 18 is needed. All these factors prove that it is highly impractical, even in the best case, to be able to sample the signal such that it can be used for beamforming.

### B.3 Beamformer gain

The purpose of this section is to investigate which performance increase can be obtained by using beamforming methods on ECG signals.

A uniform, linear array (ULA) configuration is assumed, as this eases the calculations significantly. It is assumed there are  $M = 4$  sensors, while  $K = 180$  describes the number of directions of arrival that is swept through, which can be interpreted as the number of bins in the Fourier Transform due to the similarity between spatial and temporal domains as described above. A single array output vector  $\mathbf{x}[n]$  at sampling time  $n$  is called a snapshot and is modeled as:

$$\mathbf{x}[n] = \mathbf{A}(\theta)\mathbf{s}[n] + \mathbf{v}[n] \quad (\text{B.7})$$

where:

$\theta$	is a $K \times 1$ -vector of the DoAs	[rad]
----------	---------------------------------------	-------

$\mathbf{A}(\theta)$  is an  $M \times K$  matrix of steering vectors  
 $\mathbf{s}[n]$  is a  $K$ -vector of signal waveforms  
 $\mathbf{v}[n]$  is a  $M \times 1$  noise vector

In the following the noise is ignored, and we just consider the performance of the beamformer in ideal settings. The columns  $\mathbf{a}(\theta_k)$  (steering vectors) of the steering matrix are for a ULA given as:

$$\mathbf{a}(\theta_k) = [1, \exp(2\pi j \xi d \cos(\theta_2)), \dots, \exp(2\pi j (M-1) \xi d \cos(\theta_K))]^T \quad (\text{B.8})$$

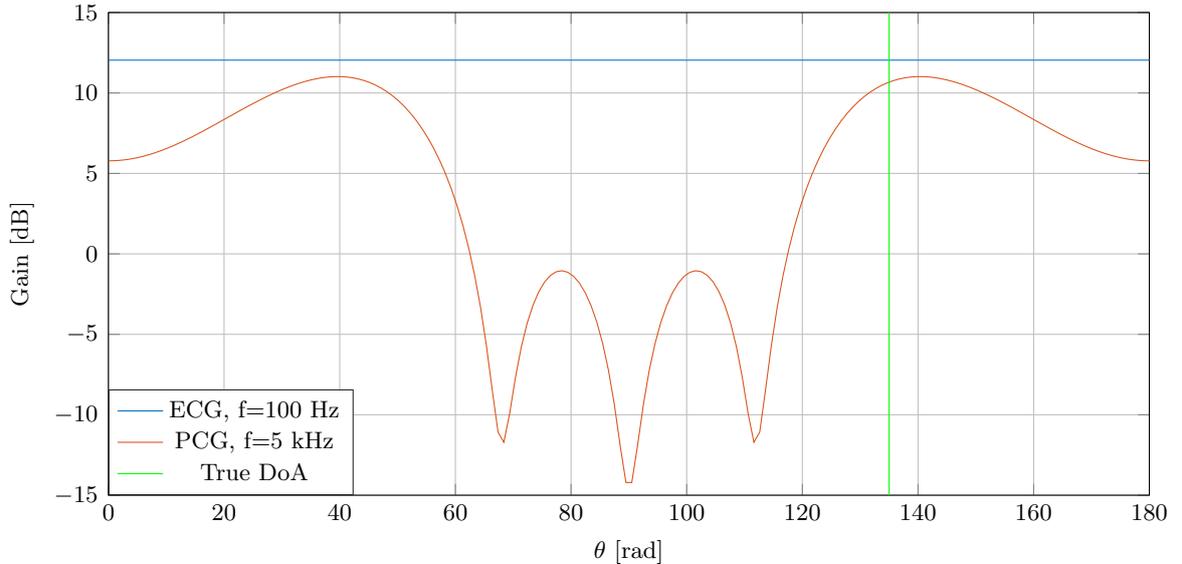
where:

$d$  is the element spacing [m]  
 $\xi = \lambda^{-1}$  is the spatial frequency [m<sup>-1</sup>]

Notice that this steering matrix is simply a matrix equivalent of the spatial DTFT as described in Equation B.1. Assuming the simple case of the transmitted signal being a sinusoid, the signal received by the the  $m$ 'th array element at time  $n$  can be found as:

$$s_m[n] = \cos(2\pi \xi_n m \cdot d \cdot \cos(\theta_0)) \quad (\text{B.9})$$

In this case, we choose to only consider a single snapshot, i.e. the case of  $n = 0$ . Setting  $d = \Delta x = 0.2$  m and using the wavelength  $\lambda = 325 \cdot 10^3 \text{ m}^{-1}$  as found above, the array output  $\mathbf{x}[n] = \mathbf{A}(\theta)\mathbf{s}[n]$  can be calculated. Converting the output to dB yields the plot seen in Figure B.5. In order to show that beamforming is a useful tool in biomedical applications, the same array is re-used, but is instead applied to phonocardiographic (PCG) signals, i.e. audio recordings of the heart. Sounds moves at a speed of  $v = 1540 \text{ m/s}$  in human tissue [Curry et al., 1990, p. 325], and for visualization purposes, an abnormal high frequency of  $f = 5 \text{ kHz}$  is chosen. The beamforming pattern emerging from these assumptions can also be seen in Figure B.5, from which it is clear that a much higher array gain is obtained.



**Figure B.5:** The array gain in dB as a function of the angle  $\theta$ , for both the ECG signal considered and a PCG signal for reference. In both cases, the DoA is  $\theta_0 = 135^\circ$ .

The reason for the near-zero array gain on the ECG signal, is due to the fact that  $d \ll \lambda$ . If one were to think of this in the time domain instead of the spacial domain, this would be equivalent to the frequency of the incoming signal being much smaller than the sampling frequency, which yields a poor resolution in the Fourier domain:

The resolution of each DFT bin is  $\Delta f = 1/(M \cdot d)$ , which for  $M = 4$  and  $d = 0.2$  m yields  $1.25 \text{ m}^{-1}$ . However, due to the high velocity of the ECG, the spatial frequency is  $\xi = \lambda^{-1} = 3.1 \cdot 10^{-6}$ . This is much smaller than the resolution of a single DFT bin, meaning that no matter the DoA, all samples will fall in the first DFT-bin, making it impossible to estimate the DoA. It would thus require a lower signal velocity (as demonstrated with the PCG signal) or more sensors or a bigger distance between them to get a higher DoA accuracy.

## B.4 Anisotropic nature of the heart

A typical assumption in beamforming applications is that the source is of isotropic nature, i.e it is a point-source which radiate with equal power in all directions. Far-field is also a typical assumption made, meaning the source is located so far away that the arrays observe a plane wave. [Johnson and Dudgeon, 1993, p. 20]

This is however not a valid assumption to pose for the human heart: As described in Appendix A, the heart can be modeled as a dipole, which does not have isotropic transmission. Likewise, far-field cannot be assumed: In fact, the whole purpose of ECG is to utilize the non-isotropic and near-field properties of the heart, such that different projection of the electric activity of the heart is observed at different leads.

This causes the problem (in a beamforming sense) that different signals are received at different array elements, due to the near-field anisotropic nature of the heart. This means the signals received are not just simple delayed versions of the original signal, even if the propagation medium was ideal (free-field), making the beamforming more difficult. The assumptions on the propagation medium is described in greater detail in the following.

## B.5 Inhomogeneous propagation medium

Typically, the medium through which the signal propagates is assumed homogeneous, i.e. having the same properties at every point in space. Inhomogeneous media "lead to refraction, which changes the direction of propagation throughout space" [Johnson and Dudgeon, 1993, p. 39], and must be taken into account by the array processing algorithm.

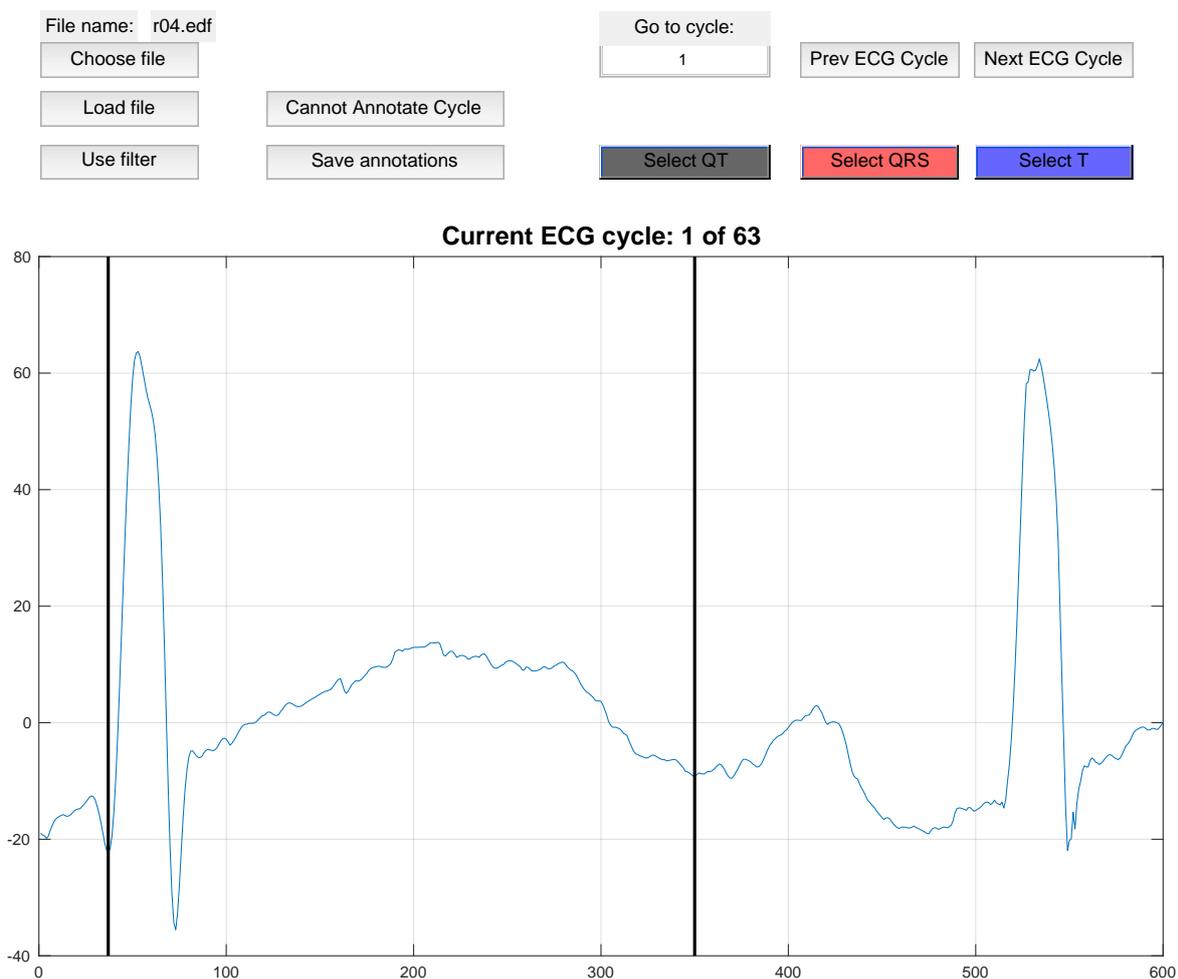
The human body cannot be considered homogeneous, as it consists of various types of tissue, both muscle, fat and skin. These tissues have different relative permittivities, which alter their conductive properties [Miklavčič et al., 2006, table 1]. As there are many changes in transmission medium (from heart to layers of muscle and fat and finally to the skin), the propagation medium becomes quite complicated to describe. This becomes even more complicated by the fact that the medium is different for different directions - the layering of muscle and fat is not the same in all directions from the heart, meaning a different propagation channels are seen by different array elements.

A final complicating factor is the inter-subject variations in the propagation medium: All humans are different, and the propagation channel will therefore depend on the amount of fat and muscle tissue in each subject, in addition to the position and movement of the fetus. Modeling the propagation medium will therefore become almost impossible, which excludes beamforming as a suitable method for extracting the fetal ECG.

## C Annotation GUI

In this appendix, the graphical user interface (GUI) which has been designed for annotating fetal ECG is presented, and can be seen in Figure C.6.

The purpose of having such a user interface is to allow annotators an fast, easy way to annotate the morphology of ECG signals. If the annotations are performed on reference signals, they can be used as a baseline for comparing the quality of the morphology from estimated fetal ECGs. One can also use the GUI for annotating the extracted signals, to allow for direct comparison with true fetal ECGs, thus avoiding the need for a segmentation algorithm.



**Figure C.6:** A screenshot of the GUI developed for annotating the morphology of (fetal) ECG. The QT interval has been annotated on this ECG cycle.

The GUI has a button for choosing which file to load, but so far only .edf files are supported, as the GUI is intended for annotating the scalp electrode data from the ADFECGDB data set, which is saved in this format. After loading a file, an optional bandpassfilter (bandwidth 1-100 Hz) can be applied. The GUI automatically presents segments of the ECG, using pre-annotated QRS locations. In order to make the

ECGs less noisy, the GUI takes the median of 10 ECG cycles - this also reduces the number of ECGs which must be annotated by a factor of 10.

As the selected performance measures are the QT-interval and the T/QRS ratio, the GUI must allow for annotating these quantities. This is done by simply clicking the designated buttons, after which one can click the start/stop or top/bottom values of the intervals to be annotated. In Figure C.6, the QT interval has been annotated using this approach.

When a new, un-annotated ECG cycle is presented, the GUI automatically suggests the annotations from the previous cycle as suggested annotations and marks the annotations as only being suggestions. The annotator can then discard the suggestion by manually annotating the cycle, or accept them by simply moving on to the next cycle.

The annotations are automatically saved every time a new ECG cycle is chosen, but there is also an option for manually saving the annotations.

## D Results of ICA on Non-ECG Signals

In this appendix, the results from the comparison of the different ICA algorithms on artificial non-ECG signals are shown in a series of tables.

Tests are performed in two cases: No noise and in 12 dB SNR. The tests are conducted on 5 synthetic signals, and are repeated 50 times using different random initializations of the de-mixing matrix  $\mathbf{B}$ , for those cases where the reference signals are not used as initialization of  $\mathbf{B}$ .

### D.1 No noise

ICA method	Sig. 1	Sig. 2	Sig. 3	Sig. 4	Sig. 5
FastICA	17.27	32.64	36.00	15.72	27.27
FastICA w. ref	17.27	32.64	36.00	15.72	27.27
InfoMax	1.74	1.27	1.38	3.17	1.47
InfoMax w. ref	0.58	0.22	0.91	0.90	0.70
Ext. InfoMax	24.49	22.22	-2.83	44.75	-2.64
Ext. InfoMax w. ref	17.95	29.61	36.27	15.03	27.63
cICA	1.60	1.65	2.78	-1.49	0.39
cICA w. ref, $\epsilon = 0.1$	41.91	41.98	48.95	182.36	50.68
cICA w. ref, $\epsilon = 0.9$	41.91	41.98	48.95	182.36	50.68
cICA w. ref, rand. init, $\epsilon = 0.1$	1.60	1.65	2.78	-1.49	0.39
cICA w. ref, rand. init, $\epsilon = 0.9$	1.60	1.65	2.78	-1.49	0.39

**Table D.3:** The SNR for all 5 non-ECG signals, for different ICA methods. Results are the found by taking the mean over 50 random starting values of the de-mixing matrix  $\mathbf{B}$ . No noise has been added to the observations.

**D.2 12 dB SNR**

ICA method	Sig. 1	Sig. 2	Sig. 3	Sig. 4	Sig. 5
FastICA	-2.79	4.81	7.88	-2.77	-0.31
FastICA w. ref	-2.79	4.81	7.88	-2.77	-0.31
InfoMax	-2.85	0.30	1.62	-2.72	0.29
InfoMax w. ref	-2.79	0.39	0.57	-2.78	-0.22
Ext. InfoMax	-2.78	3.96	8.42	-2.86	-1.00
Ext. InfoMax w. ref	-2.82	4.90	8.39	-2.78	-0.55
cICA	-2.94	0.14	0.42	-2.72	-0.92
cICA w. ref, $\epsilon = 0.1$	-2.70	5.06	8.45	-2.67	1.90
cICA w. ref, $\epsilon = 0.9$	-2.70	5.06	8.45	-2.67	1.90
cICA w. ref, rand. init, $\epsilon = 0.1$	-2.83	-0.72	1.13	-2.78	0.12
cICA w. ref, rand. init, $\epsilon = 0.9$	-2.83	-0.72	1.13	-2.78	0.12

**Table D.4:** The SNR for all 5 non-ECG signals, for different ICA methods. Results are the found by taking the mean over 50 random starting values of the de-mixing matrix  $\mathbf{B}$ . White Gaussian noise has been added to the observations at an SNR of 12 dB.

## E Results of ICA Comparison on ECG Signals

In this appendix, the results from the comparison of the different ICA algorithms is shown in a series of tables.

Tests are performed in two cases: No noise and in 6 dB SNR. The tests are conducted on 5 synthetically generated recordings i.e. on 5 different subjects. The step-size used for the cICA algorithm is 0.1 and 0.0001 for the InfoMax methods. No momentum is used.

The four algorithms tested are:

1. FastICA
2. InfoMax
3. Extended InfoMax
4. Temporally constrained ICA (ICA with reference)

All algorithms are tested with and without a reference signal - for testing purposes, the reference signals used are the true ECG signals. In the cases of FastICA and InfoMax, the reference is used as initialization for the estimated de-mixing matrix  $\mathbf{B}$ . For the constrained ICA case, the reference is used as the reference signal for the algorithm.

### E.1 No noise conditions

	FastICA	FastICA w. ref	InfoMax	InfoMax w. ref	Ext.InfoMax	Ext. InfoMax w. ref	cICA	cICA w. ref
Rec. 1	2.49	2.49	2.96	2.91	3.07	2.49	2.96	0.05
Rec. 2	2.28	2.28	3.15	3.24	2.77	2.28	3.36	0.03
Rec. 3	1.60	1.60	2.53	2.68	2.52	1.60	3.58	0.03
Rec. 4	2.03	2.03	2.97	3.10	2.91	2.03	3.49	0.06
Rec. 5	2.29	2.29	2.59	2.93	3.53	2.29	3.46	0.04
Mean	2.14	2.14	2.84	2.97	2.96	2.14	3.37	0.04

**Table E.5:** A comparison of the performance index (PI) of a selection of ICA methods in zero noise, with and without the true sources as a reference for  $\mathbf{B}$ -matrix initialization. Lower numbers are better.

	FastICA	FastICA w. ref	InfoMax	InfoMax w. ref	Ext.InfoMax	Ext. InfoMax w. ref	cICA	cICA w. ref
Rec. 1	3.60	3.60	1.93	3.81	2.39	3.60	0.19	41.93
Rec. 2	7.75	7.75	2.10	4.50	3.33	7.75	0.71	45.62
Rec. 3	8.20	8.20	6.32	6.50	6.68	8.20	1.04	44.61
Rec. 4	6.10	6.10	3.96	5.42	2.57	6.10	0.69	39.77
Rec. 5	5.14	5.14	2.41	4.32	1.85	5.14	1.97	45.14
Mean	6.16	6.16	3.34	4.91	3.37	6.16	0.92	43.42

**Table E.6:** A comparison of the signal to noise ratio (SNR) between sources for different ICA methods in zero noise, with and without the true sources as a reference for  $\mathbf{B}$ -matrix initialization. Higher numbers are better.

## E.2 6 dB SNR

	FastICA	FastICA w. ref	InfoMax	InfoMax w. ref	Ext.InfoMax	Ext. InfoMax w. ref	cICA	cICA w. ref
Rec. 1	2.55	2.55	3.06	3.46	2.81	2.55	2.94	2.56
Rec. 2	2.19	2.19	3.31	3.35	3.22	2.19	3.59	1.86
Rec. 3	2.61	2.61	3.71	3.14	2.61	2.61	3.96	2.68
Rec. 4	2.24	2.24	3.49	3.36	2.73	2.24	3.80	2.26
Rec. 5	2.69	2.69	2.86	3.14	2.54	2.69	3.37	1.31
Mean	2.46	2.46	3.28	3.29	2.78	2.46	3.53	2.14

**Table E.7:** A comparison of the performance index (PI) between sources for different ICA methods. Noise is added to the recordings, resulting in an SNR for the fetal ECG of 6 dB. Higher numbers are better.

	FastICA	FastICA w. ref	InfoMax	InfoMax w. ref	Ext.InfoMax	Ext. InfoMax w. ref	cICA	cICA w. ref
Rec. 1	4.56	4.56	2.33	2.96	3.12	4.56	-0.15	16.83
Rec. 2	5.98	5.98	1.67	3.16	1.22	5.98	0.13	9.37
Rec. 3	5.91	5.91	4.14	5.47	3.91	5.91	2.27	10.51
Rec. 4	6.03	6.03	2.94	5.67	2.06	6.03	0.73	10.44
Rec. 5	5.89	5.89	3.53	3.96	1.16	5.89	1.13	13.76
Mean	5.67	5.67	2.92	4.24	2.29	5.67	0.82	12.18

**Table E.8:** A comparison of the signal to noise ratio (SNR) between sources for different ICA methods. Noise is added to the recordings, resulting in an SNR for the fetal ECG of 6 dB. Higher numbers are better.

## F Morphology Test Results on Artificial Signals

In this appendix, the mean value of the three morphology test measures (QT-interval, T/QRS ratio and RMSE) is presented for each extraction method at each SNR level (infinite, 12 dB, 9 dB, 6 dB, 3 dB and 0 dB).

The QT interval error can be seen in Table F.9, T/QRS error in Table F.10 and RMSE in Table F.11.

### F.1 QT-interval error

Method/SNR	No noise	12 dB	9 dB	6 dB	3 dB	0 dB
<b>JADEICA</b>	11.50	3.00	11.00	6.50	9.00	10.00
<b>PCA</b>	8.00	8.00	8.00	6.50	13.00	6.00
<b>TS-C</b>	2.00	4.00	10.50	18.00	12.00	10.00
<b>TS-PCA</b>	2.00	6.00	10.00	14.00	21.00	30.00
<b>cICA</b>	16.00	26.50	26.00	18.50	23.00	12.50

**Table F.9:** Comparison of mean QT-interval error of different extraction methods at varying SNR levels.

### F.2 T/QRS-ratio error

Method/SNR	No noise	12 dB	9 dB	6 dB	3 dB	0 dB
<b>JADEICA</b>	21.88	6.78	21.03	19.61	12.57	5.95
<b>PCA</b>	6.85	6.94	7.67	5.29	6.27	4.95
<b>TS-C</b>	1.82	1.79	2.42	2.78	3.78	8.30
<b>TS-PCA</b>	0.89	0.81	1.18	3.16	1.38	8.05
<b>cICA</b>	9.46	8.19	9.22	4.84	10.49	5.41

**Table F.10:** Comparison of mean T/QRS ratio error of different extraction methods at varying SNR levels.

### F.3 RMSE

Method/SNR	No noise	12 dB	9 dB	6 dB	3 dB	0 dB
<b>JADEICA</b>	0.25	0.23	0.23	0.23	0.21	0.22
<b>PCA</b>	0.22	0.22	0.22	0.18	0.23	0.19
<b>TS-C</b>	0.04	0.05	0.06	0.09	0.12	0.21
<b>TS-PCA</b>	0.03	0.04	0.05	0.08	0.11	0.19
<b>cICA</b>	0.27	0.29	0.27	0.24	0.29	0.34

**Table F.11:** Comparison of mean RMSE of different extraction methods at varying SNR levels.

# NON-INVASIVE FETAL ECG USING CONSTRAINED ICA

Rasmus Gundorff Sæderup

Signal Processing & Computing

June 2018



Aalborg University  
Department of Electronic Systems  
Fredrik Bajers Vej 7  
9220 Aalborg Øst