UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA DE TELEINFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE TELEINFORMÁTICA
MESTRADO ACADÊMICO EM ENGENHARIA DE TELEINFORMÁTICA

LUCAS DE SOUZA ABDALAH

NONINVASIVE ANALYSIS OF PERSISTENT ATRIAL FIBRILLATION COMPLEXITY DURING CATHETER ABLATION: A TENSOR APPROACH

FORTALEZA
2023

LUCAS DE SOUZA ABDALAH

NONINVASIVE ANALYSIS OF PERSISTENT ATRIAL FIBRILLATION COMPLEXITY DURING CATHETER ABLATION: A TENSOR APPROACH

Dissertação apresentada ao Curso de Mestrado Acadêmico em Engenharia de Teleinformática do Programa de Pós-Graduação em Engenharia de Teleinformática do Centro de Tecnologia da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Engenharia de Teleinformática. Área de Concentração: Sinais e Sistemas

Orientador: Prof. Dr. Walter da C. Freitas Jr

Coorientador: Prof. Dr. Vicente Zarzoso

FORTALEZA

2023

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Sistema de Bibliotecas
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

A116n Abdalah, Lucas de Souza.
Noninvasive Analysis of Persistent Atrial Fibrillation Complexity During Catheter Ablation: A Tensor Approach / Lucas de Souza Abdalah. – 2023.
77 f. : il. color.

Dissertação (mestrado) – Universidade Federal do Ceará, Centro de Tecnologia, Programa de Pós-Graduação em Engenharia de Teleinformática, Fortaleza, 2023.
Orientação: Prof. Dr. Walter da Cruz Freitas Júnior.
Coorientação: Prof. Dr. Vicente Zarzoso Gascón-Pelegri.

1. Atrial fibrillation. 2. block term tensor decomposition. 3. electrocardiogram. 4. electrophysiological complexity. 5. tensor. I. Título.

CDD 621.38

LUCAS DE SOUZA ABDALAH

NONINVASIVE ANALYSIS OF PERSISTENT ATRIAL FIBRILLATION COMPLEXITY DURING CATHETER ABLATION: A TENSOR APPROACH

Dissertação apresentada ao Curso de Mestrado Acadêmico em Engenharia de Teleinformática do Programa de Pós-Graduação em Engenharia de Teleinformática do Centro de Tecnologia da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Engenharia de Teleinformática. Área de Concentração: Sinais e Sistemas

Aprovada em: 20 de julho de 2023

BANCA EXAMINADORA

Prof. Dr. Walter da C. Freitas Jr (Orientador)
Universidade Federal do Ceará (UFC)

Prof. Dr. Vicente Zarzoso (Coorientador)
Université Côte d'Azur (UCA)

Prof. Dr. Rodrigo Varejão (Membro Externo)
Instituto Federal do Espírito Santo (IFES)

Prof. Dr. João César Moura Mota
Universidade Federal do Ceará (UFC)

Prof. Dr. Yuri Carvalho Barbosa Silva
Universidade Federal do Ceará (UFC)

Dedico este trabalho aos meus pais pelo apoio incondicional e às cientistas, que têm trabalhado incansavelmente para que a humanidade supere seus grilhões.

ACKNOWLEDGEMENTS

I would like to express my heartfelt gratitude to my parents, Jamil and Meiry, for their unwavering support and assistance, which greatly contributed to the completion of this work.

I would like to express my deep gratitude to my advisor, Professor Walter Freitas Jr. I am grateful for the invaluable guidance and support provided throughout the entire process of my research. His insightful feedback and unwavering dedication to my academic growth have been instrumental in shaping the quality and direction of my dissertation.

I am profoundly grateful for the invaluable guidance, support, and expertise provided by my co-advisor, Professor Vicente Zarzoso. He has given me the opportunity to develop various skills, both on a personal and professional level. I am deeply grateful for his immense guidance and support throughout my studies in this field. From my debut into the research field, he has been a great mentor, whose expertise and patience have played a crucial role in the success of my internship, master's program, and now my dissertation.

I would like to express my heartfelt gratitude to my dear friend Pedro Marinho, who has been my internship co-advisor. Throughout this journey, his support, encouragement, and companionship have been invaluable to me. His constructive feedback have played a significant role in shaping both my research and personal growth.

To the examining committee of this work, I extend my sincere appreciation for the time dedicated to reading this dissertation and for providing relevant comments.

I would also like to extend my heartfelt appreciation to my friends and fellow university colleagues who have been by my side throughout this journey, offering unconditional friendship and support. Special thanks go to Vitoria, Yvo Menezes, Kenneth Brenner and Fatima (*in memorian*).

To my fellow musicians, who have encouraged me during difficult times and understood my absence as I dedicated myself to this work, especially Miguel Abdala, I extend my deepest gratitude.

I wish to express my sincere gratitude to the Ceará Foundation of Support to Scientific and Technological Development (FUNCAP) for their generous support in the form of an MSc scholarship.

"Quando a educação não é libertadora, o
sonho do oprimido é ser o opressor."
(FREIRE, 2003, p. 47)

ABSTRACT

Atrial fibrillation (AF) is the most common cardiac arrhythmia encountered in clinical practice, and it is estimated to be responsible for one-fourth of cerebrovascular accidents. The persistent form of this arrhythmia is a complex case characterized by uncoordinated and irregular cardiac activation. Although catheter ablation is increasingly used due to its lower recurrence rate compared to other treatments, widely accepted intervention protocols by the cardiology community are still being established. Surface electrocardiogram (ECG)-based analysis is relevant due to its low cost and non-invasive nature. The development of new mathematical tools to characterize AF through ECG can improve intervention guidance and increase success rates, reducing the duration of arrhythmia and the risk of complications. However, despite the growing interest in these methodologies to evaluate the complexity of persistent AF signals, their performance is still limited. To overcome these limitations, this study proposes the application of tensor decomposition techniques to quantify the complexity of ECG signals during catheter ablation procedures for AF treatment. Tensor decompositions are powerful signal processing tools. However, their application in AF signal analysis is recent. The Constrained Alternating Group Lasso (CAGL) algorithm was developed to simultaneously calculate the block term tensor decomposition (BTD) into block terms and estimate its parameters (number and rank of the blocks), for the multilinear rank-$(L_r, L_r, 1)$ particular case. This algorithm showed promise in extracting atrial activity during AF and estimating its complexity. To evaluate its performance, we compare its results with the Nondipolar Component Index (NDI) method in the context of ablation, which demonstrates the multilinear advantages over matrix methods in non-invasively extracting and quantifying AF complexity. The tensor index correlates with the reduction in AF complexity throughout the ablation steps as instinctively expected. Additionally, it has presented a significant negative correlation with AF recurrence episodes, which presents clear clinical interest, since it can assist in the development of new medical protocols.

Keywords: Atrial fibrillation. block term tensor decomposition. electrocardiogram. electrophysiological complexity. tensor.

RESUMO

A fibrilação atrial (FA) é a arritmia cardíaca mais comum encontrada na prática clínica, e estima-se que seja responsável por um quarto dos acidentes vasculares cerebrais. A forma persistente da arritmia é um caso complexo, caracterizado por ativação cardíaca descoordenada e irregular. Embora a ablação por cateter seja cada vez mais utilizada devido à sua baixa taxa de recorrência em comparação a outros tratamentos, ainda não há protocolos de intervenção amplamente aceitos pela comunidade cardiológica. As análises baseadas em eletrocardiograma (ECG) de superfície são relevantes devido ao baixo custo e à natureza não invasiva desses exames. O desenvolvimento de novas ferramentas matemáticas para caracterizar a FA por meio do ECG pode melhorar a orientação das intervenções e aumentar a taxa de sucesso, reduzindo a duração da arritmia e o risco de complicações. No entanto, apesar do crescente interesse nessas metodologias para avaliar a complexidade do sinal de FA persistente, seu desempenho ainda é limitado. Para superar essas limitações, este estudo propõe a aplicação de técnicas de decomposição tensorial para quantificar a complexidade dos sinais de ECG durante o procedimento de ablação por cateter no tratamento da FA. As decomposições tensoriais são ferramentas poderosas para análise de dados, mas sua aplicação na análise de sinais de FA é recente. O algoritmo Constrained Alternating Group Lasso (CAGL) foi desenvolvido para calcular simultaneamente a decomposição tensorial em blocos de termos (block term tensor decomposition , BTD) e estimar seus parâmetros (número e posto multilinear dos blocos), para o caso particular do posto multilinear $(L_r, L_r, 1)$. Esse algoritmo mostra-se promissor na extração da atividade atrial durante a FA e na estimativa de sua complexidade. Para avaliar o desempenho do CAGL, o método Nondipolar Component Index (NDI) é aplicado no contexto da ablação, demonstrando as vantagens das técnicas multidimensionais na extração e quantificação da complexidade da FA de forma não invasiva. O índice tensorial se correlaciona com a redução da complexidade da FA ao longo das etapas de ablação, conforme instintivamente esperado. Além disso, apresentou uma correlação negativa significativa com episódios de recorrência da FA, o que o torna clinicamente relevante, pois pode auxiliar no desenvolvimento de novos protocolos médicos.

Palavras-chave: Complexidade eletrofisiológica. decomposição tensorial em termo de blocos. eletrocardiograma. fibrilação atrial. tensor

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AA	Atrial Activity
ADMM	Alternating Direction Method of Multipliers
AF	Atrial Fibrillation
AGL	Alternating Group Lasso
ALS	Alternating Least Squares
AV	Atrioventricular
BSS	Blind Source Separation
BTD	Block Term Decomposition
CA	Catheter Ablation
CAGL	Constrained Alternating Group Lasso
DF	Dominant Frequency
ECG	Electrocardiogram
ICA	Independent Components Analysis
LA	Left Arm
LL	Left Leg
LS	Least Squares
NDI	Non-Dipolar Component Index
NLS	Nonlinear Least-Squares
PC	Principal Component
PCA	Principal Components Analysis
PV	Pulmonary Vein
PVI	Pulmonary Vein Isolation
RA	Right Arm
RL	Right Leg
SA	Sinoatrial
SC	Spectral Concentration
SVD	Singular Value Decomposition
VA	Ventricular Activity

CONTENTS

1 INTRODUCTION

This chapter introduces the topics and structure of this dissertation. In Section 1.1 we present the motivation behind our line of work. In Section 1.2 we provide a brief overview of previous work in noninvasive complexity analysis tools. In Section 1.3 we describe the dissertation objectives as well as we highlight our main contributions. Finally, Section 1.4 presents the structure of this work.

1.1 Motivation

Atrial Fibrillation (AF) is the most common sustained arrhythmia in clinical practice, with a complex and not completely understood pathophysiology. It can lead to harmful consequences such as blood clots, strokes, and, in severe cases, even death.

Given the uncertainty surrounding the mechanisms that trigger AF and the increasing number of affected individuals due to the aging world population, this heart disease has become a significant topic in public health discussions. There is a need for more effective and standardized protocols, supported by physiological signal analysis, to enhance the understanding and management of the atrial substrate. The atrial substrate plays a critical role in the pathophysiology of AF, influencing the stability, duration, and progression of the arrhythmia.

The primary objective is to improve the understanding and management of AF providing precise and relevant information using low-cost tools such as the surface Electrocardiogram (ECG), which is widely used in cardiology.

However, this task is both complex and costly. Mathematical models based on matrix and tensor decompositions have emerged to support this endeavor. These models are applied to characterize the physiological nature of the phenomena observed in signals. They allow the assessment of Atrial Activity (AA) complexity evolution of persistent AF and the investigation of the impact of therapies aimed at mitigating the effects of fibrillation. Ultimately, they may assist in the development of new protocols and support clinical decisions with minimal invasive risk.

1.2 Related Work

Previous studies have demonstrated the potential of arranging an ECG signal as a matrix, where each row contains the time samples of one lead. It allows us to estimate the sources with a Blind Source Separation (BSS) model. Singular Value Decomposition (SVD), Principal Components Analysis (PCA), and Independent Components Analysis (ICA) are the most celebrated techniques to perform the decomposition and may be effective in extracting meaningful information from ECG data and contribute to AF analysis (RIETA *et al.*, 2004; CASTELLS *et al.*, 2005; OLIVEIRA; ZARZOSO, 2019). However, they rely on strong mathematical constraints to assure decomposition uniqueness, such as mutual orthogonality between spatial factors, sufficient spectral diversity, and statistical independence (ZARZOSO, 2009).

To overcome these limitations, the application of tensor approaches to perform BSS has gained attention in the literature, as the tensor decompositions outperform matrices, with uniqueness ensured under much milder constraints. The Block Term Decomposition (BTD) has been shown as an effective method to estimate sources based on exponential models (LATHAUWER, 2011). Furthermore, the AA during AF can be modeled as a sum of complex exponentials, in a suitable manner to decompose the ECG signals into distinct components, with at least one characterizing the atrial substrate. Still, the Nonlinear Least-Squares (NLS) approach, acclaimed to compute BTD (SORBER *et al.*, 2013), imposes prior knowledge about the tensor structure, i.e, number of blocks and its multilinear rank. In scenarios where the structure of the estimated data can be controlled, this task does not impose as much difficulty as in the case of biologic signals, e.g., wireless communications (FREITAS *et al.*, 2019; HAN *et al.*, 2021). In biomedical signal processing field, usually the measured data are the only knowledge to start from, and they are already mixed and influenced by other physiological processes, artifacts, noise, etc. Hence, the estimation of these tensor parameters for biomedical signals becomes a new challenge to overcome (RIBEIRO *et al.*, 2015; ZARZOSO, 2017; OLIVEIRA; ZARZOSO, 2019).

The Alternating Group Lasso (AGL) algorithm and its constrained version, Constrained Alternating Group Lasso (CAGL), present a novel approach for computing approximate low-rank BTD with little information about the tensor parameters (GOULART *et al.*, 2020; RONTOGIANNIS *et al.*, 2021). It has been shown to be effective in simultaneously estimating AA and measuring its complexity (ABDALAH *et al.*, 2020).

Finally, taking into account the CAGL performance, the present work explore its advantages to perform the source separation and assess the AA complexity and to compare with other index proposed in the literature.

1.3 Objective and Contribution

This work proposes a discussion on the foundations of AA analysis during AF episodes. In addition, we discuss the foundations of signal processing based on tensor models and their advantages over matrix approaches, with a special focus on BSS context.

The present dissertation will focus in two main topics: persistent AF treated with stepwise-Catheter Ablation (CA) approach, detailed in the chapter 2. This theory will serve as a support for the analysis of atrial substrate modification, as well as the basis for other strategies aiming to mitigate the fibrillatory repercussions and understand atrial activation in general and reentry mechanisms. Moreover, we aim to present a study to:

- Investigate the behavior of Hankel matrices with real data to provide more evidence for their use in exponential models;
- Reinforce the effectiveness of CAGL in jointly estimating AA and quantifying its complexity from short ECG segments;
- Compare with the Non-Dipolar Component Index (NDI) (MEO *et al.*, 2018), a state-of-the-art complexity index;
- Explore the statistical correlation between the complexity indices and the response to CA therapy in a dataset comprising 20 patients with persistent AF.

During the development of this work, the following scientific contributions has been accepted and/or:

- L. Abdalah, P. M. R. Oliveira, W. Freitas Jr., V. Zarzoso, "Tensor-based noninvasive atrial fibrillation complexity index for catheter ablation", *Computing in Cardiology*, Rimni, Italy, Sep. 13-16, 2020.
- V. Zarzoso, L. Abdalah, P. M. R. Oliveira, "Décomposition tensorielle en termes blocs contrainte pour la mesure non invasive de la complexité de la fibrillation atriale persistante", *XXIXème Colloque Francophone de Traitement du Signal et des Images (GRETSI 2023)*, Grenoble, France, Sep. 2023. Final version submitted, accepted for publication and presentation.

1.4 Dissertation Structure

The rest of the manuscript is organized as follows: Chapter 2 presents the electrophysiological background and an overview of AF perspectives in western countries and Brazil. Chapter 3 presents mathematical notations and describes the tensor-based decomposition to estimate AA. Chapter 4 recalls the concept of AF complexity and its importance in characterizing CA therapy, presented in Section 2.2.3. The proposed tensor-based index to measure AF complexity and the state-of-the-art matrix-based index are also presented. Chapter 5 presents the database setup, experimental and statistical results with real AF ECG recordings, as well as a discussion on the advantages and drawbacks of the proposed method. Finally, Chapter 6 concludes the manuscript and provides some perspectives for future research on the topic.

2 ELETROPHYSIOLOGICAL BACKGROUND

A healthy heartbeat is the result of a complex and coordinated electrical and mechanical process within the heart. It functions as a pump for the human body. During the cardiac cycle, the atria relax and blood enters them, while the ventricles receive the blood from the atria. These phases are known as diastole and systole, respectively. In a healthy heart, these contractions and relaxations occur in a regular pattern called sinus rhythm. The sinus rhythm originates from the Sinoatrial (SA) or sinus node, often referred to as the "biological pacemaker" (AMBESH; KAPOOR, 2017). This node consists of a small cluster of cells that spontaneously and rhythmically initiate electrical impulses, which then propagate to the Atrioventricular (AV) node. As the atria contract, they propel blood into the ventricles. Subsequently, the ventricles contract and pump blood out of the heart as the electrical signals reach the ventricular muscle cells. After contraction, the ventricles go through a repolarization phase during which the electrical signals return to their resting state. This prepares the heart for the next heartbeat cycle. A visual representation of these phenomena will be presented in the subsequent sections (see Fig. 2.4).

2.1 Standard 12-Lead ECG Mechanisms

The standard 12-lead ECG provides a comprehensive assessment of the heart's electrical activity from different perspectives, allowing medical professionals to diagnose and monitor a wide range of cardiac conditions. It consists of recording the electrical signals generated by the heart from 12 different perspectives or "leads" by placing 10 electrodes on the body.

An ECG recording is obtained by placing electrodes at specific locations on the patient's chest, arms, and legs. The leads are divided into two categories: limb leads and precordial leads. The limb leads measure the electrical activity between the limbs, while the precordial leads measure the electrical activity in the horizontal plane of the chest (DUPRE *et al.*, 2005). In the recordings a series of waveforms known as the PQRST complex are observed. They represent the electrical activity of the heart during one cardiac cycle. The P wave corresponds to atrial depolarization, the QRS complex represents ventricular depolarization, and the T wave represents ventricular repolarization. These waveforms provide important information about the timing, duration, and morphology of

the cardiac electrical activity, allowing for the identification of various cardiac conditions and abnormalities (MAINARDI *et al.*, 2008a).

Interpreting the ECG recordings involves analyzing the shape, duration, and amplitude of the waveforms in each lead, as well as assessing the relationships between the different leads. Various parameters, such as heart rate, PR interval, QRS duration, and QT interval, are measured and analyzed to assess the overall cardiac function and detect any abnormalities or arrhythmias.

It may summarized as a fundamental tool in diagnosing and monitoring various cardiac conditions, including myocardial infarction, arrhythmias, conduction abnormalities, and electrolyte imbalances. It provides valuable insights into the electrical activity of the heart and aids in guiding appropriate treatment strategies for patients.

2.1.1 The Electrodes

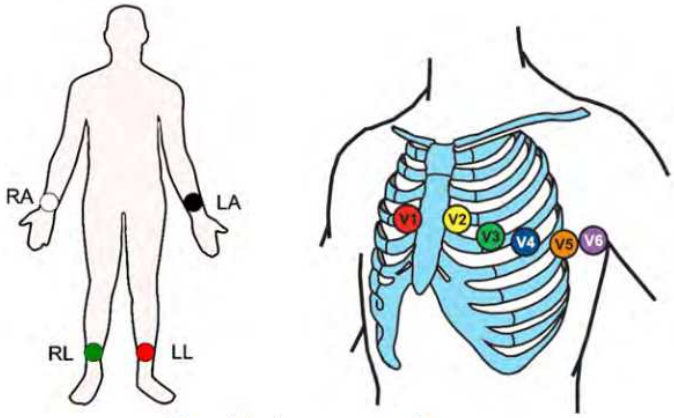
The electrodes used in an ECG typically consist of a conductive material, such as silver/silver chloride or carbon, enclosed in a gel or adhesive pad. The conductive material helps to pick up the electrical signals generated by the heart.

Before attaching the electrodes, the skin needs to be properly prepared to ensure good conductivity. The areas where the electrodes will be placed are cleaned to remove any oils, lotions, or dirt that could interfere with the electrical signal transmission. This is done by gently rubbing the skin with an alcohol swab and then allowing it to dry. Once the skin is prepared, the electrodes are attached to the designated locations. The adhesive pads or gel help to secure the electrodes in place and ensure good contact with the skin. It is important to ensure that the electrodes are firmly attached but not too tight to avoid discomfort for the patient.

The electrodes detect the electrical signals generated by the heart, which are very weak in magnitude. These signals travel through the conductive material of the electrodes and are transmitted to the ECG monitor through lead wires. The device then processes and records the waveforms. Each electrode is connected to a specific lead wire, which corresponds to a specific lead configuration (e.g., limb leads or chest leads).

The limb leads are obtained by placing electrodes on the limbs of the patient and they serve as the positive or negative terminals to these leads. Specifically, there are four limb electrodes placed on the patient's arms and legs, as illustrated in Fig. 2.1:

Figure 2.1 – 12 leads resting ECG electrode placement.



RA – right forearm or wrist
LA – left forearm or wrist
LL – left lower leg, proximal to ankle
RL – right lower leg, proximal to ankle
V1 – 4-th intercostal space, right sternal edge
V2 – 4-th intercostal space, left sternal edge
V3 – midway between V2 and V4
V4 – 5-th intercostal space, mid-clavicular line
V5 – anterior axillary line in straight line with V4
V6 – mid-axillary line in straight line with V4 and V5

Source: Norav Users Guide Personal Computer ECG 1200 (NORAV MEDICAL, 2021).

- Right Arm (RA) electrode: This electrode is placed on the right arm, usually on the wrist or lower part of the forearm.
- Left Arm (LA) electrode: This electrode is placed on the left arm, also on the wrist or lower part of the forearm.
- Right Leg (RL) electrode: This electrode is placed on the right leg, typically on the ankle or lower part of the leg.
- Left Leg (LL) electrode: This electrode is placed on the left leg, similar to the right leg electrode.

The three bipolar limb leads, known as Lead I, Lead II, and Lead III, measure the voltage between two limb electrodes. Lead I records the potential between the RA and the LA electrodes. Lead II measures the potential between the RA and the LL electrodes. Lead III records the potential between the LA and the LL electrodes. In addition, there are also three augmented unipolar limb leads: aVR, aVL, and aVF. These leads measure the potential between one limb electrode and a central reference point formed by the average of the other two limb electrodes. The aVR lead measures the potential between the RA electrode and the average of the LA and LL electrodes. The aVL lead measures the potential between the LA electrode and the average of the RA and LL electrodes. The

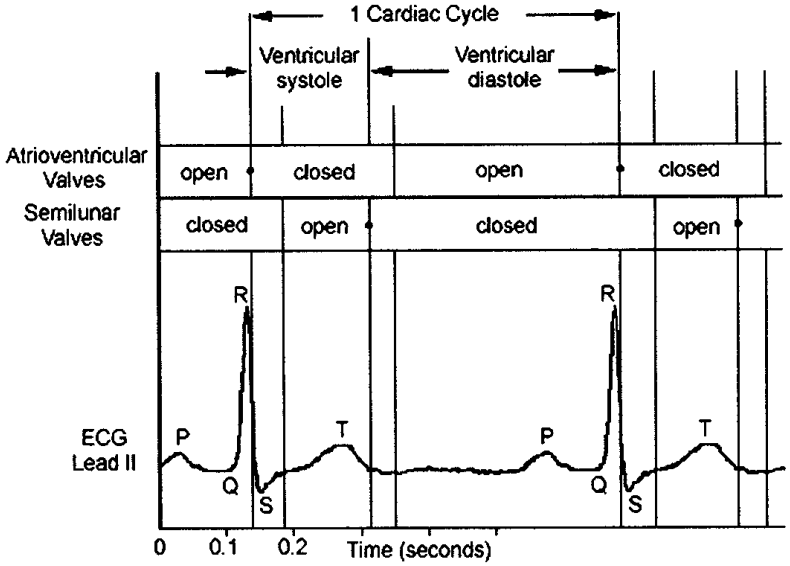
aVF lead measures the potential between the LL electrode and the average of the RA and LA electrodes.

The precordial leads V1-V6 are placed on specific locations on the chest wall (Fig. 2.1), providing a more detailed view of the electrical activity in the horizontal plane of the heart (DUPRE *et al.*, 2005). They are commonly referred to as "chest leads," and their measurements are related to the baseline established by the limb leads and augmented limb leads.

2.1.2 Electrical Signal Representation

The electrical signals obtained by combining the electrodes are recorded as voltage waveforms over time, resulting in 12 leads. These detected signals are very weak and require amplification before they can be analyzed. The recording machine amplifies, and filter out unwanted noise and interference. Fig. 2.2 illustrates how the activity of the heart is reflected in the lead II signal.

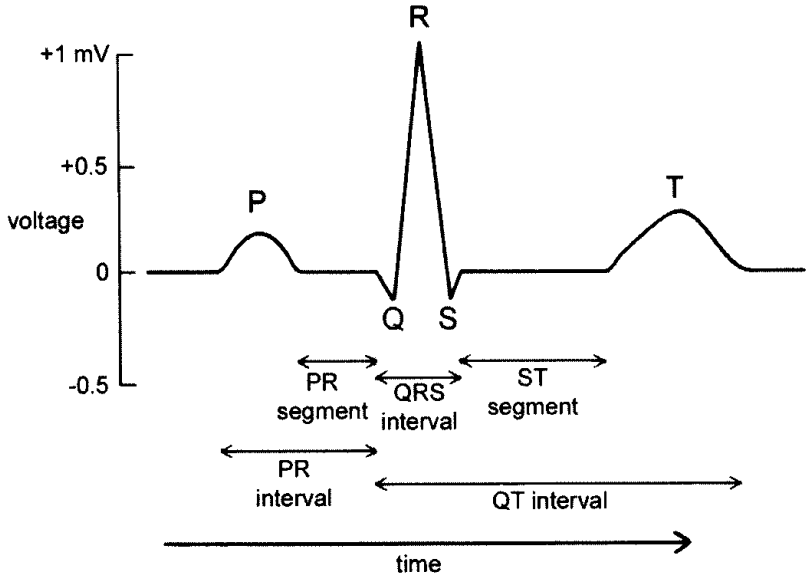
Figure 2.2 – A typical waveform for the lead II is compared to the timing of AV and valve activity, along with which segments of the cardiac cycle the ventricles are in systole/diastole.



Source: Anthony Dupre, 2005, Basic ECG Theory, Recordings, and Interpretation (DUPRE *et al.*, 2005).

These voltage waveforms represent the depolarization and repolarization of different regions of the heart during each cardiac cycle. The ECG waveform consists of several key components presented in Fig. 2.3. The main components of the cardiac cycle can be described as follows:

Figure 2.3 – A typical ECG waveform for one cardiac cycle measured from lead II. The P wave denotes atrial depolarization, the QRS indicates ventricular depolarization, and the T wave denotes ventricular repolarization. The events on the waveform occur on a scale of hundreds of milliseconds.



Source: Anthony Dupre, 2005, Basic ECG Theory, Recordings, and Interpretation (DUPRE *et al.*, 2005), modified from D.E. Mohrman and L.J. Heller, (eds.), Cardiovascular Physiology, 5th Ed., McGraw-Hill Companies, 2003.

- P Wave: Represents the depolarization of the atria, indicating the initiation of atrial contraction.
- QRS Complex: Represents the depolarization of the ventricles. It consists of three main deflections: the Q wave, the R wave (the tallest waveform), and the S wave. The QRS complex reflects the initiation of ventricular contraction.
- T Wave: Represents the repolarization of the ventricles, indicating the recovery of the ventricular muscle before the next contraction.
- PR Interval: Measures the time it takes for the electrical signal to travel from the atria to the ventricles.
- QT Interval: Measures the time between the start of ventricular depolarization (QRS complex) and the completion of ventricular repolarization (T wave). It represents the total time for ventricular depolarization and repolarization.

The presented signals are analyzed by cardiologists to assess the electrical activity of the heart and identify any abnormalities. Various aspects are examined, including the shape, duration, and intervals between different waveforms. Deviations from the normal ECG pattern can indicate cardiac conditions such as arrhythmias, ischemia, or structural abnormalities.

2.2 Atrial Fibrillation

AF is the most common sustained arrhythmia observed in clinical practice. Since its trigger and perpetuation mechanisms continue to baffle cardiologists, being referred as the last great frontier of cardiac electrophysiology. This arrhythmia, which primarily affects the elderly, is responsible for up to 1/4 of strokes (JANUARY *et al.*, 2019). The management of AF places a substantial burden on healthcare resources, including hospitalizations, medications, and interventions such as cardioversion or catheter ablation.

AF consequences are not limited to the individual level. It also has economic implications, with significant healthcare costs associated with the management of AF-related complications. The impact on healthcare systems is expected to increase as the population ages and the prevalence of AF continues to rise. Projections indicate that by 2050, between 6 and 12 million individuals in the USA and up to 18 million in Europe alone by 2060 will be affected by AF (MORILLO *et al.*, 2017).

The present work focuses on persistent AF, though it is important to note that this is only one form of AF manifestation. There is no consensus on the best way to classify atrial fibrillation. However, we present the most common classification schemes (MAINARDI *et al.*, 2008a; CALKINS *et al.*, 2017):

- **Paroxysmal AF** refers to AF that terminates spontaneously or with intervention within 7 days of onset.
- **Persistent AF** refers to AF that is sustained beyond 7 days.
- **Permanent AF** refers to AF where no attempts to interrupt it have been made, or if made, they were unsuccessful (i.e., failed restoration of sinus rhythm or recurrence of AF). In such cases, no further attempts to restore or maintain sinus rhythm will be undertaken.

The management of persistent AF typically involves a combination of strategies aimed at controlling the heart rate, restoring and maintaining normal sinus rhythm, and reducing the risk of complications associated with AF. Some commonly used therapies and protocols (JANUARY *et al.*, 2019; FAVARATO, 2021):

1. Antiarrhythmic drugs: they may be prescribed to control the heart rate and rhythm. Examples include beta-blockers, calcium channel blockers, and antiarrhythmic medications such as amiodarone, flecainide, or propafenone;
2. Anticoagulation therapy: AF increases the risk of blood clot formation, which can

lead to stroke. To mitigate this risk, anticoagulant medications such as warfarin or direct oral anticoagulants (e.g., apixaban, dabigatran) may be prescribed;

3. Electrical cardioversion: This procedure involves delivering a controlled electric shock to the heart to restore normal sinus rhythm. It is often performed under sedation or anesthesia.
4. Catheter ablation: This is a minimally invasive procedure where catheters are guided to the heart to identify and selectively destroy abnormal tissue responsible for the AF. Radiofrequency energy or cryotherapy may be used to create lesions and restore normal electrical conduction (HAïSSAGUERRE *et al.*, 1998; VERMA *et al.*, 2015; SEITZ *et al.*, 2017).

It's important to note that the choice of therapy or protocol depends on various factors, including the patient's overall health, the duration and severity of AF, presence of underlying heart conditions, and individualized assessment by a healthcare professional. Efforts are being made to raise awareness about AF, improve detection and diagnosis, and develop effective management strategies (VERMA *et al.*, 2015; ROTTNER *et al.*, 2020). Early detection, appropriate treatment, and lifestyle modifications can help manage AF and reduce the associated risks. Public health initiatives and research efforts aim to address the challenges posed by AF and improve patient outcomes.

2.2.1 Endemic In Brazil

AF is a significant concern in Brazil, as it is in many countries worldwide. Studies have shown that the prevalence of AF has been steadily increasing over the years, primarily due to factors such as an aging population, lifestyle changes, and improved detection methods.

In the Brazilian population, the overall prevalence of AF is estimated to be 1.8%, with higher rates of 8.4% among octogenarians and 11.0% among nonagenarians (MARCOLINO *et al.*, 2015). AF prevalence increases significantly with age, particularly in older adults. Other studies reinforce the exponential increase in AF with advancing age, with a 5-fold increase in the age range of 50-59 years, a 7-fold increase in the age range of 60-69 years, and a 9-fold increase in individuals over 80 years old. It is worth noting that men tend to have a higher incidence of AF (FAVARATO, 2021).

The impact of AF in South America is multifaceted. While the prevalence of

AF in the region is similar to that in high-income countries, with hypertension being the most commonly associated disease (SANTOS *et al.*, 2021), there is a specific endemic problem in the region: Chagas disease. Studies have shown that Chagas disease has the strongest association with AF in the Brazilian scenario (MARCOLINO *et al.*, 2015).

The current protocols for managing AF primarily focus on assessing the risk of thromboembolism and implementing appropriate anticoagulation strategies. For symptomatic individuals, additional approaches such as rhythm control and aggressive modification of cardiovascular risk factors are considered.

In recent years, advancements in technology, such as improved diagnostic tools and therapeutic options, have provided new opportunities for managing AF. These include the use of catheter ablation techniques, novel medications, and the development of digital health solutions for remote monitoring and follow-up care.

Overall, addressing the challenges posed by AF in Brazil requires a comprehensive approach that includes preventive measures, early detection, access to appropriate healthcare services, and ongoing management to improve patient outcomes and reduce the burden of the condition on individuals, families, and the healthcare system.

2.2.2 AF Electrophysiology

AF is characterized by rapid and irregular electrical activity in the atria of the heart. It is caused by abnormal electrical impulses originating from multiple locations within the atria, rather than the usual single location (the SA node) in a healthy heart (MAINARDI *et al.*, 2008a). This abnormal electrical activity disrupts the normal coordination and timing of the heart's electrical signals and mechanical contractions, resulting in the following effects:

1. Irregular Electrical Signal Initiation: The atria experience rapid and disorganized electrical signals firing from multiple sources. These abnormal signals can arise from various areas within the atrial tissue, such as the Pulmonary Veins (PVs) or other atrial regions. As a result, the atrial muscle fibers do not contract in a coordinated manner, leading to ineffective atrial contractions or atrial quivering (fibrillation).
2. Irregular AV Node Conduction: The irregular electrical signals generated in the atria reach the AV node, which serves as a gatekeeper between the atria and the ventricles. The AV node attempts to filter and regulate the incoming signals before

allowing them to pass to the ventricles. However, due to the chaotic and disorganized nature of the electrical impulses in AF, the AV node may conduct some signals to the ventricles while blocking others. This results in an irregular ventricular response.

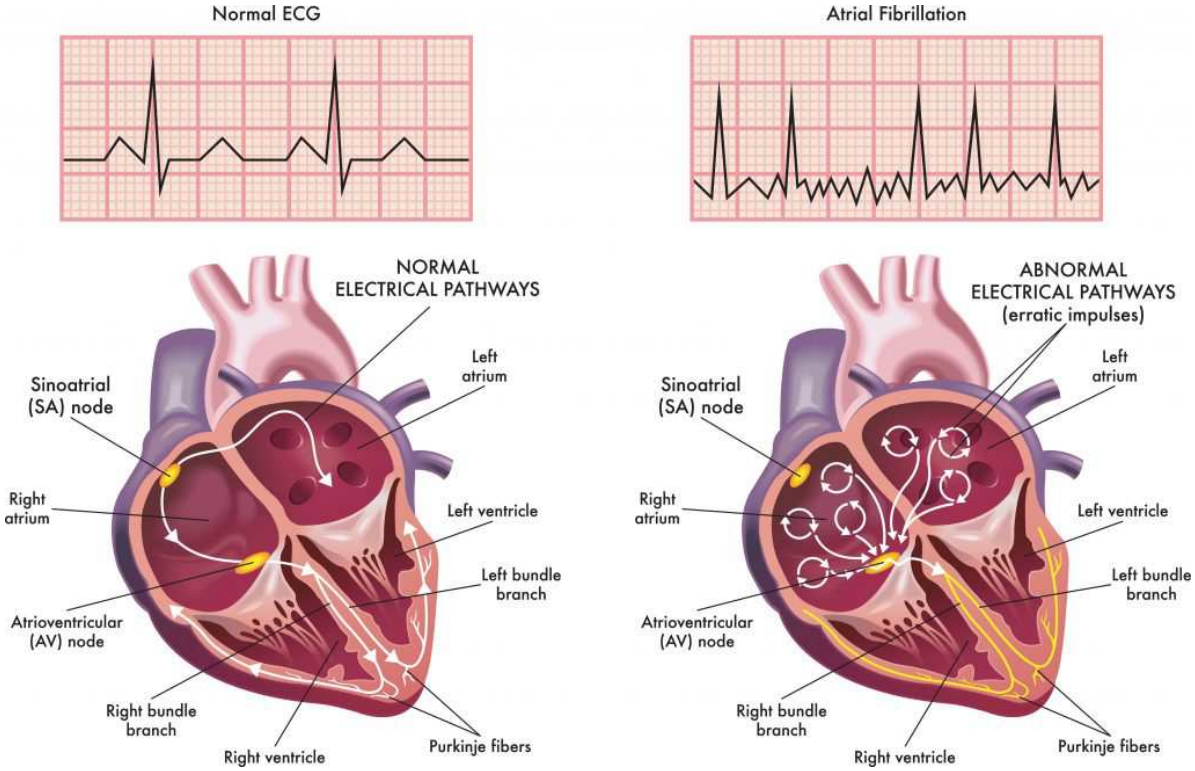
3. Irregular Ventricular Contraction: The irregularly conducted electrical signals from the atria cause the ventricles to contract in an irregular pattern. The ventricles may beat at a rapid, irregular, or both rapid and irregular rate. As a result, the pulse may feel irregular, and individuals with AF often describe the sensation of a "fluttering" heartbeat.
4. Increased Risk of Blood Clots: AF disrupts the normal blood flow within the atria. The chaotic and inefficient contractions of the atria can cause blood to pool or stagnate in certain areas. Stagnant blood increases the risk of blood clot formation within the atria. If a blood clot forms and becomes dislodged, it can travel through the bloodstream and potentially cause a blockage in a smaller blood vessel elsewhere in the body. This can result in serious complications, such as stroke, organ damage, or other complications depending on the affected area.
5. Persistent AF: Persistent AF refers to the manifestation of AF where the irregular electrical activity in the atria is sustained for a prolonged period, lasting longer than seven days (CALKINS *et al.*, 2017). In persistent AF, the abnormal electrical signals persist, and the atria fail to restore their normal coordinated contractions. This may require medical intervention to restore normal sinus rhythm.

From a signal processing perspective, AF during AA is characterized by the presence of irregular and chaotic signals. ECG recordings during AF show rapid, irregular, and often fibrillatory waves known as f waves (MEO *et al.*, 2018), which indicate the disorganized electrical activity in the atria (see Fig. 2.4).

The phenomena of AF shown in Figs. 2.4 and 2.5, compared to the healthy cardiac cycle shown in Figs. 2.2 and 2.3, exhibit four key aspects that differ from a normal heartbeat: irregular rhythm, absence of P waves, chaotic f waves, and rapid ventricular response.

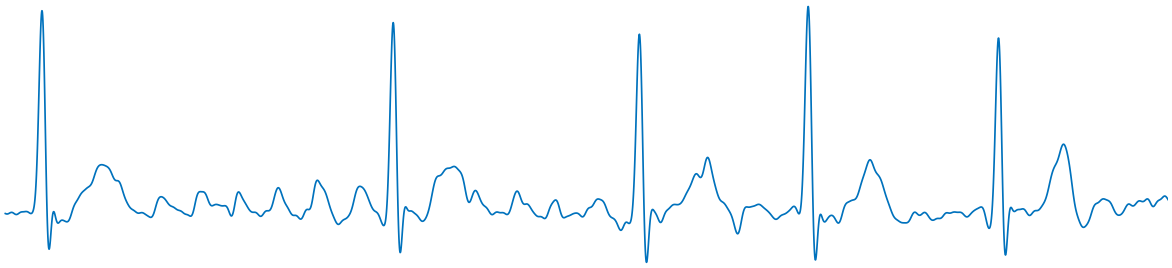
1. Irregular Rhythm: The normal regularity and pattern of the R-R intervals (the time interval between consecutive R-peaks in the ECG) observed in a healthy heart are lost during AF. Instead, the R-R intervals become highly irregular, reflecting the irregularity in atrial depolarization and subsequent ventricular response.

Figure 2.4 – Comparison of two ECG signals in lead V5: a healthy cardiac cycle (left) vs. AF (right). The waveforms demonstrate distinct electrical patterns associated with each condition. A simplified representation of the heart's structure responsible for generating each electrical activity is presented below each signal.



Source: Atrial Fibrillation - About the Heart (StopAfib.org - American Foundation for Women's Health, 2021)

Figure 2.5 – Example of AF: five cardiac cycles presenting fibrillation events in lead V5 from a standard 12-lead ECG.



Source: Elaborated by the author.

2. Absence of P waves: In a normal sinus rhythm, each heartbeat is preceded by a P wave on the ECG, representing atrial depolarization. In AF, the P waves are typically absent or obscured due to the disorganized atrial electrical activity. Instead, rapid and irregular oscillations are observed.
3. Chaotic Atrial Fibrillation Waves: The f waves seen on the ECG during AF are

characterized by irregular amplitude, frequency, and morphology. These fibrillation waves represent the superposition of multiple irregular electrical signals originating from different regions of the atria. The irregularity and variability in the shape and timing of the f waves indicate the disorganized AA during AF.

4. Rapid Ventricular Response: The irregular AA during AF affects the conduction of electrical signals to the ventricles. As a result, the ventricular response can be rapid and irregular, leading to a high heart rate. This rapid ventricular response is reflected in the irregularity of the R-R intervals on the ECG.

Signal processing techniques are used to analyze and interpret these irregularities. Various algorithms and methods, such as time-domain analysis, frequency-domain analysis, and nonlinear dynamics analysis, can be applied to quantify the irregularity, assess the complexity of the signals, and extract relevant features. These techniques contribute to the understanding of AF mechanisms, assist in diagnosis, guide treatment strategies, and facilitate the development of innovative approaches for AF management, as explored in Chapter 3.

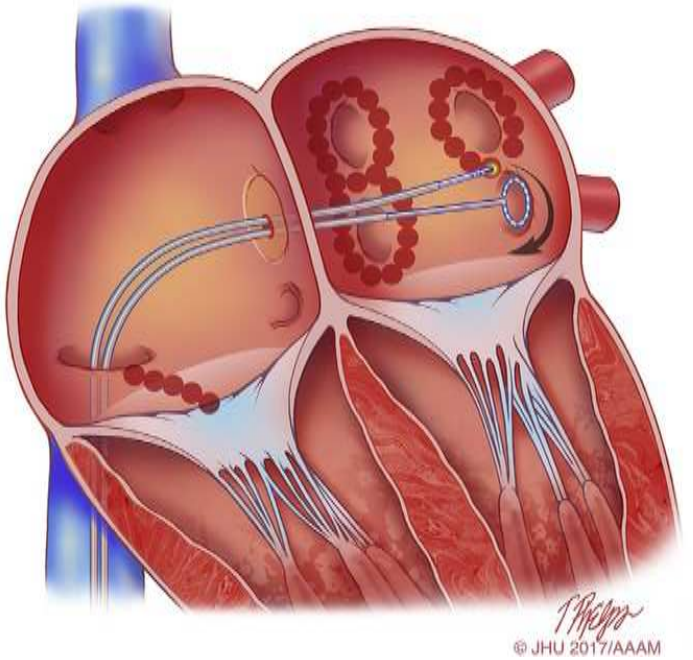
2.2.3 Catheter Ablation

Catheter ablation is currently the most effective treatment option for persistent AF, despite its cost and potential complications. This therapy involves using intracardiac catheters inserted through a blood vessel in the groin to guide them towards the heart and locate the tissue responsible for the arrhythmia. Radiofrequency energy is then applied to these specific points in the atrial tissue. The goal of catheter ablation is to terminate AF and restore sinus rhythm (cardioversion) or convert it into a more organized arrhythmia.

The widely adopted approach, known as the “sequential stepwise” procedure, is followed in most clinical centers that perform ablation (VERMA *et al.*, 2015; SEITZ *et al.*, 2017). This procedure begins with the Pulmonary Vein Isolation (PVI) technique, which electrically isolates the PVs (HAïSSAGUERRE *et al.*, 1998; CALKINS *et al.*, 2017). The PVs often contain ectopic foci triggering AF (Fig. 2.6). Subsequently, attention is directed towards the atrial substrate, targeting conduction zones responsible for the abnormal propagation patterns of AF.

The literature suggests that theAF becomes more organized throughout the ablation procedure (MCCANN *et al.*, 2021), resulting in a decrease in f wave complex-

Figure 2.6 – Schematic drawing showing catheter ablation of AF. Shows a typical wide area lesion set created using radiofrequency energy. Ablation lesions are delivered in a figure of eight pattern around the left and right PVs. A multielectrode circular mapping catheter is positioned in the left inferior PV.



Source: Tim Helps © 2017 Johns Hopkins University, AAM (CALKINS *et al.*, 2017).

ity (MEO *et al.*, 2018). However, the success rates and procedures can vary significantly across different centers, introducing a subjective element in the therapy. The development of robust and widely accepted intervention protocols remains an open challenge (VERMA *et al.*, 2015; ROTTNER *et al.*, 2020). This work aims to address such limitations by providing a new complexity index measurement, described in chapter 4.

3 MATHEMATICAL BACKGROUND

Signal processing techniques play a crucial role in extracting meaningful information from complex data, particularly in scenarios where multiple sources contribute to the observed signals. Matrix and tensor decompositions are powerful approaches that have gained significant attention in the field of signal processing. These techniques offer relevant tools for estimating and separating sources from mixed observations, enabling the recovery of underlying signals that would otherwise remain hidden.

In this chapter, we delve into the concept of a tensor and how its structure can be exploited to extract latent components in sophisticated scenarios. We will provide an overview of the underlying principles, algorithms, and applications of BSS, matrix factorization, and tensor decompositions.

Notation and Algebra Prerequisites: Scalars, vectors, matrices and tensors are represented by lower-case (a, b, ...), boldface lower-case ($\mathbf{a}$, $\mathbf{b}$, ...), boldface capital ($\mathbf{A}$, $\mathbf{B}$, ...) and calligraphic ($\mathcal{A}$, $\mathcal{B}$, ...) letters, respectively. The matrix transpose operator is represented by $(\cdot)^{\mathsf{T}}$, $||\cdot||_F$ is the Frobenius norm and $\circ$ is the outer product. Symbol $\mathbf{I}_N$ represents an identity matrix of size $N \times N$. A matrix $\mathbf{A} \in \mathbb{C}^{I_1 \times I_2}$, with scalar entries a_{i_1,i_2}, has its i_1^{th} row and i_2^{th} column represented by $\mathbf{a}_{i_1.}$ and $\mathbf{a}_{.i_2}$, respectively. Symbol $||\cdot||$ denotes the l_2-norm and $||\cdot||_{2,1}$ denotes the mixed $l_{2,1}$-norm, defined for an arbitrary matrix $\mathbf{A}$ with I_2 columns as:

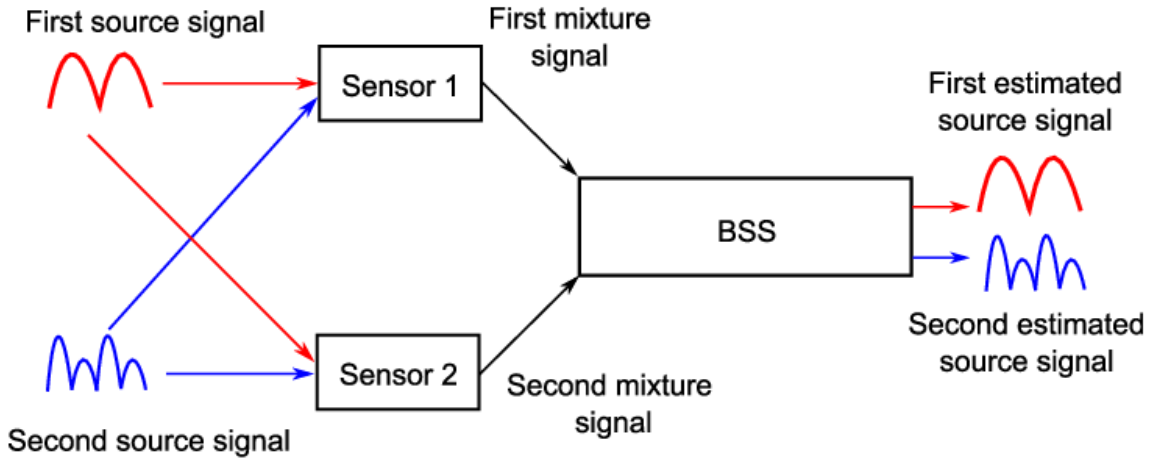
$$||\mathbf{A}||_{2,1} = \sum_{i_2=1}^{I_2} ||\mathbf{a}_{.i_2}|| \,. \tag{3.1}$$

3.1 Blind Source Separation

The cocktail party problem refers to the challenge of selectively listening to a particular sound source in a noisy and complex environment, such as a crowded cocktail party. The problem arises because the sound signals from different sources mix together, making it difficult for the listener to isolate and focus on a specific source of interest. The BSS model appears as an approach to address the cocktail party problem, aiming to separate mixed signals into their individual source components without prior knowledge of the sources or their characteristics (CARDOSO, 1998). It relies on statistical and mathematical techniques to identify the independent sources based on their statistical

properties.

Figure 3.1 – Example of the cocktail party problem. Two source signals (e.g., heart's electrical activity) are generated in two distinct locations and then recorded by two sensors (e.g. electrodes) resulting in two linear and instantaneous mixtures. The goal of this problem is to recover the original signals from the mixed signals.



Source: Adapted from Independent component analysis: An introduction (THARWAT, 2021).

BSS algorithms analyze the mixed signals received by multiple sensors to estimate the original source signals, as illustrated in Fig. 3.1. These algorithms leverage statistical dependencies and properties of the sources to separate and extract the individual sources from the mixture.

An ECG recording composed of K leads and N time samples can naturally be stored in a matrix $\mathbf{Y} \in \mathbb{R}^{K \times N}$, where the (i, j)-entry is given by the jth time sample of the ith lead $y_i(t)$:

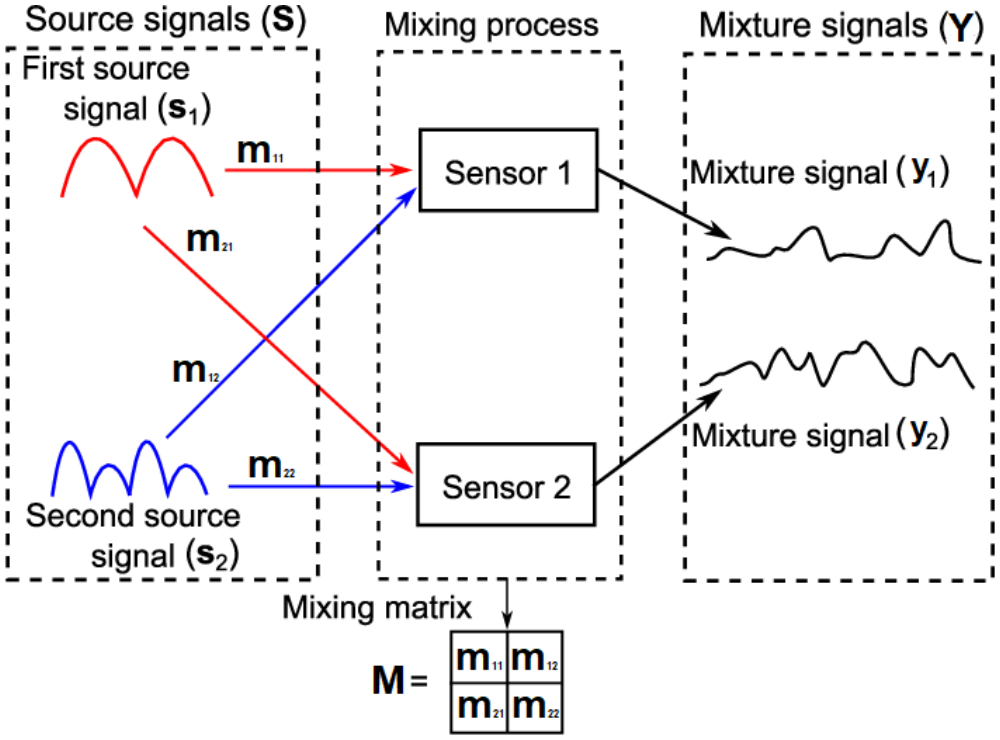
$$y_{i,j} = y_i(t_j) \qquad 1 \leq i \leq K,\ 1 \leq j \leq N.$$

This matrix may be expressed as the factorization (CARDOSO, 1998):

$$\mathbf{Y} = \mathbf{M}\mathbf{S}^{\mathsf{T}} = \sum_{r=1}^{R} \mathbf{m}_{\cdot r}\mathbf{s}_{\cdot r}^{\mathsf{T}} \tag{3.2}$$

where the columns of factor matrix $\mathbf{S} \in \mathbb{R}^{N \times R}$, denoted $\mathbf{s}_{\cdot r}$, are the temporal patterns or source signals contributing to the observed multi-lead record through the associated spatial patterns $\mathbf{m}_{\cdot r}$, the columns of mixing matrix $\mathbf{M} \in \mathbb{R}^{K \times R}$, $1 \leq r \leq R$, modeling the propagation of electrical sources from the heart to the body surface. Here, R represents the number of source signals.

Figure 3.2 – An illustrative example of the process of mixing signals, where two source signals are mixed linearly by the mixing matrix $\mathbf{M}$ to form two new mixture signals.



Source: Adapted from Independent component analysis: An introduction (THARWAT, 2021).

This electrical propagation is indicated as the mixing process in Fig. 3.2. It illustrates the source mixing behavior taken into account by the BSS model. The only prior information is that the mixture ($\mathbf{Y}$), sources, and mixture matrices (respectively $\mathbf{S}$ and $\mathbf{M}$) are unknown in real applications. Since $\mathbf{M}$ is unknown, matrix-based techniques such as PCA and ICA may be used to estimate $\hat{\mathbf{M}}^{-1}$, the inverse of the mixing matrix. Then, the estimated sources $\hat{\mathbf{S}}$ can be obtained, as presented in (ZARZOSO, 2009; OLIVEIRA; ZARZOSO, 2019).

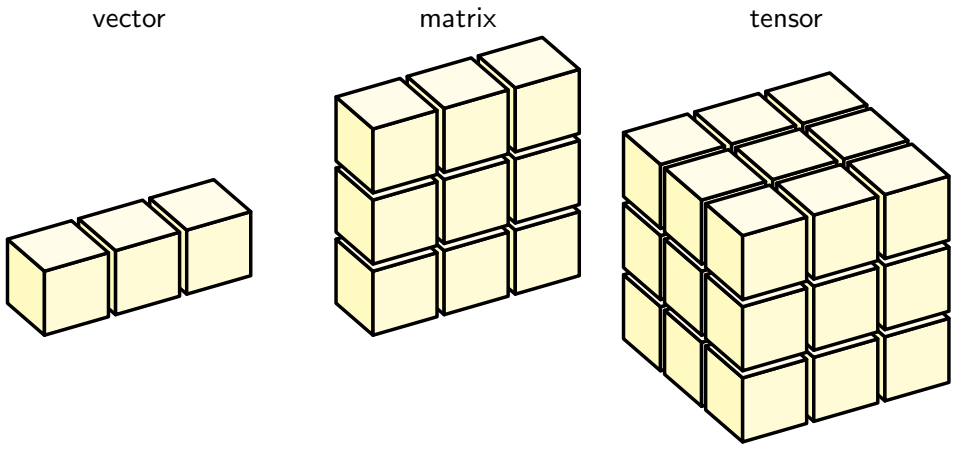
During AF, the AA and the Ventricular Activity (VA) signals are assumed uncoupled, allowing the AA extraction from an ECG to be formulated using BSS (RIETA *et al.*, 2004). Nevertheless, the factorization presented in (3.2) lacks uniqueness unless additional constraints are imposed, such as mutual orthogonality between spatial factors and statistical independence (CARDOSO, 1998; ZARZOSO, 2009).

To overcome these limitations, the ECG data can be arranged using a tensorial approach to explore milder constraints, ensuring uniqueness and outperforming matrix-based methods in terms of estimation quality as shown in (OLIVEIRA; ZARZOSO, 2019).

3.2 Block Term Decomposition

Tensor operations play an important role in tensor calculus. They allow us to rearrange the elements of a tensor into a vector or a matrix, and more generally into a reduced order tensor, to define tensor models and decompositions, and to compute the eigenvalues of a tensor. In this chapter, we will study two tensor multiplication operations, called a special case of BTD, named rank-$(L_r, L_r, 1)$ Block Term Decomposition.

Figure 3.3 – Representation of vector $\in \mathbb{R}^I$, matrix $\in \mathbb{R}^{I \times J}$ and third order tensor $\in \mathbb{R}^{I \times J \times K}$, respectively, where $I = J = K = 3$.



Source: Elaborated by the author.

3.2.1 Basic Definitions

The tensor introduces a concept that allows us to extend the concept of a matrix. While a matrix is a two-dimensional array of numbers arranged in rows and columns, a tensor is arranged in N-dimensions, where N may assume any positive integer value. Fig. 3.3 provides the geometric representation for vectors, matrices, and tensors.

We can denote a tensor as $\mathcal{X}$. More formally, an N way or Nth-order tensor is an element of the tensor product of N vector spaces, each of which has its own coordinate system (KOLDA; BADER, 2009). The third-order tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$, with scalar entries a_{i_1,i_2,i_3}, has its frontal slices represented by $\mathbf{X}_{..i_3} \in \mathbb{C}^{I_1 \times I_2}$.

Definition 3.1. Let $\mathcal{Y} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ be an N-th order tensor. A scalar component of $\mathcal{Y}$

is indicated as:

$$y_{i_1,i_2,\ldots,i_N} = [\mathcal{Y}]_{i_1,i_2,\ldots,i_N}\,, \tag{3.3}$$

where i_n is the n-th dimension of $\mathcal{Y}$, also called the mode-n of $\mathcal{Y}$.

Definition 3.2. The outer product $\mathcal{X} \circ \mathcal{Y}$ of orders M and N, respectively, where $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_M}$ and $\mathcal{Y} \in \mathbb{R}^{J_1 \times J_2 \times \cdots \times J_N}$ is tensor defined by

$$[\mathcal{X} \circ \mathcal{Y}]_{i_1,i_2,\ldots,i_M,j_1,j_2,\ldots,j_N} = x_{i_1,i_2,\ldots,i_N} y_{j_1,j_2,\ldots,j_N}, \tag{3.4}$$

For instance, the equation (3.4) is an extension of the concept applied to vectors.

Definition 3.3. The rank -1 tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times \ldots \times I_N}$ can be written as the outer product of N vectors $\mathbf{u}^{[1]} \in \mathbb{C}^{I_1}$, $\mathbf{u}^{[2]} \in \mathbb{C}^{I_2}$, ..., $\mathbf{u}^{[N]} \in \mathbb{C}^{I_N}$, i.e.:

$$x_{i_1,i_2,\ldots,i_N} = \mathbf{u}^{[1]}_{i_1} \cdot \mathbf{u}^{[2]}_{i_2} \cdot \ldots \cdot \mathbf{u}^{[N]}_{i_N}. \tag{3.5}$$

Each vector $\mathbf{u}^{[n]}$ is called a component of $\mathcal{X}$. The outer product of two vectors yields a matrix, i.e, a second-order tensor, which is important to notice since it supports to introduce the notion of rank -1 tensor as a special case.

Definition 3.4. The rank of an arbitrary tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times \ldots \times I_N}$, denoted by:

$$R = r(\mathcal{X}), \tag{3.6}$$

is the minimal number of rank -1 tensors that yield $\mathcal{X}$ in a linear combination.

Definition 3.5. The Frobenius norm of a tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times \ldots \times I_N}$ is defined as:

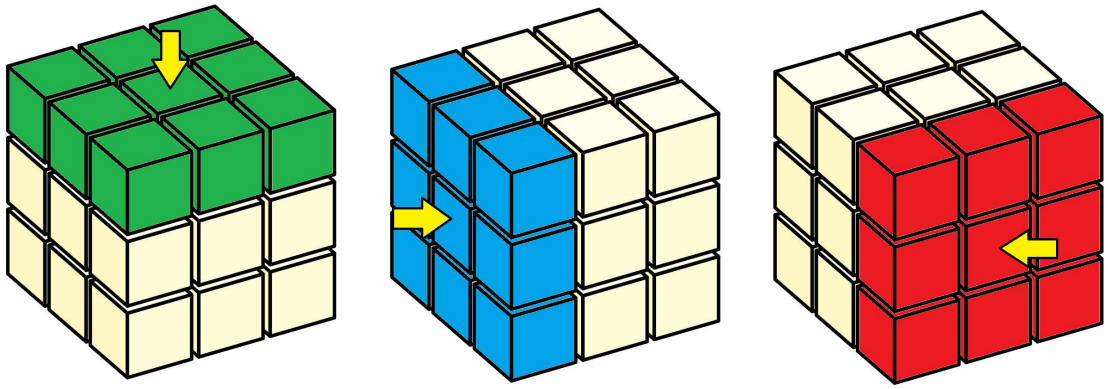
$$||X||_F = \sqrt{\langle \mathcal{X}, \mathcal{X} \rangle} = \left(\sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \cdots \sum_{i_N=1}^{I_N} |x|^2_{i_1,i_2,\ldots,i_N} \right)^{1/2}. \tag{3.7}$$

The Frobenius norm may be interpreted as the amount of energy carried in the tensor (GOLUB; LOAN, 2013; FAVIER, 2021).

Definition 3.6. Let $\mathbf{X}_n$ be the n-th mode unfolded matrix of $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times \ldots \times I_N}$. The mode-$n$ rank of $\mathcal{X}$ is the dimension of the vector space generated by the n-th mode vectors (i.e., the columns of X_n).

In the matrix case, both row and column ranks are equal, $R_{\text{row}} = R_{\text{column}} = R$. The mode-$n$ rank is a generalization of the classical concept of rank for matrices (Fig. 3.4 illustrates this for third-order tensor). However, the mode-n ranks of a higher-order tensor are not necessarily the same. Moreover, when mode-n ranks are the same, they can still differ from the rank of the tensor. The mode-n rank is always inferior or equal to the rank, i.e., $R_n \leq R$ (KOLDA; BADER, 2009).

Figure 3.4 – Representation of the unfolded versions of a third-order tensor. Respectively, first (green), second (blue), and third (red) mode slices are indicated. The arrow reinforces the unfold direction.



Source: Elaborated by the author.

3.2.2 Decomposition in multilinear rank-$(L_r, L_r, 1)$ terms

The BTD is a tensor decomposition technique that aims to decompose a higher-order tensor into a sum of simpler component tensors. Particularly useful to analyze and represent multiway data structures, where the data can be rearranged as multidimensional arrays. The decomposition results into a sum of block terms, consequently it is named a "Block Term Decomposition" (LATHAUWER, 2011). Each block term represents a component tensor that captures specific patterns or structures (i.e, underlying characteristics in original data) within the original tensor. The BTD may be regarded as a generalization of matrix-based factorization, such as the SVD, to tensors. The present work is interested in the special case where it results in the decomposition in multilinear rank-$(L_r, L_r, 1)$ terms.

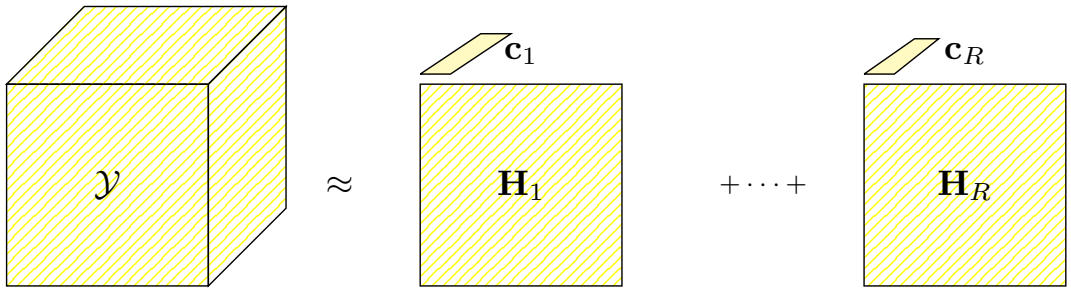
In this section we focus on the third order tensor, which represents the data structure of our model.

Definition 3.7. An arbitrary tensor $\mathcal{Y} \in \mathbb{C}^{I \times J \times K}$ decomposition into a sum of rank$(L_r, L_r, 1)$ terms is defined as:

$$\mathcal{Y} = \sum_{r=1}^{R} \mathbf{H}_r \circ \mathbf{c}_r \tag{3.8}$$

where the $(I \times J)$ matrices $\mathbf{H}_r$ are rank-L_r. Each vector $\mathbf{c}_r$ represents the scaling factors and is equivalent to a column of the mixing matrix M. We also consider the decomposition of a tensor into a sum of matrix-vector outer products, in which the different matrices do not necessarily all have the same rank. Fig. 3.5 presents the visual representation for the (3.8) model.

Figure 3.5 – Visual representation of the decomposition in rank-$(L_r, L_r, 1)$ terms of a third-order tensor into terms of the sum of R block terms for $\mathbf{H}_r$ and $\mathbf{c}_r$ outer product, for $r = 1, \ldots, R$.



Source: Elaborated by the author.

The model presented in (3.8) admits a decomposition $\mathbf{H}_r = \mathbf{A}_r \mathbf{B}_r^{\mathsf{T}}$ yielding:

$$\mathcal{Y} = \sum_{r=1}^{R} \left(\mathbf{A}_r \mathbf{B}_r^{\mathsf{T}} \right) \circ \mathbf{c}_r \tag{3.9}$$

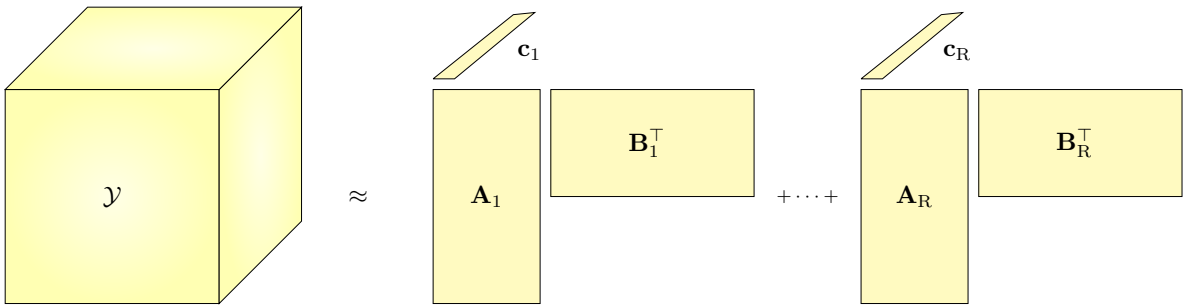
where $\mathbf{A}_r \in \mathbb{C}^{I \times L_r}$ and $\mathbf{B}_r \in \mathbb{C}^{J \times L_r}$ are the factor matrices and have rank L_r. See Fig. 3.6.

We recall the matrix model presented in (3.2), where the columns of factor matrix $\mathbf{S} \in \mathbb{R}^{N \times R}$, denoted $\mathbf{s}_{\cdot r}$, are the temporal patterns or source signals. During AF, the AA and the VA signals are assumed uncoupled (RIETA *et al.*, 2004), allowing the AA extraction from an ECG using BSS. Due to the quasi-periodic nature of AF signals, atrial sources can be well represented by an all-pole model (STRIDH; SöRNMO, 2001).

Definition 3.8. Assume that a discrete-time signal $s(n)$ is a linear combination of L_r damped complex exponentials, given by:

$$s_{r,n} = \sum_{\ell=1}^{L_r} \lambda_{r,\ell} z_{r,\ell}^{n-1} \tag{3.10}$$

Figure 3.6 – Visual representation of the decomposition in rank-$(L_r, L_r, 1)$ terms of a third-order tensor into terms of the sum of R block terms for $\mathbf{A}_r\mathbf{B}_r^\intercal$ and $\mathbf{c}_r$ outer product, for $r = 1, \ldots, R$. .



Source: Elaborated by the author.

where L_r is the number of exponential terms, $z_{r,\ell}$ is the ℓ^{th} pole of the r^{th} source, and $\lambda_{r,\ell}$ is the scaling coefficient, with $r = 1, \ldots, R$ representing the source number and $n = 1, \ldots, N$ the discrete-time index.

Definition 3.9. Given a vector $\mathbf{y}$ of length N, one can build a Hankel matrix $\mathbf{H} \in \mathbb{R}^{I \times J}$, where $I = J = \frac{N+1}{2}$ if N is odd or $I = \frac{N}{2}$ and $J = \frac{N}{2} + 1$ if N is even, with entries:

$$h_{i,j} \triangleq y_{i+j-1} \tag{3.11}$$

where $i = 1, \ldots, I$, $j = 1, \ldots, J$. Hankel-matrix construction is illustrated in Fig. 3.7.

Considering the signal model in (3.10), a data vector can be mapped onto an $M \times M$ Hankel matrix. For simplicity, we assume a vector with an odd number of samples, where each sample is placed along the anti-diagonal of $\mathbf{H}_s$. A major result in signal processing states that $\mathbf{H}_s$ can be decomposed using the Vandermonde decomposition (SCHARF, 1991; BOLEY *et al.*, 1997; LATHAUWER, 2011):

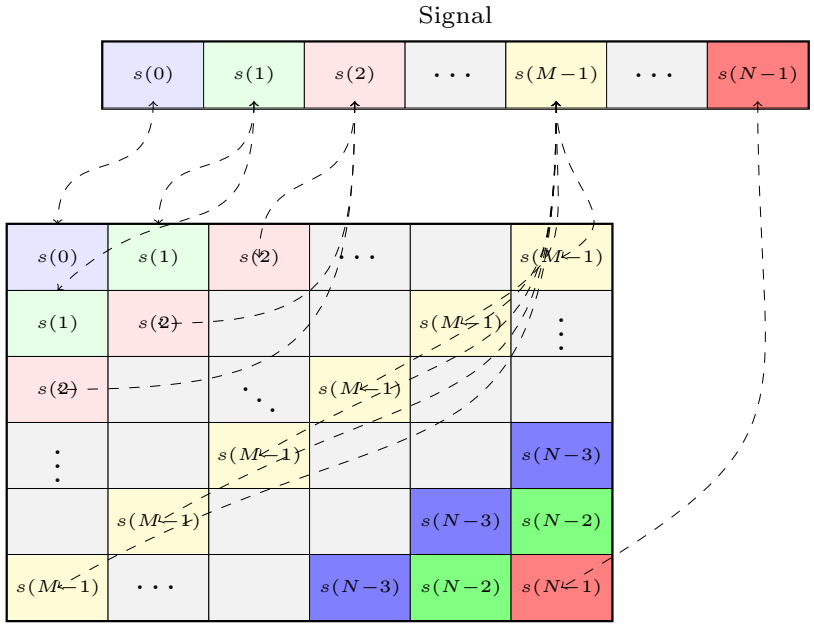
$$\mathbf{H}_s = \mathbf{V}_s \, \mathbf{D} \, \mathbf{V}_s^\intercal, \tag{3.12}$$

where $\mathbf{V}_s$ is the Vandermonde matrix:

$$\mathbf{V}_s = \begin{bmatrix} 1 & 1 & \ldots & 1 \\ z_1 & z_2 & \ldots & z_{L_r} \\ \vdots & \vdots & & \vdots \\ z_1^{M-1} & z_2^{M-1} & \ldots & z_{L_r}^{M-1} \end{bmatrix} \in \mathbb{C}^{M \times L_r}, \tag{3.13}$$

and $\mathbf{D} = \mathrm{diag}(c_1, c_1, \ldots, c_{L_r}) \in \mathbb{C}^{L_r \times L_r}$ is a diagonal matrix.

Figure 3.7 – Visual representation of a signal mapped into a Hankel matrix, $\mathbf{H}_s$, built from a signal, $s(n)$. Without loss of generality, the number of samples N is assumed to be odd.



Source: Elaborated by the author.

As a result of the Vandermonde decomposition (3.13), we see that the matrix $\mathbf{H}_s$ has rank at most $\min\{L_r, M\}$, accepts a low-rank Hankel source model, and the signal separation can be performed via BTD in rank-$(L_r, L_r, 1)$ terms (LATHAUWER, 2011; GOULART *et al.*, 2020).

Consider the ECG matrix $\mathbf{Y}$ containing the signals measured along an arbitrary segment, including VA as well as other sources of noise and interference. To transform the matrix model in (3.2) into a tensor model suitable for source separation via low-rank Hankel BTD, each lead (i.e., row) of $\mathbf{Y}$ is mapped into a Hankel matrix:

$$\mathbf{y}_{k\cdot} \in \mathbb{R}^N \mapsto \mathbf{H}_{\mathbf{Y}}^{(k)} \in \mathbb{R}^{I \times J} \tag{3.14}$$

as defined in (3.11). Then, we build a third-order tensor $\mathcal{Y} \in \mathbb{R}^{I \times J \times K}$ by stacking Hankel matrices $\mathbf{H}_{\mathbf{Y}}^{(k)}$ along its third mode (frontal) slices, as shown in Fig. 3.8:

$$\mathcal{Y}_{\cdot,\cdot,k} = \mathbf{H}_{\mathbf{Y}}^{(k)} \qquad 1 \leq k \leq K. \tag{3.15}$$

Thus the tensor, after some algebraic steps, may be reformulated as:

$$\begin{aligned}\mathcal{Y} &= \sum_{k=1}^{K} \mathbf{H}_{\mathbf{Y}}^{(k)} \circ \mathbf{e}_k = \sum_{k=1}^{K} \left(\sum_{r=1}^{R} m_{k,r} \mathbf{H}_{\mathbf{S}}^{(r)} \right) \circ \mathbf{e}_k \\ &= \sum_{r=1}^{R} \mathbf{H}_{\mathbf{S}}^{(r)} \circ \left(\sum_{k=1}^{K} m_{k,r} \mathbf{e}_k \right) \\ &= \sum_{r=1}^{R} \mathbf{H}_{\mathbf{S}}^{(r)} \circ \mathbf{m}_r \end{aligned} \tag{3.16}$$

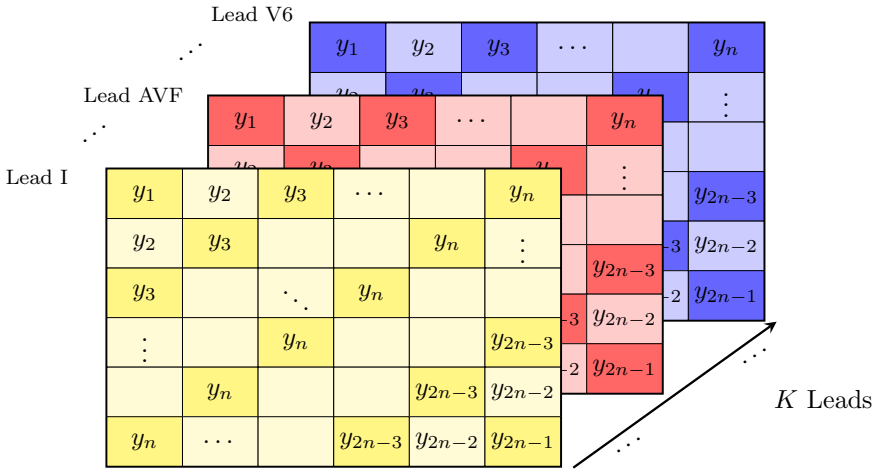
where $\mathbf{e}_k$ is the kth canonical basis vector of $\mathbb{R}^K$. Its important to notice that (3.16) is a direct implication of Hankel map's linearity, since:

$$\mathbf{H}_{\mathbf{Y}}^{(k)} = \sum_{r=1}^{R} m_{k,r} \mathbf{H}_{\mathbf{S}}^{(r)}.$$

where $\mathbf{H}_{\mathbf{S}}^{(r)}$ is the Hankel matrix associated with the rth source signal.

We consider that: rank $\left(\mathbf{H}_{\mathbf{S}}^{(r)}\right) = L_r$, $1 \leq r \leq R$, thus (3.16) defines a decomposition of tensor $\mathcal{Y}$ into blocks with multilinear rank-$(L_r, L_r, 1)$ terms, where $\mathbf{H}_{\mathbf{S}}^{(r)}$ admits a factorization $\mathbf{H}_{\mathbf{S}}^{(r)} = \mathbf{A}_r \mathbf{B}_r^{\mathsf{T}}$, with $\mathbf{A}_r \in \mathbb{R}^{I_1 \times L_r}$ and $\mathbf{B}_r \in \mathbb{R}^{I_2 \times L_r}$ presenting full column rank L_r.

Figure 3.8 – The Hankel-BTD model for ECG data: A third-order tensor $\mathcal{Y}$ built by stacking the Hankelized leads along its 3rd mode slices.



Source: Elaborated by the author.

This technique is referred as decomposition in multilinear rank-$(L_r, L_r, 1)$ terms, a special case of BTD, that arrange $\mathcal{Y}$ with Hankel structure.

Uniqueness Conditions: There is ample evidence in the literature to ensure the uniqueness of the model presented under certain assumptions (LATHAUWER, 2008;

LATHAUWER, 2011). If the following conditions are satisfied, then the BTD is essentially unique:

Theorem 3.1. *Both matrix factors* $\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_1 & \boldsymbol{A}_2 & \dots & \boldsymbol{A}_R \end{bmatrix} \in \mathbb{R}^{I_1 \times \sum_{r=1}^{R} L_r}$ *and* $\boldsymbol{B} = \begin{bmatrix} \boldsymbol{B}_1 & \boldsymbol{B}_2 & \dots & \boldsymbol{B}_R \end{bmatrix} \in \mathbb{R}^{I_2 \times \sum_{r=1}^{R} L_r}$ *are full-column rank. This condition requires that* $\sum_{r=1}^{R} L_r \leq I_1, I_2$.

Proof. Formulation proposed by (LATHAUWER, 2011) in theorem 2.2. □

Theorem 3.2. *Linear independence between the columns of the matrix* $\boldsymbol{M} = \begin{bmatrix} \boldsymbol{m}_1 & \boldsymbol{m}_2 & \dots & \boldsymbol{m}_R \end{bmatrix} \in \mathbb{R}^{I_3 \times R}$.

Proof. Formulation proposed by (LATHAUWER, 2011) in theorem 2.3. □

The decomposition is subject to the following indeterminacies:

1. The different multilinear rank-(L_r,L_r,1) terms can be arbitrarily permuted.
2. $\mathbf{A}_r$ can be postmultiplied by any nonsingular matrix, given that $\mathbf{B}_r^T$ is premultiplied by the inverse of that nonsingular matrix.
3. As long as the resulting product remains the same, the factors of the same multilinear rank-(L_r,L_r,1) term can be arbitrarily scaled.

Still, both theorems 3.1 and 3.2 are sufficient to ensure uniqueness. It implies that $I, J \approx \frac{N+1}{2} \geq \sum_{r=1}^{R} L_r$, which sets a lower bound on the length of ECG segments.

The condition for satisfying the mixing matrix $\mathbf{M}$ requires the presence of sources that are spatially localized in different regions. However, in our application, this condition is not restrictive since the atrial sources and ventricular sources are naturally associated with distinct regions of the heart. In (LATHAUWER, 2011) a more relaxed set of uniqueness conditions is addressed in the context of the BTD.

Once the block terms have been identified using an appropriate approach, the mixing matrix $\mathbf{M}$ can be estimated based on the recovered vector factors. Furthermore, the source signals may be estimated by calculating the average values along the anti-diagonals of the estimated source Hankel matrices $\mathbf{H}_{\mathbf{S}}^{(r)}$, where $1 \leq r \leq R$.

Finally, as discussed over this section, the Hankel matrix rank, rank($\mathbf{H}_{\mathbf{S}}^{(r)}$), is intrinsically linked to the original source. Since these estimated sources are mapped onto Hankel matrices, their rank matches the number of poles in the result derived from the Vandermonde decomposition (LATHAUWER, 2011; ZARZOSO, 2017). Hence, the

Hankel-BTD model is well suited to extract characteristics of AA in AF episodes. The rank L_r of the Hankel factor associated with the atrial source is directly linked with the number of poles generating the atrial signal and, therefore, can be considered as a natural measure of AF complexity in ECG records.

3.2.3 Constrained Alternating Group Lasso

In general, the algorithms that perform BTD are Least Squares (LS)-based, e.g., Gauss-Newton BTD (LATHAUWER; NION, 2008). Such approaches aim to minimize the Euclidean distance between the observed data tensor and a model of fixed structure with respect to the model components (SORBER *et al.*, 2013). These conditions, which impose prior knowledge about the model parameters, are typically unknown.

The limitations mentioned are addressed in (GOULART *et al.*, 2020) through the introduction of a novel algorithm known as the AGL. The technique aims to compute low-rank BTDs and overcomes these limitations. It achieves this by minimizing a fitting term (applying the LS) and a regularization term (sum of mixed-norms) to promote group sparsity in the factor matrices. The algorithm is provably convergent and offers a simpler alternative to existing methods, such as the Alternating Least Squares (ALS) method.

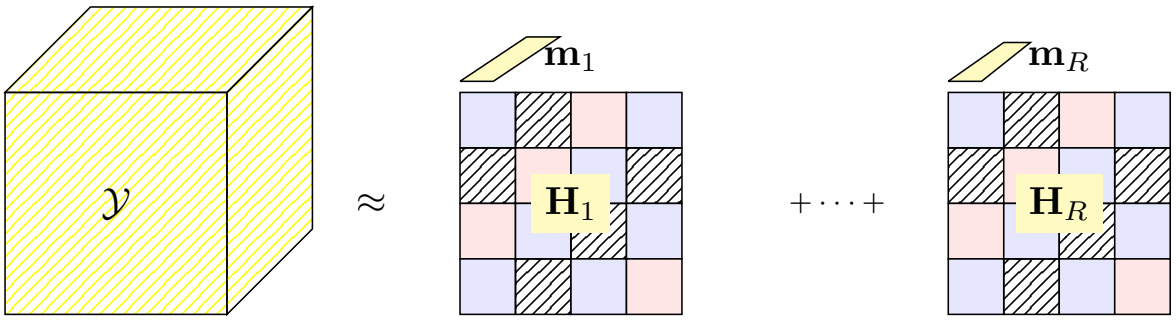
The AGL variant, CAGL, is also introduced to handle linearly constrained blocks. This technnique incorporates a structured low-rank approximation method to ensure the blocks have low rank and belong to the specified subspace, in our scenario structured as Hankel matrices.

Typically, the fixed structure used in BTD model is designed to minimize the Euclidean distance with regard to an observed data tensor $\mathcal{Y} \in \mathbb{C}^{I \times J \times K}$:

$$f(\mathbf{A}, \mathbf{B}, \mathbf{C}) \triangleq \left\| \mathcal{Y} - \sum_{r=1}^{R} \left(\mathbf{A}_r \mathbf{B}_r^{\mathsf{T}} \right) \circ \mathbf{c}_r \right\|_F^2. \tag{3.17}$$

In the approach being considered, each $\mathbf{H}_{\mathbf{S}}^{(r)}$ matrix must belong to the Hankel subspace with dimensions $(I \times J)$, denoted $\mathcal{S}_H$. The 3rd mode slices $\mathcal{Y}_{:,:,k}$ of the observed tensor are Hankel by construction, as shown in Fig. 3.9. However, due to inherent model limitations and noise, a solution $(\hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{C}})$ of (3.17) may not satisfy $\hat{\mathbf{A}}_r \hat{\mathbf{B}}_r^{\mathsf{T}} \in \mathcal{S}_H$. Besides that, techniques that rely on (3.17) are heavily influenced by its matrix factors initialization and require prior knowledge about its structure. It implies that the algorithm must use fixed values for the number of blocks R and their ranks L_r, even in the absence of this

Figure 3.9 – Visual representation of Hankel-BTD decomposed into rank-$(L_r, L_r, 1)$ terms. The third-order tensor $\mathcal{Y}$ follows the mapping presented in Fig. 3.8. The terms of the sum of R block terms for $\mathbf{H}_r$ and $\mathbf{c}_r$ outer product, for $r = 1, \ldots, R$. Each factor matrix represents an estimated source.



Source: Elaborated by the author.

information. A mismatch between the assumed parameters and the actual ones may impact estimation performance.

AGL and its constrained version address such limitations, including penalization terms promoting low-rank blocks and controlling the number of blocks, instead of using a fixed BTD structure as in (3.17), resulting in the form of:

$$F(\mathbf{A}, \mathbf{B}, \mathbf{C}) \triangleq f(\mathbf{A}, \mathbf{B}, \mathbf{C}) + \gamma\, g(\mathbf{A}, \mathbf{B}, \mathbf{C}) \tag{3.18}$$

where $\gamma > 0$ is a regularization parameter and g is a regularization function of the form:

$$g(\mathbf{A}, \mathbf{B}, \mathbf{C}) \triangleq \|\mathbf{A}\|_{2,1} + \|\mathbf{B}\|_{2,1} + \|\mathbf{C}\|_{2,1}. \tag{3.19}$$

The method takes advantage of the so-called group lasso, a generalization of the lasso estimator principle (YUAN; LIN, 2006), that presents geometric properties of the mixed $\ell_{2,1}$-norm. Thus, solutions where $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ have null columns (for values of γ that are sufficiently high) will be induced, allowing one to select the relevant low-rank blocks. It allows the algorithm to estimate the number of blocks R and their ranks L_r.

To ensure the Hankel structure of the matrix factors, the CAGL applies a structured low-rank approximation once the algorithm converges. At this stage, Cadzow's algorithm (CADZOW, 1988) is employed to perform alternating projections onto the Hankel subspace $\mathcal{S}_H$. This ensures that $\hat{\mathbf{H}}_r \approx \hat{\mathbf{A}}_r \hat{\mathbf{B}}_r^{\mathsf{T}} \in \mathcal{S}_H$. A more in-depth description of the algorithm is presented in (GOULART *et al.*, 2020).

4 COMPLEXITY ANALYSIS

In this chapter, we present the envisioned concept of complexity analysis. Since there is an absence of a universally accepted definition for AF complexity, this results in the use of this term to describe various characteristics associated with the electrical activity occurring in the fibrillating atria. In literature, some authors shed light on this problem, where atrial organization or complexity analysis may assume quite different meanings. Special emphasis is placed on highlighting the distinct characteristics of organization that are examined and captured by each method, which we refer to later in this chapter as the complexity index. Furthermore, we address the fundamental signal processing techniques used for analyzing atrial electrical activity, as discussed in (SIH, 2001; BARBARO *et al.*, 2001; MAINARDI *et al.*, 2008b). As discussed in these works, each class includes methods that share the conceptual interpretation of organization in terms of:

1. temporal regularity (rhythmic) of atrial activations;
2. regularity/complexity of single site electrograms;
3. coupling/synchronization between the electrical activity of two adjacent sites;
4. similarity of activation wave morphologies.

The approach proposed in this work explores a variety of signal processing methods to analyze atrial recordings mainly in the time domain and to characterize the level of complexity, i.e., the regularity of the fibrillating atria. Therefore, our work falls into category 1 of the classification. In addition, our framework leverages noninvasive information since the only data we have access to is from the standard 12-lead ECG.

Studies exploring these indices can be characterized also by the number of leads they consider simultaneously: single-lead or multi-lead. During single-lead analysis, each row of the ECG matrix is examined individually. This analysis can be performed as a time series or in the frequency domain using spectral representations. Commonly used complexity indices for AF include entropy-based methods, such as Shannon entropy, which quantify the unpredictability and randomness of the signal. Additionally, frequently reported methods in the literature encompass heart rate variability analysis, f-wave amplitude analysis, and techniques based on entropy and frequency tracking (PORTA *et al.*, 1999; VIKMAN *et al.*, 2003; MEO *et al.*, 2015; MCCANN *et al.*, 2021).

However, assessing the leads individually may lead to the overlooking of valuable information, disregarding the inherent spatial diversity present in the multi-lead nature of

the ECG. To address this limitation, the present study focuses exclusively on assessing multi-lead indices. Our choice is justified by the following arguments:

1. Limited information: Single-lead approaches often rely on data from a single electrode or lead, which provides limited information about the spatial distribution of signals within the atria. This limited perspective may overlook important patterns and complexities present in the overall electrical activity.
2. Incomplete representation: Atrial fibrillation is a complex arrhythmia that involves the interaction of multiple electrical sources and pathways within the atria. Analyzing a single-lead fails to capture the full extent of this spatially diverse activity, resulting in an incomplete representation of the underlying complexity.
3. Ignoring spatial relationships: The spatial diversity of signals in the atria plays a significant role in understanding the dynamics and complexity of atrial fibrillation. Single-lead approaches do not consider the interplay between different regions or the propagation patterns of electrical signals, which can limit the accuracy and comprehensiveness of complexity analysis.
4. Missed abnormalities: By not accounting for spatial diversity, single-lead approaches may overlook localized abnormalities or focal sources of electrical activity that contribute to the overall complexity. These abnormalities can have important clinical implications and may require specific diagnostic and treatment strategies.
5. Multi-lead advantages: Comparatively, a multi-lead approach allows for a comprehensive assessment of the spatial characteristics of atrial signals. By considering multiple leads simultaneously, you can capture the interactions between different regions and uncover complex patterns that would be missed by focusing solely on a single-lead.

In contrast, multi-lead indices leverage matrix techniques to exploit the spatial diversity information (RIETA *et al.*, 2004; MEO *et al.*, 2013; OLIVEIRA; ZARZOSO, 2019; CIRUGEDA *et al.*, 2020), yielding more robust results to extract AA during AF compared to the single-lead approach. Among these techniques, those based on PCA are particularly popular due to their simplicity and promising predictive capabilities for cardiac arrhythmia outcomes[1] (MAINARDI *et al.*, 2008a; BONIZZI *et al.*, 2010; MEO *et al.*, 2018; MCCANN *et al.*, 2022). However, their practical implementation in clinical

[1] Note that is this context, the use of PCA is not exactly the same as the presented in BSS model.

settings is limited as they primarily rely on body surface potential maps (BSPM) and necessitate long recordings. Additionally, TQ-interval segmentation, concatenation, or other preprocessing techniques are often employed, leading to signal discontinuity and potential loss of important data relationship.

Beyond the matrix approach, the utilization of tensor to organize ECG data, followed by factorization techniques, has demonstrated significant advantages in extracting AA from surface ECG signals. These approaches, along with its associated algorithmic advancements, has shown promising in various studies (ZARZOSO, 2017; GOULART *et al.*, 2020; OLIVEIRA *et al.*, 2022). However, the clinical potential of tensor-based methods and their ability to derive AF markers and assist in therapeutic decision-making have not been extensively investigated thus far.

The present chapter introduces the matrix-based AF complexity index, NDI, chosen for comparison with the tensor-based AF complexity index, CAGL. While both indices aim to quantify AF complexity, they employ fundamentally different mathematical frameworks. Our decision to compare them is based on the following factors: promising results (MEO *et al.*, 2018; MCCANN *et al.*, 2022), well-established theoretical foundations (MAINARDI *et al.*, 2008a; BONIZZI *et al.*, 2010), and our familiarity with the fundamental principles underlying each method. This comparative analysis enables us to evaluate the advantages and limitations of the tensor approach compared to the PCA-based approach in capturing the underlying dynamics of AF. Moreover, understanding the relative performance of these methods provides valuable insights into the robustness and reliability of the indices across various clinical conditions.

4.1 Matrix-Based AF Complexity Indices

The NDI is a complexity method that measures the presence of non-dipolar or abnormal activity in AF. The ECG signals are traditionally expected to exhibit a dipolar pattern, which arises from the coordinated electrical activity of the heart (HOLT *et al.*, 1969). This dipolar pattern corresponds to the normal activation sequence of the atria and ventricles. However, in certain conditions, such as AF or other cardiac disorders, the ECG signals deviate from this expected dipolar pattern. The deviation from a purely dipolar pattern suggests the involvement of additional sources or abnormalities in the atrial electrical activity.

These additional sources could arise from irregular and chaotic electrical signals within the atria, resulting in the absence of a clear dipole orientation. Such irregularities can occur due to abnormal conduction pathways, reentrant circuits, or the presence of focal sources of electrical activity. These deviations from the dipolar pattern can be observed as changes in the morphology, amplitude, and timing of the ECG waveforms. The P-wave, which represents atrial depolarization, may exhibit varying shapes, durations, and amplitudes, indicating the presence of spatial inhomogeneities in the atrial electrical activity. Furthermore, the irregularity in the atrial electrical signals may manifest as irregular R-R intervals on the ECG, reflecting the irregular ventricular response associated with atrial fibrillation.

The NDI is a PCA-based method that analyzes the spatial distribution of the ECG signals across multiple leads and quantifies the contribution of non-dipolar components. It aims to quantify the extent to which electrophysiological signals deviate from a purely dipolar pattern, indicating the presence of additional sources or abnormalities in the atrial electrical activity. A higher NDI value indicates increased complexity and non-dipolar activity in AF, potentially reflecting underlying pathological conditions or disturbances in atrial conduction.

The PCA is a technique that enforces orthonormality among the columns of factor matrices $\mathbf{M}$ while maximizing the L2-norm of consecutive columns of $\mathbf{S}$, which are known as Principal Components (PCs) (JOLLIFFE, 2002). The contribution of these PCs to the observed data can be measured by summing their individual power, given by $\sum_{r=1}^{k} \sigma_r^2$, where σ_r^2 represents the power associated with the rth PC. This power can be computed by squaring the rth dominant singular value of $\mathbf{Y}$, the ECG data.

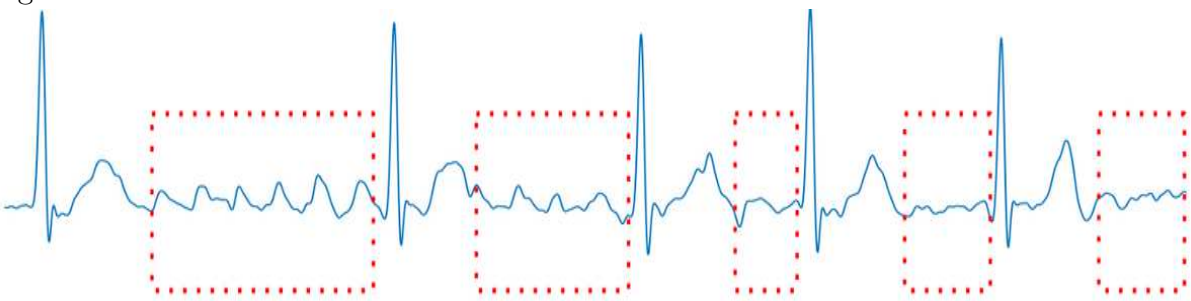
In the field of electrocardiography, it is generally assumed that the electrical activity of the heart may be a first-order approximation. It is generated by a single electric dipole with fixed position but a time-varying orientation in three-dimensional space (HOLT *et al.*, 1969; MALMIVUO; PLONSEY, 1995). Based on this assumption, the majority of the power distribution across multiple leads can be effectively approximated by the first three PCs. (LUX *et al.*, 1981). Inspired by the encouraging results obtained with the model proposed in (BONIZZI *et al.*, 2010; MARCO *et al.*, 2012), a new index for measuring the organization of AF is introduced (MEO *et al.*, 2018), based on the matrix factorization model discussed earlier.

The first step in the proposed approach involves automatically segmenting the TQ intervals, which exclusively contain AA, and storing them in the matrix $\mathbf{Y}$. It consists in extracting only the ECG segment that preponderantly presents the AF (The algorithm is illustrated in Fig. 4.1 and detailed in Algorithm 1). After extracting the segments of interest, we must concatenate them in order to get a data matrix with the AA signal. The hypothesis is that organized AA can be adequately represented by the single-dipole model. Accordingly, the multi-lead signal matrix $\mathbf{Y}$ is approximated by the 3D subspace spanned by its first 3 PCs. This approximation assumes that the majority of the variability in the observed matrix can be captured by these three PCs. Conversely, more complex patterns are expected to require a greater number of PCs to provide a more accurate description of the observed matrix. Finally, the proportion of power that cannot be explained by the three dominant PCs is referred to as the NDI:

$$\text{NDI} = 1 - \frac{\sum_{r=1}^{3} \sigma_r^2}{\sum_{r=1}^{R} \sigma_r^2}. \tag{4.1}$$

As explained earlier, the higher the complexity of the underlying AA, the lower the power ratio explained by the dipole model and the higher the value of NDI.

Figure 4.1 – This patient present abnormally short QT intervals. Hindering the TQ-segmentation.



Source: Elaborated by the author.

The computation of NDI requires the concatenation of at least three TQ segments, highlighted by red dotted rectangles in the Fig. 4.1 . However, to obtain reliable outcomes, it is recommended to use ECG segments of up to 12 seconds, as stated in (MEO *et al.*, 2018).

4.2 Tensor-Based AF Complexity Index

In this section, we describe the CAGL tensor complexity index to noninvasively quantify AF complexity. Because of the quasi-periodic nature of AF signals, the atrial sources can be effectively modeled using (3.10) (STRIDH; SöRNMO, 2001). As mentioned previously, when these signals are mapped onto Hankel matrices, the rank of these matrices is equal to the number of poles, resulting from the Vandermonde decomposition (LATHAUWER, 2011).

After performing the BTD using CAGL, a challenge remains in selecting which block $\mathbf{H}_{\mathbf{S}}^{(r)} \circ \mathbf{m}_r$, corresponds to the atrial activity source in (3.16) model. Measuring the quality of estimation is challenging due to the unavailability of ground truth. However, certain characteristics of AA during AF can be utilized to guide the selection of sources. In our proposed framework, we compute Spectral Concentration (SC), Dominant Frequency (DF), and kurtosis (ZARZOSO; COMON, 2010; OLIVEIRA; ZARZOSO, 2018; OLIVEIRA; ZARZOSO, 2019). We consider that among the estimates, only one block carries the AA source, and the selection is based on which source maximizes the SC, defined as:

$$\mathrm{SC} = \left(\sum_{f_i=0.82 f_p}^{1.17 f_p} P_{AA}(f_i) \right) \left(\sum_{f_i=0}^{F_s/2} P_{AA}(f_i) \right)^{-1}, \tag{4.2}$$

where f_p is the value of the DF, defined as $\arg\max_{f_i} P_{AA}(f_i)$, F_s is the sampling frequency, f_i is the discrete frequency and P_{AA} is the power spectrum of the AA signal computed using Welch's method as in (CASTELLS *et al.*, 2005). An AA signal during AF typically should have a DF between 3 and 9 Hz with high SC. Finally, the kurtosis of the signal in the frequency domain, acquired by a 4096-point Fast Fourier Transform, is computed as in (ZARZOSO; COMON, 2010). Since kurtosis measures peakedness (i.e, the degree to which data values are concentrated around the mean) and sparsity of a distribution, it consequently provides a quantitative measure of harmonicity of in the frequency domain of the computed signal. A high kurtosis is thus suggestive of a harmonic signal like AA during AF.

We identify the rth block that corresponds to the AA from this analysis and then compute the CAGL complexity index

$$L_{r_{\mathrm{AA}}} = \mathrm{rank}\left(\mathbf{H}_{\mathbf{S}}^{(r_{\mathrm{AA}})}\right) .$$

5 RESULTS AND DISCUSSION

This chapter presents the results of the conducted analyses. The chapter is structured as follows: In Section 5.1, we provide the patients summary statistics. In section 5.2 we present the experimental framework, presenting the details about the parameter used for each algorithm. Additionally, we present some examples of algorithm performance, indicating cases where they showed good and poor results in section 5.3. Afterwards, the computed complexity indices are summarized and assessed in terms of their evolution throughout the stepwise-CA in section 5.4. Finally, the key findings are presented in section 5.5, where a negative Pearson correlation between the CAGL and episodes of AF recurrence is observed. This promising result indicates that patients with more complex atrial activities experience relapses faster than patients who exhibit more organized activity, and it may support clinical guidelines.

5.1 AF Database

A group of $n_p = 20$ persistent AF patients was enrolled in this study. Written informed consent was obtained from each participant in accordance with the principles outlined in the Declaration of Helsinki (World Medical Association, 2013). Their baseline characteristics are reported in Table 5.1. All patients underwent stepwise CA at the Cardiology Department of Princess Grace Hospital Center, Monaco. CA ended in procedural AF termination in all cases. During the intervention the ECG was continuously acquired at a sampling rate of 977 Hz.

5.2 Experimental Setup

To standardize the assessment of AF complexity evolution throughout the ablation procedure, we propose a framework dividing the intervention into three stages:

- Initial: represents the baseline, before the procedure.
- Intermediate: after each step including the first and penultimate steps.
- Outcome: after the last CA step.

Our patient population has a total of 59 ECG records, which are grouped according to their respective stages. Each ECG is labeled according to its stage (Fig. 5.1). Depending on the number of steps performed in the CA procedure, each patient may have between

Table 5.1 – Characteristics of the study population: The analysis included a cohort of 20 male patients, among whom seven had arterial hypertension and one patient had diabetes. The symbols μ and σ represent the mean and standard deviation, respectively. Missing values (indicated by symbols ^{a} and ^{b}) were present in two and one instance, respectively, and were imputed with the mean.

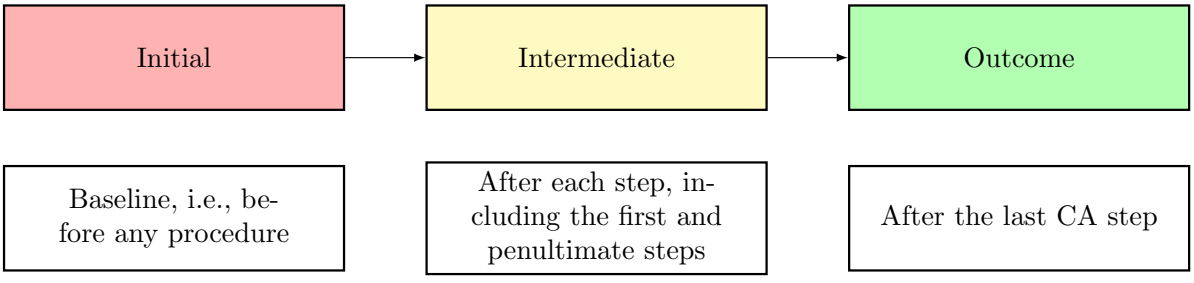
Characteristic	$\mu \pm \sigma$	Units
Age	60.8 ± 9.5	years
Weight	85.6 ± 12.7	kg
Height	177.7 ± 5.9	cm
AF History	68.6 ± 59.6	months
Current AF Episode[a]	14.2 ± 9.8	months
Left atrium diameter[b]	44.3 ± 7.4	mm
Left atrium surface[b]	28.0 ± 5.4	cm^2
Left ventricular ejection fraction (LVEF)	58.9 ± 13.5 (%)	-

Source: Elaborated by the author.

two and five ECG records. For example, a patient with two ECG records would have one initial and one outcome stage; a patient with five ECG records would have one initial, three intermediate, and one outcome stage. All ECG records are preprocessed using a zero-phase forward-backward type-II Chebyshev bandpass filter with cutoff frequencies of 0.5 and 40 Hz to suppress high-frequency noise and baseline wandering.

The simulations were conducted on Intel(R) Core(TM) i7-10510U CPU @ 1.80GHz/2.30 GHz with 20.0 GB of RAM.

Figure 5.1 – The number of steps performed for each patient can vary significantly, depending on stepwise CA progression. To standardize the assessment, we present a framework to organize each ECG record into one of the following stages.



Source: Elaborated by the author.

5.2.1 NDI

Following the guidelines presented in (MEO *et al.*, 2018), we apply TQ-segmentation using fixed 12-second ECG segments per ablation stage to compute the NDI, as the method requires sufficiently long segments to produce consistent results. First, to detect accurately the segment TQ intervals, we begin by performing The Pan-Tompkins algorithm (PAN; TOMPKINS, 1985) to locate the R-wave time instants in lead V1. For the Q-wave onsets, we subtract by 40 ms from the R peak, representing a typical ventricular activation time, as in (ZARZOSO *et al.*, 2016). To estimate the T-wave offsets, we perform an adapted Woody's method (CABASSON; MESTE, 2008). Finally, the segments were mean-centered and concatenated. For the implementation of the NDI, Algorithm 1 presents pseudocode outlining the steps involved.

In section 5.3 we illustrate algorithm performance to compute the complexity index.

Algorithm 1: NDI Complexity Index Pseudocode

Require: ECG matrix $\mathbf{Y}$
$r = 2$
$t_{\text{NDI}} = 0$
while $t_{\text{NDI}} < 12$ s **do**
    R-wave$^{(r)}$ $\leftarrow$ PanTompkins algorithm in lead V1
    Q-wave$^{(r-1)}$ $\leftarrow$ R-wave$^{(r)}$ - 40 ms
    T-wave$^{(r-1)}$ $\leftarrow$ offsets were estimated by an adapted Woody's method
    r = r + 1;
    Segments mean centred
    X $\leftarrow$ Concatenate segmented intervals
end while
$\Sigma \leftarrow$ Compute PCA(X)
$\text{NDI} = 1 - \dfrac{\sum_{r=1}^{3} \sigma_r^2}{\sum_{r=1}^{R} \sigma_r^2}$
Ensure: NDI.

5.2.2 CAGL

For CAGL, after each ablation step of each patient, we find the QQ segment with the largest TQ interval in the ECG, and map it into a data tensor as an input to the CAGL method. Segments length range from 0.72 to 1.42 seconds (1.06 ± 0.20 s).

We downsample the signals by a factor of 10 before applying the tensor de-

composition in order to reduce its computing time, with no information loss since the frequency content of interest in AA is below 10 Hz (HOLM *et al.*, 1998).

Example: an observation window with length 1.06 s yields 1037 samples and a direct row-Hankelization of this matrix results in a tensor of dimensions $519 \times 519 \times 12$, whose approximate BTD demands a large computing time. Downsampling leads to a resulting tensor $\mathcal{Y}$ with dimensions $52 \times 53 \times 12$.

CAGL is applied to the ECG recording after each CA step with a γ-sweeping procedure, inspired by solution path techniques (HASTIE *et al.*, 2015; GOULART *et al.*, 2020), by taking 50 equispaced values in the interval $[8 \times 10^{-4}, 0.5 \times 10^{-2}]$ and keeping the last solution. We start the algorithm with $R = 6$ random blocks and rank $L_r = 40$, $r = 1, 2, \ldots, R$ as initial guess (GOULART *et al.*, 2020). The task of measuring estimation quality is challenging since the ground truth is unknown, as discussed previously. We take into account the parameters used to evaluate AA extraction as detailed in chapter 4. For the present work, we consider that only one AA source exists and the selection is based on which source maximizes the SC.

The algorithm 2 shows a pseucode for the CAGL index implementation.

Algorithm 2: CAGL Complexity Index Pseudocode

Require: Data tensor $\mathcal{Y}$, penalty parameter γ, proximal term weight τ, initial point $\mathbf{A}^{(0)}, \mathbf{B}^{(0)}, \mathbf{C}^{(0)}$

t = 1;
while stopping criteria not met **do**
  Estimate $\mathbf{A}^{(t)}$ from $\mathbf{A}^{(t-1)}$, $\mathbf{B}^{(t-1)}$ and $\mathbf{C}^{(t-1)}$ via group lasso
  Estimate $\mathbf{B}^{(t)}$ from $\mathbf{A}^{(t)}$, $\mathbf{B}^{(t-1)}$ and $\mathbf{C}^{(t-1)}$ via group lasso
  for $r = 1, \ldots, R$ **do**
    $L_r^{(t)} \leftarrow \text{rank}\left(\mathbf{A}_r^{(t)}(\mathbf{B}_r^{\mathsf{T}})\right)$
    $(\mathbf{A}_r^{(t)}, \mathbf{B}_r^{(t)}) \leftarrow$ Cadzow algorithm to enforce Hankel structure
  end for
  Estimate $\mathbf{C}^{(t)}$ from $\mathbf{A}^{(t)}$, $\mathbf{B}^{(t)}$ and $\mathbf{C}^{(t-1)}$
end while
$\mathbf{H}_{r_{\text{AA}}} \leftarrow$ Select Atrial Source (r_{AA})
$L_{r_{\text{AA}}} \leftarrow \text{rank}\,(\mathbf{H}_{r_{\text{AA}}})$

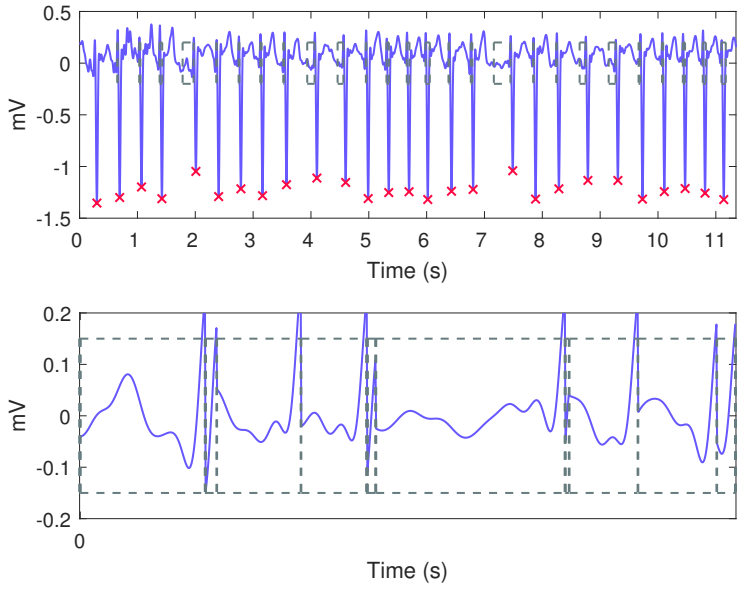
Ensure: $L_{r_{\text{AA}}}$, $\hat{\mathbf{A}}$, $\hat{\mathbf{B}}$, $\hat{\mathbf{C}}$

5.3 Real ECG Scenario: AF Complexity

We present four patients results to illustrate both algorithms' performance, indicating the cases where some constraints may impair the quality of the complexity estimation. The x symbol present in the patients A and B exams indicate the detected peaks, i.e, R-waves.

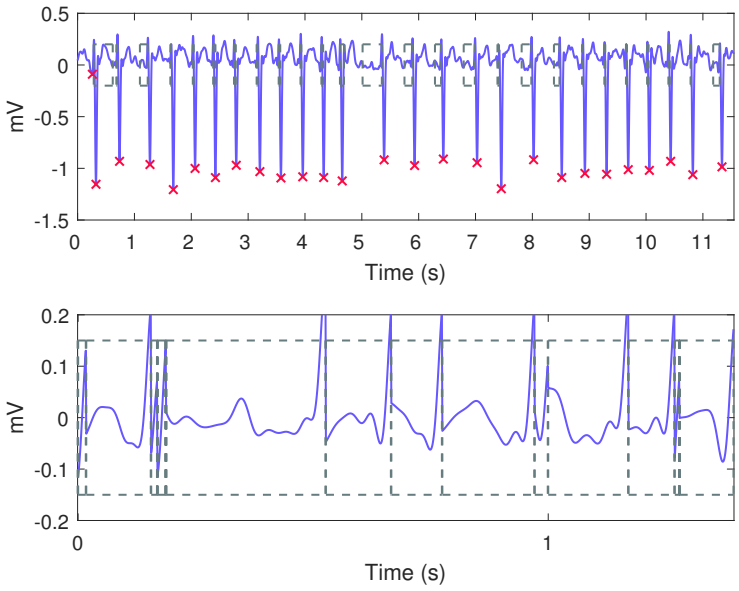
Patient A: The patient presents abnormally short QT intervals in the ECG at all the CA stages. However, the most severe case is shown in Fig. 5.2, at the initial stage. This shortened QT interval hampers the TQ-segmentation, therefore, the NDI algorithm fails to estimate the AA complexity in this case, whereas the CAGL estimates the Hankel matrix rank equal to 15. NDI is able to estimate the subsequent steps, as shown in figures 5.3 and 5.4. Normally, the QT intervals are expected to have a specific minimal duration. However, in some cases, they may appear shorter than usual. This observation can indicate various possibilities, ranging from a normal variation within the patient's cardiac physiology to certain medical conditions, abnormalities, medication effects, or an underlying heart condition. Further evaluation by a cardiologist is necessary to determine the exact cause in the specific patient's case.

Figure 5.2 – **Patient A:** Initial stage - NDI fails to measure complexity. Top plot: ECG Lead V1 presenting abnormally short QT intervals and the detected R-waves. Bottom plot: Hindering the TQ-segmentation. CAGL is 15.



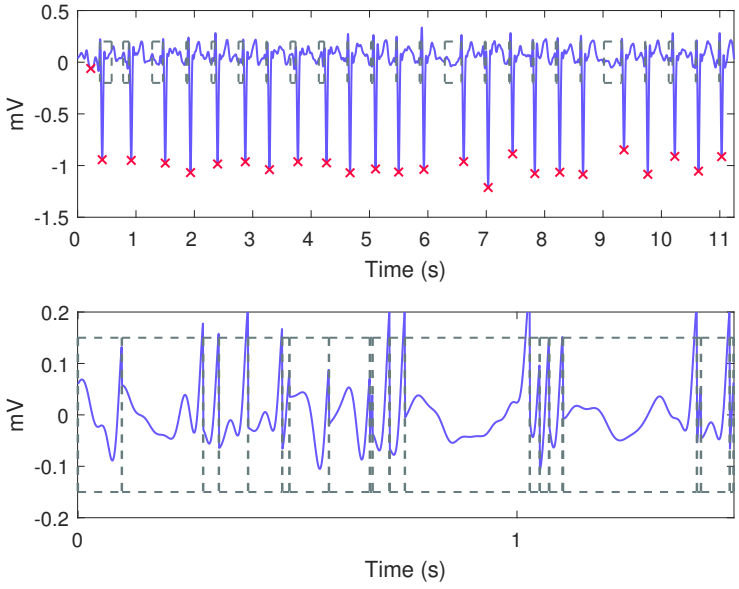
Source: Elaborated by the author.

Figure 5.3 – **Patient A:** Intermediate stage - Even with the abnormally short QT intervals, we compute the TQ-segmentation successfully. Top plot: ECG Lead V1 presenting abnormally short QT intervals and the detected R-waves. Bottom plot: The concatenation result for the Lead V1. NDI = 0.0088, CAGL = 11.



Source: Elaborated by the author.

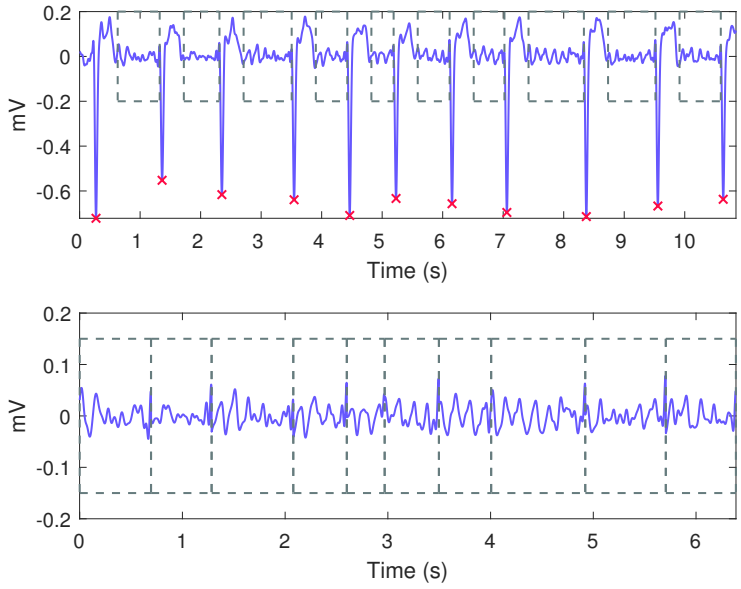
Figure 5.4 – **Patient A:** Outcome stage - Even with the abnormally short QT intervals, we compute the TQ-segmentation successfully. Top plot: ECG Lead V1 presenting abnormally short QT intervals and the detected R-waves. Bottom plot: The concatenation result for the Lead V1. NDI = 0.0072, CAGL = 11.



Source: Elaborated by the author.

Patient B: The patient presents normal QT intervals, what makes the TQ-segmentation easier to compute. Both NDI and CAGL algorithms are able to estimate the AA complexity for all stages (The values are indicated on the Figs. 5.5 and 5.6). We can see that as the ablation was performed, the complexity index indicated a decrease on the AA complexity.

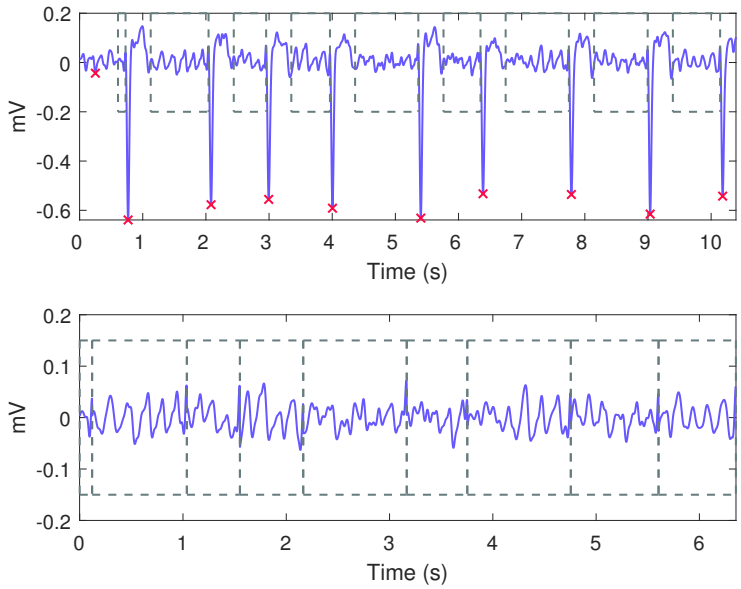
Figure 5.5 – **Patient B:** Initial stage - the exam presents normal QT intervals and we compute the TQ-segmentation successfully. Top plot: ECG Lead V1 presenting normal QT intervals with visible AF signal and the detected R-waves. Bottom plot: The concatenation result for the Lead V1. NDI = 0.00542, CAGL = 25.



Source: Elaborated by the author.

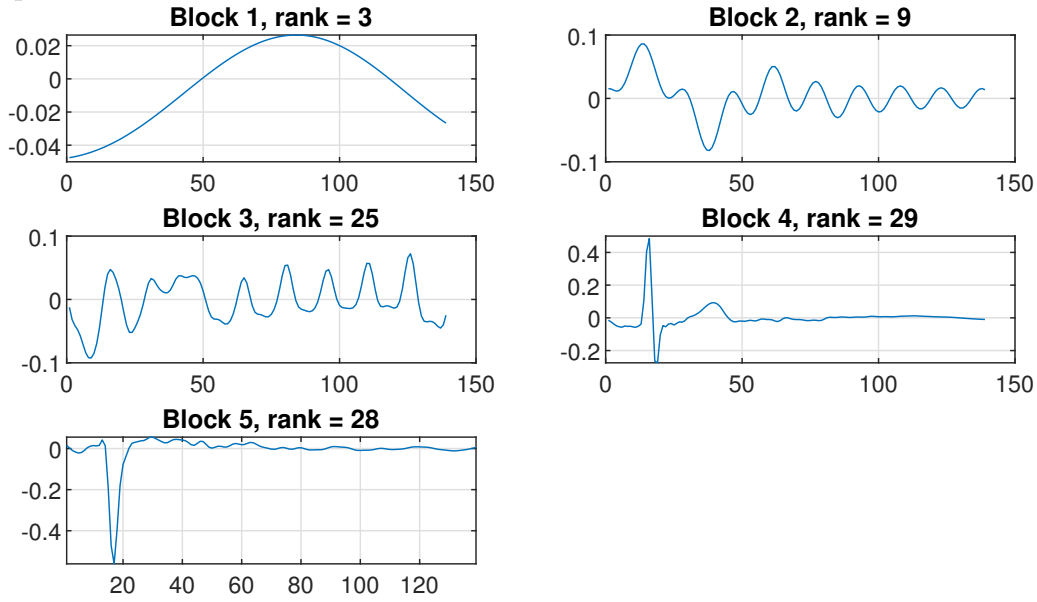
Patient C: The CAGL estimates 5 sources at the initial stage, where the 3rd block presents the AA signal, as shown in Fig. 5.7. The comparison between the estimated source and the ECG Lead V1 illustrates how accurate the decomposition is performed (see Fig. 5.8). In the outcome stage, 4 sources are estimated, where the 4th block presents the AA signal, as shown in Fig. 5.9. Once more, the plot of Fig. 5.10, the estimated source vs. the ECG Lead V1, illustrates how accurate the decomposition is performed.

Figure 5.6 – **Patient B:** Outcome stage - the exam presents normal QT intervals and we compute the TQ-segmentation successfully. Top plot: ECG Lead V1 presenting normal QT intervals with visible AF signal and the detected R-waves. Bottom plot: The concatenation result for the Lead V1. NDI = 0.029, CAGL = 6.



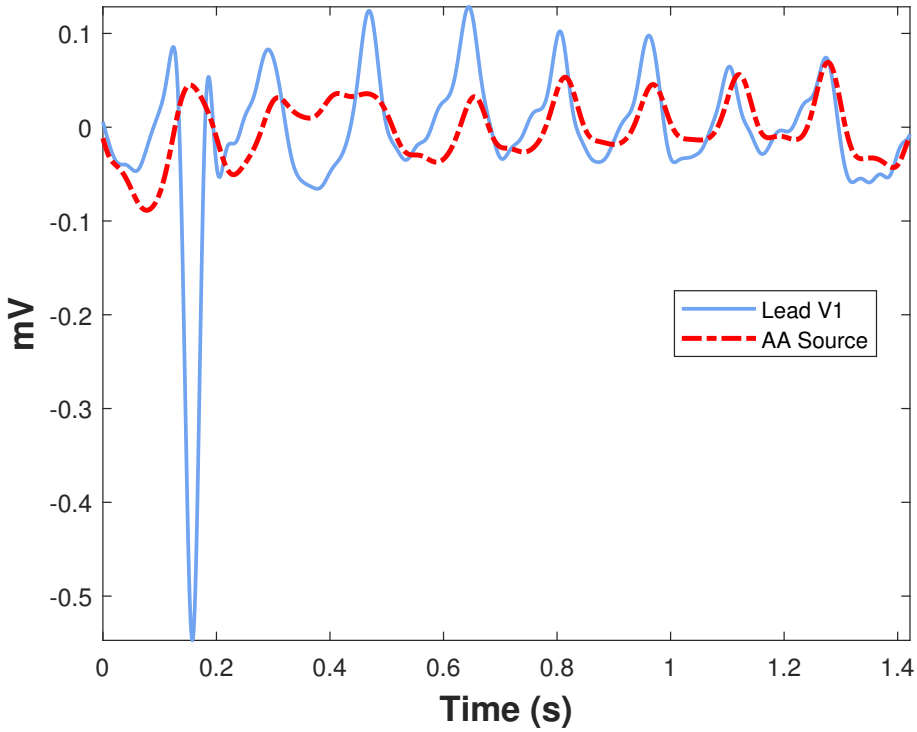
Source: Elaborated by the author.

Figure 5.7 – **Patient C:** Initial Stage - each plot represents a source estimated by CAGL. AA is present 3rd block.



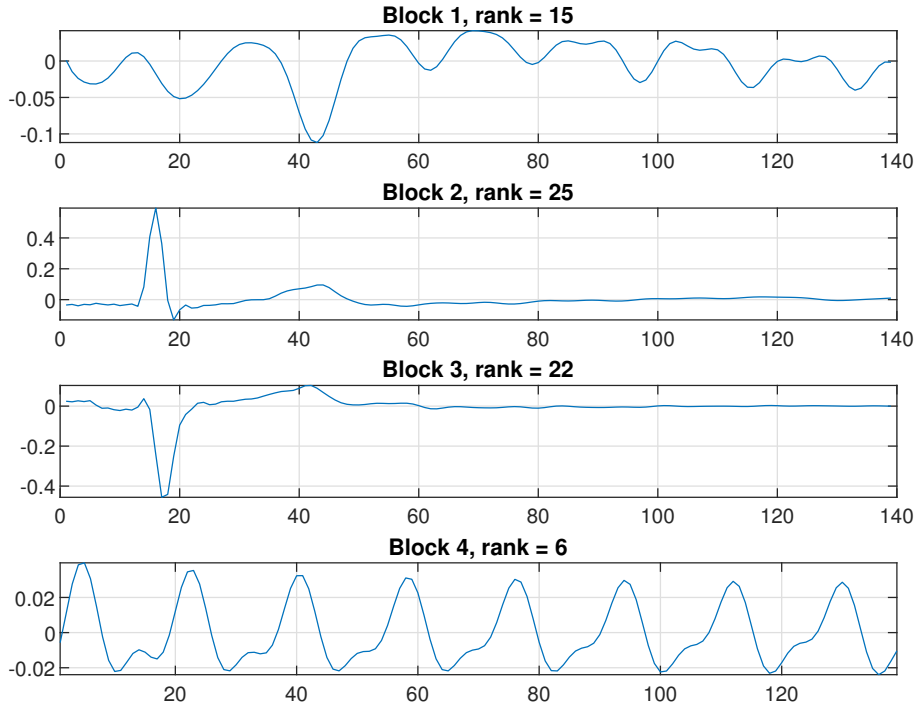
Source: Elaborated by the author.

Figure 5.8 – **Patient C:** Initial Stage - ECG Lead V1 presenting a heartbeat and the estimated AA source. Parameters: NDI = 0.0542, CAGL = 25, SC = 63.46, DF = 6.20, P(r) = 1.07e-03, kurtosis = 126.63.



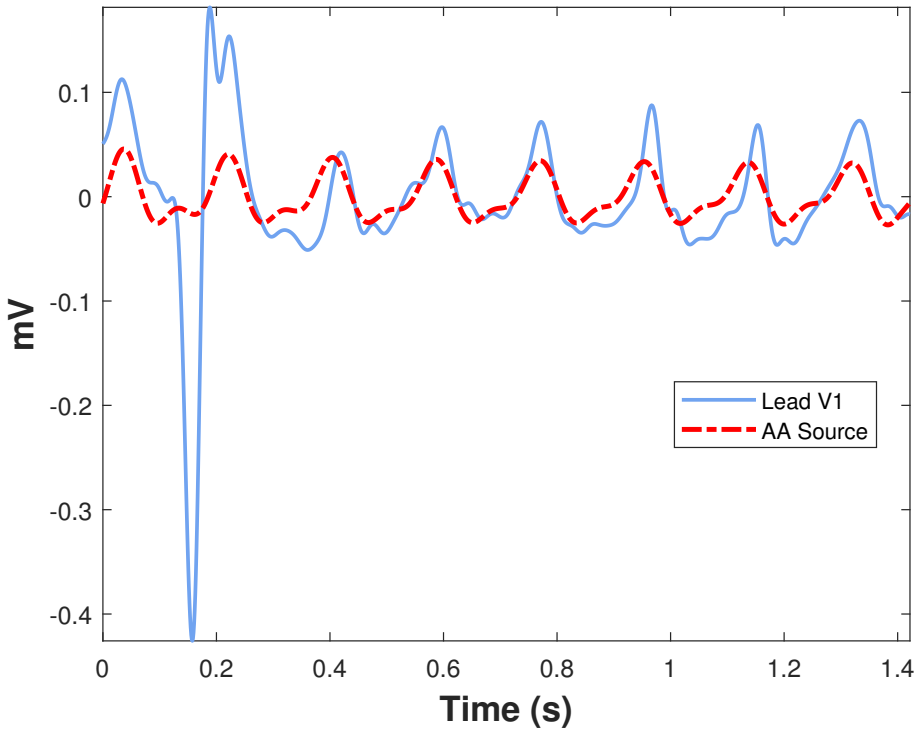
Source: Elaborated by the author.

Figure 5.9 – **Patient C:** Outcome Stage - CAGL estimates 4 sources, AA is present 4th block.



Source: Elaborated by the author.

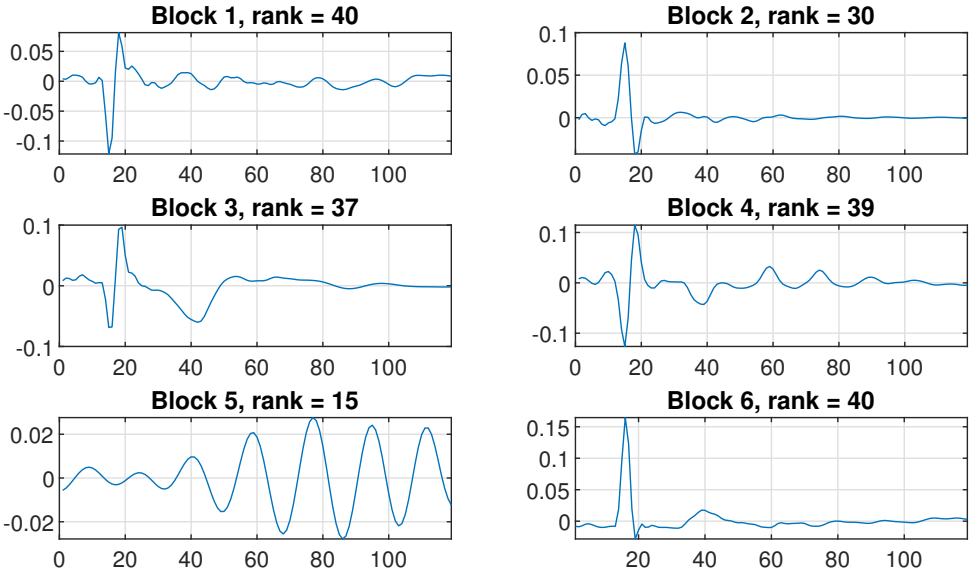
Figure 5.10 – **Patient C:** Outcome Stage - ECG Lead V1 presenting a heartbeat and the estimated AA source. Parameters: NDI = 0.0282, CAGL = 6, SC = 81.52, DF = 5.49, P(r) = 3.77e-04, kurtosis = 333.52.



Source: Elaborated by the author.

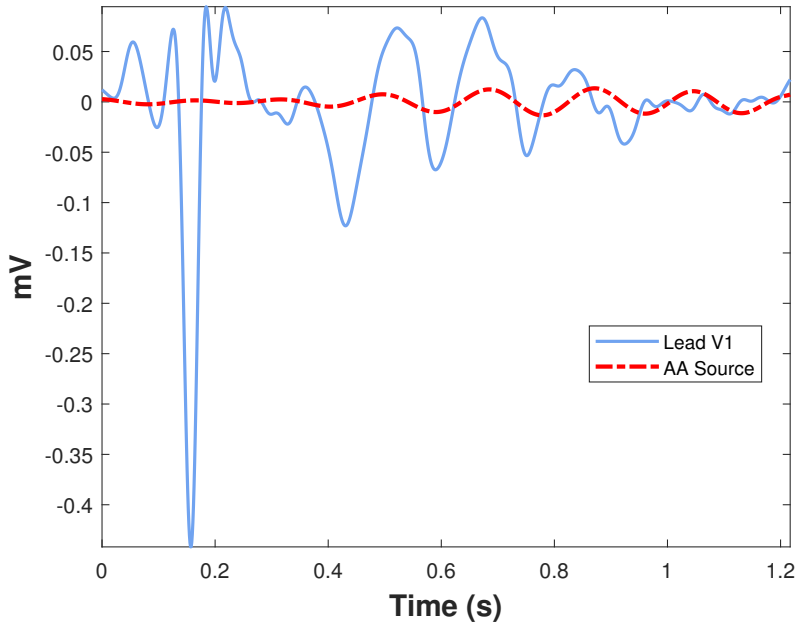
Patient D: The CAGL estimates 6 sources at the initial stage, where the 5th block presents the AA signal, as shown in Fig. 5.11 In the present case, both NDI and CAGL algorithms can estimate AA complexity in the intermediate stage. However, CAGL did not present a good performance to estimate the source, see Fig. 5.12. Although the visual representation deviates from the expected instinctive pattern of FA, the other parameters and even the AA signal representation in frequency domain (Fig. 5.13) reveals results commonly associated with atrial source.

Figure 5.11 – **Patient D:** Intermediate Stage - CAGL estimates 6 sources, AA is present 5th block.



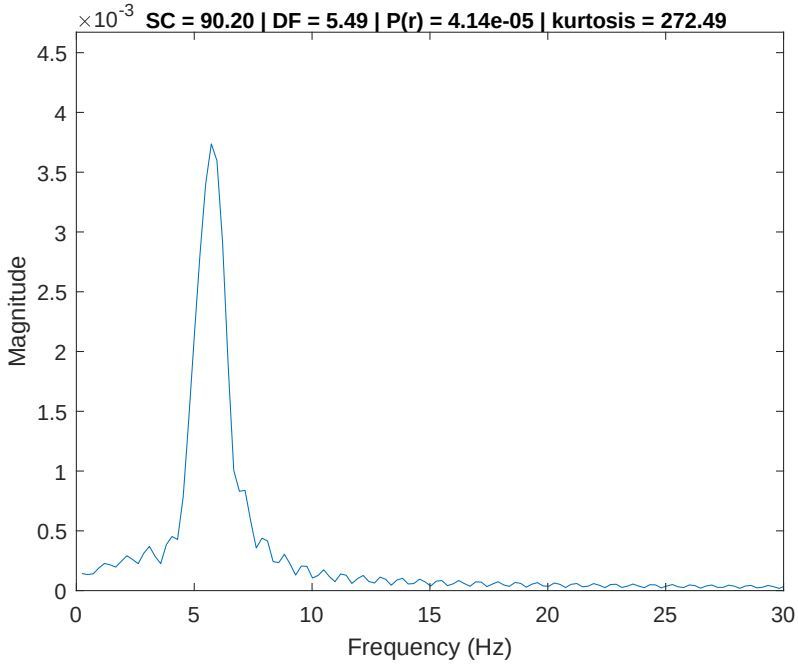
Source: Elaborated by the author.

Figure 5.12 – **Patient D:** Intermediate Stage - ECG Lead V1 presenting a heartbeat and the estimated AA source. Parameters: NDI = 0.0529, CAGL = 15, SC = 90.20, DF = 5.49, P(r) = 4.14e-05, kurtosis = 272.49.



Source: Elaborated by the author.

Figure 5.13 – **Patient D:** Intermediate Stage - AA estimated signal in the frequency domain acquired by a 4096-point Fast Fourier Transform. The indicated parameters and the plot presents AA source characteristics.



Source: Elaborated by the author.

5.4 Complexity Evolution During CA

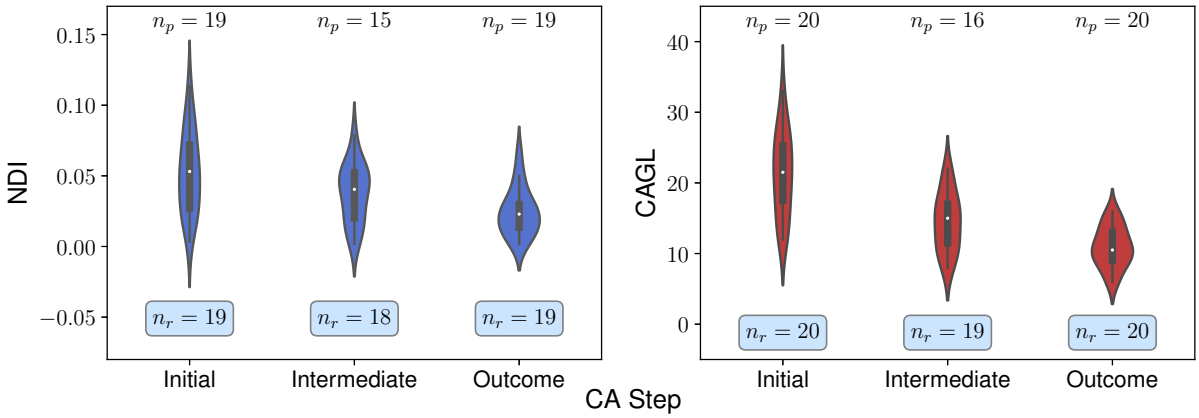
To assess the impact of ablation on AF complexity, Fig. 5.14 presents violin plots of the computed indices grouped by stage for the two non-invasive indices. The violin plot is a data visualization that combines aspects of a box plot and kernel density plot. It displays the distribution of a continuous variable or numeric data across categories or groups. The width of each violin corresponds to the data frequency in that category. A white dot or central line represents the median. The width of the violin at any point indicates the density or concentration of data. Wider sections indicate higher density, while narrower sections indicate lower density. This plot facilitates intuitive comparison of variable distributions across groups, revealing central tendency, spread, and skewness. It aids in identifying group differences, similarities, outliers, and unusual patterns.

The evolution of complexity for both NDI and CAGL is shown in Fig. 5.14. The indices decrease throughout the stages of cardiac ablation: initial (before ablation), intermediate (after each step, including the first and penultimate steps), and outcome (after the last step).

This assessment includes the entire population under study, comprising 20 patients. The symbols n_p and n_r represent the number of patients and ECG records contributing to each stage, respectively. It is important to highlight that NDI failed to estimate the complexity of three ECG segments, as indicated by the smaller n_r values shown in the violin plot. As discussed in Section 5.3, patients with very short QT intervals or other cardiac abnormalities that impair TQ-segmentation may result in performance loss for NDI.

In contrast to CAGL, which can compute the decomposition for all 20 patients, it is important to emphasize that both complexity indices decrease throughout the CA procedure, reaching their lowest values at the end of the intervention (outcome). While both methods demonstrate a trend towards more organized signals as the procedure approaches AF termination, CAGL exhibits a notable advantage with a more significant decrease in complexity per stage

Figure 5.14 – Evolution of NDI (left) and CAGL (right) complexity indices along CA stages: initial (before ablation), intermediate (after each step including first and penultimate steps); outcome (after the last step). Symbols n_p and n_r represent the number of patients and ECG records contributing to each stage, respectively.



Source: Elaborated by the author.

5.5 AF Recurrence

To assess the relationship between the features of the population, we compute the data correlation matrix. The Pearson correlation coefficients (r) are arranged as a heatmap for all available features.

A correlation matrix is a table that summarizes the strength and direction of relationships between multiple variables. Each cell in the matrix represents the correlation coefficient between two variables, ranging from -1 to $+1$. Negative values indicate a negative correlation, positive values indicate a positive correlation, and 0 indicates no correlation.

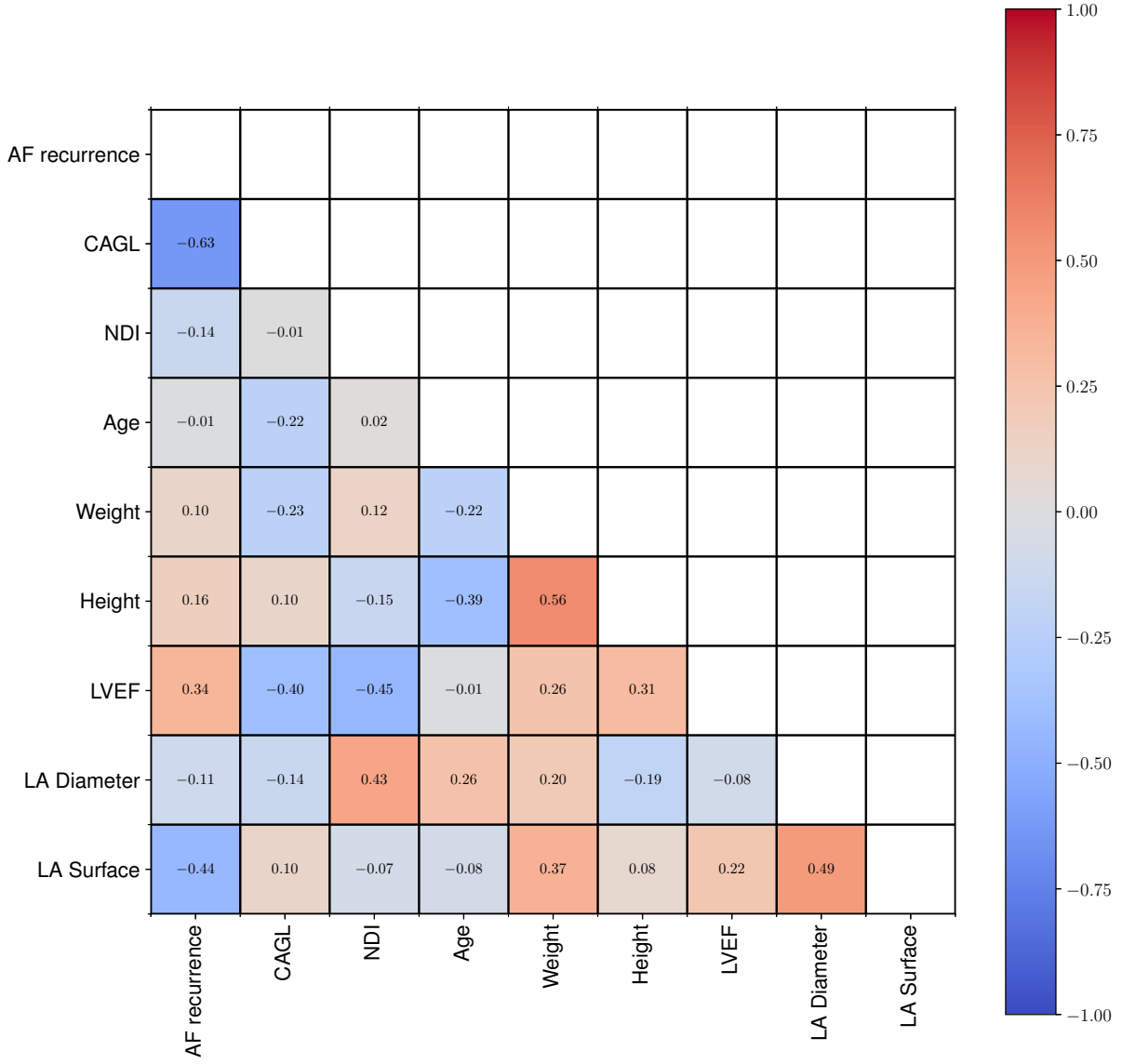
A heatmap is a graphical representation of the correlation matrix using color gradients. It visually displays the correlation patterns between variables, with colors representing the strength and direction of the correlation. Cool colors like blue represent negative correlations, while warm colors like red represent positive correlations.

Correlation matrices and their heatmaps are valuable tools for analyzing relationships and identifying patterns in data. They help identify variables that are strongly or weakly correlated and reveal significant positive or negative relationships among variables.

The correlation matrix heat map shown in Fig. 5.15 shows that the correlation between the initial rank estimated by CAGL, i.e., before patients had undergone any CA procedures, and AF recurrence, i.e., the time that each patient remained in sinus

rhythm before AF relapse is −0.63. A statistical relevant value of negative correlation well illustrated by its scatter plot and linear regression in Fig. 5.17 seems to indicate an influence of initial rank on AF recurrence.

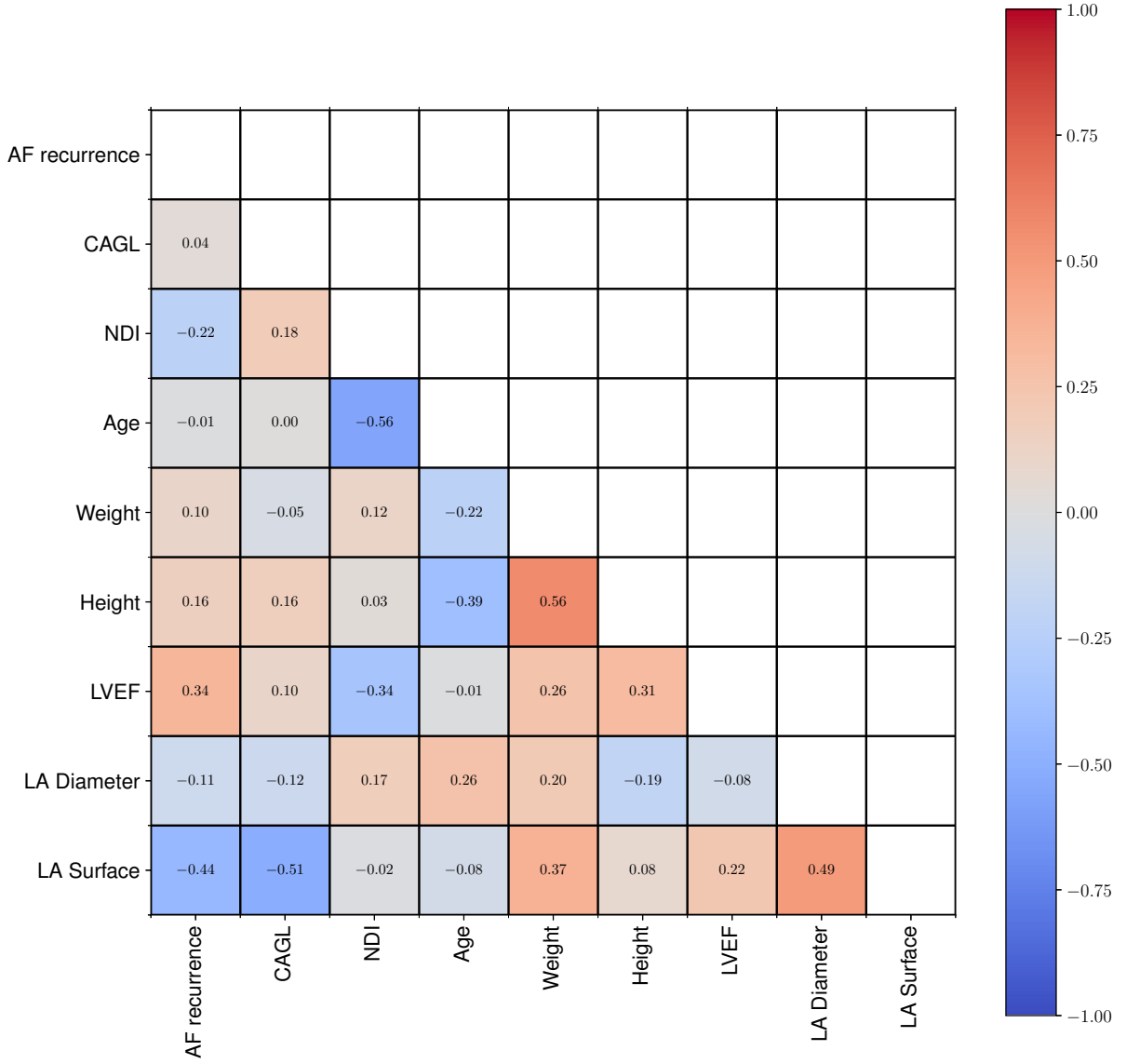
Figure 5.15 – Correlation matrix plot comparing the complexity indices at the initial stage.



Source: Elaborated by the author.

Neither complexity index shows a relevant correlation with the outcome stage results, as indicated in Fig. 5.16.

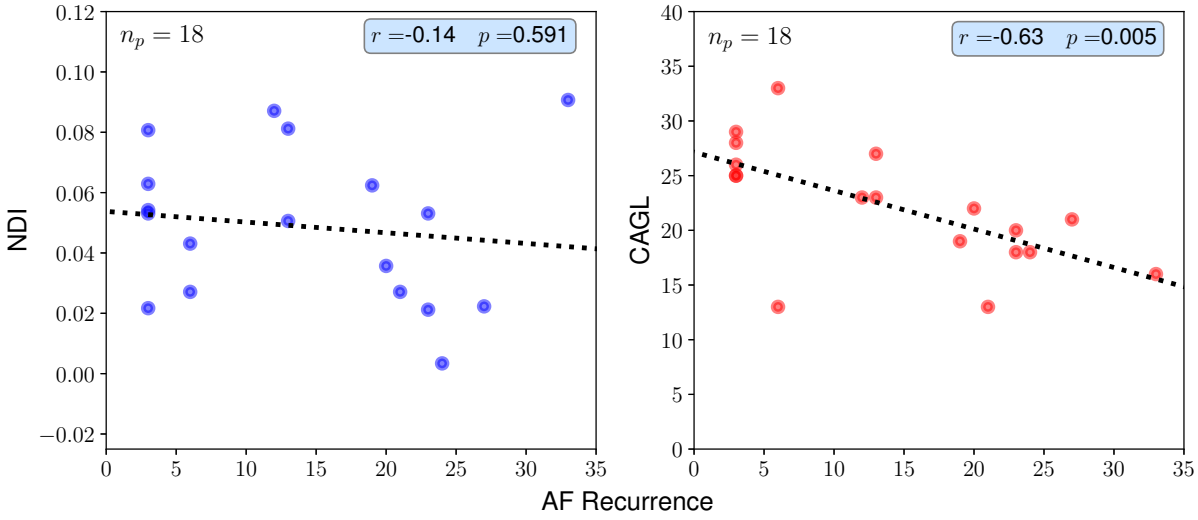
Figure 5.16 – Correlation matrix plot comparing the complexity indices at the outcome stage.



Source: Elaborated by the author.

We are interested in the relationship between pre-ablation AA complexity and AF recurrence, which refers to the time elapsed between the end of the procedure and the recurrence of the arrhythmia. In the scatter plots of Fig. 5.17, CAGL exhibits a statistically significant Pearson correlation ($r = -0.63$, $p = 0.005$). This negative correlation is consistent with the concept of complexity: the more complex the AF is before ablation, the shorter the period during which the patient remains free of arrhythmia before recurrence after ablation. On the other hand, the matrix-based index does not show a significant correlation ($r = -0.14$, $p = 0.591$).

Figure 5.17 – Linear Regression



Source: Elaborated by the author.

Finally, it is important to note that two patients lacked information regarding AF recurrence after the CA procedure, as they dropped out of the study. Consequently, they were excluded from this assessment. For these patients, CAGL complexity indices were 12 and 18, whereas NDI complexity index was 0.1135 for the first patient, but it was not available for the second patient.

6 CONCLUSION

This work introduces a novel complexity index for noninvasive measurement of AF complexity utilizing tensor techniques. The proposed index is based on the CAGL algorithm, which employs Hankel-based BTD. This algorithm enables the simultaneous extraction of the AA signal and quantification of AF complexity, even from very short ECG recordings (1.06 ± 0.20 s).

Comparisons with the NDI complexity index revealed limitations in the accuracy of this matrix-based method. Nevertheless, both were able to characterize AA activity, particularly in the cases for ECG recordings with short QT intervals. This is due to the TQ-concatenation technique, which the Hankel-BTD model does not depend on, what becomes an advantage to the tensor-based index in very complex/disorganized scenarios.

The conducted experiments on a database comprising 20 patients with persistent AF highlighted the impact of each step in stepwise-CA on AF complexity. It was observed that the CAGL algorithm exhibited a more significant and noticeable decrease in complexity throughout the ablation procedure, indicating its potential as an effective tool for quantifying changes in AA complexity during ablation.

The assessment of the correlation between pre-ablation complexity and AF recurrence revealed a statistically significant negative correlation for the CAGL algorithm. This suggests that patients with higher complexity in their AA before ablation experienced a shorter duration of freedom from arrhythmia before AF recurrence. The NDI algorithm, however, did not exhibit a significant correlation in this regard.

This finding suggests that the rank parameter derived from the BTD could potentially enhance the outcomes of CA procedures and improve the patient care. Overall, our findings suggest that the tensor-based index has the potential to assist in clinical decisions, as it presents a correlation with CA success before the intervention. In addition, the CAGL approach has the potential to guide CA in real-time by quantifying the overall AF complexity and organization, and could prove useful to define more effective ablation protocols.

Overall, this work highlights the potential of complexity indices, particularly the CAGL algorithm, in assessing AF complexity and predicting AF recurrence after ablation. However, further research and validation are necessary to better understand and utilize these complexity indices in clinical practice. The findings of this study contribute

to the growing body of knowledge in the field of AF complexity assessment and may have implications for improving patient outcomes and optimizing treatment strategies for AF patients undergoing ablation procedures.

6.1 Main Contributions

The tensor approach is shown to:

1. evolve in coherence with the decrease in complexity expected along the different steps of the intervention;
2. correlate significantly with AF recurrence after the intervention. In addition, the proposed index can be computed from ECG records as short as a single heartbeat, thus offering the possibility of noninvasively monitoring catheter ablation in real time.

An outcome of this work resulted on a paper named "*Décomposition tensorielle en termes blocs contrainte pour la mesure non invasive de la complexité de la fibrillation atriale persistante*" in the **XXIXème Colloque Francophone de Traitement du Signal et des Images (GRETSI 2023)**, accepted for publication and presentation.

6.2 Further Work

This study acknowledges the challenges posed by the limited number of patients and the heterogeneity of steps across CA intervention. To provide a more comprehensive step-by-step complexity assessment, future work should focus on increasing the number of patients and implementing a more homogeneous CA approach for all participants.

Comparing the tensor-based index with other state-of-the-art algorithms is very relevant to reinforce its relevance among the methods that extract or measure AF complexity.

It is relevant to consider scenarios that involve multiple atrial sources in order to address the limitations of the present model, which assumes a single source. Ignoring the presence of multiple sources may result in the omission of relevant information. By addressing this limitation and accounting for multiple atrial sources, it becomes possible to achieve a more accurate estimation of the AA and a comprehensive measurement of AF complexity. This improvement is crucial for enhancing our understanding of the underlying

mechanisms and dynamics of AF.

Explore other relevant factorization and optimization methods: e.g., perform the CAGL with the Alternating Direction Method of Multipliers (ADMM), that solves convex optimization problems by breaking them into smaller pieces. It has recently found wide application in a number of areas.

Additionally, incorporating machine learning techniques could enhance the analysis of AF complexity. Machine learning algorithms, particularly those specialized in pattern recognition, can be trained on an expanded dataset to identify subtle patterns and correlations that may not be evident through traditional analysis, especially given the limited number of patients.

By addressing these aspects, a more robust and detailed understanding of the complexity dynamics during the procedure can be achieved.

BIBLIOGRAPHY

ABDALAH, L. S.; OLIVEIRA, P. M. R. de; JR, W. F.; ZARZOSO, V. Tensor-based noninvasive atrial fibrillation complexity index for catheter ablation. In: **Proc. Computing in Cardiology**. Rimini, Italy: [*S. n.*], 2020. v. 47.

AMBESH, P.; KAPOOR, A. Biological pacemakers: Concepts and techniques. **The National Medical Journal of India**, v. 30, p. 324, 11 2017.

BARBARO, V.; BARTOLINI, P.; CALCAGNINI, G.; CENSI, F. Extraction of physiological and clinical information from intra-atrial electrograms during atrial fibrillation: review of methods. **Ann Ist Super Sanita**, v. 37, n. 3, p. 319–324, 2001.

BOLEY, D.; LUK, F.; VANDEVOORDE, D. A General Vandermonde Factorization of a Hankel Matrix? In: . Hong Kong: [*S. n.*], 1997. p. 10.

BONIZZI, P.; GUILLEM, M. de la S.; CLIMENT, A. M.; MILLET, J.; ZARZOSO, V.; CASTELLS, F.; MESTE, O. Noninvasive assessment of the complexity and stationarity of the atrial wavefront patterns during atrial fibrillation. **IEEE Transactions on Biomedical Engineering**, v. 57, n. 9, p. 2147–2157, 2010.

CABASSON, A.; MESTE, O. Time Delay Estimation: A New Insight Into the Woody's Method. **IEEE Signal Processing Letters**, v. 15, p. 573–576, 2008.

CADZOW, J. A. Signal enhancement-a composite property mapping algorithm. **IEEE Transactions on Acoustics, Speech, and Signal Processing**, v. 36, n. 1, p. 49–62, 1988.

CALKINS, H.; HINDRICKS, G. *et al.* 2017 HRS/EHRA/ECAS/APHRS/SOLAECE expert consensus statement on catheter and surgical ablation of atrial fibrillation: Executive summary. **Journal of Arrhythmia**, v. 33, n. 5, p. 369–409, 2017.

CARDOSO, J. Blind signal separation: statistical principles. **Proceedings of the IEEE**, v. 86, n. 10, p. 2009–2025, 1998.

CASTELLS, F.; RIETA, J.; MILLET, J.; ZARZOSO, V. Spatiotemporal blind source separation approach to atrial activity estimation in atrial tachyarrhythmias. **IEEE Transactions on Biomedical Engineering**, v. 52, n. 2, p. 258–267, 2005.

CIRUGEDA, E.; CALERO, S.; PLANCHA, E.; ENERO, J.; RIETA, J. J.; ALCARAZ, R. Multidimensional Characterization of the Atrial Activity to Predict Electrical Cardioversion Outcome of Persistent Atrial Fibrillation. In: . [*S. n.*], 2020. Disponível em: http://www.cinc.org/archives/2020/pdf/CinC2020-377.pdf.

DUPRE, A.; VINCENT, S.; IAIZZO, P. A. **Basic ECG Theory, Recordings, and Interpretation**. Totowa, NJ: Humana Press, 2005. 191–201 p.

FAVARATO, D. Brazilian Population Presents Prevalence of Atrial Fibrillation Similar to Higher Income Countries, and a Low Use of Anticoagulation Therapy. **Arquivos Brasileiros de Cardiologia**, v. 117, p. 435–436, set. 2021.

FAVIER, G. **Matrix and Tensor Decompositions in Signal Processing, Volume 2**. Volume 2 edition. Hoboken: Wiley-ISTE, 2021. ISBN 978-1-78630-155-0.

FREITAS, W. d. C.; FAVIER, G.; ALMEIDA, A. L. F. de; HAARDT, M. Two-Way MIMO Decode-and-Forward Relaying Systems with Tensor Space-Time Coding. In: **2019 27th European Signal Processing Conference (EUSIPCO)**. [*S. l.*: *s. n.*], 2019. p. 1–5.

GOLUB, G. H.; LOAN, C. F. V. **Matrix computations**. Fourth edition. Baltimore: The Johns Hopkins University Press, 2013. (Johns Hopkins studies in the mathematical sciences). OCLC: ocn824733531. ISBN 978-1-4214-0794-4.

GOULART, J. H. d. M.; OLIVEIRA, P. M. R. de; FARIAS, R. C.; ZARZOSO, V.; COMON, P. Alternating group lasso for block-term tensor decomposition and application to ECG source separation. **IEEE Transactions on Signal Processing**, v. 68, p. 2682–2696, 2020.

HAïSSAGUERRE, M.; JAïS, P.; SHAH, D. C.; TAKAHASHI, A.; HOCINI, M. *et al.* Spontaneous initiation of atrial fibrillation by ectopic beats originating in the pulmonary veins. **New England Journal of Medicine**, v. 339, n. 10, p. 659–666, 1998.

HAN, X.; ZHAO, X.; ALMEIDA, A. L. F. de; FREITAS, W. d. C.; BAI, W. Enhanced Tensor-Based Joint Channel and Symbol Estimation in Dual-Hop MIMO Relaying Systems. **IEEE Communications Letters**, v. 25, n. 5, p. 1655–1659, maio 2021.

HASTIE, T.; TIBSHIRANI, R.; WAINWRIGHT, M. **Statistical Learning with Sparsity: The Lasso and Generalizations**. New York: Chapman and Hall/CRC, 2015.

HOLM, M.; PEHRSON, S.; INGEMANSSON, M.; SöRNMO, L.; JOHANSSON, R.; SANDHALL, L.; SUNEMARK, M.; SMIDEBERG, B.; OLSSON, C.; OLSSON, S. Non-invasive assessment of the atrial cycle length during atrial fibrillation in man: introducing, validating and illustrating a new ECG method. **Cardiovascular Research**, v. 38, n. 1, p. 69–81, 04 1998.

HOLT, J. H.; BARNARD, A. C. L.; LYNN, M. S.; SVENDSEN, P. A study of the human heart as a multiple dipole electrical source. **Circulation**, v. 40, n. 5, p. 687–696, 1969.

JANUARY, C. T.; WANN, L. S.; CALKINS, H.; CIGARROA, J. E. *et al.* 2019 AHA/ACC/HRS focused update of the 2014 AHA/ACC/HRS guideline for the management of patients with atrial fibrillation. **Journal of the American College of Cardiology**, v. 74, n. 1, p. 104–132, 2019.

JOLLIFFE, I. **Principal component analysis**. New York: Springer Verlag, 2002. (Springer Series in Statistics).

KOLDA, T. G.; BADER, B. W. Tensor Decompositions and Applications. **SIAM Review**, v. 51, n. 3, p. 455–500, ago. 2009.

LATHAUWER, L. D. Decompositions of a higher-order tensor in block terms – Part II: Definitions and uniqueness. **SIAM Journal on Matrix Analysis and Applications**, v. 30, n. 3, p. 1033–1066, 2008.

LATHAUWER, L. D. Blind separation of exponential polynomials and the decomposition of a tensor in rank-$(l_r, l_r, 1)$ terms. **SIAM Journal on Matrix Analysis and Applications**, v. 32, n. 4, p. 1451–1474, 2011.

LATHAUWER, L. D.; NION, D. Decompositions of a Higher-Order Tensor in Block Terms—Part III: Alternating Least Squares Algorithms. **SIAM Journal on Matrix Analysis and Applications**, v. 30, n. 3, p. 1067–1083, jan. 2008. ISSN 0895-4798. Publisher: Society for Industrial and Applied Mathematics. Disponível em: https://epubs.siam.org/doi/abs/10.1137/070690730.

LUX, R. L.; EVANS, A. K.; BURGESS, M. J.; WYATT, R. F.; ABILDSKOV, J. A. Redundancy reduction for improved display and analysis of body surface potential maps. i. spatial compression. **Circulation Research**, v. 49, n. 1, p. 186–196, 1981.

MAINARDI, L.; SöRNMO, L.; CERUTTI, S. **Understanding Atrial Fibrillation: The Signal Processing Contribution**. [*S. l.*]: Morgan & Claypool Publishers, 2008.

MAINARDI, L.; SöRNMO, L.; CERUTTI, S. **Understanding Atrial Fibrillation: The Signal Processing Contribution, Part II**. Cham: Springer International Publishing, 2008. (Synthesis Lectures on Biomedical Engineering).

MALMIVUO, J.; PLONSEY, R. **Bioelectromagnetism – Principles and Applications of Bioelectric and Biomagnetic Fields**. New York: Oxford University Press, 1995.

MARCO, L. Y. D.; BOURKE, J. P.; LANGLEY, P. Spatial complexity and spectral distribution variability of atrial activity in surface ECG recordings of atrial fibrillation. **Medical Biol. Eng. Comput.**, v. 50, n. 5, p. 439–446, 2012.

MARCOLINO, M. S.; PALHARES, D. M.; BENJAMIN, E. J.; RIBEIRO, A. L. Atrial fibrillation: prevalence in a large database of primary care patients in Brazil. **Europace**, v. 17, n. 12, p. 1787–1790, dez. 2015.

MCCANN, A.; LUCA, A.; PASCALE, P.; PRUVOT, E.; VESIN, J.-M. Novel spatiotemporal processing tools for body-surface potential map signals for the prediction of catheter ablation outcome in persistent atrial fibrillation. **Frontiers in Physiology**, v. 13, 2022.

MCCANN, A.; VESIN, J.-M.; PRUVOT, E.; ROTEN, L.; STICHERLING, C.; LUCA, A. ECG-based indices to characterize persistent atrial fibrillation before and during stepwise catheter ablation. **Frontiers in Physiology**, v. 12, 2021.

MEO, M.; HIDALGO-MUNOZ, A. R.; ZARZOSO, V.; MESTE, O.; LATCU, D. G.; SAOUDI, N. F-wave amplitude stability on multiple electrocardiogram leads in atrial fibrillation. In: **2015 Computing in Cardiology Conference (CinC)**. Nice: IEEE, 2015. p. 505–508.

MEO, M.; PABRUM, T.; DERVAL, N.; DUMAS-POMIER, C.; PUYO, S.; DUCHÂTEAU, J.; JAÏS, P.; HOCINI, M.; HAÏSSAGUERRE, M.; DUBOIS, R. Noninvasive assessment of atrial fibrillation complexity in relation to ablation characteristics and outcome. **Frontiers in Physiology**, v. 9, n. 929, 2018.

MEO, M.; ZARZOSO, V.; MESTE, O.; LATCU, D. G.; SAOUDI, N. Spatial Variability of the 12-Lead Surface ECG as a Tool for Noninvasive Prediction of Catheter Ablation Outcome in Persistent Atrial Fibrillation. **IEEE Transactions on Biomedical Engineering**, v. 60, n. 1, p. 20–27, jan. 2013. ISSN 1558-2531.

Conference Name: IEEE Transactions on Biomedical Engineering. Disponível em: https://ieeexplore.ieee.org/document/6311437.

MORILLO, C. A.; BANERJEE, A.; PEREL, P.; WOOD, D.; JOUVEN, X. Atrial fibrillation: the current epidemic. **Journal of Geriatric Cardiology**, v. 14, n. 3, p. 195–203, mar. 2017.

NORAV MEDICAL. **Norav Users Guide PC-ECG 1200**. Mainz-Kastel, Germany: [*S. n.*], 2021. Available on: <https://www.noravmedical.com/wp-content/uploads/2023/02/PC-ECG-Operation-Manual.pdf>. Accessed on: 06/07/2023.

OLIVEIRA, P. M. R. de; GOULART, J. H. d. M.; FERNANDES, C. A. R.; ZARZOSO, V. Blind Source Separation in Persistent Atrial Fibrillation Electrocardiograms Using Block-Term Tensor Decomposition With Löwner Constraints. **IEEE Journal of Biomedical and Health Informatics**, v. 26, p. 1538–1548, abr. 2022. Conference Name: IEEE Journal of Biomedical and Health Informatics.

OLIVEIRA, P. M. R. de; ZARZOSO, V. Source analysis and selection using block term decomposition in atrial fibrillation. In: **14th International Conference on Latent Variable Analysis and Signal Separation (LVA/ICA)**. Guildford, UK: [*S. n.*], 2018. p. 46–56.

OLIVEIRA, P. M. R. de; ZARZOSO, V. Block term decomposition of ECG recordings for atrial fibrillation analysis: Temporal and inter-patient variability. **Journal of Communication and Information Systems**, v. 34, p. 111–119, 2019.

PAN, J.; TOMPKINS, W. J. A Real-Time QRS Detection Algorithm. **IEEE Transactions on Biomedical Engineering**, BME-32, n. 3, p. 230–236, mar. 1985.

PORTA, A.; BASELLI, G.; LOMBARDI, F.; MONTANO, N.; MALLIANI, A.; CERUTTI, S. Conditional entropy approach for the evaluation of the coupling strength. **Biological Cybernetics**, v. 81, n. 2, p. 119–129, ago. 1999.

RIBEIRO, L. N.; HIDALGO-MUñOZ, A. R.; FAVIER, G.; MOTA, J. C. M.; ALMEIDA, A. L. F. de; ZARZOSO, V. A tensor decomposition approach to noninvasive atrial activity extraction in atrial fibrillation ECG. In: **2015 23rd European Signal Processing Conference (EUSIPCO)**. [*S. l.*: *s. n.*], 2015. p. 2576–2580.

RIETA, J. J.; CASTELLS, F.; SáNCHEZ, C.; ZARZOSO, V.; MILLET, J. Atrial activity extraction for atrial fibrillation analysis using blind source separation. **IEEE Transactions on Biomedical Engineering**, v. 51, n. 7, p. 1176–1186, 2004.

RONTOGIANNIS, A. A.; KOFIDIS, E.; GIAMPOURAS, P. V. Block-Term Tensor Decomposition: Model Selection and Computation. **IEEE Journal of Selected Topics in Signal Processing**, v. 15, n. 3, p. 464–475, abr. 2021.

ROTTNER, L.; BELLMANN, B.; LIN TINA REISSMANN, B. *et al.* Catheter ablation of atrial fibrillation: State of the art and future perspectives. **Cardiology and Therapy**, v. 9, p. 45–58, 2020.

SANTOS, I. S.; LOTUFO, P. A.; BRANT, L.; PINTO, M. M.; PEREIRA, A. D. C.; BARRETO, S. M.; RIBEIRO, A. L.; THOMAS, G. N.; LIP, G. Y. H.; BENSENOR,

I. M. Diagnóstico de Fibrilação Atrial na Comunidade Utilizando Eletrocardiograma e Autorrelato: Análise Transversal do ELSA-Brasil. **Arquivos Brasileiros de Cardiologia**, v. 117, n. 3, p. 426–434, set. 2021.

SCHARF, L. L. **Statistical Signal Processing: Detection, Estimation, and Time Series Analysis**. 1st edition. ed. Reading, Mass: Pearson, 1991.

SEITZ, J.; BARS, C.; THéODORE, G.; BEURTHERET, S. *et al.* AF ablation guided by spatiotemporal electrogram dispersion without pulmonary vein isolation: A wholly patient-tailored approach. **Journal of the American College of Cardiology**, v. 69, n. 3, p. 303–321, 2017.

SIH, H. J. Measures of organization during atrial fibrillation. **Annali dell'Istituto Superiore Di Sanita**, v. 37, n. 3, p. 361–369, 2001.

SORBER, L.; BAREL, M. V.; LATHAUWER, L. D. Optimization-Based Algorithms for Tensor Decompositions: Canonical Polyadic Decomposition, Decomposition in Rank-$(L_r,L_r,1)$ Terms, and a New Generalization. **SIAM Journal on Optimization**, v. 23, n. 2, p. 695–720, jan. 2013.

StopAfib.org - American Foundation for Women's Health. **Atrial Fibrillation - About the Heart**. 2021. Available on: <https://www.stopafib.org/learn-about-afib/what-is-afib/about-the-heart/>. Accessed on: 06/07/2023.

STRIDH, M.; SöRNMO, L. Spatiotemporal QRST cancellation techniques for analysis of atrial fibrillation. **IEEE Transactions on Biomedical Engineering**, v. 48, n. 1, p. 105–111, 2001.

THARWAT, A. Independent component analysis: An introduction. **Applied Computing and Informatics**, v. 17, n. 2, p. 222–249, abr. 2021.

VERMA, A.; JIANG, C.-y.; BETTS, T. R.; CHEN, J.; DEISENHOFER, I. *et al.* Approaches to catheter ablation for persistent atrial fibrillation. **New England Journal of Medicine**, v. 372, n. 19, p. 1812–1822, 2015.

VIKMAN, S.; MäKIKALLIO, T. H.; YLI-MäYRY, S.; NURMI, M.; AIRAKSINEN, K. J.; HUIKURI, H. V. Heart rate variability and recurrence of atrial fibrillation after electrical cardioversion. **Annals of Medicine**, Taylor & Francis, v. 35, n. 1, p. 36–42, 2003.

World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. **JAMA**, v. 310, n. 20, p. 2191–2194, nov. 2013.

YUAN, M.; LIN, Y. Model selection and estimation in regression with grouped variables. **Journal of the Royal Statistical Society: Series B (Statistical Methodology)**, v. 68, n. 1, p. 49–67, 2006.

ZARZOSO, V. Extraction of ECG Characteristics Using Source Separation Techniques: Exploiting Statistical Independence and Beyond. In: NAïT-ALI, A. (Ed.). **Advanced Biosignal Processing**. Berlin, Heidelberg: Springer, 2009. p. 15–47.

ZARZOSO, V. Parameter estimation in block term decomposition for noninvasive atrial fibrillation analysis. In: **IEEE 7th Int. Workshop Comput. Adv. Multi-Sensor Adapt. Process. (CAMSAP)**. Curaçao, Dutch Antilles: [*S. n.*], 2017.

ZARZOSO, V.; COMON, P. Robust independent component analysis by iterative maximization of the kurtosis contrast with algebraic optimal step size. **IEEE Transactions on Neural Networks**, v. 21, n. 2, p. 248–261, 2010.

ZARZOSO, V.; LATCU, D. G.; HIDALGO-MUñOZ, A. R.; MEO, M.; MESTE, O.; POPESCU, I.; SAOUDI, N. Non-invasive prediction of catheter ablation outcome in persistent atrial fibrillation by fibrillatory wave amplitude computation in multiple electrocardiogram leads. **Archives of Cardiovascular Diseases**, v. 109, n. 12, p. 679–688, dez. 2016.