5-24-2010

Modeling And Detection Of Uterine Contractions Using Magnetomyography

Patricio La Rosa
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST. LOUIS

School of Engineering and Applied Science

Department of Electrical & Systems Engineering

Dissertation Examination Committee:
Dr. Arye Nehorai, Chair
Dr. Hiro Mukai
Dr. Hubert Preissl
Dr. Jung-Tsung Shen
Dr. Nan Lin
Dr. R. Martin Arthur

MODELING AND DETECTION OF UTERINE CONTRACTIONS USING MAGNETOMYOGRAPHY

by

Patricio Salvatore La Rosa Araya

A dissertation presented to the Graduate School of Arts and Sciences
of Washington University in partial fulfillment of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

August 2010
Saint Louis, Missouri

ABSTRACT OF THE DISSERTATION

Modeling and Detection of Uterine Contractions using Magnetomyography

by

Patricio Salvatore La Rosa Araya

Doctor of Philosophy in Electrical Engineering

Washington University in St. Louis, August 2010

Research Advisor: Dr. Arye Nehorai

In this dissertation, we develop a novel mathematical framework for modeling and analyzing uterine contractions using biomagnetic measurements. The study of myometrium contractility during pregnancy is relevant to the field of reproductive assessment. Its clinical importance is grounded in the need for a better understanding of the bioreproduction mechanisms. For example, in the last decade the number of preterm labors has increased significantly. Preterm birth can cause health problems or even be fatal for the fetus if it happens too early, and, at the same time, it imposes significant financial burdens on health care systems. Therefore, it is critical to develop models and statistical tools that help to monitor non-invasively the uterine activities during pregnancy.

We derive a forward electromagnetic model of uterine contractions during pregnancy. Existing models of myometrial contractions approach the problem either at an organ level or lately at a cellular level. At the organ level, the models focus on generating contractile forces that closely resemble clinical measurements of normal intrauterine

pressure during contractions in labor. At the cellular level, the models focus on predicting the changes of ionic concentrations in a uterine myocyte during a contraction, and, as a consequence, on modeling the transmembrane potential evolution as a function of time. In this work, we propose an electromagnetic modeling approach taking into account electrophysiological and anatomical knowledge jointly at the cellular, tissue, and organ levels. Our model aims to characterize myometrial contractions using magnetomyography (MMG) and electromyography (EMG) at different stages of pregnancy. In particular, we introduce a four-compartment volume conductor geometry, and we use a bidomain approach to model the propagation of the myometrium transmembrane potential on the human uterus. The bidomain approach is given by a set of reaction-diffusion equations. The diffusion part of the equations governs the spatial evolution of the transmembrane potential, and the reaction part is given by the local ionic current cell dynamics. Here we introduce a modified version of the Fitzhugh-Nagumo (FHN) equation for modeling ionic currents in each myocyte, assuming a plateau-type transmembrane potential. We incorporate the anisotropic nature of the uterus by considering conductivity tensors in our model. In particular, we propose a general approach to design the conductivity-tensor orientation and to estimate the conductivity-tensor values in the extracellular and intracellular domains for any uterine shape. We use finite element methods (FEM) to solve our model, and we illustrate our approach by presenting a numerical example to model a uterine contraction at term. Our results are in good agreement with the values reported in the experimental technical literature, and these are potentially important as a tool for helping in the characterization of contractions and for predicting labor.

We propose an automatic, robust, single-channel statistical detector of uterine MMG contractions. One common restriction of previous techniques is that algorithm parameters, such as the detection threshold and the window length of analysis need to be calibrated experimentally, based on a particular data set. Therefore, the detection performance might change from patient to patient, for example, because of differences in the pregnancy stage and tissue conductivities. In contrast, the proposed algorithm does not require the use of a sliding window of analysis, and the detection threshold is determined analytically; thus, it does not need to be calibrated. Our detection algorithm consists of two stages: In the first stage, we segment the measurements using a multiple change-point estimation algorithm and assuming a piecewise constant time-varying autoregressive model of the measurements; In the second stage, we apply the non-supervised K-means cluster algorithm to classify each time segment, using the RMS and FOZC as candidate features. As a result a discrete-time binary decision signal is generated indicating the presence of a contraction. Moreover, since each single channel detector provides local information regarding the presence of a contraction, we propose a spatio-temporal estimator of the magnetic activity generated by uterine contractions. The algorithm, when evaluated with real MMG measurements, detects uterine activity much earlier than the patient begins to sense it. It also enables visualizing the relative location of the origin of uterine contraction and quantifying the amount of energy delivered during a contraction. These results are important in obstetrics, e.g., as a tool for helping to characterize contractions and to predict labor.

For the aforementioned problem of multiple change-point estimation, a class of one-dimensional segmentation, we also compute fundamental mathematical results for

minimal bounds on mean-square error estimation. Indeed, if an estimator is available, the evaluation of its performance depends on knowing whether it is optimal or if further improvement is still possible. In our segmentation problem the parameters are discrete therefore the conventional Cramer-Rao bound does not apply. Hence, we derive Barankin-type lower bounds, the greatest lower bound on the covariance of any unbiased estimator, which are applicable to discrete parameters. The computation of the bound is challenging, as it requires finding the supremum on a finite set of symmetric matrices with respect to the Loewner ordering, which is not a lattice order. Therefore, we discuss the existence of the supremum, propose a minimal upper-bound by using tools from convex geometry, and compute closed-form solutions for the Barankin information matrix for several distributions. The results have broad biomedical applications, such as DNA sequence segmentation, MEG and EEG segmentation, and uterine contraction MMG detection, and they also have applications for signal segmentation in general, such as speech segmentation and astronomical data analysis.

Acknowledgments

I am very grateful for being blessed with very joyful years of PhD studies in the United States of America. I am happy to look back on the path that I have walked, and remember all the people who, in many different ways, helped me in this journey.

My deepest thanks go to my research advisor and mentor, Professor Arye Nehorai, for providing me with an exceptional PhD training experience. I am very thankful to him for his excellent and dedicated guidance in the pursuit of high quality and state-of-the-art research experiences. During these past years, I have been very much inspired by his pioneering research work in statistical array signal processing and I am very grateful for his leading me in performing multidisciplinary research, having both breadth and depth. I am also thankful for providing me with fruitful teaching opportunities with graduate and undergraduate students, and especially for all his constructive feedback and caring which helped me greatly to develop my teaching skills. His words of wisdom and advice have been an strong help in my process of personal maturity. I feel fortunate to have been part of his excellent research groups at the University of Illinois at Chicago (UIC) and later at Washington University in Saint Louis (WUSTL).

I would like to thank my dissertation committee members, including Dr. Martin Arthur, Dr. Nan Lin, Dr. Hiro Mukai, Dr. Hubert Preissl, and Dr. Jung-Tsung Shen, for carefully revising and providing constructive suggestions to this dissertation. Also, I would like to thank our scientific collaborators, particularly, I thank Dr. Hary Eswaran and Dr. Hubert Preissl for their feedback and discussions on several technical aspects of uterine magnetomyography, for kindly hosting my visits to their research facilities, and for providing us with the experimental measurements used in this work. I am very grateful to Dr. Carlos Muravchik for our discussions on several technical aspects of our work about lower bounds on mean-squared error estimates, and for carefully reviewing several parts of this dissertation. I am grateful for his encouragement and advice at different stages of my PhD studies. I thank Dr. Alexandre Renaux for presenting the subject of performance bounds and discussing its application to the problem of multiple change-point estimation. I enjoyed working and learning from his expertise during his postdoctoral work at Professor Nehorai's

lab at WUSTL. I especially thank Prof. Jim Ballard for all his editorial suggestions on this dissertation.

I convey my sincere gratitude to my UIC instructors, specially to Dr. Dibyen Majumdar for providing me with a deep foundation in linear and multivariate statistics. His excellent lectures were very inspiring and beneficial for performing my research work. I also thank for their dedicated instruction to Dr. Daniel Graupe, Dr. Derong Liu, Dr. Klaus Miescke, Dr. T. E. S. Raghavan, and Dr. Milos Zefran.

My warm thanks go to my past and present labmates at UIC and WUSTL. Their friendships and positive spirit towards life in general made our lab a welcoming and pleasant place to work. In particular, I thank Dr. Tong Zhao, Dr. Gang Shi, Dr. Jian Wang, and Dr. Nannan Cao for their detailed comments on several parts of my work and for sharing with me about China and their rich culture. I appreciate very much their continuous encouragement. I thank Dr. Pinaki Sarder for his long lasting and warm friendship during all this journey. I am very grateful for his interesting discussions on several topics and for sharing with me his delicious culinary creations from India. I cannot forget to thank Mr. Heeralal Choudhary for always looking to cheer me up with creative jokes and for teaching me several Indian recipes, including the outstanding masala tea. His friendship and support are very valuable to me. I also thank Dr. Martin Hurtado, Dr. Marija Nikolic, Mr. Simone Ferrera, Mr. Murat Akcakaya, Mr. Satyabrata Sen, Ms. Vanessa Tidwell, Mr. Gongguo Tang, Mr. Tao Li, Mr. Kofi Adu-Labi, Mr. Sandeep Gogineni, and Mr. Phani Chavali, for their scientific discussions with me and for being very good labmates during all these years.

I would like to thank all the staff members of the Electrical & Systems Engineering Department: Ms. Rita Drochelman, Ms. Sandra Devereaux, Ms. Elaine Murray, Ms. Shauna Dollison, Mr. Ed Richter, and Mr. David Goodbary for their time and help.

Last but not least, I offer my sincere appreciation to my parents and my wife for all their support and love. Their care and encouragement through this journey have been a vital source of inspiration.

Patricio Salvatore La Rosa Araya

Washington University in Saint Louis
August 2010

AMGD - Dedicated to Salvador and Rosa, my parents, and to Alicia, my wife.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Assessment of fetal health is important to reduce perinatal mortality and morbidity. It is based on the prediction, detection and management of fetal malformations, disorders of growth in utero, and premature labor. Labor is the physiologic process that results in the expulsion of the fetus and placenta from the uterus via the cervix and vagina [2]. The occurrence of labor begins with the appearance of periodic contractions which, in general, change the intrauterine-pressure to the point that cervix dilatation is manifested. However, from clinical experience, not all uterine contractions lead to a completion of labor, in which case the process is referred to as false labor. Labor is expected to occur after the $37th$ week of pregnancy, but in the last decade, the number of preterm labors has increased significantly. Preterm birth can cause health problems or even be fatal for the fetus if it happens too early, and, at the same time, it imposes significant financial burdens on health care systems [3].

In general, it is well accepted that monitoring the frequency and intensity of the uterine activity provides sensitive information for distinguishing between false and true labor [4]. However, there are no objective methods for consistently assessing the efficiency of contractions and thus reliably predicting pre-term labor. Therefore, a

better understanding of the mechanisms behind uterine contractions would allow for developing more effective ways to predict and control the occurrence of labor. We claim that this understanding can be achieved by developing the following:

- Non-invasive sensing devices with high spatio-temporal resolution for imaging functionalities of the uterus
- Physical models that interrelate system properties at the organ, tissue, and cellular imaging levels
- Efficient statistical algorithms for solving the imaging problems at each level.

In the following we will describe non-invasive techniques for sensing uterine activity. Our contributions will be specific to physical modeling and statistical algorithms for analyzing uterine contractions.

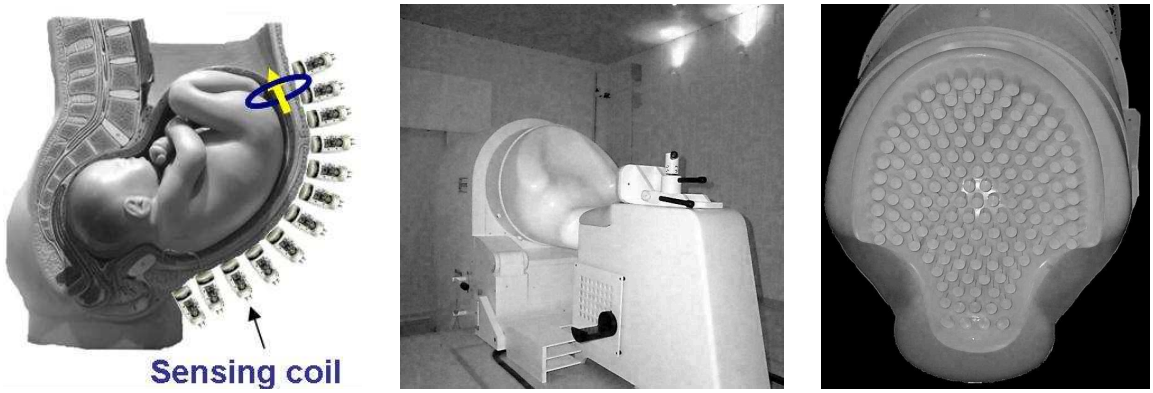
1.1 Noninvasive contraction sensing

Uterine contractions can be described by their mechanical and electrophysiological aspects. A mechanical contraction is manifested as a result of stimulation, which results in propagation of electrical activities in the uterine muscle, and appears as an intrauterine pressure change. Different techniques have been developed to quantify uterine contractions, such as tocography (TOCO), electrohysterography (EHG) or electromyography (EMG), and magnetomyography (MMG). TOCO measures the strength of the force the uterine muscle exerts on the abdominal wall, using an external mechanical method, and the contractions are recorded using tensometric transducers attached to the patient's abdomen. This technique is attractive because it

is noninvasive and simple, but it is of limited value due to its low sensitivity and accuracy [5].

EMG and MMG are functional imaging techniques of the bioelectromagnetic type. They reconstruct and image current density distributions associated with the electrophysiological activity of muscles, using either electrical potential or magnetic field measurements or both. The uterine EMG measures the action potentials of the myometrium cells, using either internal electrodes or abdominal surface electrodes [6, 7]. This technique has a high temporal resolution and has captured attention in the past decade, in particular filtering techniques have been developed for noise and artifacts suppression and for time-frequency characterization of the EMG waveforms [6–11]. However, because of differences in the conductivities of tissue layers, the uterine EMG signals get filtered during their propagation to the surface of the maternal abdomen.

The uterine MMG is a non-invasive technique that measures the magnetic fields associated with muscle action potentials. The first MMG recordings were reported by Eswaran et al. in 2002 [5], using a 151-channel non-invasive device, known as the superconducting quantum interference device array for reproductive assessment (SARA). They established the feasibility of recording uterine contractile activities with a spatial-temporal resolution high enough to determine localized regions of activation and propagation through the uterus. Unlike electrical recordings, magnetic recordings are independent of any references, thus ensuring that each sensor mainly records localized activities. However, MMG is of limited applicability because it is expensive (in the range of $ 2 - 3 million dollars) and non-portable.



(a) **(b)** **(c)**

Figure 1.1: (a) A simplified illustration of the sensing array and the uterine MMG field. (b) The SARA system installed at the University of Arkansas for Medical Sciences (UAMS) Hospital. (c) 151-channel sensor array embedded under the concave surface upon which the patient leans her abdomen. The sensor coils are placed 3 cm apart, covering a total area of approximately 1350 cm^2.

1.2 Our contributions

In this dissertation, we propose a forward electromagnetic model of uterine contractions during pregnancy [12, 13], derive statistical algorithms for automatic detection of uterine contractions based on time-series segmentation and an unsupervised clustering approach [14, 15], and derive performance bounds for the class of unbiased model-based segmentation algorithms [16, 17].

Physical model: Having an electromagnetic model of uterine contractions is relevant to predicting and interpreting uterine activity using MMG and EMG measurements. In particular, our model aims to describe the electrophysiological aspects of uterine contractions during pregnancy at both the cellular and the organ levels. We introduce a four-compartment volume conductor geometry and use a bidomain

approach [18, 19] to model the propagation of the myometrium transmembrane potential. The bidomain approach is given by a set of reaction-diffusion equations. The diffusion part of the equations governs the spatial evolution of the transmembrane potential, and the reaction part is given by the ionic current cell dynamics locally. Assuming a plateau-type transmembrane potential, we introduce a modified version of the Fitzhugh-Nagumo (FHN) equation [20–22] for modeling ionic currents in each myocyte, and we incorporate the anisotropic nature of the uterus by designing conductivity-tensor fields. In particular, we propose a general approach to design the conductivity-tensor orientation and to estimate the conductivity-tensor values in the extracellular and intracellular domains for any uterine shape. We use finite element methods (FEM) to solve our model, and we illustrate our approach by presenting a numerical example to model a uterine contraction at term. Our results are in good agreement with the values reported in the experimental technical literature, and these are potentially important as a tool for helping in the characterization of contractions and for predicting labor.

Statistical detection of uterine contractions: Biomagnetic measurements obtained using the aforementioned SARA system contain the electrophysiological activities of several organs in the vicinity of the abdomen, as well as the fetus. Therefore, to analyze uterine contractions, it is necessary to first filter out non-desired signals from the measurements, and also to detect the time span in which the contraction takes place. In our work, we propose a distributed processing framework to process the measurements from an array of magnetometers. Our method is based on a single-channel, two-stage statistical detector of uterine contractions that is robust and automatic. Unlike in previous approaches, the proposed detection algorithm does not require the use of a sliding window of analysis, and the detection threshold is determined

analytically. In the first stage, we propose a model-based segmentation procedure, which detects multiple change-points in the parameters of a piecewise constant time-varying autoregressive model using a robust formulation of the Schwarz information criterion (SIC) and a binary search approach. We compute and evaluate the relative energy variation [root mean square (RMS)] in discriminating between time segments with and without contractions. Thus, in the second stage, we apply a nonsupervised K-means cluster algorithm to classify the detected time segments using the RMS values. We validate our method using real MMG measurements and compare the detected time intervals with the patients' feedback. This method proves to be helpful in understanding the uterine MMG contraction activity spatially and temporally.

Performance bounds on model-based segmentation algorithms: The literature is abundant concerning estimation algorithms for change-point estimation (see, *e.g.*, [23–25]). However, less work has been done concerning the ultimate performance of such algorithms in terms of mean-square error (MSE). Indeed, if an estimator is available, the evaluation of its performance depends on knowing whether it is optimal or if further improvement is still possible. Unfortunately, for discrete time-measurement models, as in our aforementioned time-series segmentation problem, the change-point location parameter is discrete, therefore the Cramér-Rao bound [26] is not applicable. Consequently, we focus on computing the Barankin bound (BB) [27], the greatest lower bound on the covariance of any unbiased estimator, which is still valid for discrete parameters. To the best of our knowledge, performance bounds have never been derived in the multiple change-point context. In our work, we compute the multi-parameter version of the Hammersley-Chapman-Robbins, which is a Barankin-type lower bound in the context of an independent vector sequence. The computation of the BB requires finding the supremum of a finite set of positive definite matrices with respect

to the Loewner partial ordering. Although each matrix in this set of candidates is a lower bound on the covariance matrix of the estimator, the existence of a unique supremum for this set, i.e., the tightest bound, might not be guaranteed. To overcome this problem, we compute a suitable minimal-upper bound on this set given by the matrix associated with the Lowner-John Ellipsoid of the set of hyper-ellipsoids associated to the set of candidate lower-bound matrices. We present numerical examples to compare the proposed approximated BB with the performance achieved by the maximum likelihood estimator.

The organization of this dissertation is as follows: In Chapter 2, basic uterine anatomy is discussed, and the electromagnetic modeling of uterine contractions is presented. In Chapter 3, our automatic algorithm for detecting uterine contraction using MMG measurements is derived. In Chapter 4, the performance bounds for time-series segmentation algorithms are computed. Finally, in Chapter 5, we summarize the contributions of this dissertation and discuss future work.

Chapter 2

Forward Electromagnetic Modeling of Uterine Contractions During Pregnancy

2.1 Introduction

In the following chapter we will describe the details of our physical model of the electromagnetic activity associated to uterine contractions. We begin by discussing briefly the uterine microanatomy and previous uterine contraction models.

2.1.1 Uterine microanatomy:

The adult uterus is a thick walled, hollow, muscular organ formed by three layers: the external serous perimetrium, the myometrium, and the inner mucous endometrium [28]. The myometrium is responsible for contractions and it is formed by fasciculi which comprise sheet-like and cylindrical bundles of myocytes embedded in

a connective tissue matrix [1]. The myocytes in a cylindrical bundle contract, thus shortening the smooth tissue and increasing uterus wall tension, hence increasing the intrauterine pressure. Fig. 2.1 illustrates the microanatomy of the pregnant human myometrium.

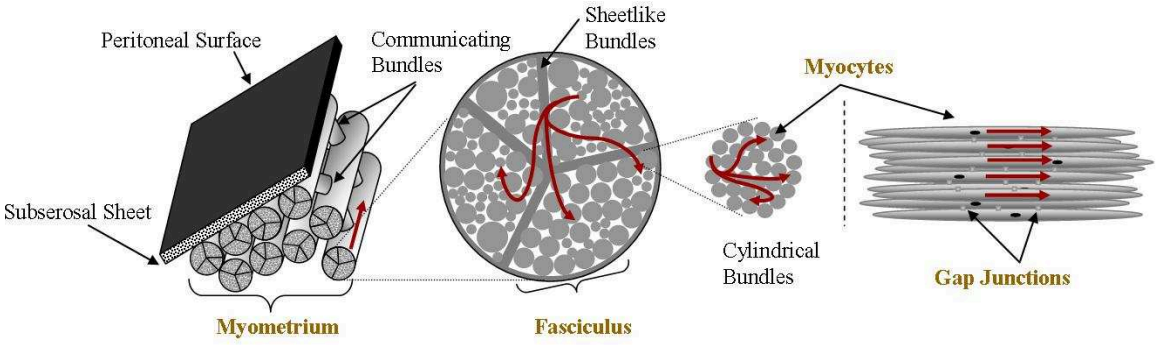


Figure 2.1: Diagram of microanatomy of pregnant human myometrium [1]. Red lines represent current flows.

The uterine microanatomy is consistent with action potential propagation [1]: (i) myocytes are densely packed within a bundle, (ii) bundles are contiguous within a fasciculus, and (iii) fasciculi are contiguous via communicating bridges formed with myocytes. In addition, the uterine changes during gestation is accompanied by the formation of gap junctions which are one of the mechanisms for coordinated transmission of contractile activity from cell to cell [1, 28]. The structure of the fasiculata within the uterus has not yet been well defined, but generally it makes the propagation of the action potential anisotropic [29, 30].

2.1.2 Uterine contraction models

Uterine contractions can be described by their mechanical and electrophysiological aspects. A mechanical contraction is manifested as a result of the excitation as well

as the propagation of electrical activities in the uterine muscle, and appears in the form of an intrauterine pressure increase.

Existing models approach the problem separately at organ level [31, 32], or lately at a cellular level [33–35]. At the organ level, the models focus on predicting the contractile forces that closely resemble clinical measurements of normal intrauterine pressure during contractions in labor. In [31], the authors assume that the uterus is a hollow ovoid formed by discrete contractile elements that propagate electrical impulses, generate tension, and have defined contracting and refractory periods. The envisioned mechanism for intercellular communication is based on action potential propagation, which is simulated by using a discrete state model for each cell. In [32], the author uses a discrete state model for combining two mechanisms of intercellular communication, namely, action potential propagation and intercellular calcium wave propagation. However, in both [31] and [32], mathematical and physical descriptions of their models are not provided. On the other hand, at a cellular level, the models focus on predicting the changes of ionic concentrations in the intracellular and extracellular mediums during a contraction, and, as a consequence, on modeling the transmembrane potential evolution of a myocyte as a function of time. In [33, 34] a model is developed to simulate the complete process of a single myometrial smooth muscle contraction, which is initiated by depolarization. The model is based on the electrophysiological properties of a myocyte, and on the cellular mechanisms that relate the rise in concentration levels of intracellular ion calcium C_a^{2+} to stress production.

In this work, we propose a forward electromagnetic model of human myometrial contractions during pregnancy taking into account electrophysiological and anatomical knowledge jointly at the cellular, tissue, and organ levels. Our model aims to helping

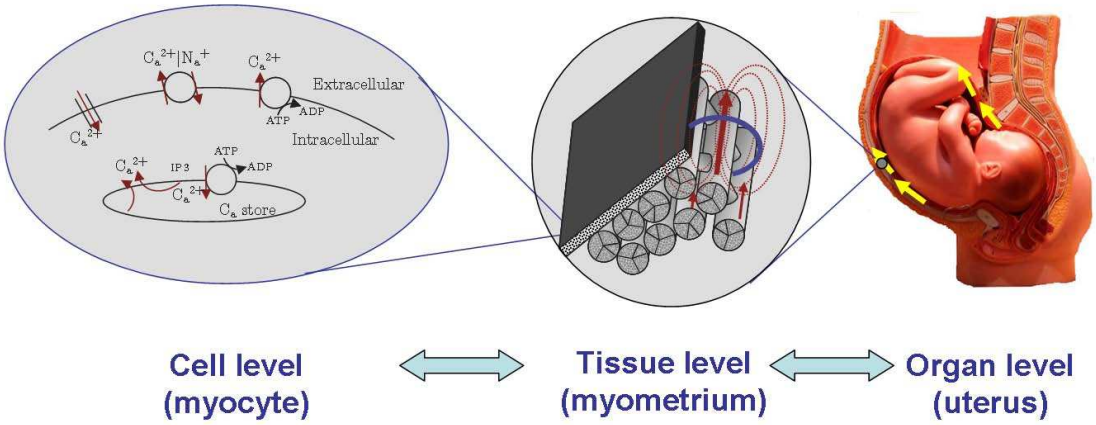


Figure 2.2: Illustration of the proposed modeling approach.

in the characterization of contractions and for predicting labor using MMG [5] and EMG [36]. Here we extend our partial results presented in [12]. Fig. 2.2 illustrates the different levels considered in our modeling approach. In particular, our approach is two fold: first, we model the current source density at the myometrium, using models of myocyte electrophysiological activity and anisotropic conductivity; And second, we solve the forward electromagnetic problem, namely, we compute the magnetic field and the action potential at the abdominal surface generated by the myometrial current-source density, using Maxwell's equations subject to a volume conductor geometry. To model the current source density at the myometrium we propose to apply a bidomain approach. The bidomain equations is a set of reaction-diffusion equations derived first for modeling the current sources of the myocardium as a function of the cardiac-myocyte transmembrane potential, and these equations proved to be a successful approach to study heart functioning [18,19]. The diffusion part of the equations governs the spatial evolution of the transmembrane potential, and the reaction part is given by the ionic current cell dynamics locally. Here we introduce a modified version of the FitzHugh-Nagumo (FHN) equation for modeling ionic currents in each

myocyte. Though FHN does not consider explicitly the ${C_a}^{2+}$ dynamics, the simplicity of the FHN model makes it an attractive candidate for modeling the propagation of depolarization waves in large 2D and 3D simulations as in the numerical examples presented in this work. We propose a general approach to design conductivity-tensor orientation for any uterine shape. We estimate the conductivity-tensor values in the extracellular and intracellular domains, using Archie's law [37] and an analytical expression of the transmembrane potential propagation speed, derived in this work, as a function of the model parameters.

The notational convention adopted in this paper is as follows: italic font indicates a scalar quantity, as in a; lowercase boldface indicates a vector quantity, as in $\boldsymbol{a}$, except for vector fields used in Maxwell's equations such as electric field $\boldsymbol{E}$, magnetic field $\boldsymbol{B}$, and current density $\boldsymbol{J}$; upper case italic indicates a matrix quantity, as in A. The matrix transpose is indicated by a superscript "T" as in A^T, and the identity matrix of size $n \times n$ is denoted I_n. The set $\mathbb{S}^n$ denotes the vector space of symmetric $n \times n$ matrices and the subsets of nonnegative definite matrices and positive definite matrices are denoted by $\mathbb{S}^n_+$ and $\mathbb{S}^n_{++}$, respectively. The inner product and norm defined in the Euclidean space is denoted by $\langle \cdot, \cdot \rangle$ and $\|\cdot\|$, respectively.

This chapter is organized as follows: Section 2.2, describes the volume conductor geometry of the problem and the forward electromagnetic model; Section 2.3, presents the current source model based on the bidomain equations; Section 2.4 describes our approach for modeling the myometrial conductivity tensors. Section 2.5 , introduces the monodomain approximation, boundary conditions, and numerical computations of our model; Section 2.6 presents the numerical examples and discussion; And in Section 2.7 a summary of the chapter is provided.

2.2 Forward electromagnetic model

In this section we discuss the forward electromagnetic model of myometrial contractions. We will introduce a four-compartment volume conductor model formed by an anisotropic bidomain myometrium, and we will present the expressions for the extrauterine electrical potential and magnetic field, respectively.

2.2.1 Volume conductor model

Fig. 2.3 illustrates the four-compartment volume conductor geometry for our problem, where $\mathcal{A}$ represents the abdominal cavity and $\partial\mathcal{A}$ the boundary surface defined by the abdomen, $\mathcal{M}$ represents the myometrium, and $\partial\mathcal{M}$ and $\partial\mathcal{U}$ are its external and internal boundary surfaces, respectively. The volume denoted by $\mathcal{U}$ represents the space filled with amniotic fluid which exists between the internal uterine wall $\partial\mathcal{U}$ and the boundary $\partial\mathcal{F}$ defined by the fetus volume $\mathcal{F}$. The vectors $\boldsymbol{r}$ and $\boldsymbol{r}'$ indicate the positions of the observation point and source, respectively, with respect to the main axis of reference.

2.2.2 Extrauterine magnetic field and electrical potential

The electromagnetic analysis of uterine contractions can be derived by solving a set of Maxwell's equations [38] subject to given boundary conditions given by the volume conductor geometry. Moreover, since common bioelectrical phenomena contain mostly frequencies below 1 KHz and the characteristic length scale is much larger than the diameter of the uterus, it is suitable to use the quasi-static approximation of

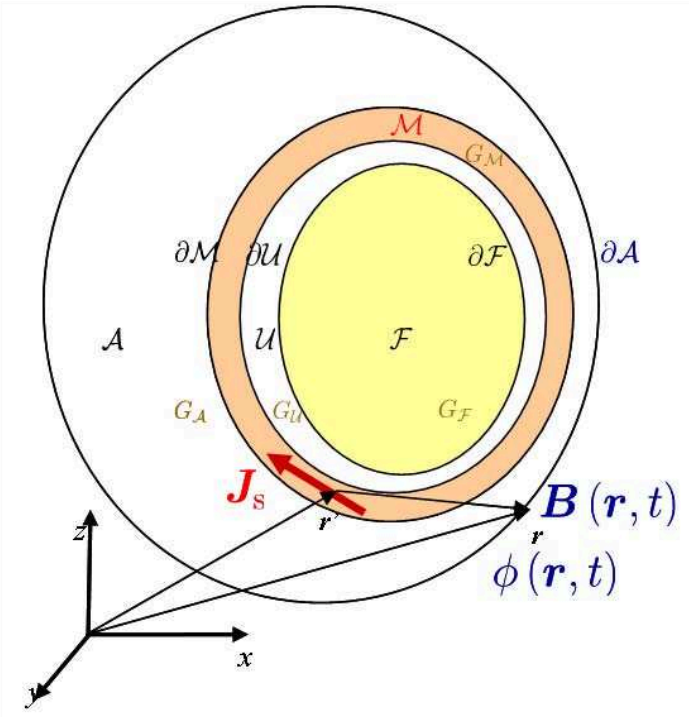


Figure 2.3: Representation of the four-compartment volume conductor geometry and the forward electromagnetic problem of uterine contractions.

Maxwell's equations. Therefore, the extrauterine magnetic field $\boldsymbol{B}(\boldsymbol{r},t)$ at a position $\boldsymbol{r}$ and instant t is given as follows:

$$\nabla \times \frac{1}{\mu_{\mathrm{o}}} \boldsymbol{B}(\boldsymbol{r},t) = \boldsymbol{J}(\boldsymbol{r},t), \tag{2.1}$$

where μ_{o} is the permeability of the free space and $\boldsymbol{J}(\boldsymbol{r},t)$ is the total current density (in $\mathrm{A/m^2}$). $\boldsymbol{J}(\boldsymbol{r},t)$ is given by

$$\boldsymbol{J}(\boldsymbol{r},t) = \boldsymbol{J}_{\mathrm{s}}(\boldsymbol{r},t) + G(\boldsymbol{r})\,\boldsymbol{E}(\boldsymbol{r},t), \tag{2.2}$$

where $\boldsymbol{J}_{\mathrm{s}}(\boldsymbol{r},t)$ is the uterine current density source, and $G(\boldsymbol{r})\,\boldsymbol{E}(\boldsymbol{r},t)$ is the conduction current density (or return currents), as described by Ohm's law, with $\boldsymbol{E}(\boldsymbol{r},t)$

the electric field established by $\boldsymbol{J}_{\mathrm{s}}(\boldsymbol{r},t)$ and $G(\boldsymbol{r}) \in \mathbb{S}^3_{++}$ is the conductivity tensor defined by each compartment. Then from the quasi-static conditions, $\nabla \cdot \boldsymbol{J}(\boldsymbol{r},t) = 0$, so $\nabla \cdot G(\boldsymbol{r})\,\boldsymbol{E}(\boldsymbol{r},t) = -\nabla \cdot \boldsymbol{J}_{\mathrm{s}}(\boldsymbol{r},t)$. Moreover, since $\nabla \times \boldsymbol{E}(\boldsymbol{r},t) = 0$, it follows that $\boldsymbol{E}(\boldsymbol{r},t) = -\nabla\phi(\boldsymbol{r},t)$, where $\phi(\boldsymbol{r},t)$ denotes the potential. Thus, the equation that governs the relationship between the electromyogram potentials and uterus current sources is

$$\nabla \cdot G(\boldsymbol{r})\,\nabla\phi(\boldsymbol{r},t) = \nabla \cdot \boldsymbol{J}_{\mathrm{s}}(\boldsymbol{r},t)\,. \tag{2.3}$$

Therefore, solving the forward electromagnetic problem of uterine contractions implies computing $\boldsymbol{B}(\boldsymbol{r},t)$ and $\phi(\boldsymbol{r},t)$ at $\partial\mathcal{A}$ using Eqs. (2.1) and (2.3) assuming known $\boldsymbol{J}_{\mathrm{s}}(\boldsymbol{r},t)$ in $\mathcal{M}$ and $G(\boldsymbol{r})$ in all the domain defined by the volume conductor geometry (see Fig. 2.3).

The biological current sources $\boldsymbol{J}_{\mathrm{s}}(\boldsymbol{r},t)$ in the myometrium are the transmembrane ionic fluxes, due to concentration gradients, which flow across the surface membrane of the myocyte (smooth cells) from the extracellular medium into the intracellular medium and vice versa. The density of these ionic currents is also referred to as impressed current density since its origin is non electrical in nature, and it is the primary cause for the establishment of an electric field which induces secondary density currents in a conductive domain. We will model $\boldsymbol{J}_{\mathrm{s}}(\boldsymbol{r},t)$ using a bidomain approach, which has proved to be a successful method to study electrophysiological activity in the myocardium [18, 19].

2.2.3 Current source model

In the myometrium both the intracellular and extracellular domains are physically connected through membrane gates, and the intracellular domain is connected though gap junctions [28, 36]. Therefore, we model the myometrium using the bidomain modeling approach. This approach represents the tissue (myometrium) as two interpenetrating extra-intracellular continuous domains, with different conductivity values along and across the direction of the fiber [18, 19], and it models the tissue using the generalized-passive cable equation. The bidomain modeling approach was originally derived for modeling the propagation of the transmembrane potential of the myocardium and proved to be a successful approach to study heart functioning [18, 19]. Fig. 2.4 shows a simplified illustration of the tissue and the bidomain approach, where $\phi_{\mathrm{i}}(\boldsymbol{r},t)$ and $\phi_{\mathrm{e}}(\boldsymbol{r},t)$ are the intracellular and interstitial potentials, respectively, and $v_{\mathrm{m}}(\boldsymbol{r},t)=\phi_{\mathrm{i}}(\boldsymbol{r},t)-\phi_{\mathrm{e}}(\boldsymbol{r},t)$ is transmembrane potential. The conductivity tensors in the intracellular and extracellular domains are denoted by $G_{\mathrm{i}}^{'}$ and $G_{e}^{'}$ (in S/m), and, using Ohm's law, the current densities in each domain are given by $\boldsymbol{J}_{\mathrm{i,e}}(\boldsymbol{r},t)=-G_{\mathrm{i,e}}^{'}\nabla\phi_{\mathrm{i,e}}(\boldsymbol{r},t)$. The transmembrane volume current density in (A/m^3) is denoted by $j_{\mathrm{m}}(\boldsymbol{r},t)$ and is given by

$$
\begin{aligned}
j_{\mathrm{m}}(\boldsymbol{r},t) &= a_{\mathrm{m}}\, c_{\mathrm{m}}\frac{\partial v_{\mathrm{m}}}{\partial t}+j_{\mathrm{ion}}-j_{\mathrm{stim}}, && (2.4)\\
&= a_{\mathrm{m}}\left(c_{\mathrm{m}}\frac{\partial v_{\mathrm{m}}}{\partial t}+\boldsymbol{J}_{\mathrm{ion}}-\boldsymbol{J}_{\mathrm{stim}}\right), && (2.5)
\end{aligned}
$$

where $j_{\mathrm{ion}}(\boldsymbol{r},t)$ is the ionic volume current density (in A/m^3) of a myocyte, $j_{\mathrm{stim}}(\boldsymbol{r},t)$ is the stimulus volume current density (in A/m^3), c_{m} is the membrane capacitance per unit area (in F/m^2), and a_{m} is the surface-to-volume ratio of the membrane (in 1/m).

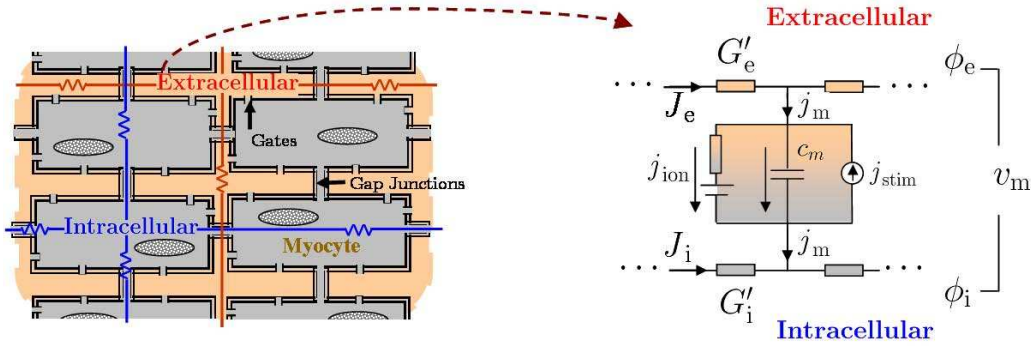


Figure 2.4: Illustration of the bidomain modeling approach.

Applying conservation of charges to both domains, we obtain the following relationships:

$$\nabla \cdot \boldsymbol{J}_{\mathrm{e}}(\boldsymbol{r},t) = j_{\mathrm{m}}(\boldsymbol{r},t), \text{ and} \tag{2.6}$$

$$\nabla \cdot \boldsymbol{J}_{\mathrm{i}}(\boldsymbol{r},t) = -j_{\mathrm{m}}(\boldsymbol{r},t). \tag{2.7}$$

Adding (2.6) and (2.7), we have that $\nabla \cdot (\boldsymbol{J}_{\mathrm{i}}(\boldsymbol{r},t) + \boldsymbol{J}_{\mathrm{e}}(\boldsymbol{r},t)) = 0$. Hence, the total current density in the myometrium is given by

$$\boldsymbol{J}(\boldsymbol{r},t) = -G'_{\mathrm{i}} \nabla \phi_{\mathrm{i}}(\boldsymbol{r},t) - G'_{\mathrm{e}} \nabla \phi_{\mathrm{e}}(\boldsymbol{r},t), \; r \in \mathcal{M}, \tag{2.8}$$

which can be expressed in terms of $v_{\mathrm{m}}(\boldsymbol{r},t)$ and $\phi_{\mathrm{e}}(\boldsymbol{r},t)$ as follows:

$$\boldsymbol{J}(\boldsymbol{r},t) = -G'_{\mathrm{i}} \nabla v_{\mathrm{m}}(\boldsymbol{r},t) - G'_{\mathcal{M}} \nabla \phi_{\mathrm{e}}(\boldsymbol{r},t), \; r \in \mathcal{M}, \tag{2.9}$$

where $G'_{\mathcal{M}} = \left(G'_{\mathrm{e}} + G'_{\mathrm{i}}\right) \in \mathbb{S}^3_{++}$ is the bulk myometrium conductivity tensor. Since spatial variations of $v_{\mathrm{m}}(\boldsymbol{r},t)$ depend on the local establishment of a transmembrane current density, $j_{\mathrm{m}}(\boldsymbol{r},t) \neq 0$, then we define the impressed current-density source

as $\boldsymbol{J}_{\mathrm{s}}(\boldsymbol{r},t) = -G_{\mathrm{i}}'\nabla v_{\mathrm{m}}(\boldsymbol{r},t)$. Note that $\boldsymbol{J}_{\mathrm{s}}(\boldsymbol{r},t)$ exists only when the spatial gradient exists, i.e., only in a region where the myometrium is undergoing depolarization (excitation) or repolarization.

The total current at the myometrium $\boldsymbol{J}\left(\boldsymbol{r},t\right)$ depends on the spatio-temporal variations of $v_{\mathrm{m}}\left(\boldsymbol{r},t\right)$ and $\phi_{\mathrm{e}}\left(\boldsymbol{r},t\right)$, which are governed by the system of equations formed by Eqs. (2.4), (2.6) and (2.7). Using simple algebraic manipulations, the aforementioned system of equations can be written in terms of $v_{\mathrm{m}}\left(\boldsymbol{r},t\right)$ and $\phi_{\mathrm{e}}\left(\boldsymbol{r},t\right)$ only, obtaining the following equivalent expressions:

$$\nabla \cdot G_{\mathrm{i}}'\nabla(v_{\mathrm{m}}\left(\boldsymbol{r},t\right)+\phi_{\mathrm{e}}\left(\boldsymbol{r},t\right)) = a_{\mathrm{m}}\left(c_{\mathrm{m}}\frac{\partial v_{\mathrm{m}}\left(\boldsymbol{r},t\right)}{\partial t}+\boldsymbol{J}_{\mathrm{ion}}\left(\boldsymbol{r},t\right)-\boldsymbol{J}_{\mathrm{stim}}\left(\boldsymbol{r},t\right)\right), \tag{2.10}$$

$$\nabla \cdot (G_{\mathrm{i}}'+G_{\mathrm{e}}')\nabla\phi_{\mathrm{e}}\left(\boldsymbol{r},t\right) = -\nabla \cdot G_{\mathrm{i}}'\nabla v_{\mathrm{m}}\left(\boldsymbol{r},t\right). \tag{2.11}$$

This set of reaction-diffusion equations is also known as bidomain equations [18, 19]. The diffusion part of the equations governs the spatial evolution of both the transmembrane and extracellular potentials, and the reaction part is given by the local ionic current cell dynamics. The solutions for $v_{\mathrm{m}}\left(\boldsymbol{r},t\right)$ and $\phi_{\mathrm{e}}\left(\boldsymbol{r},t\right)$ depend on $\boldsymbol{J}_{\mathrm{ion}}\left(\boldsymbol{r},t\right)$, $\boldsymbol{J}_{\mathrm{stim}}\left(\boldsymbol{r},t\right)$, and the conductivity tensors, in addition to boundary and initial conditions. Since our goal is to model the propagation of the electrical activity in the myometrium, we are interested in the class of traveling wave solutions of these equations which waveform depends on $\boldsymbol{J}_{\mathrm{ion}}\left(\boldsymbol{r},t\right)$ and its initiation depends on $\boldsymbol{J}_{\mathrm{stim}}\left(\boldsymbol{r},t\right)$. In what follows, we describe the models for both current densities $\boldsymbol{J}_{\mathrm{ion}}\left(\boldsymbol{r},t\right)$ and $\boldsymbol{J}_{\mathrm{stim}}\left(\boldsymbol{r},t\right)$, respectively.

2.2.4 Ionic current model

The predominant type of transmembrane-potential waveforms measured in the human myometrium are spike and plateau [28, 39, 40]. In this work, we focus on modeling the plateau-type transmembrane potential, as it has been more frequently observed [28,39–41]. Therefore, we model $\boldsymbol{J}_{\text{ion}}(\boldsymbol{r}, t)$, using a variation of the FitzHugh-Nagumo (FHN) equations [20–22], as follows:

$$\boldsymbol{J}_{\text{ion}}(\boldsymbol{r}, t) = -\frac{1}{\epsilon_1}\left(k\ \left(v_{\text{m}} - v_1\right)\left(v_2 - v_{\text{m}}\right)\left(v_{\text{m}} - v_3\right) - w\right), \text{ and} \tag{2.12}$$

$$\frac{\partial w}{\partial t} = \epsilon_2\left(\beta v_{\text{m}} - \gamma w + \delta\right), \tag{2.13}$$

where ϵ_1, ϵ_2, k, v_1, v_2, v_3, δ, γ, and β are model constants, and w (in V) is a state variable of the model. The parameter ϵ_1 (in Ωm^2) controls the sharpness of the leading and trailing edges of the action potential waveform: the smaller ϵ_1 is, the more vertical the edge is. Note that ϵ_1 has unit of resistivity, therefore the smaller its value the larger the membrane permeability to ionic flux. The parameter ϵ_2 (in s^{-1}) controls the action potential duration: the smaller ϵ_2, the longer it takes a cell to recover. The parameters v_1, v_2, v_3 (in V), and k (in $1/\text{V}^2$) control the range of $v_{\text{m}}(\boldsymbol{r}, t)$. Note that for a given set of parameters values k, v_1, v_2, v_3 and γ, the parameters β and δ (in V) control the excitability threshold of the cell. The larger β, the lower the excitability threshold setting the cell dynamic to an oscillatory stable behavior between resting and exciting states. Over a certain value, the cell dynamic becomes bistable; namely, if the cell starts from a resting potential, it changes to an excited state and remains there. On the other hand, a very negative β value results in a permanent resting state. In Section Results and Discussion we select the model parameters using phase-space analysis, and using as a reference the transmembrane

potentials recorded from isolated human myometrial strips at term [33, 40]. This model does not consider explicitly the C_a^{2+} dynamics, and, moreover, it assumes that changes in the intra- and extra-cellular ion concentrations are insignificant even after several depolarizations. However, its simplicity facilitates modeling the propagation of depolarization waves in large 2D and 3D domains.

2.2.5 Stimulus current model

We also introduce a temporal-spatial model for $\boldsymbol{J}_{\text{stim}}$, representing the stimulus due to pacemaker areas [28, 36], as follows:

$$\boldsymbol{J}_{\text{stim}}(r,t) = \frac{1}{\epsilon_1} \sum_{i=1}^{N_p} \nu_i h_i(r,t), \tag{2.14}$$

where $h_i(r,t)$ is a spatio-temporal function with range in $[0,1]$, ν_i is the amplitude (in V), and N_p is the number of pacemaker areas. Intuitively, the former should modify the excitability of the cell at a certain instant of time based on the threshold value. In particular, our model assumes that the uterine myocyte can act as either a pacemaker or pace-follower, namely, the spontaneous electrical behavior exhibited by the myometrium is an inherent property of the uterine myocyte (see [36] for more details.) Note that the size, duration, and intensity of the pacemaker area need to be chosen such that a stable traveling waveform solution to the bidomain equations on the myometrium is granted.

2.3 Myometrial conductivity tensors

The structure of the fasiculata within the uterus has not been well-defined, but it generally runs in a structured organization [29, 30]. In [29] the authors investigated the global fiber architecture of the non-pregnant uterus by magnetic resonance (MR) diffusion tensor imaging (DTI). From the ex-vivo analysis of five non-pregnant uteri, the authors identified an inner circular layer around the uterine cavity on slices orthogonal to the long axis of the organ. In the regions outside the inner circular layer, they could not identify a global structure, but did find several locally aligned groups of fibers. At the level of the cervix, they found an outer circular layer and an inner region with mostly longitudinal components. In the following we will introduce an approach for designing the conductivity tensors in the myometrium.

Assume that the conductivity tensors are diagonal in a local coordinate system which is defined with respect to each myocyte and characterized by the unit vectors $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$. In particular, $G_{\rm i}$ and $G_{\rm e}$ are diagonal matrices $\in \mathbb{S}^3_{++}$ given by

$$G_{\rm i} = \begin{pmatrix} \sigma_{\rm ix} & 0 & 0 \\ 0 & \sigma_{\rm iy} & 0 \\ 0 & 0 & \sigma_{\rm iz} \end{pmatrix}, \quad G_{\rm e} = \begin{pmatrix} \sigma_{\rm ex} & 0 & 0 \\ 0 & \sigma_{\rm ey} & 0 \\ 0 & 0 & \sigma_{\rm ez} \end{pmatrix}. \tag{2.15}$$

In order to take into account variable fiber orientation in the myometrium, we need to describe it in a global Cartesian coordinate system in which the local basis is defined at any point $\boldsymbol{r}$ as $A = [\boldsymbol{a}_1(\boldsymbol{r}),\ \boldsymbol{a}_2(\boldsymbol{r}),\ \boldsymbol{a}_3(\boldsymbol{r})]$ where $\boldsymbol{a}_3(\boldsymbol{r})$ is parallel to the main fiber axis. The representation of the tensors $G_{\rm i}$ and $G_{\rm e}$ in terms of a global coordinate

system is given by

$$G'_{\mathrm{i}} = AG_{\mathrm{i}}A^T, \quad \mathrm{G}'_{\mathrm{e}} = AG_{\mathrm{e}}A^T. \tag{2.16}$$

Assuming that the myocyte fiber conductivities in both domains have a cylindrical symmetry, then $\sigma_{\mathrm{ex}} = \sigma_{\mathrm{ey}} = \sigma_{\mathrm{et}}$, $\sigma_{\mathrm{ix}} = \sigma_{\mathrm{iy}} = \sigma_{\mathrm{it}}$, $\sigma_{\mathrm{iz}} = \sigma_{\mathrm{il}}$, and $\sigma_{\mathrm{ez}} = \sigma_{\mathrm{el}}$. Therefore, the conductivity tensors can be expressed as follows [42]:

$$G'_{\mathrm{i}} = (\sigma_{\mathrm{il}} - \sigma_{\mathrm{it}})\, \boldsymbol{a}_3(\boldsymbol{r})\boldsymbol{a}_3^T(\boldsymbol{r}) + \sigma_{\mathrm{it}}\ I_3, \text{ and} \tag{2.17}$$

$$G'_{\mathrm{e}} = (\sigma_{\mathrm{el}} - \sigma_{\mathrm{et}})\, \boldsymbol{a}_3(\boldsymbol{r})\boldsymbol{a}_3^T(\boldsymbol{r}) + \sigma_{\mathrm{et}}\ I_3. \tag{2.18}$$

Hence, to construct the conductivity tensors as a function of $\boldsymbol{r}$, because of the cylindrical symmetry assumption, it is enough to define the vector field $\boldsymbol{a}_3(\boldsymbol{r})$ in each location of the anisotropic domain, as well as the conductivity values σ_{il} and σ_{el}. To design $\boldsymbol{a}_3(\boldsymbol{r})$ at each point $\boldsymbol{r}$, we represent the uterus as a hollow volume with uniform thickness, and we describe it by the union of mutually disjoint closed surfaces or layers. We use the implicit definition of a surface, namely, the set of points $\boldsymbol{r}$ satisfying $f(\boldsymbol{r}) = 0$. Then, at each point $\boldsymbol{r}$, we define a set of local orthonormal coordinates axes given by $\{\widehat{\boldsymbol{n}}(\boldsymbol{r}), \widehat{\boldsymbol{t}}_1(\boldsymbol{r}), \widehat{\boldsymbol{t}}_2(\boldsymbol{r})\}$, where $\widehat{\boldsymbol{n}}(\boldsymbol{r}) = \frac{\nabla f(\boldsymbol{r})}{\|\nabla f(\boldsymbol{r})\|}$ is the normal vector to the layer containing $\boldsymbol{r}$, and $\widehat{\boldsymbol{t}}_1(\boldsymbol{r})$ and $\widehat{\boldsymbol{t}}_2(\boldsymbol{r})$ are mutually orthogonal vectors which belong to the tangent plane of the respective layer at point $\boldsymbol{r}$. We define $\widehat{\boldsymbol{t}}_1(\boldsymbol{r})$ and $\widehat{\boldsymbol{t}}_2(\boldsymbol{r})$, using as a reference the curve of symmetry of the uterine inner-circular layer [29]. This curve goes from the fundus to the cervix, and it coincides with the long axis of the non-pregnant uterus (see Appendix A for more details on the computation of $\widehat{\boldsymbol{t}}_1(\boldsymbol{r})$ and $\widehat{\boldsymbol{t}}_2(\boldsymbol{r})$). Hence, given $\widehat{\boldsymbol{t}}_1(\boldsymbol{r})$ and $\widehat{\boldsymbol{t}}_2(\boldsymbol{r})$, we define $\boldsymbol{a}_3(\boldsymbol{r})$ as follows:

$$\boldsymbol{a}_3(\boldsymbol{r}) = \widehat{\boldsymbol{t}}_1(\boldsymbol{r}) \cos(\alpha) + \widehat{\boldsymbol{t}}_2(\boldsymbol{r}) \sin(\alpha), \tag{2.19}$$

where α is the fiber orientation angle with respect to $\widehat{\boldsymbol{t}}_1(\boldsymbol{r})$.

To the best of our knowledge, values of the intracellular and extracellular conductivity tensors have not been reported for the human myocyte, and therefore, these have to be estimated. To estimate the extracellular conductivity values σ_{el} and σ_{et}, we assume a grid-type distribution of myocytes in the myometrium and use an estimate of the extracellular conductivity the human myometrium obtained by applying Archie's law [37]. Human myocytes can be best be described as long cylinders with diameter d_{cell} and axis length l_{cell}, such that $d_{\mathrm{cell}} \ll l_{\mathrm{cell}}$. Assuming that myocytes are uniformly arranged in a cubical grid whose length $l_{\mathrm{T}} = l_{\mathrm{cell}} + 2\Delta_{\mathrm{e}}$ and whose cross section has sides $d_{\mathrm{T}} = d_{\mathrm{cell}} + 2\Delta_{\mathrm{e}}$, then we have that σ_{el} and σ_{et} are given as follows:

$$\sigma_{\mathrm{el}} = \widetilde{\sigma}_{\mathrm{e}} \left(1 - \frac{\pi \left(\frac{d_{\mathrm{cell}}}{2} \right)^2}{d_{\mathrm{T}}^2} \right), \text{ and} \tag{2.20}$$

$$\sigma_{\mathrm{et}} = \widetilde{\sigma}_{\mathrm{e}} \left(1 - \frac{d_{\mathrm{cell}} l_{\mathrm{cell}}}{d_{\mathrm{T}} l_{\mathrm{T}}} \right), \tag{2.21}$$

where $\widetilde{\sigma}_{\mathrm{e}}$ is the conductivity of the extracellular medium in the myometrium. $\widetilde{\sigma}_{\mathrm{e}}$ can be computed using the effective myometrium conductivity $\sigma_{\mathcal{M}}$, available in the literature, and Archie's law [37] as follows:

$$\widetilde{\sigma}_{\mathrm{e}} = \frac{\sigma_{\mathcal{M}}}{(1-p)^m}, \tag{2.22}$$

with p the volume fraction occupied by the myocytes and collagenous fibers in the tissue, and m the so-called cementation factor, which depends on the shape and orientation of the myocyte in the tissue.

To compute the intracellular conductivities σ_{il} and σ_{it}, we assume that the intracellular and extracellular domains have equal anisotropy ratios, i.e.,

$$G_{\mathrm{i}} = \varsigma G_{\mathrm{e}}, \tag{2.23}$$

and thus we need to compute ς. We obtain an analytical expression for ς, using reported values of the propagation speed of a transmembrane potential waveform traveling on isolated tissue strips from pregnant human myometrium at term [41]. In particular, replacing (2.23), (2.12), and (2.13) in the bidomain equations (2.10) and (2.11), and solving $v_{\mathrm{m}}(\boldsymbol{r}, t)$ for a traveling wave solution $v_{\mathrm{m}}(\boldsymbol{\xi} \cdot \boldsymbol{r} - c\, t)$, with $\boldsymbol{\xi}$ a unitary vector pointing along the main axis of the myocyte and c the speed of propagation, we obtain the following expression for ς:

$$\varsigma = g\left(\frac{2\ c^2 \epsilon_1 a_{\mathrm{m}}\, c_{\mathrm{m}}^2}{\sigma_{\mathrm{el}}\ k\ \left(v_1^* - 2\ v_2^* + v_3^*\right)^2}\right), \tag{2.24}$$

where $g(x) = \frac{x}{1-x}$. Further, v_1^*, v_2^*, and v_3^* are the roots of the following polynomial in v_m:

$$f(v_m) = (v_{\mathrm{m}} - v_1)(v_2 - v_{\mathrm{m}})(v_{\mathrm{m}} - v_3) - \frac{1}{k\gamma}(\beta v_{\mathrm{mr}} + \delta), \tag{2.25}$$

with v_{mr} the resting transmembrane potential of the human myocyte. Note that in order to $\varsigma \geq 0$, then ϵ_1 has to satisfy the following inequality:

$$0 < \epsilon_1 < \frac{\sigma_{\mathrm{el}}\ k\ \left(v_1^* - 2\ v_2^* + v_3^*\right)^2}{2\ c^2 a_{\mathrm{m}}\, c_{\mathrm{m}}^2}. \tag{2.26}$$

2.4 Monodomain approximation and boundary conditions

The equal anisotropy ratio assumption, Eq. (2.23), simplifies the solution of the bidomain equations (2.10) and (2.11) by decoupling them as follows:

$$\nabla \cdot \frac{\varsigma}{(\varsigma+1)} G'_{\mathrm{e}} \nabla v_{\mathrm{m}}(\boldsymbol{r},t) = a_{\mathrm{m}} \left(c_{\mathrm{m}} \frac{\partial v_{\mathrm{m}}(\boldsymbol{r},t)}{\partial t} + \boldsymbol{J}_{\mathrm{ion}}(\boldsymbol{r},t) - \boldsymbol{J}_{\mathrm{stim}}(\boldsymbol{r},t) \right), \text{ in } \mathcal{M}, \quad (2.27)$$

$$\nabla \cdot \frac{(\varsigma+1)}{\varsigma} G'_{\mathrm{e}} \nabla \phi_{\mathrm{e}}(\boldsymbol{r},t) = -\nabla \cdot G'_{\mathrm{e}} \nabla v_{\mathrm{m}}(\boldsymbol{r},t), \text{ in } \mathcal{M}. \quad (2.28)$$

The above simplification is also known as the monodomain approximation of the bidomain equations, which, under suitable boundary conditions, allows to computing v_{m} and, thus, $\boldsymbol{J}_{\mathrm{s}}$, independent from ϕ_{e}.

To set boundary conditions for computing electrical potentials, we need to take into account the volume conductor geometry (see Fig. 2.3). In particular, we have two bidomain-monodomain interfaces: One between the myometrium $\mathcal{M}$ and abdominal volume $\mathcal{A}$, and one between the myometrium and the intrauterine cavity $\mathcal{U}$. Therefore, we have the following boundary conditions, namely, (i) continuity of the interstitial potential ϕ_{e} at the perimetrium surface $\partial\mathcal{M}$ to the abdomen potential $\phi_{\mathcal{A}}$, (ii) flow of the normal component of $\boldsymbol{J}$ that crosses over from the uterus to the abdominal medium, (iii) no flow of the normal component of $\boldsymbol{J}_{\mathrm{s}}$ to the abdominal medium, (iv) continuity of the interstitial potential ϕ_{e} at the endometrium surface $\partial\mathcal{U}$ to the intrauterine cavity potential $\phi_{\mathcal{U}}$ (v) flow of the normal component of $\boldsymbol{J}$ that crosses over from the uterus to the intrauterine cavity filled with amniotic fluid, (vi) no flow of the normal component of $\boldsymbol{J}_{\mathrm{s}}$ to the intrauterine cavity, (vii) no flow of the normal component of $\boldsymbol{J}$ that crosses over from the abdominal cavity to air, and (viii)

no flow or flow of the normal component of $\boldsymbol{J}$ that crosses over from the intrauterine cavity, filled with amniotic fluid, to the fetus depending if it is covered with vernix caseosa ($\lambda = 0$) or not ($\lambda = 1$) [43]. These boundaries conditions are summarized as follows:

$$\begin{aligned}
\phi_{\mathrm{e}}(r,t) &= \phi_{\mathcal{A}}(r,t), \text{ in } \partial\mathcal{M}, && (2.29)\\
\widehat{\boldsymbol{n}}_{\mathcal{M}} \cdot (G_{\mathrm{i}}' \nabla \phi_{\mathrm{i}}(r,t) + G_{\mathrm{e}}' \nabla \phi_{\mathrm{e}}(r,t)) &= \widehat{\boldsymbol{n}}_{\mathcal{M}} \cdot G_{\mathcal{A}} \nabla \phi_{\mathcal{A}}(r,t), \text{ in } \partial\mathcal{M}, && (2.30)\\
\widehat{\boldsymbol{n}}_{\mathcal{M}} \cdot G_{\mathrm{i}}' \nabla v_{\mathrm{m}}(r,t) &= 0, \text{ in } \partial\mathcal{M}, && (2.31)\\
\phi_{\mathrm{e}}(r,t) &= \phi_{\mathcal{U}}(r,t), \text{ in } \partial\mathcal{U} && (2.32)\\
\widehat{\boldsymbol{n}}_{\mathcal{U}} \cdot (G_{\mathrm{i}}' \nabla \phi_{\mathrm{i}}(r,t) + G_{\mathrm{e}}' \nabla \phi_{\mathrm{e}}(r,t)) &= \widehat{\boldsymbol{n}}_{\mathcal{U}} \cdot G_{\mathcal{U}} \nabla \phi_{\mathcal{U}}(r,t), \text{ in } \partial\mathcal{U}, && (2.33)\\
\widehat{\boldsymbol{n}}_{\mathcal{U}} \cdot G_{\mathrm{i}}' \nabla v_{\mathrm{m}}(r,t) &= 0, \text{ in } \partial\mathcal{U}, && (2.34)\\
\widehat{\boldsymbol{n}}_{\mathcal{A}} \cdot G_{\mathcal{A}} \nabla \phi_{\mathcal{A}}(r,t) &= 0, \text{ in } \partial\mathcal{A} && (2.35)\\
\widehat{\boldsymbol{n}}_{\mathcal{F}} \cdot G_{\mathcal{U}} \nabla \phi_{\mathcal{U}}(r,t) &= \lambda \left(\widehat{\boldsymbol{n}}_{\mathcal{F}} \cdot G_{\mathcal{F}} \nabla \phi_{\mathcal{F}}(r,t) \right), \text{ in } \partial\mathcal{F}, && (2.36)
\end{aligned}$$

where $\widehat{\boldsymbol{n}}_j$ is the normal vector to the surface j in each case.

2.5 Numerical computation

The computation of $v_{\mathrm{m}}(\boldsymbol{r},t)$, $\phi(\boldsymbol{r},t)$, and $\boldsymbol{B}(\boldsymbol{r},t)$ are given by the following procedure:

- **Step 1**: Solve for $v_{\rm m}(r,t)$ using Eqs. (2.27), (2.12), (2.13), and (2.14) subject to boundary conditions (2.31) and (2.34), and to initial conditions given by

$$\begin{aligned} v_{\rm m}(r,0) &= v_{\rm mr}, && (2.37)\\ w_{\rm m}(r,0) &= \left(\frac{\beta}{\gamma}v_{\rm mr}+\frac{\delta}{\gamma}\right), && (2.38)\\ \frac{\partial v_{\rm m}(r,0)}{\partial t} &= 0, \text{ and} && (2.39)\\ \frac{\partial w_{\rm m}(r,0)}{\partial t} &= 0. && (2.40) \end{aligned}$$

- **Step 2**: Solve for $\phi_{\rm e}(r,t)$ in $\mathcal{M}$ and $\phi(r,t)$ in $\mathcal{A}$ and $\mathcal{U}$, using the solution of $v_{\rm m}(r,t)$, computed in Step1, and the following expressions:

$$\begin{aligned} \nabla\cdot\frac{(\varsigma+1)}{\varsigma}G'_{\rm e}\nabla\phi_{\rm e}(r,t) &= -\nabla\cdot G'_{\rm e}\nabla v_{\rm m}(r,t), \text{ in } \mathcal{M}, && (2.41)\\ \nabla\cdot G_{\mathcal{A}}\nabla\phi(r,t) &= 0, \text{ in } \mathcal{A}, && (2.42)\\ \nabla\cdot G'_{\mathcal{U}}\nabla\phi(r,t) &= 0, \text{ in } \mathcal{U}, && (2.43) \end{aligned}$$

  subject to boundary conditions (2.29), (2.30), (2.32), (2.33), (2.35), and (2.36).

- **Step 3**: Solve for $\boldsymbol{B}(\boldsymbol{r},t)$ using Eq. (2.1), and computing the total current density $\boldsymbol{J}(\boldsymbol{r},t)$ in all the domain using the solutions of $v_{\rm m}(r,t)$, $\phi_{\rm e}(r,t)$, and $\phi(r,t)$, obtained in Steps 1 and 2.

To compute the solution in each of the above steps, we use the FEM solver COMSOL Multiphysics running on a server with 8 64-bit processors at 2.3GHz, with 32 Gb RAM.

2.6 Results and discussion

In the following we illustrate our modeling approach by considering the electrophysiological and anatomical characteristics of the uterus at term. In Figure 2.5 we illustrate the four-compartment volume conductor geometry used in the numerical examples. We defined a spherical myometrium of 16 cm radius measured from the center to $\partial\mathcal{M}$ and the uterine wall has a uniform thinness of 1 cm. We also consider an spherical fetus of 12 cm radius concentric to the myometrium fully covered with vernix caseosa, i.e., $\lambda = 0$ in (2.36). The adnominal compartment is also spherical with 21 cm radius shifted -3 cm from the center of the myometrium in the x axis. We set the coordinate axis of reference at the center of the myometrium.

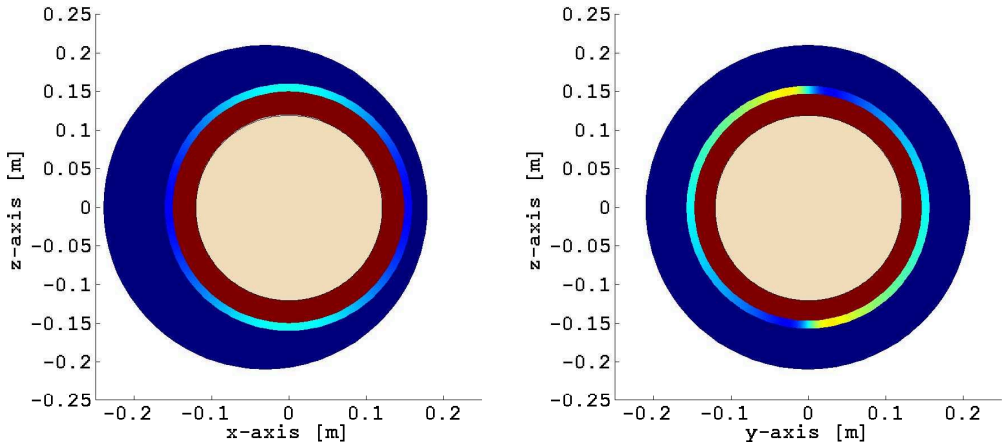


Figure 2.5: Four-compartment volume conductor geometry used in the numerical examples. (a) View of z-x plane, and (b) z-y plane. Each compartment is assigned a different color. The myometrium has a non-uniform color to denote that its conductivity is anisotropic.

The conductivity values for each compartment are given in Table 2.1. In particular, to compute the extracellular myometrial conductivity tensors, we use average values

for the uterine myocyte dimensions at term based on data reported in [1, 28, 32] (see Table 2.2). The average human myocyte can be best described as a long cylinder with a small cross section, therefore, we use a cementation factor $m = 4/3$ (see [37] for more details on the computation of this factor). The volume fraction p occupied by myocytes and collagenous fibers in the myometrium is set equal to 0.6. In order to consider an average myometrial fiber architecture, which ranges between circular and oblique fibers, we choose the fiber orientation angle α to be 45^o. Figure 2.6 illustrates the global structural of the myometrial fiber orientation for this angle.

Table 2.1: Conductivity values of the volume conductor geometry.

Symbol	Value	Reference
$G_{\mathcal{A}}$	0.2 S/m	[43]
$G_{\mathcal{U}}$	1.74 S/m	[43]
$G_{\mathcal{F}}$	0.2 S/m	[43]
σ_{el}	0.68 S/m	Eq. (2.20)
σ_{et}	0.22 S/m	Eq. (2.21)
ς	0.8	Eq. (2.24)

Table 2.2: Myocyte dimensions and Archie's law parameters.

Symbol	Value	Reference
d_{cell}	7 μm	[28]
l_{cell}	450 μm	[28]
d_{T}	8 μm	
l_{T}	451 μm	
$\sigma_{\mathcal{M}}$	0.5 S/m	[43]
m	4/3	[1, 37]
p	0.6	[29, 37]

We select the model parameters of the ionic current model using phase-space analysis, and using as a reference the average plateau-type transmembrane potentials recorded from isolated tissue strips of human myometrium at term [33, 40]. In particular, the

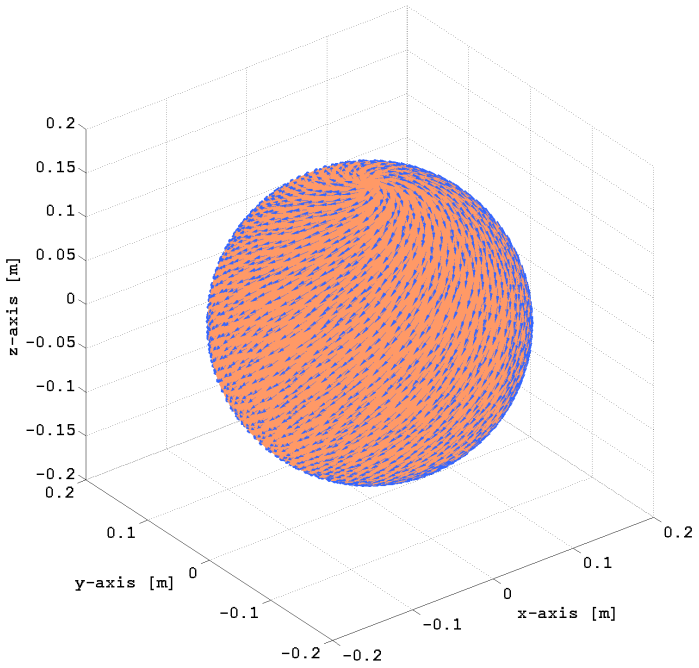


Figure 2.6: Geometry and fiber orientation in spherical myometrium given by $\alpha = 45^o$.

average resting potentials, considering the results reported from the 37th weeks of pregnancy onwards, is approximately -56 mV. The plateau has an average depolarization of -27 ± 1 mV that terminates in 0.9 ± 0.2 minutes by an abrupt repolarization to the resting level [33, 40]. Table 2.3 has the parameter values used in the numerical example. Note that we compute the surface-area to volume-ratio a_{m} using the myocyte dimensions in Table 2.2.

In [36] has been indicated that in the human uterus there may be a preferential direction of propagation of contractions, and thus of transmembrane potential propagation, from the fundus toward the isthmus, which could aid in the expulsion of the fetus. Therefore, in order to study this assumption with our model, we consider j_{stim}

Table 2.3: Ionic current model parameters.

Symbol	Value	Reference
c_m	0.01 F/m^2	[43]
v_{mr}	-0.056 V	[40]
a_m	5.758710^5 m^{-1}	Table 2.2
ϵ_1	200 ω m^2	
ϵ_2	0.09 1/s	
v_1	-0.02 V	
v_2	-0.04 V	
v_3	-0.065 V	
k	10^4 1/V^2	
δ	0.0520 V	
γ	0.1	
β	1	
c	1.15 cm/sec	[41]

with $N_p = 1$, $\nu_1 = 2$ V, and $h_1(r,t)$ as follows:

$$h_1(r,t) = \{1, \text{ if } 0 \leq t \leq 0.1 \ , \ 0.15 \ \leq \|r\| \leq 0.16 \text{ and } z \geq 0.15 \ ; \ 0, \text{ otherwise.}$$

The size and intensity of the pacemaker area are chosen in order to obtain a stable traveling waveform solution to the bidomain equations on the spherical myometrium.

Figs. 2.7 show several snapshots of the FEM solution for one pacemaker on the fundus of a spherical myometrium, assuming anisotropy given by Fig. 2.6. Figs. 2.7 (a)-(c) illustrate the transmembrane potential and source current density distribution at the myometrium, Figs. 2.7 (d)-(f) the electrical potential at the abdominal surface, and Figs. 2.7 (g)-(i) the magnetic field density at the abdominal surface. The magnetic field measured at the abdominal surface, $\boldsymbol{B}_{MMG}$, is proportional to $\boldsymbol{B}_{\boldsymbol{n}_{\mathcal{A}}}$, the projection of $\boldsymbol{B}$ onto the normal vector of the abdominal surface, $\boldsymbol{n}_{\mathcal{A}}$. Note that, because of the anisotropy in the conductivity, the direction of the current density $\boldsymbol{J}_{\mathrm{s}}$ is rotated

on a certain angle from the main direction of the propagating transmembrane potential v_{m}, and it is the transversal component of this current, parallel to the x-y plane, which generates the magnetic field $\boldsymbol{B}_{\boldsymbol{n}_{\mathcal{A}}}$. This observation is in agreement with the analysis presented in [44] and it is important to take it into account when interpreting the magnetic field measurements generated by uterine contractions in the presence of volume conductor geometry. Therefore, the spatial signature of $\boldsymbol{B}_{\boldsymbol{n}_{\mathcal{A}}}$ is highly dependent on the fiber orientation of the myometrium. Because of the proximity between the sensors and the myometrium, it is not strictly applicable to assume a moving dipole parallel to the direction of propagation of the transmembrane potential, as the main model for the current source generated the measured magnetic field. This last interpretation might be suitable in case the transversal length of the transmembrane potential front is short in com-parison to the area covered by the array of sensors. In contrast, if the transversal length is larger and thus no covered by the measured area, for example when several cells are recruited, then it suitable to consider a moving line source (stretched ring) model instead.

In Fig. 2.8 (a) we illustrate the temporal response of FEM solutions for the transmembrane potential at different elevations at times. It can be seen that a stable traveling waveform has been established as the shape remains same. Also, the maximum depolarization is -16 mV, the average potential in the plateau area is of -25 mV, and the transmembrane potential duration, before hyperpolarization, is around 35 s, which is a fair approximation with respect to the average recorded transmembrane potentials discussed in [33, 40]. Note that our ionic current model introduces hyperpolarization, which constrains the excitability of the cell and, thus, consecutive contractions can only take place until v_{m} reaches resting potential. In our case, the minimum time between two consecutive contractions is 240 s. In Fig. 2.8(b) we illustrate the percentage

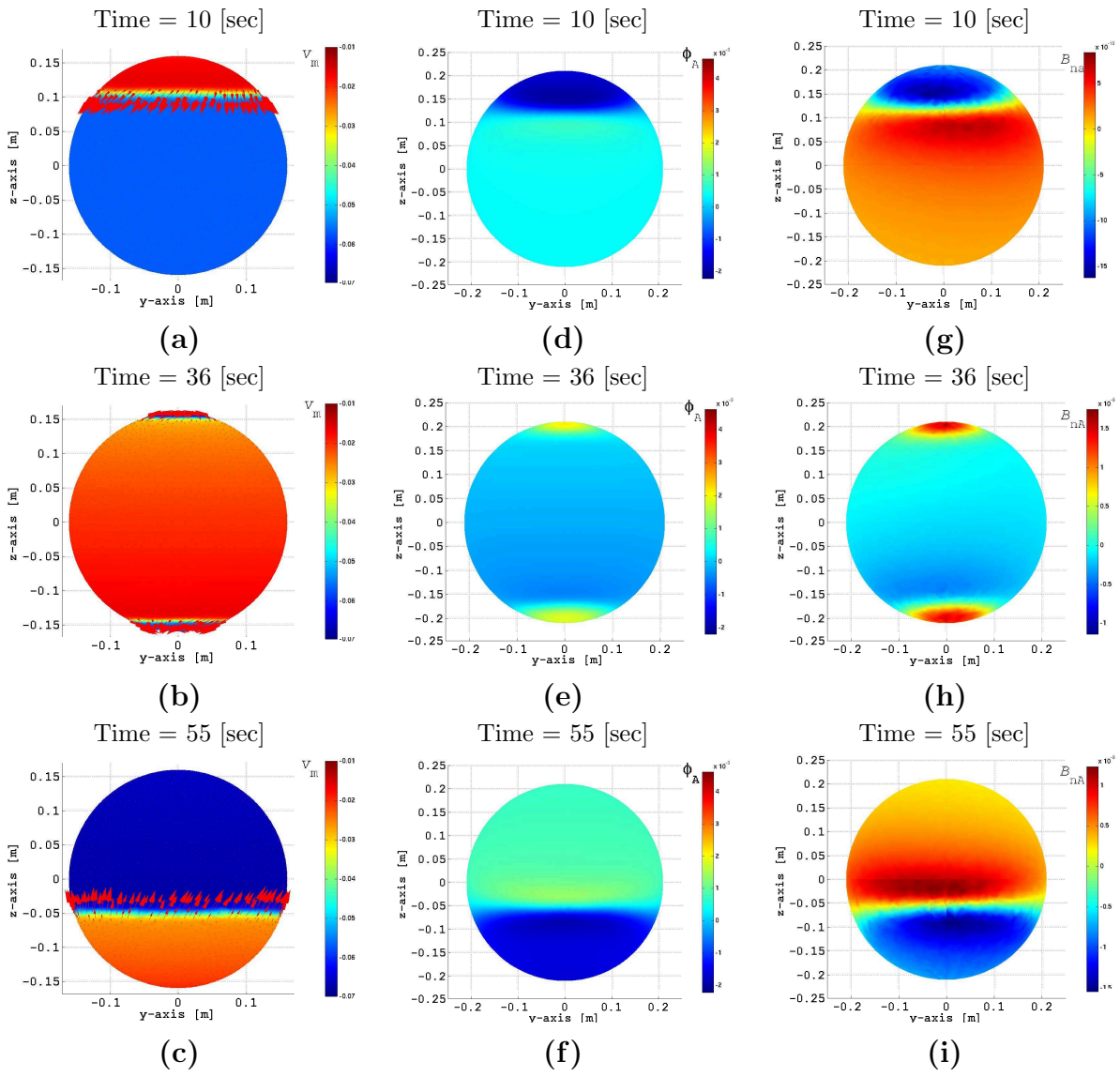


Figure 2.7: FEM solution at time instants t = 10 [s], 36 [s], 55 [s] for one pacemaker on the fundus of a spherical myometrium, assuming anisotropy. (a)-(c) transmembrane potential and source current density distribution at the myometrium, (d)-(f) electrical potential at the abdominal surface, and (g)-(i) magnetic field density at the abdominal surface.

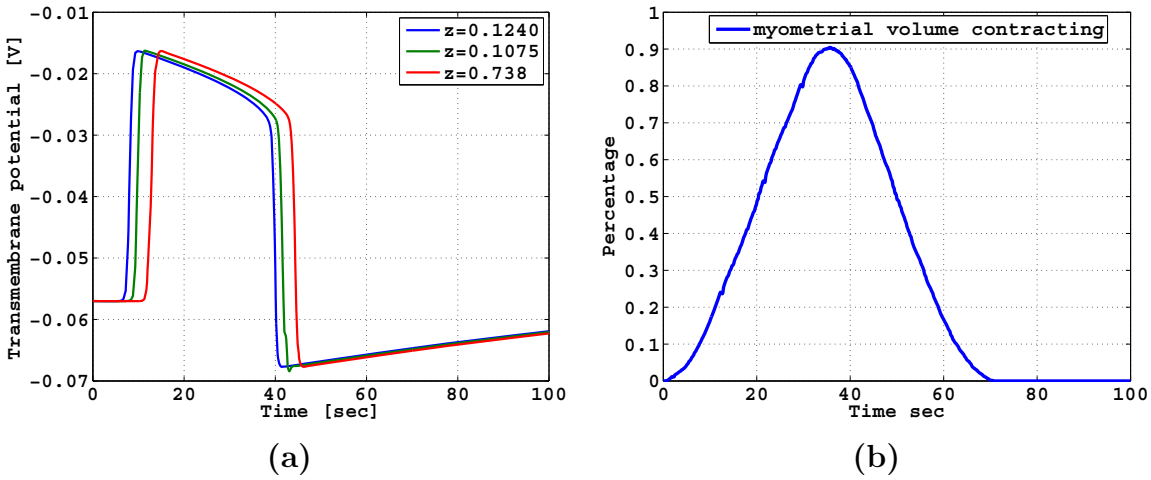


Figure 2.8: (a) Temporal response of FEM solutions for transmembrane potential at different elevations; (b) Percentage of contracting myometrial volume as a function of time.

of contracting myometrial volume as function of time, which in [31, 32] was used as a reference to compute the changes in the intrauterine pressure due to a contraction. Interestingly, we observe that the percentage of myometrial cell contracting has the symmetric properties and length of the intrauterine pressure waveforms of human pregnant myometrium at term, as discussed in [32]. Note that a larger ϵ_2 value can extend the transmembrane potential duration to values closer to the average duration reported on [33, 40], however, it also extends the duration of hyperpolarization and the plateau of the curve describing the percentage of contracting myometrial volume.

2.7 Summary

We proposed a forward electromagnetic model of uterine contractions during pregnancy. Our model incorporates knowledge of the electrophysiological aspects of the uterine contractions during pregnancy at both the cellular and organ levels. We applied a bidomain approach for modeling the propagation of the myometrium transmembrane potential v_{m} on the uterus and used this to compute the action potential ϕ and the magnetic field $\boldsymbol{B}$ at the abdominal surface. We introduced a modified version of the FitzHugh-Nagumo equation for modeling the ionic currents in each cell. Though our ionic current model does not consider explicitly C_a^{2+}dynamics, the simplicity of the FitzHugh-Nagumo allows for capturing the nuances of the uterine myocyte response, but it can be used to model the propagating action potential under well defined conditions as shown in this paper. We also proposed a general approach to design conductivity tensors in the myometrium and to estimate the conductivity tensor values in the extracellular and intracellular domains. We introduced a simplified geometry for the problem and proposed a discretized model solution based on a finite element method approach. Finally, we illustrated our modeling approach through a numerical example modeling a uterine contractions at term. Our model is potentially important as a tool for helping in the characterization of contractions and for predicting labor using MMG and EMG.

Chapter 3

Detection of Uterine Contractions Using MMG[1]

3.1 Introduction

As we elaborated in Chapter 1, the analysis of uterine contractions during pregnancy is clinically important for predicting labor [6,45]. Therefore, an automated method to detect uterine contractions can be very helpful in the clinical evaluation of a patient. Detection of uterine contractions has been performed in the past using either a single EHG channel [9–11,46] or an array of MMG channels [47]. Among these references, we find time domain and multiresolution domain analyses. In the time-domain analysis, the uterine contractions are detected by applying a discrimination rule (threshold) on the values of a function (feature) evaluated in a sliding time window. For example, using EHG measurements, the authors in [48] discriminate the root mean square (RMS) values using an experimental threshold. In [9] the authors discuss the feasibility of using the first-order zero-crossing (FOZC) as a feature. However, no threshold

[1]Based on P. S. La Rosa, A. Nehorai, H. Eswaran, C. Lowery, and H. Preissl, "Detection of uterine MMG contractions using a multiple change point estimator and the K-means cluster algorithm," *IEEE Trans. on Biomedical Engineering,* Vol. 55, pp. 453-467, Feb. 2008.

for detection is provided. In [11], the authors discriminate the RMS values using an adaptive threshold level, determined by applying a moving-media average filter on an overlapping sliding window of EHG measurements. In [47] the authors propose to use the generalized synchronization index as an indicator of uterine contractions, using an array of MMG channels.

In a multiresolution analysis, a single EHG measurement is represented in a higher dimensional linear space, which is formed by decomposing the signal into several subspaces spanned by a set of basis functions [49]. For example, in [46] a single channel EHG measurement is decomposed using a set of biorthogonal wavelet functions, and the contraction is detected using a two-stage algorithm. In the first stage, the authors design a multivariable sequential cumulative likelihood ratio test to sequentially detect changes in the covariance matrix of the decomposed signal. Then, in the second stage, the diagonal parameters of the covariance matrices are classified using a neural network trained by an expert. As with the time-domain techniques, to detect changes in the parameters, the segmentation stage of this multiresolution approach requires setting up a time window as well as a threshold level. Also, the basis functions (wavelets) are selected based on the average performance obtained in the classification stage.

One common restriction of the above techniques is that the time-window length, as well as the threshold level, is determined experimentally based on a particular data set. In general the selection of the time-window length depends on the length of the shortest event to be detected, on the global detection delay, and on the number of samples required to properly estimate the hypothesis parameters [46]. Therefore, the detection performance might change from patient to patient because of, for example, differences in the pregnancy stage, tissue conductivities, etc.

In this chapter we propose a single channel two-stage time-segment discriminator of uterine-contraction magnetomyograms (MMGs). In the first stage, we propose a model-based segmentation procedure, which detects multiple change-points in the parameters of a piecewise constant time-varying autoregressive model using a robust formulation of the Schwarz information criterion (SIC) and a binary search approach. This novel segmentation technique does not require the use of a sliding window as in previous work. In particular, we propose a test statistic that depends on the SIC, derive its asymptotic distribution, and obtain closed-form optimal detection thresholds in the sense of the Neyman-Pearson criterion; therefore, we control the probability of false alarm and maximize the probability of change-point detection in each stage of the binary search algorithm. Then, in the second stage, we apply the non-supervised K-means cluster algorithm to classify each time segment, using the RMS and FOZC as candidate features. Finally, we provide a discrete-time binary decision signal indicating the presence of a contraction. We validate our method using real MMG measurements and compare the detected time intervals with the patient's feedback. Since each single channel detector provides local information regarding the presence of a contraction, we also analyze the fusion of the decision signals from all the sensors, as in the parallel distributed detection approach. As we illustrate with real data, this approach proves to be helpful in understanding uterine-contraction MMGs spatially and temporally. Figure 3.1 illustrates the proposed single channel scheme to estimate the contraction intervals.

This chapter is organized as follows: in Section II we present the model-based time-domain segmentation method and analyze its performance; in Section III we define the feature space and describe the classification method. We discuss the performance

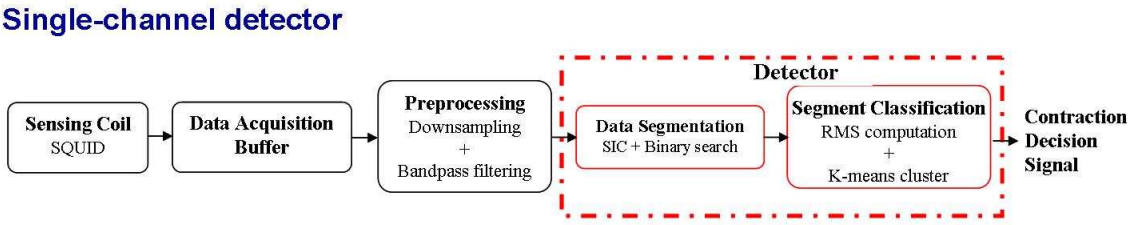


Figure 3.1: A diagram of the proposed single channel detector to detect uterine contractions, using MMG.

of our algorithm using real MMG data in Section IV, and summarize our results in Section V.

3.2 Model-based time-domain segmentation

Let $y(t)$, $t = 1, \ldots, n$ be the samples of an MMG measurement acquired from a position close to the abdominal surface of the patient, where n is the number of total samples. Let $x(t) = f(y(1), y(2), \ldots, y(l))$ for $l \leq n$, $t = 1, \ldots, n$, be the preprocessed measurement, where $f : \mathbb{R}^l \mapsto \mathbb{R}$ is a function defined over the MMG samples. In our case, f represents the downsampling followed by a bandpass filtering. Using the central limit theorem, we assume that the distribution of the preprocessed samples' amplitudes is Gaussian with unknown mean and variance. Also, as shown in [50,51], EMG measurements can be considered as a series of stationary segments with Gaussian distributed amplitudes. Therefore, we design a time-series segmentation algorithm that detects multiple change points in the distributions parameters.

3.2.1 Detection principle

The statistical problem of segmentation can be formulated as follows: assuming that (i) $x(t)$ is a piecewise stationary real process so that there exist instants $\{t'_i,\, i > 0\}$ such that $(x(t'_i + 1), \ldots, x(t'_{i+1}))$ is stationary $\forall\, i \in \mathbb{N}$. Assume that the probability distribution of the samples between $t'_i + 1$ and t'_{i+1} belongs to the same family of distributions $\mathbb{P}(\boldsymbol{\theta})$, where $\boldsymbol{\theta} \in \mathbb{R}^m$ are the unknown parameters. Then our problem consists of detecting changes in the distribution of $\{x(t), 1 \leq t \leq n\}$ associated with changes in $\boldsymbol{\theta}$. Therefore, we apply a statistical test looking for changes in $\boldsymbol{\theta}$. Usually the statistical inference about change points has two aspects: the first is to detect if there is any change in the sequence of random variables observed, and the second is to estimate the number of change points and their corresponding locations. In particular, the statistical inference of the change-point problem consists of testing the following hypotheses:

$$\mathrm{H}_0 : \boldsymbol{\theta}(1) = \boldsymbol{\theta}(2) = \cdots = \boldsymbol{\theta}(n) = \boldsymbol{\theta}_0, \tag{3.1}$$

$$\mathrm{H}_1 : \boldsymbol{\theta}(1) = \cdots = \boldsymbol{\theta}(t'_1) \neq \boldsymbol{\theta}(t'_1 + 1) = \cdots = \boldsymbol{\theta}(t'_2) \neq \cdots \neq \boldsymbol{\theta}(t'_q + 1) = \cdots = \boldsymbol{\theta}(n), \tag{3.2}$$

where $\boldsymbol{\theta}$ is the unknown parameter vector, q is the unknown number of the change points, and $1 < t'_1 < t'_2 < \cdots < t'_q < n$ are the unknown positions of the change points. Therefore, our problem is to estimate the set of change points $t'_1, \ldots, t'_q$. Note that in our case, we assume that all the data are available, and thus our detection approach is offline or, at best, has a delay of n.

The problem of detecting multiple change points has been addressed by means of a binary segmentation procedure in [52, 53]. Essentially, this approach simplifies the

general problem by evaluating iteratively the alternative hypothesis of detecting a single change point in the parameter. It can be described as follows. First, detect a single change. If there is no change, then the null hypothesis is accepted. If there is a change, then the estimated change point divides the original sequence into two subsequences. For each subsequence, detect a change as in the first step, and continue the process until no more changes can be found in any of the subsequences. Using the above approach, the alternative hypothesis becomes

$$\mathrm{H}_1 : \boldsymbol{\theta}(1) \;=\; \cdots = \boldsymbol{\theta}(t_0') = \boldsymbol{\theta}_1 \neq \boldsymbol{\theta}(t_0'+1) = \cdots = \boldsymbol{\theta}(n) = \boldsymbol{\theta}_2. \tag{3.3}$$

Here we estimate the change point t_0' using the Schwarz information criterion (SIC)[2] [55], which is defined as

$$\mathrm{SIC}(\tau) = -2\,\ln(L(\hat{\boldsymbol{\theta}}_\tau)) + m_\tau \ln(n),\ \tau = 1, \ldots, n \tag{3.4}$$

where $L(\cdot)$ is the likelihood function of the samples, $\hat{\boldsymbol{\theta}}_\tau$ is the maximum likelihood estimation of $\theta_\tau = [\boldsymbol{\theta}_1, \boldsymbol{\theta}_2]$ (assuming a change point at sample $t = \tau$), and $m_\tau = 2m$ is the number of unconstrained parameters in the model under H_1 of (3.3), with $t_0' = \tau$. The case $\tau = n$ corresponds to the situation without a change point, thus, $\theta_\tau = [\boldsymbol{\theta}_0]$ and $m_\tau = m$. The SIC has been applied in [53] to estimate the change points of the variance of a normal distribution. They preferred using the SIC over the AIC, because the minimum of the SIC among the possible models has been proved to be an asymptotically unbiased estimator of the model order of the true model [53, 55], which makes this criterion appropriate for designing a change point test [53]. Hence, the decision to accept H_0 or H_1 is based on the principle of the minimum information criterion. Thus H_0 is accepted if $\mathrm{SIC}(n) \leq \min_\tau \mathrm{SIC}(\tau)$, and

[2] SIC is a modification of the Akaike information criterion (AIC) proposed by Akaike in 1973 [54].

H_0 is rejected if $\mathrm{SIC}(n) > \mathrm{SIC}(\tau)$ for some τ. In the latter case the change point t_0' is estimated by $\widehat{t'}_0$ such that

$$\mathrm{SIC}(\widehat{t_0'}) = \min_{m<\tau<n-m-1} \mathrm{SIC}(\tau), \tag{3.5}$$

where $\mathrm{SIC}(n)$ is the SIC under H_0, $\mathrm{SIC}(\tau)$ is the SIC under H_1 for $\tau = m+1, \ldots, n-m-1$, and m is the total number of unconstrained parameters under H_0. Note that τ is constrained to the set $m < \tau < n-m-1$ so that we can compute the maximum likelihood estimates of the parameters, and $n > 2m+1$.

The SIC approach for testing change points does not require resorting to the distribution of the test statistic [53] as in the classical testing procedures, such as the generalized likelihood ratio test. If the SIC values are very close, then it is of interest to test if this difference is caused by data fluctuation when there is actually not change. To avoid misleading change-point detections, in [53] a robust SIC-based change-point detector is proposed. We will derive a test for the same purpose but based on a time-varying AR-model of the preprocessed measurements $x(t)$.

3.2.2 AR-modeling based segmentation

Assume that we model the time series $x(t)$ using a stable autoregressive (AR) model of fixed known order d, with time-varying coefficients; driven by a zero-mean white Gaussian process with time-varying variance. Let $\{x(t), 1 \leq t \leq n\}$ be the sequence of preprocessed measurements. Then

$$x(t) = \varphi_1(t)x(t-1) + \ldots + \varphi_d(t)x(t-d) + \varphi_0(t) + w(t), \tag{3.6}$$

or equivalently

$$w(t) = x(t) - \Phi^T(t)\,\tilde{\boldsymbol{x}}_{t-d,t}, \tag{3.7}$$

where $\tilde{\boldsymbol{x}}_{\beta,\upsilon} = [x(\beta), x(\beta+1), \ldots, x(\upsilon), 1]^T, \beta < \upsilon$, $\Phi(t) = [\varphi_d(t), \ldots, \varphi_1(t), \varphi_0(t)]^T$, and $w(t)$ is a zero-mean Gaussian process with variance $\sigma^2(t)$. Then, we test the following hypothesis,

$$\mathrm{H}_0 : \boldsymbol{\phi}(1) = \boldsymbol{\phi}(2) = \cdots = \boldsymbol{\phi}(n) = \boldsymbol{\phi}_0, \tag{3.8}$$

$$\mathrm{H}_1 : \boldsymbol{\phi}(1) = \cdots = \boldsymbol{\phi}(t_1') \neq \boldsymbol{\phi}(t_1'+1) = \cdots = \boldsymbol{\phi}(t_2') \neq \cdots \neq \boldsymbol{\phi}(t_q'+1) = \cdots = \boldsymbol{\phi}(n), \tag{3.9}$$

where $\boldsymbol{\phi}(t) = [\varphi_d(t), \cdots, \varphi_1(t), \varphi_0(t), \sigma^2(t)]^T$ is unknown, q is the unknown number of change points, and $\{1 < t_1' < t_2' \cdots < t_q' < n\}$ are the unknown instants of the change points. Hence, our problem is to estimate the set of change points $t_1', \cdots, t_q'$. Accordingly, using the binary search procedure, the alternative hypothesis becomes

$$\mathrm{H}_1 : \boldsymbol{\phi}(1) = \cdots = \boldsymbol{\phi}(t_0') = \boldsymbol{\phi}_1 \neq \boldsymbol{\phi}(t_0'+1) = \cdots = \boldsymbol{\phi}(n) = \boldsymbol{\phi}_2. \tag{3.10}$$

Let $\boldsymbol{x}_{\beta,\upsilon} = [x(\beta), x(\beta+1), \ldots, x(\upsilon)]^T$, where $\beta < \upsilon$. We approximate the likelihood function of the n samples under both hypotheses by considering the $n-d$ samples conditioned on the d initial values. This approximation allows us to obtain closed form solutions for the maximum likelihood estimates of the parameters and to reduce the computational complexity of the algorithm. Therefore, under H_0, the approximate

likelihood function is given by

$$\begin{aligned} L(\boldsymbol{\phi}_0|\mathrm{H}_0) &= \Pr[\boldsymbol{x}_{d+1,n}|\boldsymbol{x}_{1,d}] \\ &= \frac{1}{(2\,\pi\,\sigma_0^2)^{\frac{n-d}{2}}} \exp\left\{\frac{-1}{2\,\sigma_0^2}\left\{\sum_{t=d+1}^{n}(x(t) - \Phi_0^T\tilde{\boldsymbol{x}}_{t-d,t})^2\right\}\right\}. \end{aligned} \tag{3.11}$$

Then the approximate log likelihood is

$$\begin{aligned} \ln L(\boldsymbol{\phi}_0|\mathrm{H}_0) &= -\frac{n-d}{2}\ln(2\pi\sigma_0^2) \\ &\quad -\frac{1}{2\,\sigma_0^2}\left\{\Phi_0{}^T\mathbf{R}_{\tilde{\boldsymbol{x}},\tilde{\boldsymbol{x}}}(d+1,n)\,\Phi_0 - 2\,\Phi_0{}^T\boldsymbol{r}_{\tilde{\boldsymbol{x}},x}(d+1,n) + \sum_{t=d+1}^{n}x^2(t)\right\}, \end{aligned} \tag{3.12}$$

where $\boldsymbol{r}_{\tilde{\boldsymbol{x}},x}(d+1,n) \in \mathbb{R}^{d+1\times 1}$ with $\boldsymbol{r}_{\tilde{\boldsymbol{x}},x}(d+1,n) = \sum_{t=d+1}^{n}\tilde{\boldsymbol{x}}_{t-d,t}x(t)$, and $\mathbf{R}_{\tilde{\boldsymbol{x}},\tilde{\boldsymbol{x}}}(d+1,n) \in \mathbb{R}^{d+1\times d+1}$ with $\mathbf{R}_{\tilde{\boldsymbol{x}},\tilde{\boldsymbol{x}}}(d+1,n) = \sum_{t=d+1}^{n}\tilde{\boldsymbol{x}}_{t-d,t}\tilde{\boldsymbol{x}}_{t-d,t}^T$. Then, the maximum likelihood estimates $\widehat{\sigma_0^2}$ and $\widehat{\Phi}_0$ are obtained by equating to zero the derivatives of (3.12) with respect to $\widehat{\sigma_0^2}$ and $\widehat{\Phi}_0$, respectively, and are given by

$$\widehat{\sigma_0^2} = \frac{\sum_{t=d+1}^{n}(x(t) - \widehat{\Phi}_0^T\tilde{\boldsymbol{x}}_{t-d,t})^2}{n-d} \quad \text{and} \quad \widehat{\Phi}_0 = \{\mathbf{R}_{\tilde{\boldsymbol{x}},\tilde{\boldsymbol{x}}}(d+1,n)\}^{-1}\{\boldsymbol{r}_{\tilde{\boldsymbol{x}},x}(d+1,n)\}. \tag{3.13}$$

Now, replacing the concentrated likelihood (i.e., the likelihood function written, when possible, as a function of a particular parameter) as a function of σ_0^2 in (3.4) with $m_n = d+2$, we have the SIC(n) given by

$$\mathrm{SIC}(n) = (n-d)\,\ln(2\pi) + (n-d)\,\ln(\widehat{\sigma_0^2}) + (n-d) + (d+2)\,\ln(n). \tag{3.14}$$

Under H_1, considering a change point at time τ, we obtain the likelihood function as

$$\begin{aligned} L(\boldsymbol{\phi}_1, \boldsymbol{\phi}_2|\mathrm{H}_1) &= \Pr[\boldsymbol{x}_{d+1,n}|\boldsymbol{x}_{1,d}] \\ &= \Pr[\boldsymbol{x}_{d+1,\tau}|\boldsymbol{x}_{1,d}]\ \Pr[\boldsymbol{x}_{\tau+1,n}|\boldsymbol{x}_{\tau-d+1,\tau}], \end{aligned} \tag{3.15}$$

where

$$\Pr[\boldsymbol{x}_{d+1,\tau}|\boldsymbol{x}_{1,d}] = \frac{1}{(2\,\pi\,\sigma_1^2)^{\frac{\tau-d}{2}}} \exp\left\{\frac{-1}{2\,\sigma_1^2}\left\{\sum_{t=d+1}^{\tau}(x(t) - \boldsymbol{\Phi}_1^T\tilde{\boldsymbol{x}}_{t-d,t})^2\right\}\right\}, \tag{3.16}$$

and

$$\Pr[\boldsymbol{x}_{\tau+1,n}|\boldsymbol{x}_{\tau-d,\tau}] = \frac{1}{(2\,\pi\,\sigma_2^2)^{\frac{n-\tau}{2}}} \exp\left\{\frac{-1}{2\,\sigma_2^2}\left\{\sum_{t=\tau+1}^{n}(x(t) - \boldsymbol{\Phi}_2^T\tilde{\boldsymbol{x}}_{t-d,t})^2\right\}\right\}. \tag{3.17}$$

Then, applying the natural logarithm to (3.15), we have

$$\begin{aligned} \ln L(\boldsymbol{\phi}_1, \boldsymbol{\phi}_2|\mathrm{H}_1) &= -\frac{n-d}{2}\ln(2\pi) - \frac{\tau-d}{2}\ln(\sigma_1^2) - \frac{n-\tau}{2}\ln(\sigma_2^2) - \\ &\quad \frac{1}{2\,\sigma_1^2}\left\{\boldsymbol{\Phi}_1{}^T\mathbf{R}_{\tilde{\boldsymbol{x}},\tilde{\boldsymbol{x}}}(d+1,\tau)\,\boldsymbol{\Phi}_1 - 2\,\boldsymbol{\Phi}_1{}^T\boldsymbol{r}_{\tilde{\boldsymbol{x}},x}(d+1,\tau) + \sum_{t=d+1}^{\tau}x^2(t)\right\} - \\ &\quad \frac{1}{2\,\sigma_2^2}\left\{\boldsymbol{\Phi}_2{}^T\mathbf{R}_{\tilde{\boldsymbol{x}},\tilde{\boldsymbol{x}}}(\tau+1,n)\,\boldsymbol{\Phi}_1 - 2\,\boldsymbol{\Phi}_1{}^T\boldsymbol{r}_{\tilde{\boldsymbol{x}},x}(\tau+1,n) + \sum_{t=\tau+1}^{n}x^2(t)\right\}, \end{aligned} \tag{3.18}$$

where the maximum likelihood estimates $\widehat{\sigma_1^2}$ and $\widehat{\boldsymbol{\Phi}}_1$ are given by (3.13) by substituting $n = \tau$, and $\widehat{\sigma_2^2}$ and $\widehat{\boldsymbol{\Phi}}_2$ are given by

$$\widehat{\sigma_2^2} = \frac{\sum_{t=\tau+1}^{n}(x(t) - \widehat{\boldsymbol{\Phi}}_2^T\tilde{\boldsymbol{x}}_{t-d,t})^2}{n-\tau} \quad \text{and} \quad \widehat{\boldsymbol{\Phi}}_2 = \{\mathbf{R}_{\tilde{\boldsymbol{x}},\tilde{\boldsymbol{x}}}(\tau+1,n)\}^{-1}\{\boldsymbol{r}_{\tilde{\boldsymbol{x}},x}(\tau+1,n)\}. \tag{3.19}$$

Then replacing the concentrated likelihood as a function of $\widehat{\sigma_1^2}$ and $\widehat{\sigma_2^2}$ in (3.4) with $m_n = 2(d+2)$, the $\mathrm{SIC}(\tau)$ is given as follows,

$$\mathrm{SIC}(\tau) = (n-d)\,\ln(2\pi) + (\tau-d)\,\ln(\widehat{\sigma_1^2}) + (n-\tau)\,\ln(\widehat{\sigma_2^2}) + (n-d) + 2\,(d+2)\,\ln(n), \tag{3.20}$$

and the change point t_0' is estimated by $\widehat{t'}_0$ such that

$$\mathrm{SIC}(\widehat{t'}_0) = \min_{d+2<\tau\leq n-d-3} \mathrm{SIC}(\tau). \tag{3.21}$$

Then, for improving the robustness in detecting a change-point, the hypothesis H_0 is accepted when $\mathrm{SIC}(n) < \min_{d+2<\tau\leq n-d-3} \mathrm{SIC}(\tau) + \gamma$, where γ is a threshold value to be determined. Let λ_{SIC} be the test statistic defined as

$$\begin{aligned} \lambda_{\mathrm{SIC}} &= \mathrm{SIC}(n) - \min_{d+2<\tau\leq n-d-3} \mathrm{SIC}(\tau) \\ &= \max_{d+2<\tau\leq n-d-3} \{\mathrm{SIC}(n) - \mathrm{SIC}(\tau)\}. \end{aligned} \tag{3.22}$$

Then H_0 or H_1 is accepted if

$$\lambda_{\mathrm{SIC}} \underset{\mathrm{H}_0}{\overset{\mathrm{H}_1}{\gtrless}} \gamma. \tag{3.23}$$

Hence, if we know the probability function of λ_{SIC}, we can determine an optimal $\gamma = C_\alpha$ using, for example, the Neyman-Pearson criterion as follows,

$$\Pr\left[\lambda_{\mathrm{SIC}} < C_\alpha | \mathrm{H}_0\right] = 1 - \alpha, \tag{3.24}$$

where α is the significance level of the test. One possibility to determine the probability distribution of λ_{SIC} under H_0 would be to estimate it empirically (histogram)

using Monte Carlo simulations. However, the test statistic depends on the sample size, which is not fixed, and it also depends on the binary search procedure; therefore, we would have to compute a histogram for all possible sample sizes, which is impractical. Here, for $2d+1 < \tau \leq n-d-3$, we obtain and apply an approximation of the distribution of λ_{SIC} under H_0 as follows (see Appendix B for details on the derivation),

$$\Pr\left[\lambda_{\text{SIC}} < C_\alpha | \text{H}_0\right] \;\cong\; \exp\left(-2\,\exp-\left(\frac{C_\alpha + (d+2)\,\ln(n) - b_n(d+2)}{2\,a_n(d+2)}\right)\right), \tag{3.25}$$

where

$$b_n(x) \;=\; \frac{\{2\ln\ln(n) + x/2\ln\ln\ln(n) - \ln(\Gamma(x/2))\}^2}{2\ln\ln(n)} \tag{3.26}$$

$$a_n(x) \;=\; \sqrt{\left(\frac{b_n(x)}{2\ln\ln(n)}\right)}, \tag{3.27}$$

and

$$\lambda_{\text{SIC}} = \max_{2d+1\,<\tau\leq\,n-d-3}(n-d)\,\ln(\widehat{\sigma_0^2}) - (\tau-d)\,\ln(\widehat{\sigma_1^2}) - (n-\tau)\,\ln(\widehat{\sigma_2^2}) - (d+2)\ln(n). \tag{3.28}$$

Therefore, applying the above approximation in (3.24) we find that α and C_α are related as follows,

$$C_\alpha(m,n) \approx b_n(m) - 2a_n(m)\left\{\ln(2) - \ln\ln((1-\alpha)^{-1})\right\} - (m)\ln(n), \tag{3.29}$$

where $m = d+2$. The constant $C_\alpha(m,n)$ depends on the sample size n, the significance level α, and the number of parameters m. Note that, if $n < 3d+4$, then the given

segment is not tested for changes and H_0 is accepted. The binary search procedure is finalized when the hypothesis H_0 is accepted for all time segments.

Remark - Here, we derived a robust SIC change-point test to detect changes in the distribution of the preprocessed measurements $x(t)$. In particular, we assumed that the preprocessed samples are modelled by a piecewise time-varying autoregressive (AR) model of order d with input given by a white Gaussian noise with time-varying variance. We also consider the special case assuming that the AR model order is zero, and thus $\{x(t), 1 \leq t \leq n\}$ is a sequence of independent zero-mean normal random variables.

Special case

Assume that $d = \phi_0 = 0$ in (3.6). Then $\{x(t),\, 1 < t < n\}$ is a sequence of independent zero-mean normal random variables with parameters $\sigma^2(1)$, $\sigma^2(2)$, ..., $\sigma^2(n)$. Hence, our problem is reduced to looking for changes in the variance of a sequence of independent normal random variables. Similarly, the multiple change-point problem is given by (3.8) with $\boldsymbol{\phi}(t) = [\sigma^2(t)]$, and the alternative hypothesis becomes as in (3.10). The SIC for H_0 is

$$\mathrm{SIC}(n) = n\, \ln(2\pi) + n\, \ln(\widehat{\sigma_0^2}) + n + \ln n, \tag{3.30}$$

and for H_1 is

$$\mathrm{SIC}(\tau) = n\, \ln(2\pi) + \tau\, \ln(\widehat{\sigma_1^2}) + (n-\tau)\, \ln(\widehat{\sigma_2^2}) + n + 2\, \ln n, \tag{3.31}$$

where $\widehat{\sigma^2} = \sum_{i=1}^{n} x_i^2/n$, $\widehat{\sigma_1^2} = \sum_{i=1}^{\tau} x_i^2/\tau$, and $\widehat{\sigma_2^2} = \sum_{i=\tau+1}^{n} x_i^2/(n-\tau)$ are the maximum likelihood estimators under H_0 and H_1, respectively. The test statistic for this problem is

$$\lambda_{\mathrm{SIC}} = \max_{1<\tau\leq n-2} \{\mathrm{SIC}(n) - \mathrm{SIC}(\tau)\}\,. \tag{3.32}$$

The approximate distribution of λ under H_0 is given by (3.25) with $m = 1$, and the approximate threshold value is $C_\alpha(1, n)$.

Model order estimation

The model order d in (3.6) is typically unknown and needs to be estimated. A possible information criterion approach to estimate the model order is by minimizing the overall SIC, which is the SIC computed on $\{x(t), 1 \leq t \leq n\}$ using the estimated change points for a given set of d values. It is given by

$$\mathrm{SIC}_T = -2\,\ln(L(\widehat{\boldsymbol{\theta}}_d(t'_1), \ldots, \widehat{\boldsymbol{\theta}}_d(t'_u + 1))) + (u+1)\,(d+2)\,\ln(n), \tag{3.33}$$

where $\widehat{\boldsymbol{\theta}}_d(t'_i)$, $i = 1, \ldots, u+1$ is the maximum likelihood estimate of $\boldsymbol{\theta}_d(t'_i)$ using the estimated change points $\{\widehat{t'}_1, \ldots, \widehat{t'}_u\}$ and assuming a model order d. Therefore, d can be estimated by,

$$\hat{d} = \arg\left\{\min_d \mathrm{SIC}_T\right\}. \tag{3.34}$$

3.2.3 Performance analysis

Using terminology from signal detection theory[3], we analyze the performance of the change-point detector by computing the probability of detection (P_{D}) and the probability of false alarm (P_{FA}). P_{D} of our problem is given by

$$P_{\mathrm{D}} = \Pr(\lambda_{\mathrm{SIC}} > \gamma | \mathrm{H}_1), \tag{3.35}$$

and P_{FA} is

$$P_{\mathrm{FA}} = \Pr(\lambda_{\mathrm{SIC}} > \gamma | \mathrm{H}_0), \tag{3.36}$$

where $\gamma \in \mathbb{R}$ is the threshold that defines the decision regions. Using (3.25), we approximate P_{FA} as follows:

$$P_{\mathrm{FA}} \cong 1 - \exp\left\{-2 \exp\left\{-\frac{\gamma + (m) \ln(n) - b_n(m)}{2\, a_n(m)}\right\}\right\}. \tag{3.37}$$

To analyze P_{D} as a function of P_{FA}, we plot the receiver operating characteristic (ROC), which is a parametric plot of P_{D} versus P_{FA}. We compute the P_{D} and P_{FA} as a function of γ for a sample size n equal to 100 and 1,000, and $d = \phi_0 = 0$ using 1,000 Monte Carlo simulations. Figure 3.2(a) illustrates the P_{FA} computed through Monte Carlo simulations and the P_{FA} using (3.37). As expected, when the sample size n increases, the P_{FA} given by (3.37) approaches the P_{FA} computed numerically, and both decrease their values as γ increases.

[3]In signal detection theory the probability of a type I error is called the probability of false alarm (P_{FA}), whereas the probability of a type II error is called the probability of a miss (P_{M}). The quantity 1 - P_{M} is the probability of detection (P_{D}).

The test statistic λ under H_1 depends on $\widehat{\sigma_1^2}$, $\widehat{\sigma_2^2}$, $\hat{t}$, and n. For the purpose of analysis, we set ${\sigma_2}^2 = {\sigma_1}^2 + \epsilon$, $-{\sigma_1}^2 < \epsilon$. Then, we study the sensitivity of ROC for different combinations of n, τ, and ϵ. Figure 3.2(b) shows the ROC curves for $\tau = \{10, 30, 60, 85\}$, $n = 100$, and $\epsilon = 1$. As intuitively expected, it can be seen that for a given P_{FA}, the P_{D} increases as the change point approaches the middle of the time segment and the P_{D} decreases when the change points are towards the end points of the time segment. The latter occurs because of the poor performance (higher variance) of the maximum likelihood estimate obtained from the shorter segment due to the reduced number of samples. As a consequence, the presence of a change point might pass unnoticed when ϵ is small. For example, in Figure 3.2(c) we observe that for $\tau = 15$, $n = 100$, $\epsilon = \{0.2, 1, 2, 4\}$, and a given P_{FA} value, the P_{D} increases with ϵ. Finally, in Figure 3.2(d) for $\tau = 15$, $n = \{50, 100, 200, 500\}$, $\epsilon = 2$, and a given P_{FA}, we cannot observe significant differences between the P_{D}s as n increases.

Next, under H_1 and for $\lambda > C_\alpha$, we study the estimation bias (detection delay) $\delta_t = \hat{t} - t$, where $\hat{t} = \arg\{\min_{2 \leq \tau \leq n-2} \mathrm{SIC}(\tau)\}$. In our setup $\delta_t \in \mathbb{N}$, i.e., we can expect negative delays. Note that in a sequential detection setup a negative delay will be considered as a false alarm. Here, we are interested in studying the mean value of the detection delay:

$$\mathrm{E}[\delta_t | \lambda > C_\alpha, \mathrm{H}_1]. \tag{3.38}$$

We compute (3.38) using 1000 Monte Carlo simulations. Figures 3.3(a) and 3.3(b) illustrate the average value of δ_t in (3.38) as a function of the change point t, in (a) for $n = 100$, $\alpha = 0.01$, $\epsilon = [1, 2, 4]$; and in (b) for $n = 100$, $\alpha = [0.005, 0.01, 0.02, 0.03]$, $\epsilon = 2$. As expected, the average delay decreases for larger ϵ values. Also, for change points near the origin of the time segment, the expected delay is large compared with the one obtained for change points located near the center of the time window.

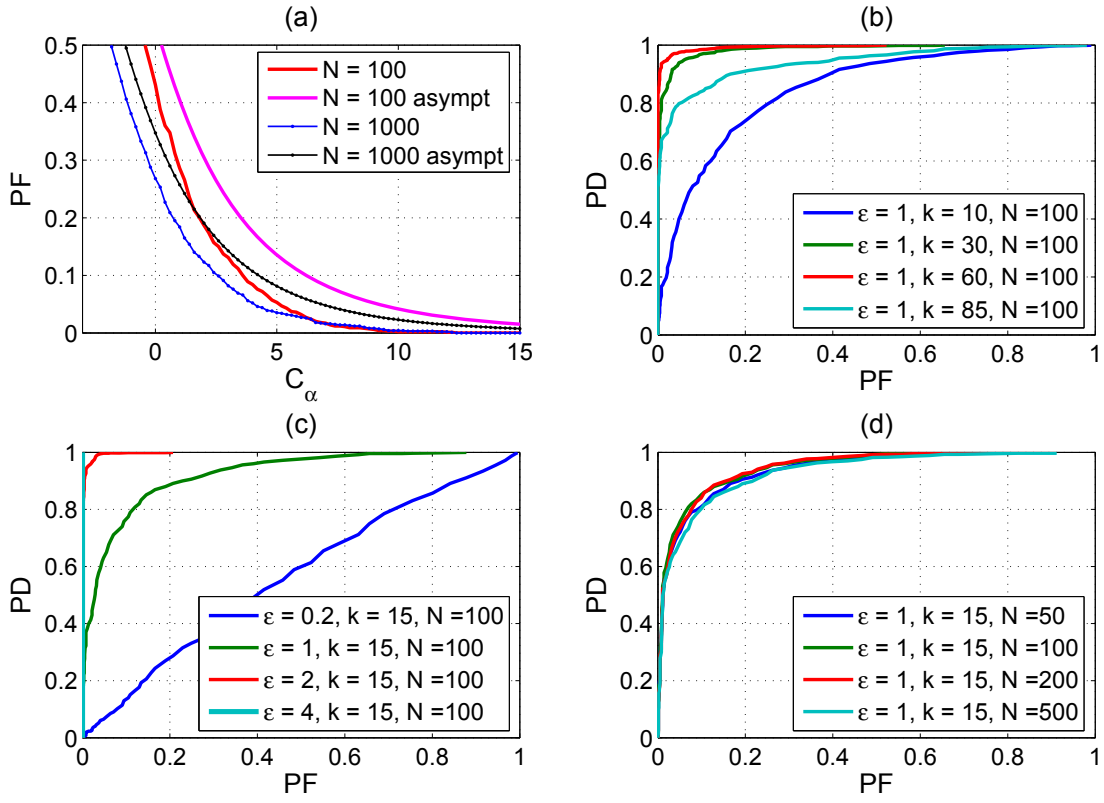


Figure 3.2: (a) Computed probability of false alarm (P_{FA}) and asymptotic P_{FA} as a function of γ for $n = 100$ and $n = 1000$ samples. (b) The receiver operating characteristic (ROC) for different change points. (c) ROC for different ϵ values. (d) ROC for different sample sizes. ©[2008] IEEE.

As a reference in Figures 3.3(c) and (d) we illustrate the P_{D} for the respective α values. Note that for a given $P_{\text{FA}} = \alpha$, the P_{D} for change points at the end of the window is smaller than the P_{D} for change points in the center. Therefore, if ϵ is large enough, then $P_{\text{D}} = 1$ over a large number of change points, and the expected delay is approximately uniform over the same range of change points.

The mean-square error (MSE) performance of change-point estimation algorithms is also an interesting and useful criterion to study. However, because finding analytical expressions of the MSE is usually intractable, an alternative procedure is to derive

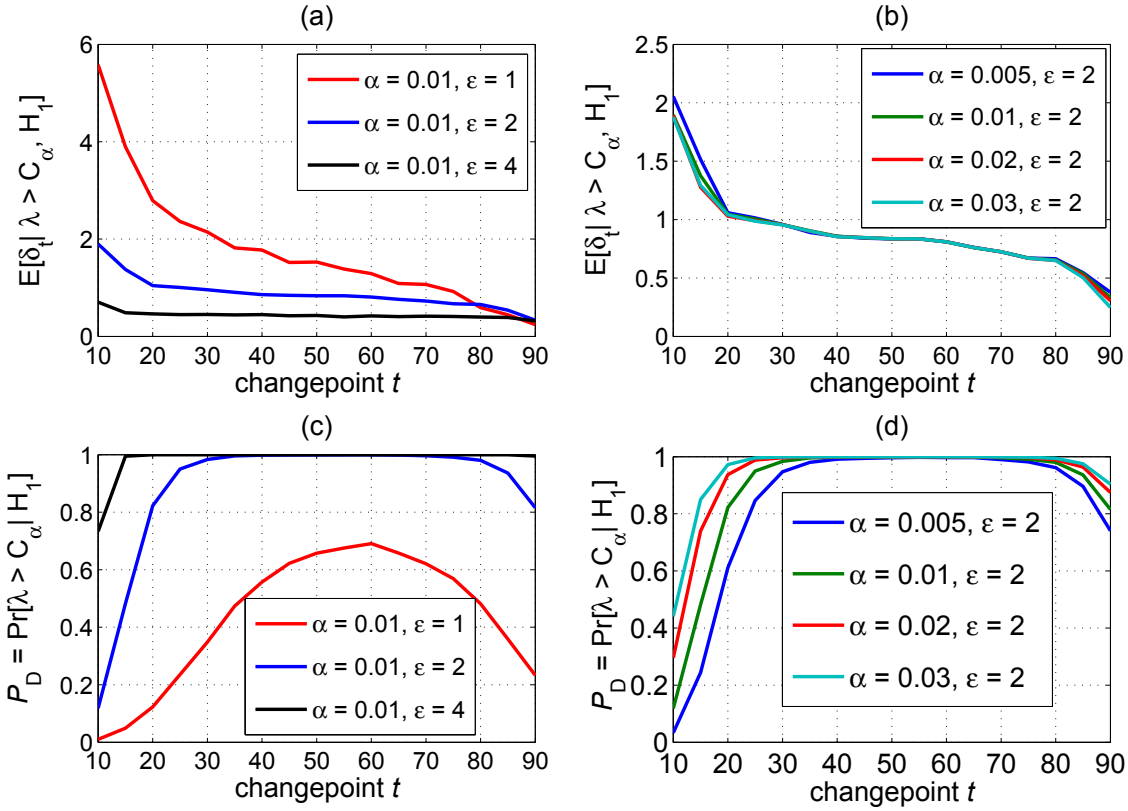


Figure 3.3: (a) Average value of the detection delay as a function of the change point for $n = 100$, $\alpha = 0.01$, and $\epsilon = 1$, 2, and 4. (b) Average value of the detection delay as a function of the change point for $n = 100$, $\epsilon = 2$, and $\alpha = 0.005$, 0.01, 0.02, and 0.03. (c) Probability of detection (P_{D}) as a function of the change point for $n = 100$, $\alpha = 0.01$, and $\epsilon = 1$, 2, and 4. (d) P_{D} as a function of the change point for $n = 100$, $\epsilon = 2$, and $\alpha = 0.005$, 0.01, 0.02, and 0.03. ©[2008] IEEE.

lower bounds on the MSE. In Chapter 4, we will study in detail the derivation of lower bounds on the MSE for the general case of multiple change-point estimates.

So far we have provided a statistical segmentation technique based on some prior knowledge of the data distribution. In particular, we have assumed that $x(t)$, $t = 1, \ldots, n$ is piecewise stationary and that the probability distribution of $x(t)$ belongs to same family $\mathbb{P}(\boldsymbol{\theta})$, $\boldsymbol{\theta} \in \mathbb{R}^m$ with piecewise constant time-varying parameters. A uterine contraction time segment can be modelled by several piecewise stationary

processes. Thus, in our approach, it can be modeled by a subset of $\mathbb{P}(\boldsymbol{\theta})$, so we need to define features that characterize each $\mathbb{P}(\boldsymbol{\theta})$ in order to be able to classify the time segments that belong to a uterine contraction. In the next section we describe the candidate features and the K-means cluster algorithm.

3.3 Classification

The classification of each detected time segment is based on the values of certain features. In general, a classification problem requires the definition of an appropriate feature space and a classification algorithm. In this section we introduce the RMS and FOZC as candidate features for discriminating whether a time segment belongs to a uterine contraction or background activities. We describe the K-means cluster algorithm to discriminate the feature values. Then, we define the binary decision signal and introduce the criterion to evaluate the performance of detection.

3.3.1 Candidate features

Let $t_{cp} = \{\widehat{t'}_1, \ldots, \widehat{t'}_j, \widehat{t'}_{j+1}, \ldots, \widehat{t'}_u\}$ be the time index sequence of the estimated change points, and define $t'_{cp} = \{1, t_{cp}, n\}$. The samples of a time segment are defined by $x(t) : \widehat{t'}_j \leq t \leq \widehat{t'}_{j+1}$. Let $\boldsymbol{x}(\widehat{t'}_{j+1}) = [x(\widehat{t'}_j), x(\widehat{t'}_j + 1), \ldots, x(\widehat{t'}_{j+1} - 1)]^T$, be a vector that contains the samples between the change points $\widehat{t'}_j$ and $\widehat{t'}_{j+1}$. Let $\zeta_i(\cdot) : \mathbb{R}^q \mapsto \mathbb{R}$, $i = 1, \ldots, m$, be a mapping function, where $q = \widehat{t'}_{j+1} - \widehat{t'}_j$ and m is the total number of feature mappings. The selection of $\zeta_i(\cdot)$ is associated with the particular application. In our case, in order to detect uterine contractions, we need

to find $\zeta_i(\cdot)$ whose values differentiate a uterine contraction from background activity coexisting in the same frequency bandwidth and time segment of analysis.

In [9,48] the application of RMS and FOZC have been effectively used to discriminate uterine contractions in EHG measurements. Uterine contractions have been shown to appear as high amplitude variations in the recorded MMG [5]. Thus it is intuitively reasonable to evaluate RMS and FOZC as features in MMG recordings, defined as follows:

- *RMS* of a sample vector $\boldsymbol{x}(\widehat{t'}_{j+1})$ is defined as

$$\mathrm{RMS}(\boldsymbol{x}(\widehat{t'}_{j+1})) = \sqrt{\frac{1}{\widehat{t'}_{j+1} - \widehat{t'}_{j}} \sum_{t=\widehat{t'}_{j}}^{\widehat{t'}_{j+1}-1} x(t)^2}, \quad \forall\, \widehat{t'}_{j}, \widehat{t'}_{j+1} \in t'_{cp}. \tag{3.39}$$

- *FOZC* is the first-order count of the number of zero crossings on the time series in a given time segment [56]. It is given by

$$\mathrm{FOZC}(\boldsymbol{x}(\widehat{t'}_{j+1})) = \sum_{t=\widehat{t'}_{j}}^{\widehat{t'}_{j+1}-2} \frac{1}{4}\{\mathrm{sgn}(\nabla(x(t+1))) - \mathrm{sgn}(\nabla(x(t)))\}^2, \quad \forall\, \widehat{t'}_{j}, \widehat{t'}_{j+1} \in t'_{cp}, \tag{3.40}$$

  where $\nabla(x(t)) = x(t+1) - x(t)$ and $\mathrm{sgn}(\cdot)$ denotes the sign function defined as $sgn(w) = 1$ if $w \geq 0$, and 0 otherwise. In general, zero crossing analysis is an effective tool for modeling the spectral characteristics of stochastic processes [56] and it has been applied for signal detection and estimation [56]. For example, in [9] the authors show that the FOZC is applicable in detecting uterine contractions in EHG measurements. In particular, the dominant frequency principle [56] says that the normalized expected zero-crossing rate $\pi \mathrm{E}[\mathrm{FOZC}/(\widehat{\mathrm{t}'}_{\mathrm{j+1}} - \widehat{\mathrm{t}'}_{\mathrm{j}})]$, which is a weighted average of the spectral mass, tends to

admit values in the neighborhood of a significantly dominant frequency. Therefore, it is a practical way to discriminate changes in the spectrum between time segments.

3.3.2 Classification algorithm

To discriminate the features, we use the K-means clustering algorithm [57], which classifies the time segments into K groups based on a set of features (RMS or FOZC, or both). K-means is one of the simplest unsupervised learning algorithms (although it requires a priori K) that solve the well-known clustering problem. This algorithm finds a partition in which objects within a cluster are as close to each other as possible, and as far from objects in other clusters as possible. The centroid of each cluster is the point at which the sum of the distances from all objects in that cluster is minimized. The grouping is done by minimizing the criterion J, which is the sum of the squared distances between the feature vectors and the corresponding cluster centroid:

$$J = \sum_{j=1}^{K} \sum_{i \in t'_{cp}} ||\boldsymbol{\zeta}^j(\boldsymbol{x}(i)) - \underline{\boldsymbol{\zeta}}^j||^2, \tag{3.41}$$

where $\boldsymbol{\zeta}^j(\boldsymbol{x}(i)) = [\zeta_1(\boldsymbol{x}(i)), \ldots, \zeta_j(\boldsymbol{x}(i)), \ldots, \zeta_q(\boldsymbol{x}(i))]^T$ for $i \in t'_{cp}$ is the time-segment feature vector, $\underline{\boldsymbol{\zeta}}^j$ is the centroid for the j^{th} cluster, K is the total number of clusters fixed a priori, and $||\cdot||$ is the Euclidean distance. The minimization of J is performed using an iterative algorithm, which essentially moves objects between clusters until the sum cannot be decreased further. The algorithm is briefly described as follows [58]: the first step consists of randomly selecting K time-segment feature vectors from the population. These features represent initial centroids. In the second step, it assigns each feature to the cluster that has the closest centroid. After all the features have

been assigned, as a third step, the positions of the K centroids are recalculated. The second and third steps are repeated until the K centroids no longer move. Note that if the number of segments is less than the number of clusters K, then no classification can be performed and we assume that no contraction has taken place. For example, if the total number of clusters is $K = 3$ and the total number of time segments detected is equal to 2, we have only two feature values to be classified, so no classification is performed and it is assumed that there is no contraction in the measurements.

3.3.3 Cluster labelling and binary decision signal

An advantage of the K-means cluster algorithm is that it is a non-supervised data classification technique. However, in order to interpret the results, namely cluster labelling, it is required to have some knowledge of the feature space. In this sense, for example, if the RMS is used as a feature, then we might expect the centroid with the largest RMS value to be labelled as a contraction. Similarly, if the FOZC is used, we might assign as a contraction the centroid with the lowest FOZC value. In practice, as we discuss in the experimental results section, the feasibility of discriminating a uterine contraction using the RMS and FOZC depends also on the segmentation stage.

Let $N_K := \{1, \ldots, k, \ldots, K\}$ be the set of clusters labels, where $1 \leq k \leq K$ is the label assigned arbitrarily by the classification algorithm to a resulting cluster. Also, let $N_C \subset N_K$ be the set of clusters assigned as a contraction based on a certain criterion. Therefore, a time segment $i \in t'_{cp}$ is classified as a contraction if its feature

$\boldsymbol{\zeta}^j(\boldsymbol{x}(i)) \in N_C$. Then, the binary decision signal $s(t)$ is defined as follows,

$$s(t) := \begin{cases} 1 & \text{if } \boldsymbol{\zeta}(\boldsymbol{x}(\widehat{t'}_{j+1})) \in N_C, \widehat{t'}_j < t < \widehat{t'}_{j+1}, \\ 0 & \text{otherwise,} \end{cases} \tag{3.42}$$

where $\widehat{t'}_j, \widehat{t'}_{j+1} \in t'_{cp}$.

3.3.4 Performance evaluation

Let $r(t), 1 \leq t \leq n$ be the binary reference signal according, for example, to the patient's feedback, i.e., $r(t) = 1$ indicates the presence of a contraction at sample t. Then, we evaluate the performance of the detection algorithm by computing the detection ratio (DR), false alarm ratio (FAR), and correlation coefficient (CORR) which are defined as follows,

$$\text{DR} = \frac{\sum_{t=1}^{n} s(t)r(t)}{\sum_{t=1}^{n} r(t)}, \tag{3.43}$$

$$\text{FAR} = \frac{\sum_{t=1}^{n} s(t)(1-r(t))}{\sum_{t=1}^{n}(1-r(t))}, \tag{3.44}$$

$$\text{CORR} = \begin{cases} \frac{\sum_{t=1}^{n} s(t)r(t)}{\sqrt{\sum_{t=1}^{n}(r(t))^2 \sum_{l=1}^{n}(s(l))^2}} & \\ 0 & \text{if } \sum_{t=1}^{n} r(t) = 0 \text{ or } \sum_{t=1}^{n} s(t) = 0. \end{cases} \tag{3.45}$$

The DR computes the percentage of agreements in detection between $s(t)$ and $r(t)$ for $1 \leq t \leq n$, given the total number of time-samples with contractions according to $r(t)$. The FAR computes the percentage of disagreements in detection (false alarms) between $s(t)$ and $r(t)$ for $1 \leq t \leq n$, given the total number of time-samples without

contractions according to $r(t)$. The CORR indicates how similar the sequences $s(t)$ and $r(t)$ are in the time interval $1 \leq t \leq n$.

3.4 Experimental results

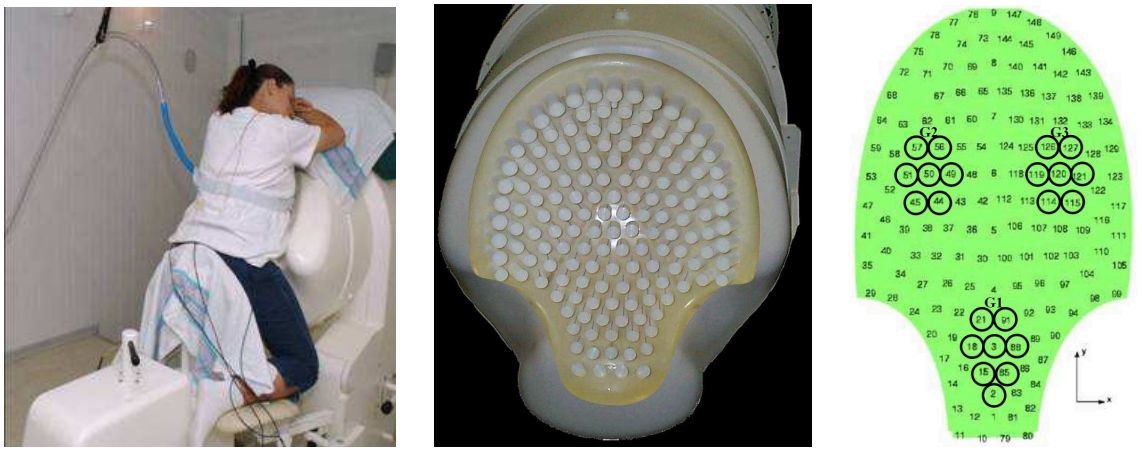
In this section we show the performance of the proposed detection scheme using real MMG data. In particular we evaluate the performance of the pre-segmentation techniques as well as the discrimination capabilities of RMS and FOZC. Also, we analyze the fusion of the decision signals from all the sensors as in the parallel distributed detection approach, allowing us to study the temporal and spatial distribution of the uterine contraction activities.

3.4.1 Data acquisition and preprocessing

The uterine MMG recordings were obtained using a 151-channel magnetic field sensor array system named SARA[4](SQUID Array for Reproductive Assessment). SARA is a passive, stationary, floor-mounted instrument on which the patient sits and leans her abdomen against a concave surface which contains an array of sensors (Figure 3.4a). The sensor array covers a region of approximately 45 cm high and 33 cm wide, with a surface of 1300 cm^2 inclined at 45^o. Figure 3.4(c) shows the 151-channel array embedded under the concave surface upon which the patients leans her abdomen. The whole system is in a 3-layer magnetically shielded room (MSR) and is equipped with high-order synthetic gradiometer noise cancellation, which effectively eliminates the vibrational noise transmitted by the mother. Recordings were performed on

[4]SARA was built in collaboration with VSMMedTech Ltd., Canada and is installed at the University of Arkansas for Medical Sciences (UAMS) Hospital.

10 patients who presented themselves in the labor and delivery unit complaining of contractions, at gestational ages ranging from 31 to 40 weeks. Table 3.1 illustrates the individual gestation ages of the patients used in this study. The recording session was 10 minutes long, with a sampling rate of 250 Hz. The records also contain an additional channel that registers the beginning and end of the contraction according to the patient's feedback, which we used here to evaluate the performance of our method. Note that the perception of contraction by the mother is limited, because it is well known that this subjective report is not exact. However, if a mother reports contraction, we can be assured that a contraction occurred. In this respect, we may miss some contractions, but it can be assured that the reported contractions are real ones.



(a) **(b)** **(c)**

Figure 3.4: (a) SARA system installed at the University of Arkansas for Medical Sciences (UAMS) Hospital. (b) 151-channel sensor array embedded under the concave surface upon which the patient leans her abdomen. The sensor coils are placed 3 cm apart, covering a total area of approximately 1350 cm^2. (c) Diagram of sensor array with channels identification numbers. The circles indicate the groups of channel G1, G2, and G3. ©[2008] IEEE.

We remove the presence of artifacts, such as fetal and maternal magnetocardiogram MCG and maternal breathing, by first downsampling the data to 5 Hz and then applying a bandpass filter focusing on the primary uterine magnetic activity. This activity is represented by a low frequency, typically between 0.1 and 0.4 Hz [47, 59]. The MMG activity in this range likely represents the plateau and repolarization phase of the action potentials [47,59]. Note that the biomagnetic signal also includes possible contributions from motions of the fetus or intestines. However based on the investigated frequency band, these contributions are limited, and, in addition, there is no hypotheses of their occurrence during contractions.

In Figure 3.5 we illustrate the normalized power spectral density (PSD), computed using Welch's method on samples from channels 2, 50, and 120 and obtained from six different patients. The PSDs were computed using all samples from the 10 minutes of measurements from each patient, thus, they do contain contractions. We chose the above channels (sensor positions) to illustrate three different abdominal areas. Figure 3.4(c) illustrates the spatial location of the selected channels. We chose three pairs of patients at 38, 39, and 40 weeks of pregnancy, respectively. It can be seen that we could not distinguish a common PSD pattern among the different patients. Therefore, we studied the performance of the detection algorithm in the frequency ranges from 0.1 to 0.4 Hz, and 0.2 to 0.4 Hz, looking for an adequate band range that maximized the average of the detection of contractions. In each case, we applied equiripple FIR bandpass filters to the downsampled measurement, with lower and upper pass frequency limits f_{lp} and f_{up}, respectively, given by the limits of the frequency ranges defined above. The upper and lower stop frequencies were $f_{\mathrm{ls}} = f_{\mathrm{lp}} - \delta_{\mathrm{f}}$ and $f_{\mathrm{us}} = f_{\mathrm{up}} + \delta_{\mathrm{f}}$, respectively, with $\delta_{\mathrm{f}} = 0.03$. The gains in the pass and stop band

are constrained to 0 dB and -60 dB, respectively, and the maximum allowed ripple in both bands was 0.05 dB.

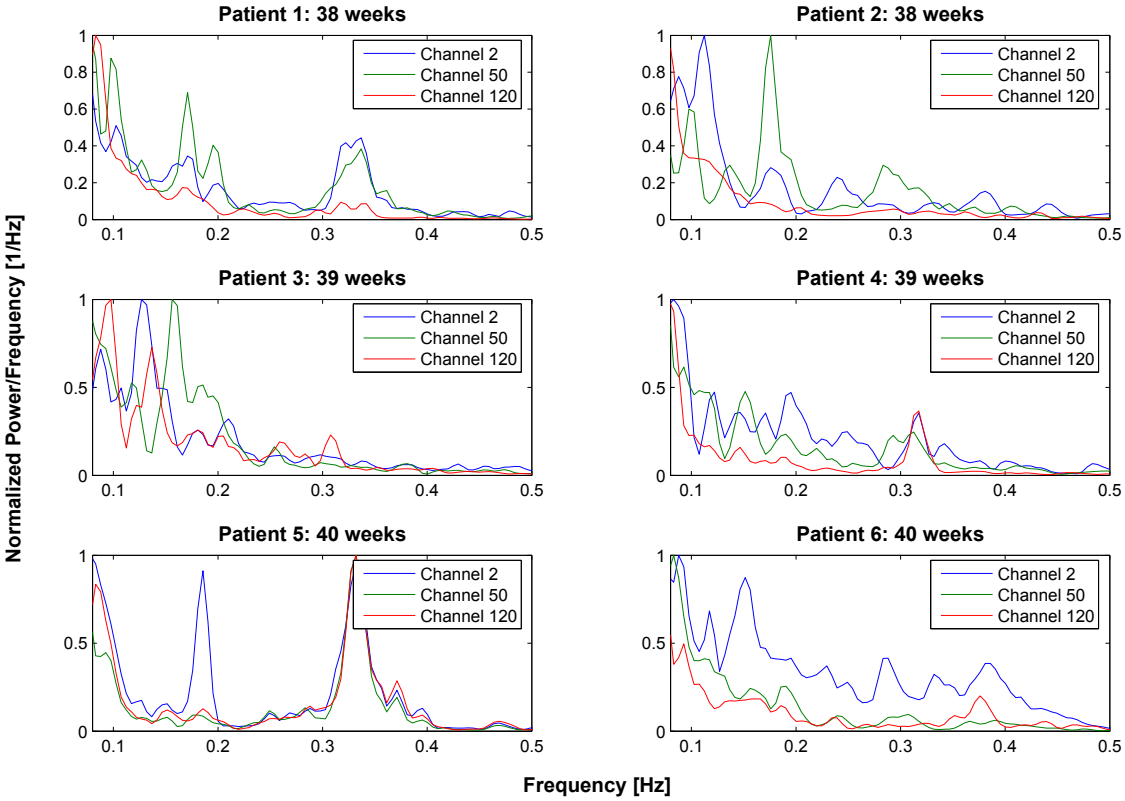


Figure 3.5: Power spectral density (PSD) computed using Welch's method on samples from channels 2, 50, and 120 obtained from six different patients with gestational ages between 38 and 40 weeks. ©[2008] IEEE.

3.4.2 Model order estimation, feature evaluation, and cluster labelling

The segmentation stage requires the estimation of the model order d, which can be estimated, for example, using the approach proposed in section 3.2.2. In practice, we found that this approach performed well in only a portion of channels, while in the

remaining part, it did not lead to good performance. Namely, in channels signaling the clear presence of contractions, we found that the model order that minimized the SIC_T had a large value and did not lead to detection of changes. However, we also found that for all the channels the number of change points detected tended to decrease as the model order increased. Therefore, in practice, the model order can be thought as a resolution parameter in terms of the number of change points detected.

We evaluated the discriminating capabilities of the RMS and FOZC by computing them on the time segments detected in both frequency ranges for $d = 0, \ldots, 5$ and with a significance value $\alpha = 0.01$. We found that the RMS is a good candidate feature to discriminate between time segments with and without contractions according to the patient's feedback. In particular, we found that in both frequency ranges the time segments with contractions usually have larger RMS values than time segments without contractions. In the case of the FOZC, we were not able to find a consistent pattern for discrimination. One possible reason might be the reduced length of some estimated time segments, especially when a contraction is taking place. Therefore, in our examples below we use the RMS as a feature to identify the presence of a contraction in a particular time segment.

Assuming that the preprocessed records in the analyzed frequency ranges contain only the presence of uterine contractions, then RMS values are related to the energy evolution of the burst of uterine activities. Therefore, if we divide the RMS value into three groups, the largest values of RMS correspond to the peak phase of the burst of activities, the second largest ones correspond to the rising and falling phase, and the smallest values represent the resting state. Using the assumption above, we classify the RMS values in all the examples below using the K-means cluster algorithm by fixing the number of clusters (K) at three. Thus, we label as a contraction the time

segments during which the RMS values belong to the two clusters with the largest centroids. In practice, we observed that the RMS values are more likely to be classified in three groups rather than two, specially due to the presence of large RMS values or peaks. In this sense the classification using only two groups assigns centroids around the peaks, and as a consequence, very often assigns the middle level RMS values as background noise.

As an example, we illustrate in Figure 3.6 the records from channel 2 of patient 6 at 40 weeks of pregnancy, bandpass filtered between 0.2 and 0.4 Hz. Figure 3.6(a) shows the downsampled and filtered signals. The vertical grid lines indicate the estimated change points according to an AR model with order $d = 1$ and $\alpha = 0.01$. We choose $d = 1$ since, as we discuss in the next subsection, the detection methods achieves on average the maximum DR and CORR in the given frequency range. It can be seen that the number of change points estimated increases at the time intervals when the patient indicates the presence of a contraction. This observation suggests that the samples that belong to the contraction interval might be modelled by an AR model with a larger model order. Figures 3.6(b) and 3.6(c) illustrate the RMS and FOZC computed on the estimated time segments. In general we observe that in this frequency range the FOZC does not identify the presence of a contraction very well according to the patient's feedback. Figure 3.6(d) shows the cluster groups obtained after applying the K-means algorithm on the RMS values. The cluster labels are given in ascending order according to the centroid RMS values, i.e., label 3 is the centroid with largest RMS value. Figure 3.6(e) shows the binary decision signal $s(t)$ amplified by 1.2 times the maximum value of the preprocessed measurements. In this case, the RMS values that belong to the two clusters with the largest centroid values

were classified as a contraction. Figure 3.6(f) illustrates the time-intervals in which the patient acknowledged the presence of a contraction.

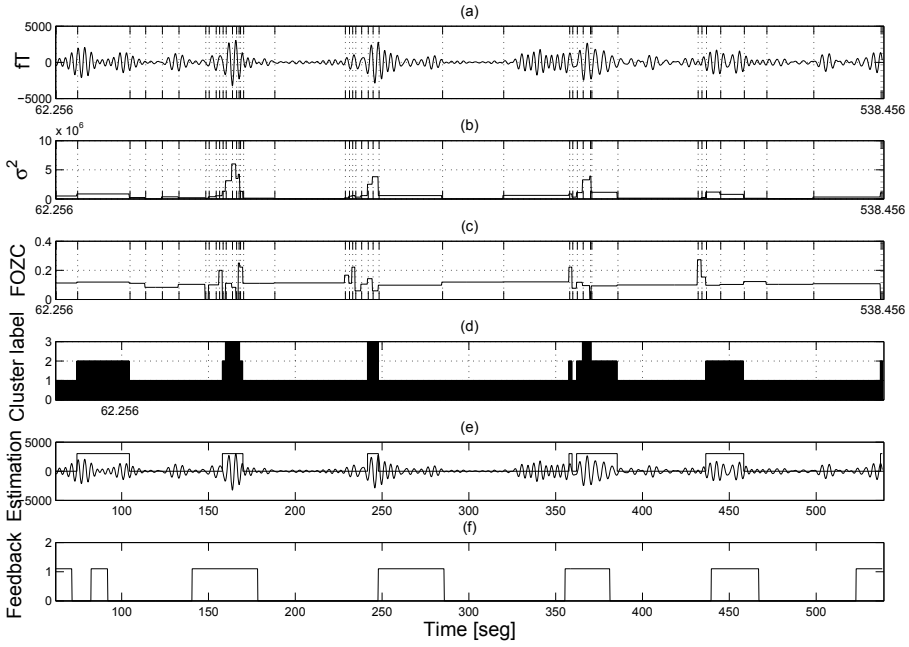


Figure 3.6: Bandpass filtered records from channel 2 of patient 6 with a gestational age of 40 weeks: (a) Preprocessed channel with grid lines indicating the estimated change points for $d = 1$ and $\alpha = 0.01$; (b) RMS in each time segment; (c) FOZC in each time segment; (d) cluster groups using RMS features; (e) estimated contraction segments; (f) time segments with contractions according to the patient feedback. ©[2008] IEEE.

3.4.3 Performance analysis and discussion

We evaluated the DR, FAR, and CORR of the detector algorithm in 10 patients in 4 groups of channels defined in Table 3.1. We selected the groups according to their relative position in the array. Figure 3.4(c) illustrates the array with the selected group of channels.

Table 3.1: Dataset Summary ©[2008] IEEE.

Patient	1	2	3	4	5	6	7	8	9	10
Gestation age	38	38	39	39	40	40	31	37	38	38
Channels G1	{ 2, 3, 15, 18, 21, 85, 88, 91}									
Channels G2	{ 44, 45, 49, 50, 51, 56, 57}									
Channels G3	{ 114, 115, 119, 120, 121, 126, 127}									
Channels G4	{G1, G2, G3, 4, 5, 6, 7 }									

Table 3.2 illustrates the average and standard deviation of DR, FAR, and CORR computed in each group of channels G1, G2, G3, and G4 for the 10 patients used in this study, for $d = 0, 1, \ldots, 5$ with $\alpha = 0.01$. The case $d = 0$, assumes that the mean of the process is zero, therefore, we looked for changes in variance. Here, to simplify our analysis below, we applied the same d value to all channels from the same patient. From the results obtained, the detector performance with respect to the frequency band and with respect to the sensor positions. We observe that the maximum average DR and average CORR are achieved consistently in all groups of channels for d equal to 0 and 1, in the frequency range 0.2-0.4 Hz. Also, in the same frequency range for $d = 0, 1$, the lowest average FAR are obtained in the group of channels G1. We further see that in the frequency range 0.1-0.4 Hz, the maximum average DR and average CORR are obtained for $d = 0, 3$.

Table 3.2: Average DR, FAR, and CORR computed in channel groups 1, 2, and 3, bandpass filtered in the frequencies ranges 0.1 to 0.4 Hz, and 0.2 to 0.4 Hz for $d = 0, 1, 2, \ldots, 5$ with $\alpha = 0.01$ ©[2008] IEEE.

			f: 0.1 - 0.4 [Hz]						f: 0.2 - 0.4 [Hz]					
Channels		d :	0	1	2	3	4	5	0	1	2	3	4	5
Group 1	DR	avg	0.46	0.28	0.27	0.37	0.28	0.19	0.55	0.58	0.38	0.46	0.25	0.31
		std	0.21	0.15	0.12	0.22	0.19	0.16	0.11	0.14	0.12	0.14	0.12	0.21
	FAR	avg	0.33	0.14	0.20	0.24	0.23	0.15	0.36	0.39	0.22	0.27	0.19	0.20
		std	0.11	0.05	0.06	0.08	0.15	0.14	0.12	0.14	0.09	0.12	0.10	0.14
	Corr	avg	0.16	0.13	0.12	0.16	0.10	0.07	0.21	0.22	0.18	0.19	0.09	0.08
		std	0.09	0.08	0.08	0.10	0.10	0.06	0.09	0.10	0.06	0.08	0.05	0.04
Group 2	DR	avg	0.40	0.33	0.26	0.31	0.24	0.28	0.57	0.56	0.41	0.31	0.38	0.23
		std	0.22	0.26	0.19	0.24	0.29	0.23	0.19	0.20	0.13	0.11	0.32	0.30
	FAR	avg	0.35	0.27	0.22	0.27	0.27	0.28	0.52	0.51	0.35	0.33	0.38	0.20
		std	0.18	0.20	0.11	0.20	0.28	0.24	0.24	0.24	0.16	0.19	0.29	0.26
	Corr	avg	0.15	0.11	0.11	0.14	0.11	0.08	0.19	0.19	0.17	0.13	0.13	0.07
		std	0.10	0.06	0.08	0.11	0.12	0.08	0.10	0.10	0.09	0.09	0.11	0.08
Group 3	DR	avg	0.31	0.21	0.25	0.40	0.34	0.08	0.59	0.65	0.38	0.41	0.33	0.10
		std	0.20	0.22	0.18	0.27	0.27	0.13	0.27	0.25	0.15	0.26	0.34	0.22
	FAR	avg	0.27	0.19	0.20	0.33	0.35	0.08	0.45	0.56	0.26	0.31	0.26	0.11
		std	0.13	0.18	0.11	0.17	0.28	0.11	0.25	0.28	0.16	0.18	0.24	0.17
	Corr	avg	0.14	0.10	0.13	0.16	0.11	0.05	0.21	0.23	0.18	0.17	0.10	0.02
		std	0.10	0.09	0.09	0.10	0.11	0.11	0.11	0.15	0.10	0.10	0.13	0.04
Group 4	DR	avg	0.39	0.27	0.26	0.38	0.29	0.19	0.57	0.59	0.38	0.41	0.31	0.24
		std	0.15	0.13	0.11	0.18	0.19	0.13	0.13	0.11	0.11	0.14	0.22	0.21
	FAR	avg	0.32	0.20	0.22	0.31	0.28	0.17	0.45	0.48	0.28	0.31	0.28	0.19
		std	0.10	0.11	0.07	0.09	0.15	0.12	0.13	0.16	0.09	0.09	0.16	0.15
	Corr	avg	0.15	0.11	0.12	0.15	0.11	0.07	0.21	0.21	0.17	0.17	0.10	0.06
		std	0.09	0.06	0.07	0.09	0.10	0.06	0.10	0.11	0.08	0.08	0.07	0.04

Consider that each single channel detector decides the presence of a contraction based on local spatial information. Thus, a detection classified as FA might be indicating the occurrence of a burst of activity in the myometrium; however its power and surface distribution might not be sufficient to be detected by the patient. Therefore, we also analyzed the temporal-spatial detection of the sensor array by fusing the decision signals as in the parallel distributed detection approach. In particular, we proposed and analyzed the following criteria: the percentage of active sensors as a function of time, which is the result of adding all binary decision signals, obtained from processing every channel, normalized by the total number of channels; the total RMS as function of time, which is the result of adding the RMS value computed on the time segment estimated as contraction in each sensor; and the spatio-temporal propagation of uterine contractions and their RMS values by keeping only the sensor measurements detected as contractions (masking). In Figure 3.7 we illustrate our distributed processing approach for analyzing uterine contractions using an array of magnetometers.

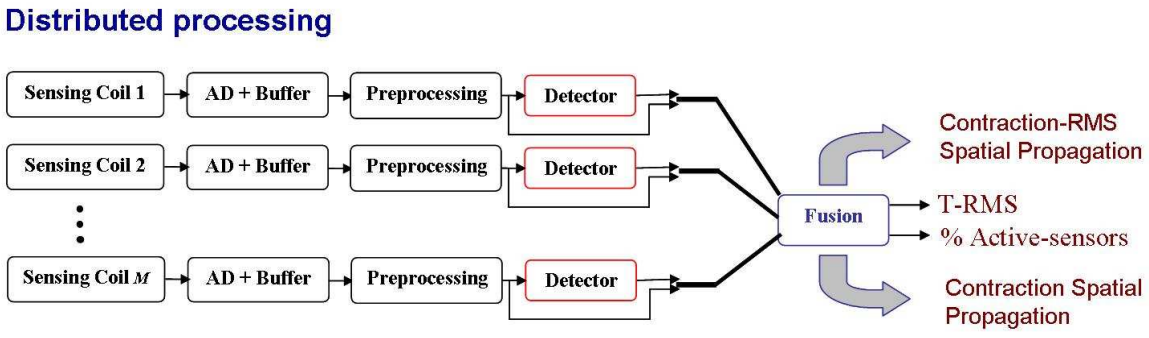


Figure 3.7: Illustration of our distributed processing approach for analyzing the array of temporal-spatial detection signals obtained after applying the algorithm for detecting uterine contractions in each MMG channel.

The distributed processing approach proved to be helpful in understanding the uterine MMG contraction activity spatially and temporally. As an example, in Figure 3.8

we illustrate the total RMS and the percentage of active sensors as a function of time for patients 1, 5, and 6 using $d = 1$ and $\alpha = 0.01$ in the frequency band 0.2 to 0.4 Hz. The time interval of a contraction according to the patient's feedback is indicated by the high level of the pulse train in each figure. The pulse amplitude is just for purposes of illustration. In the case of patient 5, we detected only the first two contractions. Interestingly, we found that the power variation associated with the other three contractions were detectable in the frequency band 0.4 to 0.45 Hz (see the PSD of patient 5 in Figure 3.5). We also observed that in patient 6, the maximum agreement in detection, as well as in RMS values, coincided with the starting point of 3 contractions, rather than the 5 evidenced in this graph according the patient's feedback. However, comparison with intrauterine pressure is required in order to quantify the above observations.

We also illustrate the spatial and temporal distribution of the activity detected based on each sensor decision in the frequency band 0.2 to 0.4 Hz. In Figure 3.9, we illustrate snapshots of the reconstructed measurement surface for patient 6 using the array of measurements $z_i(t) = s_i(t)x_i(t), i = 1, \ldots, 151$, which are the preprocessed measurements masked with their corresponding binary decision signals. The time interval illustrated contains a contraction, based on patients' feedback between 246 and 286 seconds (see also Figure 3.8, patient 6). It can be seen that the maximum number of activated sensors occurs between snapshots at 252 to 260 seconds. It is interesting to note that the percentage of activated sensors increases from 15% to 80% in 3 seconds, just a couple of seconds before the patient starts acknowledging the presence of a contraction. On the other hand, the percentage of activated sensors decreases from 85% to 50% in 25 seconds. This difference might lead us to the conclusion that the contractions rise faster than they fall in terms of spatial distribution;

however, we could not generalize such observation in other contraction segments in the same patient.

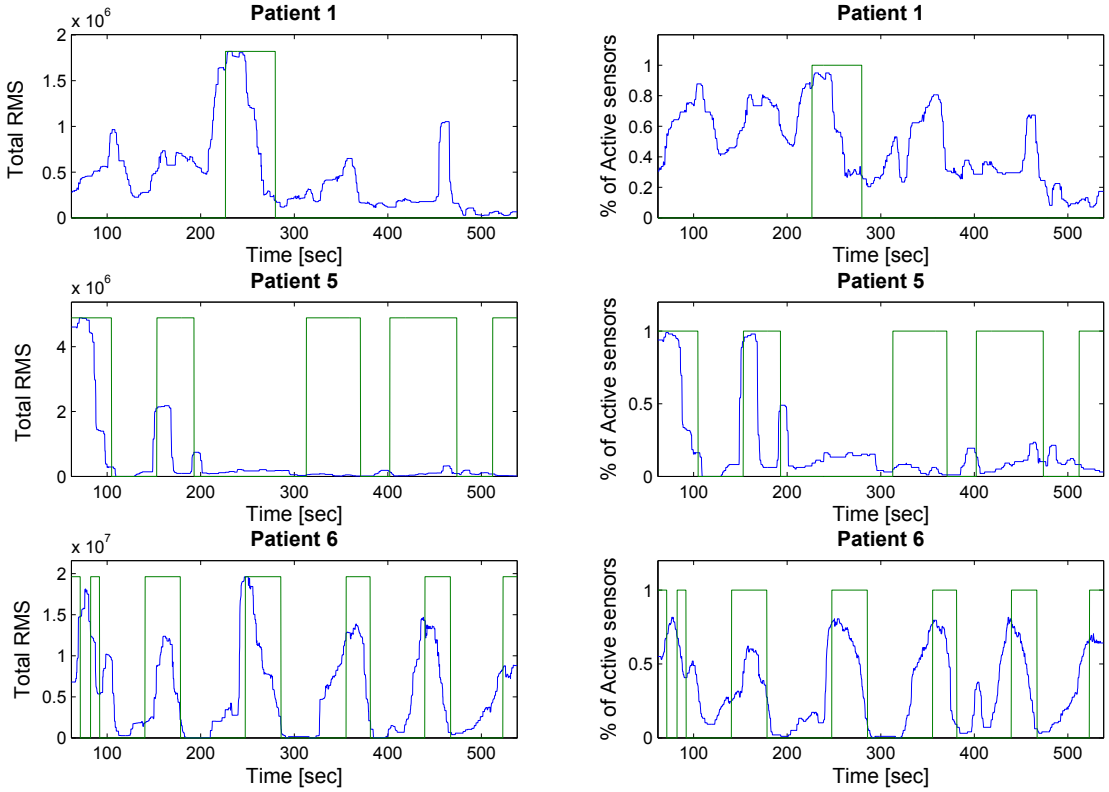


Figure 3.8: From left to right: total RMS and percentage of active sensors as a function of time for patients 1, 5, and 6. The time interval of a contraction according to the patients' feedback is indicated by the high level of the pulse train in each figure, respectively. Note that the pulse amplitude is just for illustration purposes. ©[2008] IEEE.

3.5 Summary

In this chapter we have developed a general analysis for the detection of uterine MMG contractions. In particular, we have proposed a two-stage statistical time-segment discriminator using a single channel of MMG measurements. We assumed that the

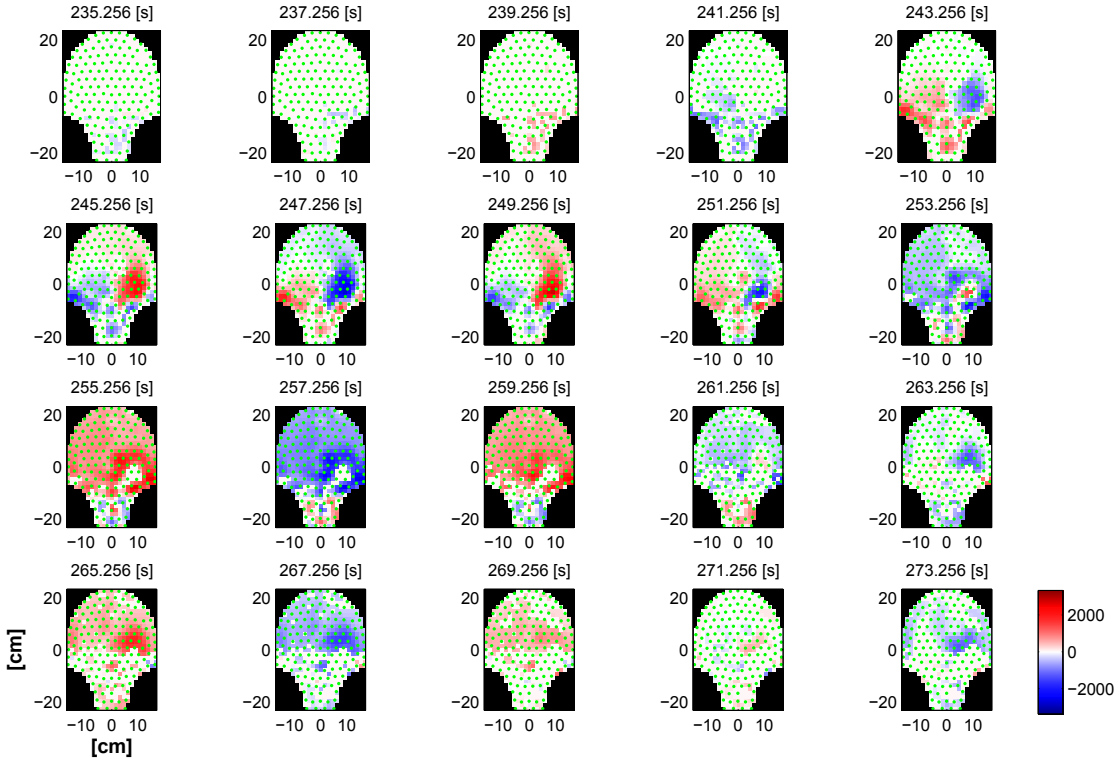


Figure 3.9: Snapshots of the reconstructed measurement surface from patient 6 masked with the binary decision signals in the frequency band 0.2 to 0.4 Hz. The time selected coincides with the presence of a contraction, according to the patient's feedback. ©[2008] IEEE.

preprocessed channels are modeled by a piecewise time-varying AR model of a certain order with an input given by a white Gaussian noise with time-varying variance. Therefore, we first designed a statistical model-based segmentation algorithm based on the SIC to estimate the time-instants of changes in the parameters. To discriminate time segments that contain a contraction, we evaluated features such as the time segment power (RMS) and the dominant frequency component (FOZC). Then, we applied the non-supervised cluster algorithm K-means to classify the RMS values, obtaining then a discrete-time binary decision signal indicating the presence of a contraction. Since each single channel detector provided local information regarding

the presence of a burst of activities, we also analyzed the fusion of the decision signals from all the sensors as in the parallel distributed detection approach. This approach proved helpful in understanding the uterine MMG contraction activity both spatially and temporally.

Chapter 4

Performance Bounds on Multiple Change-point Estimation[5]

4.1 Introduction

In Chapter 3 we developed a robust detector of uterine contractions based on a multiple-change point estimation algorithm. We also analyzed the performance of detection for a single change point in a data segment. In this chapter, we study the global mean-square error (MSE) performance for the class of unbiased estimator of change points in an independent sequence, including the case for a Gaussian distribution with changes of variance ($d = 0$ in our previous chapter).

Estimation of changes in time series is an important and active research area with several applications, for example, in fault detection, medical imaging, genetics, and econometrics. Many estimation algorithms for change-point estimation (see, *e.g.*, [23–25]) have been proposed in past. However, less work has been done concerning

[5]Based on P. S. La Rosa, A. Renaux, C. Muravchik, and A. Nehorai,, "Barankin-type Lower Bound on Multiple Change-point Estimation," *to appear in IEEE Trans. on Signal Processing,* 2010.

the ultimate performance of such algorithms in terms of MSE. Indeed, if an estimator is available, the evaluation of its performance depends on knowing whether it is optimal or if further improvement is still possible. Note that some other criteria of performance in the context of sequential detection of a change-point are available in the literature, see, *e.g.*, [60, 61] and references therein.

The classic way to analyze the performance of an estimator in terms of MSE is to compute the well-known Cramér-Rao bound (CRB) [26]. Unfortunately, for discrete time-measurement models the change-point location parameter is discrete; therefore the CRB, which is a function of the derivative of the likelihood of the observations w.r.t. the parameters, is not defined.

Several authors have proposed solutions to this problem. Indeed, in the change-point estimation framework, the CRB has already been studied using approximations (see, *e.g.*, [62–67]). Depending on the particular parametrization of the data likelihood, two main challenges have been addressed concerning the CRB computation on the change-point time index: (i) the discrete nature of the aforementioned parameter and (ii) the regularity conditions of the likelihood of the observation. The former implies that the parameter does not have a defined derivative because of its discrete nature [65], and the latter implies that the likelihood of the observations has to be smooth (details are given in [26] and [68]), which is not the case for signal parameters with sudden changes. To overcome the discrete nature of the change-point time index, a continuous parametrization has been proposed (see, *e.g.*, [67, 69]). To satisfy the regularity conditions of the data likelihood, the step-like function, which represents a change in parameter, is generally approximated by another function with smooth properties (*e.g.*, the so-called sigmoidal function introduced in [64] and [67], or a Heaviside function filtered by a Gaussian filter, as in [62]). This new function depends

on parameters that have to be adjusted, and it tends to the step-like function when the appropriate values of these parameters are used. The main problem that appears when using this technique is that the CRB tends to zero when the approximate function tends to the step-like function [63, 67].

Moreover, it is noteworthy that these previous works concerning change-point estimation were always done in the framework of a single change point. To the best of our knowledge, performance bounds have never been derived in a multiple change-point context. The latter is important in off-line estimation of change points where batch-data are available, for example, in biomedical applications, such as DNA sequence segmentation [70], rat EEG segmentation (see [25], Chapter 2), detection of uterine contraction using MMG [14,15], and in signal segmentation in general, such as speech segmentation [71] and astronomical data analysis [72].

In this paper, we analyze the Barankin bound (BB) [27] for multiple change-point estimation in the context of an independent vector sequence. The Barankin bound is the greatest lower bound for any unbiased estimator. Moreover, in contrast to the CRB, its computation is not limited by the discrete nature of the parameter and the regularity assumptions on the likelihood of the observations [68,73]. However, the BB requires the use of parameters called test points. The choice of these test points is left to the user, and, in order to obtain the best (*i.e.*, the tightest) bound, a nonlinear maximization over these test points has to be performed. This requirement explains why this bound is so much less used and less known than the CRB; nevertheless, the BB is often a practical bound for realistic scenarios, see *e.g.*, [74].

To the best of our knowledge, minimal bounds other than the CRB have been proposed in the context of change-point estimation only in the foundational communication of Ferrari and Tourneret [75]. A simplified and practical version of the BB (*i.e.*, one test point per parameter), the so-called Hammersley-Chapman-Robbins (HCR) bound, [73, 76], is studied in that paper. As in the previous works on the CRB, only one change point is considered.

In this chapter we extend the results presented in [75] to the case of multiple change points. We consider the multi-parameter HCR bound, and we show that the so-called Barankin information matrix (BIM), which has to be inverted, has an interesting structure (*viz.*, a block diagonal matrix structure). We show that the estimation of one change point is corrupted by its neighboring change points, and we give the details of the computation for the two change-point case. This case facilitates the derivation of a closed-form expression for the inverse of the BIM. Note that it is possible to find tighter bounds by using more test-points per parameter; however, such an approach does not allow for obtaining closed-form expressions of the BIM and its inverse as derived here. We also discuss on the existence of the supremum of the finite set formed by all possible BB solutions and, following ideas from [77] and from convex optimization, we compute a suitable minimal-upper bound to this candidate set with respect to the Loewner cone, the set of semipositive definite matrices. In particular, we show that its computation is given by the matrix associated with the Lowner-John ellipsoid of the candidate set, which is the minimum-volume hyper-ellipsoid covering the set of hyper-ellipsoids associated to each matrix in the candidate set. We apply the bounds to the case of changes in the parameters of Gaussian and Poisson observations. We finally present numerical examples for comparing our bound to the performance achieved by the maximum likelihood estimator (MLE).

The notational convention adopted in this chapter is as follows: italic indicates a scalar quantity, as in A; lowercase boldface indicates a vector quantity, as in $\mathbf{a}$; uppercase boldface indicates a matrix quantity, as in $\mathbf{A}$. The matrix transpose is indicated by a superscript ^{T}, as in $\mathbf{A}^T$. The m^{th}-row and n^{th}-column element of the matrix $\mathbf{A}$ is denoted by $[\mathbf{A}]_{mn}$. The identity matrix of size $N \times N$ is denoted $\mathbf{I}_N$. By $\mathbf{1}_{M\times N}$, we define the matrix such that $[\mathbf{1}]_{mn} = 1$, $\forall m = 1 \dots M$ and $\forall n = 1 \dots N$, and $\mathbf{D}(\mathbf{a})$ is a diagonal matrix formed by the elements of the row vector $\mathbf{a}$. The trace operator is defined as $Tr\{.\}$. The determinant of a matrix is denoted by $|.|$ and cardinality when applying to a set. $\mathbb{S}^n$ denotes the vector space of symmetric $n \times n$ matrices, and the subsets of nonnegative definite matrices and positive definite matrices are denoted by $\mathbb{S}^n_+$ and $\mathbb{S}^n_{++}$, respectively. The notation $\mathbf{A} \succeq \mathbf{B}$ means that for $\mathbf{A}$, $\mathbf{B} \in \mathbb{S}^n$, $\mathbf{A} - \mathbf{B} \in \mathbb{S}^n_+$, also known as Loewner partial ordering of symmetric matrices [78, 79]. The absolute value is denoted by $abs(.)$. The indicator function of a set S is denoted by $I_S(.)$ and the expectation operator is denoted by $\mathbb{E}[.]$. The observation space and the parameter space are denoted, respectively, by Ω and Θ.

The remainder of this chapter is organized as follows: In Section 4.2, we present the signal model and the assumptions, and we introduce the general structure of the Barankin bound for the signal model parameters. The computation and analysis of the Barankin bound for the change-point localization parameters are provided in Section 4.3. In Section 4.4, we analyze the cases of changes in the parameters of Gaussian and Poisson distributions. To illustrate our results, simulations are presented in Section 4.5. Finally, in Section 4.6 we summarize this chapter.

4.2 Problem formulation

4.2.1 Observation model

We consider the general case of N independent vector observations $\mathbf{X} = [\mathbf{x}_1,\ \mathbf{x}_2,\ \ldots,\ \mathbf{x}_N] \in \mathbb{R}^{M\times N}$, which can be obtained, for example, by a multiple sensor system and are modeled as follows:

$$\begin{cases} \mathbf{x}_i \sim p_1\left(\mathbf{x}_i;\boldsymbol{\eta}_1\right) \text{ for } i = 1,\ldots,t_1, \\ \mathbf{x}_i \sim p_2\left(\mathbf{x}_i;\boldsymbol{\eta}_2\right) \text{ for } i = t_1+1,\ldots,t_2, \\ \vdots \\ \mathbf{x}_i \sim p_{q+1}\left(\mathbf{x}_i;\boldsymbol{\eta}_{q+1}\right) \text{ for } i = t_q+1,\ldots,N, \end{cases} \tag{4.1}$$

where M is the size of the sample vector (*e.g.*, the number of sensors), q is the number of change-points, and p_j is a probability density function (or mass function for discrete random variables) with parameters $\boldsymbol{\eta}_j \in \mathbb{R}^L$. In other words,

$$\begin{aligned} \mathbf{x}_i \ \sim\ p_j\left(\mathbf{x}_i;\boldsymbol{\eta}_j\right) &\text{ for } i = t_{j-1}+1,\ \ldots,\ t_j, \\ &\text{with } j = 1,\ \ldots,\ q+1, \end{aligned} \tag{4.2}$$

where we define $t_0 = 0$ and $t_{q+1} = N$. Note that if $M = 1$, the problem is reduced to the estimation of changes in a time series. We assume that all probability density functions p_j belong to a common distribution. The unknown parameters of interest are the change-point locations $\{t_1, t_2, \ldots, t_q\}$ with $\{t_k \in \mathbb{N} - \{0\},\quad k = 1,\ldots,\ q\}$, $1 < t_1 < t_2 < \cdots < t_q < N$, and $q < N-2$. The observations between two consecutive change points are assumed to be stationary. Consequently, the $q \times 1$

vector of unknown true parameters for this model is

$$\mathbf{t} = [t_1,\ t_2,\ \ldots,\ t_q]^T. \tag{4.3}$$

The observation model (4.2) is useful in signal processing; several examples were mentioned in the Introduction. Note that, since we focus on the change-point estimation, we assume that the parameters $\boldsymbol{\eta}_j$ are known. The resulting bound will still be useful if these parameters are unknown, but overly optimistic. Moreover, the complexity of the bound derivation increases for unknown $\boldsymbol{\eta}_j$ and therefore we do not consider this case in this work.

4.2.2 Barankin bound

The P-order BB of a vector $\boldsymbol{\theta}_0 \in \mathbb{R}^q$, denoted by $\mathbf{BB}_P(\boldsymbol{\theta}_0, \mathbf{H}_{q\times P})$, is given as follows (see [80–83] for more details):

$$\mathbf{Cov}(\widehat{\boldsymbol{\theta}}) \succeq \mathbf{BB}_P(\boldsymbol{\theta}_0, \mathbf{H}_{q\times P}) = \mathbf{H}_{q\times P}(\boldsymbol{\Phi} - \mathbf{1}_{P\times P})^{-1}\mathbf{H}_{q\times P}^T, \tag{4.4}$$

where $\mathbf{Cov}(\widehat{\boldsymbol{\theta}})$ is the covariance matrix of an unbiased estimator $\widehat{\boldsymbol{\theta}}$ of the parameter vector $\boldsymbol{\theta}_0$. The matrix $\mathbf{H} = [\boldsymbol{\theta}_1 - \boldsymbol{\theta}_0,\ \ldots,\ \boldsymbol{\theta}_P - \boldsymbol{\theta}_0]$ is a function of the set $\{\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_P\}$, the so-called "test points", which are left to the user. We define $\mathbf{h}_i = \boldsymbol{\theta}_i - \boldsymbol{\theta}_0$ such that the matrix $\mathbf{H} \in \mathbb{R}^{q\times P}$ becomes $\mathbf{H} = [\mathbf{h}_1,\ \ldots,\ \mathbf{h}_P]$. Moreover, note that $\boldsymbol{\theta}_0 + \mathbf{h}_j \in \Theta$. In the following, for simplicity, we use the term "test point" for the vectors $\mathbf{h}_i$. Finally, $\boldsymbol{\Phi}$ is a $\mathbb{R}^{P\times P}$ matrix whose elements $[\boldsymbol{\Phi}]_{kl}$ are given by

$$[\boldsymbol{\Phi}]_{kl} = \mathbb{E}[L(\mathbf{X}, \boldsymbol{\theta}_0, \mathbf{h}_k)L(\mathbf{X}, \boldsymbol{\theta}_0, \mathbf{h}_l)], \tag{4.5}$$

where $L(\mathbf{X}, \boldsymbol{\theta}_0, \mathbf{h}_j)$ is defined by,

$$L(\mathbf{X}, \boldsymbol{\theta}_0, \mathbf{h}_j) = \frac{p(\mathbf{X}; \boldsymbol{\theta}_0 + \mathbf{h}_j)}{p(\mathbf{X}; \boldsymbol{\theta}_0)}, \tag{4.6}$$

where $p(\mathbf{X}; \boldsymbol{\varphi})$ is the likelihood of the observations with parameter vector $\boldsymbol{\varphi}$. Note that the matrix $\boldsymbol{\Phi} - \mathbf{1}_{P \times P}$ is sometimes referred to as the Barankin information matrix (BIM) [84].

As already stated, the choice of test points is left to the user, since any set of test points in $\mathbf{BB}_P(\boldsymbol{\theta}_0)$ satisfies the inequality (4.4). Thus, the tightest BB, denoted by $\mathbf{BB}(\boldsymbol{\theta}_0)$, is given as follows:

$$\mathbf{BB}(\boldsymbol{\theta}_0) = \lim_{P \longrightarrow |\Theta|} \sup_{\{\mathbf{h}_1, \ldots, \mathbf{h}_P\}} \mathbf{BB}_P(\boldsymbol{\theta}_0, \mathbf{H}_{q \times P}) \succeq \mathbf{CRB}(\boldsymbol{\theta}_0), \tag{4.7}$$

where $|\Theta|$ is the cardinality of the set Θ formed by all possible parameter values, and $\mathbf{CRB}(\boldsymbol{\theta}_0)$ is the CRB of $\boldsymbol{\theta}_0$, which, assuming that it exists, is smaller than the $\mathbf{BB}(\boldsymbol{\theta}_0)$ in the Loewner ordering sense. The computation of $\mathbf{BB}(\boldsymbol{\theta}_0)$ is costly, since the limit on P usually implies that a large, possibly infinite, number of test points needs to be considered, e.g. for $\Theta \subseteq \mathbb{R}^q$, a nonlinear maximization over the test points has to be performed, and the inverse of the BIM has to be computed.

Concerning the BB for the parameter vector given in (4.3), *i.e.*, $\boldsymbol{\theta}_0 = \mathbf{t}$, and $|\Theta|$ depends on the number of samples N and change points q as follows:

$$\begin{aligned} |\Theta| &= \sum_{t_1=1}^{N-q} \sum_{t_2=t_1+1}^{N-q+1} \cdots \sum_{t_{q-1}=t_{q-2}+1}^{N-1} (N - t_{q-1} - 1), \\ &= \binom{N-1}{q}. \end{aligned} \tag{4.8}$$

Note that $|\Theta| \to \infty$ as $(N-q) \to \infty$, and for N finite then $|\Theta|$ is finite. In practice, the number of test points and the particular structure of the matrix $\mathbf{H}$ is usually chosen based on the analytical and computational complexity associated with it, which lead to approximated versions of the BB. In the latter case it would be useful to have some knowledge of how different Barankin bound approximations compare among each other *w.r.t. Loewner partial ordering.* In the following proposition we provide with a general guideline for this purpose:

Lemma 1 *Let* $\mathbf{A} \in \mathbb{S}_{++}^q$, $\mathbf{B} \in \mathbb{S}_{+}^q$ *with* $rank\,(\mathbf{B}) = m < q$, *and let* $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m > 0$ *and* $\lambda_{m+1} = \cdots = \lambda_q = 0$ *be the roots of the characteristic equation* $|\mathbf{B} - \lambda\mathbf{A}| = 0$. *If* $\lambda_1 \leq 1$, *then* $\mathbf{A} \succ \mathbf{B}$, *otherwise* $\mathbf{A}$ *and* $\mathbf{B}$ *are not mutually comparable.* ***Proof.*** *See Appendix C* ■

If $rank\,(\mathbf{H}_{q\times P}) = q$, then $\mathbf{BB}_P(\boldsymbol{\theta}_0, \mathbf{H}_{q\times P}) \in \mathbb{S}_{++}^q$, since $(\boldsymbol{\Phi} - 1)^{-1} \in \mathbb{S}_{++}^q$ by construction, and if $rank\,(\mathbf{H}_{q\times P'}) < q$ then $\mathbf{BB}_P(\boldsymbol{\theta}_0, \mathbf{H}_{q\times P'}) \in \mathbb{S}_{+}^q$. The lemma can now be used with $\mathbf{A} = \mathbf{BB}_P(\boldsymbol{\theta}_0, \mathbf{H}_{q\times P})$ and $\mathbf{B} = \mathbf{BB}_{P'}(\boldsymbol{\theta}_0, \widetilde{\mathbf{H}}_{q\times P'})$, provided $rank\,(\mathbf{H}_{q\times P}) = q > rank\left(\widetilde{\mathbf{H}}_{q\times P'}\right)$. Note that $rank\left(\widetilde{\mathbf{H}}_{q\times P'}\right) < q$ implies that the number of test-points $P' < q$, therefore, a matrix bound $\mathbf{BB}_{P'}(\boldsymbol{\theta}_0, \widetilde{\mathbf{H}}_{q\times P'})$ cannot be larger, *w.r.t. Loewner partial ordering,* than any matrix bound given by a test-point matrix $\mathbf{H}_{q\times P}$ consisting of $P = q$ independent test-point vectors. Consequently, in the following we will use an approximate version of the BB that allows us to derive efficiently computed closed-form expressions for the BIM and its inverse in the context of our multiple change-point estimation problem. In particular, we will compute the multi-parameter HCR bound with the classical assumption of one test point per parameter

$(P = q)$, i.e., $\mathbf{h}_j = [0,\ \ldots,\ \alpha_j,\ \ldots,\ 0]^T$. Then, $\mathbf{H}$ is a diagonal matrix given by

$$\begin{aligned} \mathbf{H} &= [\mathbf{h}_1,\ \ldots,\ \mathbf{h}_q] \\ &= \mathbf{D}(\boldsymbol{\alpha}), \end{aligned} \tag{4.9}$$

where the vector $\boldsymbol{\alpha} = [\alpha_1,\ \ldots,\ \alpha_q]^T$ corresponds to the set of test points associated to the parameters $\mathbf{t} = [t_1,\ t_2,\ \ldots,\ t_q]^T$. Note that $\alpha_j \neq 0$ is defined such that $t_j + \alpha_j$ ranges over all possible values of t_j, for $j = 1, \ldots, q$. Thus, $\alpha_j \in \{\mathbb{Z} \cap [t_{j-1} - t_j + 1,\ t_{j+1} - t_j - 1] - \{0\}\}$. Let $S \subset \mathbb{Z}^q$ be a set formed by all possible values of $\boldsymbol{\alpha}$. The set S is finite given that t_{q+1} is finite.

The matrix $\boldsymbol{\Phi} - \mathbf{1}_{q\times q}$ corresponds to the BIM for change-point locations $\mathbf{t}$, denoted here by $\mathbf{BIM_t}$. The approximated BB, $\mathbf{BB}_{\mathbf{t},\,q}$, is then obtained from

$$\begin{aligned} \mathbf{BB}_{\mathbf{t},\,q} &= \sup_{[\mathbf{h}_1,\ldots,\mathbf{h}_q]} \mathbf{BB}_q(\mathbf{t}, \mathbf{H}_{q\times q}), \\ &= \sup_{\boldsymbol{\alpha}\,\in\,S} \mathbf{D}(\boldsymbol{\alpha})\ \mathbf{BIM}_{\mathbf{t}}^{-1}\mathbf{D}(\boldsymbol{\alpha})^T. \end{aligned} \tag{4.10}$$

By construction, the finite set $C := \{\mathbf{BB}_q(\mathbf{t}, \mathbf{D}(\boldsymbol{\alpha})),\ \boldsymbol{\alpha} \in S\}$ is a subset of the partially ordered set $(\mathbb{S}^q, \preceq)$ with partial order "$\preceq$" given by the *Loewner ordering* [78,79]. This partial order is not a lattice ordering, *i.e.*, each finite subset of $\mathbb{S}^q$ may not be closed under the least-upper (infimum) and greatest-lower bounds (supremum) [79]. In other words, the notion of a unique supremum or an infimum of C might not exist with respect to the Loewner ordering. The supremum does not exist if there is no upper bound to the set, or if the set of upper bounds does not have a least element. If the supremum exists, it does not need to be defined in the set, but if it belongs

to it, then it is the greatest element[6] in the set. Note that a set with respect to the partially order set $(\mathbb{S}^q, \succeq)$ may have several maximal[7] and minimal elements without having a greatest and least element in the set, respectively. If the set has a greatest or least element, then it is the unique maximal or minimal element, and therefore it is the supremum or infimum of the set. Here, we will approach the computation of the supremum by computing a suitable minimal element of the set of upper bounds of C, namely, a minimal-upper bound $\mathbf{B}_q \in \mathbb{S}^q_{++}$ such that $\mathbf{B}_q \succeq C$ and which is minimal in the sense that there is not smaller matrix $\mathbf{B}'_q \preceq \mathbf{B}_q$ such that $\mathbf{B}'_q \succeq C$. From Eq. (4.4), $\mathbf{Cov}(\widehat{\boldsymbol{\theta}})$ belongs to set of upper bounds of C, therefore if the set of upper bounds has a unique minimal element, i.e., a least element, then $\mathbf{Cov}(\widehat{\boldsymbol{\theta}}) \succeq \mathbf{B}_q$. However, if the set of upper bounds has several minimal elements, then in general we can expect that $\mathbf{Cov}(\widehat{\boldsymbol{\theta}}) \succeq \mathbf{B}_q$, or that $\mathbf{Cov}(\widehat{\boldsymbol{\theta}})$ and $\mathbf{B}_q$ are not mutually comparable.

Having a closed form for $\mathbf{BIM}_\mathbf{t}^{-1}$ makes the task of computing $\mathbf{B}_q$ much less computationally demanding than having to invert $\mathbf{BIM}_\mathbf{t}$ for every $\boldsymbol{\alpha} \in S$. In the following section, we will first derive the elements of $\mathbf{BIM}_\mathbf{t}$ and obtain closed-form expressions for $\mathbf{BIM}_\mathbf{t}^{-1}$. Then, we will introduce the approach for computing the minimal-upper bound $\mathbf{B}_q$.

[6] $\mathbf{B}_i \in C$ is the greatest element of C *w.r.t.* $(\mathbb{S}^q, \preceq)$ if $\mathbf{B}_i \succeq \mathbf{Y}$ for all $\mathbf{Y} \in C$. If the greatest element exists it is an upper-bound of C contained in it. The least element of C is defined similarly, considering $\mathbf{B}_i \preceq \mathbf{Y}$.

[7] $\mathbf{B}_i \in C$ is a maximal element of C *w.r.t.* $(\mathbb{S}^q, \preceq)$ if there is not $\mathbf{Y} \in C$ such that $\mathbf{Y} \succeq \mathbf{B}_i$ and is a minimal element if there is not $\mathbf{Y} \in C$ such that $\mathbf{B}_i \succeq \mathbf{Y}$.

4.3 Barankin bound type for multiple change-point estimation

To compute the BB for the change point localization parameters, we first need to compute $\mathbf{BIM_t}$, which depends on the matrix $\mathbf{\Phi}$. From Equations (4.5) and (4.6), the elements of $[\mathbf{\Phi}]_{kl}$, for k, $l = 1, \ldots, q$ are given by

$$[\mathbf{\Phi}]_{kl} = \int_{\Omega} \frac{p\left(\mathbf{X}; \mathbf{t} + \mathbf{h}_k\right) p\left(\mathbf{X}; \mathbf{t} + \mathbf{h}_l\right)}{p\left(\mathbf{X}; \mathbf{t}\right)} d\mathbf{X}, \tag{4.11}$$

where $p\left(\mathbf{X}; \mathbf{t}\right)$ is given by

$$p\left(\mathbf{X}; \mathbf{t}\right) = \prod_{i=1}^{t_1} p_1(\mathbf{x}_i; \boldsymbol{\eta}_1) \cdots \prod_{i=t_{k-1}+1}^{t_k} p_k(\mathbf{x}_i; \boldsymbol{\eta}_k) \cdots \prod_{i=t_q+1}^{N} p_{q+1}(\mathbf{x}_i; \boldsymbol{\eta}_{q+1}), \tag{4.12}$$

and $p\left(\mathbf{X}; \mathbf{t} + \mathbf{h}_k\right)$ is given by

$$p\left(\mathbf{X}; \mathbf{t} + \mathbf{h}_k\right) = \prod_{i=1}^{t_1} p_1(\mathbf{x}_i; \boldsymbol{\eta}_1) \cdots \prod_{i=t_{k-1}+1}^{t_k+\alpha_k} p_k(\mathbf{x}_i; \boldsymbol{\eta}_k) \cdots \prod_{i=t_q+1}^{N} p_{q+1}(\mathbf{x}_i; \boldsymbol{\eta}_{q+1}), \tag{4.13}$$

and where $p\left(\mathbf{X}; \mathbf{t} + \mathbf{h}_l\right)$ is same as Equation (4.13) ($k = l$).

In order to study and to simplify $\mathbf{\Phi}$, we will analyze its diagonal and non-diagonal elements separately.

4.3.1 Diagonal elements of $\mathbf{\Phi}$

Replacing $k = l$ in (4.11) and using (4.13), we obtain the following expression:

$$\begin{aligned}[\mathbf{\Phi}]_{kk} &= \int_{\Omega} \frac{p^2\left(\mathbf{X}; \mathbf{t} + \mathbf{h}_k\right)}{p\left(\mathbf{X}; \mathbf{t}\right)} d\mathbf{X} \\ &= \int_{\Omega} \frac{\prod_{i=1}^{t_1} p_1^2(\mathbf{x}_i; \boldsymbol{\eta}_1) \cdots \prod_{i=t_{k-1}+1}^{t_k+\alpha_k} p_k^2(\mathbf{x}_i; \boldsymbol{\eta}_k) \cdots \prod_{i=t_q+1}^{N} p_{q+1}^2(\mathbf{x}_i; \boldsymbol{\eta}_{q+1})}{\prod_{i=1}^{t_1} p_1(\mathbf{x}_i; \boldsymbol{\eta}_1) \cdots \prod_{i=t_q+1}^{N} p_{q+1}(\mathbf{x}_i; \boldsymbol{\eta}_{q+1})} d\mathbf{X}. \quad (4.14)\end{aligned}$$

This equation can be further simplified by considering the the cases $\alpha_k > 0$ and $\alpha_k < 0$, obtaining the following expression (see Appendix D.0.1 for details on its derivation):

$$[\mathbf{\Phi}]_{kk} = \begin{cases} \left(\int_{\Omega} \frac{p_k^2(\mathbf{x}; \boldsymbol{\eta}_k)}{p_{k+1}(\mathbf{x}; \boldsymbol{\eta}_{k+1})} d\mathbf{x}\right)^{\alpha_k}, & \text{if } \alpha_k > 0, \\ \left(\int_{\Omega} \frac{p_{k+1}^2(\mathbf{x}; \boldsymbol{\eta}_{k+1})}{p_k(\mathbf{x}; \boldsymbol{\eta}_k)} d\mathbf{x}\right)^{-\alpha_k}, & \text{if } \alpha_k < 0. \end{cases} \quad (4.15)$$

4.3.2 Non-diagonal elements of $\mathbf{\Phi}$

The computation of the off-diagonal elements of $\mathbf{\Phi}$ can be simplified by using the fact that the matrix $\mathbf{\Phi}$ is symmetric; therefore, we can focus on either the upper or lower triangular part of $\mathbf{\Phi}$. In our derivations below we consider the upper triangular part, *i.e.*, $k < l$, then by using (4.11) and (4.13), we obtain the following expression for the

elements of $\mathbf{\Phi}$:

$$
\begin{aligned}
[\mathbf{\Phi}]_{kl} \;=\; & \int_{\Omega} \frac{\prod_{i=1}^{t_1} p_1(\mathbf{x}_i;\boldsymbol{\eta}_1)\cdots \prod_{i=t_{k-1}+1}^{t_k+\alpha_k} p_k(\mathbf{x}_i;\boldsymbol{\eta}_k)}{\prod_{i=1}^{t_1} p_1(\mathbf{x}_i;\boldsymbol{\eta}_1)\cdot\cdots\cdot \prod_{i=t_q+1}^{N} p_{q+1}(\mathbf{x}_i;\boldsymbol{\eta}_{q+1})} \\
& \times \prod_{i=t_k+\alpha_k+1}^{t_{k+1}} p_{k+1}(\mathbf{x}_i;\boldsymbol{\eta}_{k+1})\cdots \prod_{i=t_q+1}^{N} p_{q+1}(\mathbf{x}_i;\boldsymbol{\eta}_{q+1}) \\
& \times \prod_{i=1}^{t_1} p_1(\mathbf{x}_i;\boldsymbol{\eta}_1)\cdots \prod_{i=t_{l-1}+1}^{t_l+\alpha_l} p_l(\mathbf{x}_i;\boldsymbol{\eta}_l)\cdots \prod_{i=t_q+1}^{N} p_{q+1}(\mathbf{x}_i;\boldsymbol{\eta}_{q+1}) d\mathbf{X}. \quad (4.16)
\end{aligned}
$$

Following the same idea as for the diagonal elements, $[\mathbf{\Phi}]_{kl}$ can be simplified by analyzing the four possible combinations of test-point ranges, namely,

$$
\begin{cases}
\text{Case 1: } \alpha_k > 0 \text{ and } \alpha_l > 0, \\
\text{Case 2: } \alpha_k < 0 \text{ and } \alpha_l < 0, \\
\text{Case 3: } \alpha_k < 0 \text{ and } \alpha_l > 0, \\
\text{Case 4: } \alpha_k > 0 \text{ and } \alpha_l < 0.
\end{cases} \quad (4.17)
$$

For the last case, *i.e.* $\alpha_k > 0$ and $\alpha_l < 0$, two subcases have to be analyzed: (i) $t_k + \alpha_k < t_l + \alpha_l$ and (ii) $t_k + \alpha_k > t_l + \alpha_l$. These two cases correspond to non-overlapping and overlapping test points, respectively (see Fig. 4.1). Note that since $k < l$, $t_k < t_l$ and since $\alpha_j \in \{\mathbb{Z} \cap [t_{j-1} - t_j + 1, t_{j+1} - t_j - 1] - \{0\}\}$, the case $t_k + \alpha_k > t_l + \alpha_l$ which corresponds to an overlapping between two test points, can appear only when $l = k+1$, or, in other words, when we are analyzing two neighboring change points.

Then, for Cases 1-3 and subcase (i), Equation (4.16) becomes (see Appendix D.0.2)

$$
[\mathbf{\Phi}]_{kl} = 1, \text{ for } l > k \quad (4.18)
$$

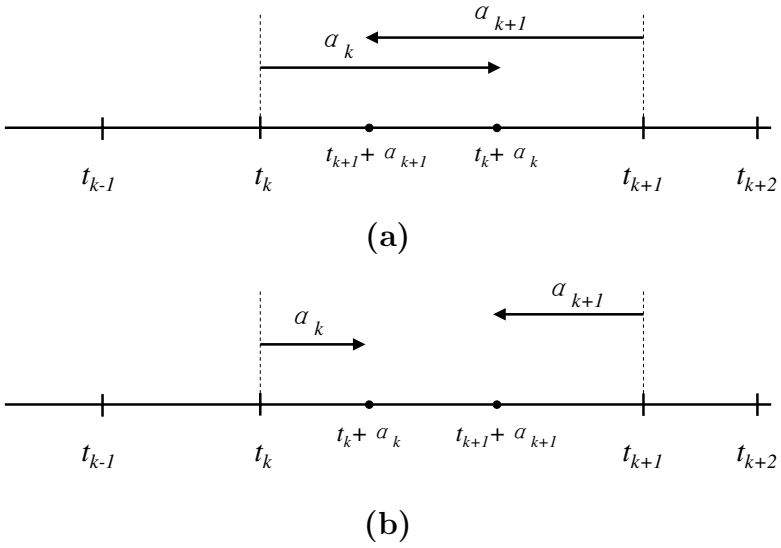


Figure 4.1: (a) Case of overlapping; (b) Case of no overlapping.

and for subcase (ii), keeping in mind that $\alpha_k > 0$ and $\alpha_{k+1} < 0$, Equation (4.16) becomes

$$[\mathbf{\Phi}]_{kl} = \begin{cases} \left(\int_{\Omega} \frac{p_k(\mathbf{x};\boldsymbol{\eta}_k) p_{k+2}(\mathbf{x};\boldsymbol{\eta}_{k+2})}{p_{k+1}(\mathbf{x};\boldsymbol{\eta}_{k+1})} d\mathbf{x} \right)^{\beta_k}, & \text{for } l = k+1, \\ 1, & \text{for } l > k+1, \end{cases} \tag{4.19}$$

where $\beta_k = (t_k + \alpha_k) - (t_{k+1} + \alpha_{k+1})$.

Remark: This last result is fundamental because it proves the natural intuition that the estimation of q change points is not equivalent to q times the estimation of one change point. In other words, it means that the estimation of one change point is perturbed by its two neighbors. We now summarize the previous results.

4.3.3 Barankin information matrix $\mathbf{\Phi} - \mathbf{1}_{q\times q}$

Using Equations (4.15), (4.18), and (4.19), it is clear that $\mathbf{BIM_t}$ has at least a tri-diagonal structure:

$$\mathbf{BIM_t} = \begin{bmatrix} A_1 & B_1 & 0 & \cdots & 0 \\ B_1 & A_2 & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & A_{q-1} & B_{q-1} \\ 0 & \cdots & 0 & B_{q-1} & A_q \end{bmatrix}, \tag{4.20}$$

where

$$\begin{aligned} A_k &= [\mathbf{\Phi}]_{kk} - 1, \quad \text{for } k = 1, \ldots, q \\ &= \begin{cases} \left(\int_\Omega \frac{p_k^2(\mathbf{x};\boldsymbol{\eta}_k)}{p_{k+1}(\mathbf{x};\boldsymbol{\eta}_{k+1})} d\mathbf{x}\right)^{\alpha_k} - 1 & \text{if } \alpha_k > 0, \\ \left(\int_\Omega \frac{p_{k+1}^2(\mathbf{x};\boldsymbol{\eta}_{k+1})}{p_k(\mathbf{x};\boldsymbol{\eta}_k)} d\mathbf{x}\right)^{-\alpha_k} - 1 & \text{if } \alpha_k < 0, \end{cases} \end{aligned} \tag{4.21}$$

and

$$\begin{aligned} B_k &= [\mathbf{\Phi}]_{kk+1} - 1, \quad \text{for } k = 1, \ldots, q-1 \\ &= \begin{cases} 0, & \text{if } \beta_k < 0, \\ \left(\int_\Omega \frac{p_k(\mathbf{x};\boldsymbol{\eta}_k) p_{k+2}(\mathbf{x};\boldsymbol{\eta}_{k+2})}{p_{k+1}(\mathbf{x};\boldsymbol{\eta}_{k+1})} d\mathbf{x}\right)^{\beta_k} - 1, & \\ & \text{if } \beta_k > 0. \end{cases} \end{aligned} \tag{4.22}$$

In the case of one change-point estimation, $\mathbf{BIM_t}$ is reduced to a scalar A_1, and by replacing $\alpha_1 = \alpha$ we re-obtain the result proposed by Ferrari and Tourneret (see

Equations (5) and (6) in [75]):

$$A_1 = \begin{cases} \left(\int_\Omega \frac{p_1^2(\mathbf{x};\boldsymbol{\eta}_1)}{p_2(\mathbf{x};\boldsymbol{\eta}_2)} d\mathbf{x}\right)^{\alpha} - 1 & \text{if } \alpha > 0, \\ \left(\int_\Omega \frac{p_2^2(\mathbf{x};\boldsymbol{\eta}_2)}{p_1(\mathbf{x};\boldsymbol{\eta}_1)} d\mathbf{x}\right)^{-\alpha} - 1 & \text{if } \alpha < 0. \end{cases} \tag{4.23}$$

Note also that the diagonal elements of $\mathbf{BIM_t}$ can be computed numerically in one step (*i.e.*, $\forall \alpha_k \gtrless 0$) as follows:

$$A_k = \left(\int_\Omega \left(\frac{p_k(\mathbf{x};\boldsymbol{\eta}_k)}{p_{k+1}(\mathbf{x};\boldsymbol{\eta}_{k+1})}\right)^{\epsilon_k} p_{k+1}(\mathbf{x};\boldsymbol{\eta}_{k+1}) d\mathbf{x}\right)^{abs(\alpha_k)} - 1, \tag{4.24}$$

where $\epsilon_k = \frac{1}{2}\left(3\frac{\alpha_k}{abs(\alpha_k)} + 1\right)$.

The next step of our analysis is to compute $(\mathbf{BIM_t})^{-1}$. For a given set of test points, it is clear that $t_k + \alpha_k > t_{k+1} + \alpha_{k+1} \Longrightarrow t_{k+1} + \alpha_{k+1} < t_{k+2} + \alpha_{k+2}$, since $\alpha_j \in \{\mathbb{Z} \cap [t_{j-1} - t_j + 1, t_{j+1} - t_j - 1] - \{0\}\}$. In other words, $\forall k$, if $B_k \neq 0$, then $B_{k+1} = B_{k-1} = 0$; therefore, $\mathbf{BIM_t}$ is block diagonal and the maximum size of one block is 2×2. Since the problem is reduced to finding, at worst, the inverse of several 2×2 matrices with the same structure, we will have a straightforward inversion. In this section, we detail the case of two change points, we give the generalization to two neighboring points, and we use this to derive a closed-form expression for the inverse of $\mathbf{BIM_t}$ and thus $\mathbf{BB}_q(\mathbf{t}, \mathbf{D}(\boldsymbol{\alpha}))$.

The case of two change points

In this case we have $q = 2$, $\mathbf{t} = [t_1,\ t_2]^T$, and $\mathbf{BB}_q(\mathbf{t}, \mathbf{D}(\boldsymbol{\alpha}))$ becomes

$$\begin{aligned} \mathbf{BB}_2(\mathbf{t}, \mathbf{D}(\boldsymbol{\alpha})) &= \begin{bmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{bmatrix} \begin{bmatrix} A_1 & B_1 \\ B_1 & A_2 \end{bmatrix}^{-1} \begin{bmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{bmatrix} \\ &= \frac{1}{A_1 A_2 - B_1^2} \begin{bmatrix} \alpha_1^2 A_2 & -\alpha_1\alpha_2 B_1 \\ -\alpha_1\alpha_2 B_1 & \alpha_2^2 A_1 \end{bmatrix}, \end{aligned} \tag{4.25}$$

with

$$A_1 = \begin{cases} \left(\int_\Omega \frac{p_1^2(\mathbf{x};\boldsymbol{\eta}_1)}{p_2(\mathbf{x};\boldsymbol{\eta}_2)} d\mathbf{x}\right)^{\alpha_1} - 1 & \text{if } \alpha_1 > 0, \\ \left(\int_\Omega \frac{p_2^2(\mathbf{x};\boldsymbol{\eta}_2)}{p_1(\mathbf{x};\boldsymbol{\eta}_1)} d\mathbf{x}\right)^{-\alpha_1} - 1 & \text{if } \alpha_1 < 0, \end{cases} \tag{4.26}$$

$$A_2 = \begin{cases} \left(\int_\Omega \frac{p_2^2(\mathbf{x};\boldsymbol{\eta}_2)}{p_3(\mathbf{x};\boldsymbol{\eta}_3)} d\mathbf{x}\right)^{\alpha_2} - 1 & \text{if } \alpha_2 > 0, \\ \left(\int_\Omega \frac{p_3^2(\mathbf{x};\boldsymbol{\eta}_3)}{p_2(\mathbf{x};\boldsymbol{\eta}_2)} d\mathbf{x}\right)^{-\alpha_2} - 1 & \text{if } \alpha_2 < 0, \end{cases} \tag{4.27}$$

$$B_1 = \begin{cases} 0 & \text{if } \beta_1 < 0, \\ \left(\int_\Omega \frac{p_1(\mathbf{x};\boldsymbol{\eta}_1) p_3(\mathbf{x};\boldsymbol{\eta}_3)}{p_2(\mathbf{x};\boldsymbol{\eta}_2)} d\mathbf{x}\right)^{\beta_1} - 1 & \\ & \text{if } \beta_1 > 0, \end{cases} \tag{4.28}$$

where $\beta_1 = (t_1 + \alpha_1) - (t_2 + \alpha_2)$.

Consequently, depending on the given set of test points, the following five combinations, corresponding respectively to cases (1), (2), (3), and (4) in (4.17), are possible

for $\mathbf{BB}_2(\mathbf{t}, \mathbf{D}(\boldsymbol{\alpha}))$:

$$\left\{ \begin{bmatrix} \frac{\alpha_1^2}{\Delta_{112}^{\alpha_1}-1} & 0 \\ 0 & \frac{\alpha_2^2}{\Delta_{223}^{\alpha_2}-1} \end{bmatrix}, \begin{bmatrix} \frac{\alpha_1^2}{\Delta_{221}^{abs(\alpha_1)}-1} & 0 \\ 0 & \frac{\alpha_2^2}{\Delta_{332}^{abs(\alpha_2)}-1} \end{bmatrix}, \right.$$
$$\begin{bmatrix} \frac{\alpha_1^2}{\Delta_{221}^{abs(\alpha_1)}-1} & 0 \\ 0 & \frac{\alpha_2^2}{\Delta_{223}^{\alpha_2}-1} \end{bmatrix}, \begin{bmatrix} \frac{\alpha_1^2}{\Delta_{112}^{\alpha_1}-1} & 0 \\ 0 & \frac{\alpha_2^2}{\Delta_{332}^{abs(\alpha_2)}-1} \end{bmatrix}, \tag{4.29}$$
$$\left. \kappa^{-1} \begin{bmatrix} \alpha_1^2 \left(\Delta_{332}^{abs(\alpha_2)} - 1\right) & \alpha_1\alpha_2 \left(1 - \Delta_{132}^{\beta_1}\right) \\ \alpha_1\alpha_2 \left(1 - \Delta_{132}^{\beta_1}\right) & \alpha_2^2 \left(\Delta_{112}^{\alpha_1} - 1\right) \end{bmatrix} \right\}$$

where we define

$$\Delta_{ijk} = \int_\Omega \frac{p_i(\mathbf{x}; \boldsymbol{\eta}_i) p_j(\mathbf{x}; \boldsymbol{\eta}_j)}{p_k(\mathbf{x}; \boldsymbol{\eta}_k)} d\mathbf{x},$$

and $\kappa = \left(\Delta_{112}^{\alpha_1} - 1\right)\left(\Delta_{332}^{abs(\alpha_2)} - 1\right) - \left(\Delta_{132}^{\beta_1} - 1\right)^2$.

Generalization to q change points

Note that for more change points the process is the same, except that the inversion has to be computed because of the increase of possibilities. However, the matrix to be inverted is block diagonal, with block of size 1×1 or 2×2, as stated in the previous section. In particular, depending on the values of $\boldsymbol{\alpha}$, the elements of $\left[\mathbf{BIM}_{\mathbf{t}}^{-1}\right]_{kl}$ for $1 < k < q$ and $l = \{k,\, k+1\}$, with $\mathbf{BIM}_{\mathbf{t}}$, A_k, and B_k given by Equations (4.20), (4.21), and (4.22), respectively, and $\alpha_0 = \alpha_q = 0$ and $B_q = 0$, have the following possible values:

If $t_{k+1}+\alpha_{k+1}<t_k+\alpha_k$, then $\alpha_k>0$, $B_k\neq 0$ and $B_{k-1}=B_{k+1}=0$, thus

$$\left[\mathbf{BIM}_{\mathbf{t}}^{-1}\right]_{kl}=\begin{cases}\frac{A_{k+1}}{A_kA_{k+1}-B_k^2}, & \text{for } l=k,\\ -\frac{B_k}{A_kA_{k+1}-B_k^2}, & \text{for } l=\ k+1.\end{cases} \tag{4.30}$$

If $t_k+\alpha_k<t_{k-1}+\alpha_{k-1}$, then $\alpha_k<0$, $B_{k-1}\neq 0$ and $B_{k-2}=B_k=0$, thus

$$\left[\mathbf{BIM}_{\mathbf{t}}^{-1}\right]_{kl}=\begin{cases}\frac{A_{k-1}}{A_kA_{k-1}-B_{k-1}^2}, & \text{for } l=k,\\ 0, & \text{for } l=\ k+1.\end{cases} \tag{4.31}$$

If $t_{k-1}+\alpha_{k-1}<t_k+\alpha_k<t_{k+1}+\alpha_{k+1}$, then $B_{k-1}=B_k=0$, thus

$$\left[\mathbf{BIM}_{\mathbf{t}}^{-1}\right]_{kl}=\begin{cases}\frac{1}{A_k}, & \text{for } l=k,\\ 0, & \text{for } l=\ k+1.\end{cases} \tag{4.32}$$

Therefore, the elements of $\left[\mathbf{BIM}_{\mathbf{t}}^{-1}\right]_{kl}$ for $k,l=1,\ldots,q$, which is a symmetric matrix, are given by

$$\left[\mathbf{BIM}_{\mathbf{t}}^{-1}\right]_{kl}=\begin{cases}\frac{A_{k-1}I_{[-\infty,-1]}(\alpha_k)+A_{k+1}I_{[1,\infty]}(\alpha_k)}{A_k\left(A_{k-1}I_{[-\infty,-1]}(\alpha_k)+A_{k+1}I_{[1,\infty]}(\alpha_k)\right)-\left(B_{k-1}^2+B_k^2\right)}, & \\ \qquad\qquad\qquad\qquad \text{for } l=k,\\ \frac{-B_k}{A_kA_{k+1}-B_k^2}, \qquad \text{for } l=\ k+1,\\ 0, \qquad\qquad \text{for } l>k+1.\end{cases} \tag{4.33}$$

Since the matrix $\mathbf{BB}_q(\mathbf{t},\mathbf{D}\left(\boldsymbol{\alpha}\right))$ is given by

$$\mathbf{BB}_q(\mathbf{t},\mathbf{D}\left(\boldsymbol{\alpha}\right))=\mathbf{D}\left(\boldsymbol{\alpha}\right)\ \mathbf{BIM}_{\mathbf{t}}^{-1}\mathbf{D}\left(\boldsymbol{\alpha}\right)^T, \tag{4.34}$$

then, $[\mathbf{BB}_q(\mathbf{t}, \mathbf{D}(\boldsymbol{\alpha}))]_{kl}$ for $k, l = 1, \ldots, q$ is given as follows:

$$[\mathbf{BB}_q(\mathbf{t}, \mathbf{D}(\boldsymbol{\alpha}))]_{kl} = \begin{cases} \frac{\alpha_k^2\left(A_{k-1}I_{[-\infty,-1]}(\alpha_k)+A_{k+1}I_{[1,\infty]}(\alpha_k)\right)}{A_k\left(A_{k-1}I_{[-\infty,-1]}(\alpha_k)+A_{k+1}I_{[1,\infty]}(\alpha_k)\right)-\left(B_{k-1}^2+B_k^2\right)}, \\ \qquad\qquad\qquad\qquad \text{for } l = k, \\ -\frac{\alpha_k\alpha_{k+1}B_k}{A_kA_{k+1}-B_k^2}, \quad \text{for } l = k+1, \\ 0, \qquad\qquad \text{for } l > k+1, \end{cases} \tag{4.35}$$

where $[\mathbf{BB}_q(\mathbf{t}, \mathbf{D}(\boldsymbol{\alpha}))]_{kl} = [\mathbf{BB}_q(\mathbf{t}, \mathbf{D}(\boldsymbol{\alpha}))]_{lk}$. If for a given set of test points there is no overlap with the neighboring change-points t_{k-1} and t_{k+1}, then $B_{k-1} = B_k = 0$ in (4.35) and we obtain the particular result $[\mathbf{BB}_q(\mathbf{t}, \mathbf{D}(\boldsymbol{\alpha}))]_{kk} = \alpha_k^2/A_k$ and $[\mathbf{BB}_q(\mathbf{t}, \mathbf{D}(\boldsymbol{\alpha}))]_{kk+1} = [\mathbf{BB}_q(\mathbf{t}, \mathbf{D}(\boldsymbol{\alpha}))]_{k+1k} = 0$. This is equivalent to the bound obtained using the same set of test points and assuming one change-point located in the time interval between t_{k-1} and t_{k+1} with total numbers of time-samples $N = t_{k+1} - t_{k-1}$.

4.3.4 Computation of the supremum

To obtain the tightest bound from the finite set $C := \{\mathbf{BB}_q(\mathbf{t}, \mathbf{D}(\boldsymbol{\alpha})),\ \boldsymbol{\alpha} \in S\} \subset \mathbb{S}_{++}^q$, we need to compute the supremum of C with respect to the partially order set $(\mathbb{S}^q, \succeq)$. The partial order is given by the Loewner ordering, which is defined via the cone of positive semidefinite matrices [78, 79]. In general, this problem is indeed very complex since it requires to look for $\boldsymbol{\alpha}^* \in S$ such that $\mathbf{BB}_q(\mathbf{t}, \mathbf{D}(\boldsymbol{\alpha}^*)) \succeq \mathbf{BB}_q(\mathbf{t}, \mathbf{D}(\boldsymbol{\alpha}))$ for all $\boldsymbol{\alpha} \in S$. To the best of our knowledge, no formal approach for solving this problem has been proposed in the technical literature of minimal bounds. For example, in [81, 85] the choice of the test point $\boldsymbol{\alpha}$ is guided by some physical considerations of the model being studied. Also, from an optimal design context [78],

an approximation for solving this problem is to compute the matrix in C with the largest trace, $\mathbf{B}\mathbf{B}_{\text{tr}}$. However, the fact that $Tr\{\mathbf{B}\mathbf{B}_{\text{tr}}\} > Tr\{\mathbf{B}_i\}$ for $\mathbf{B}_i \in C$, does not imply that $\mathbf{B}\mathbf{B}_{\text{tr}} \succ \mathbf{B}_i$, only the converse statement is valid. In fact, only if C has a greatest element, i.e., the supremum of the set, then it is given by the matrix in C with the largest trace. Let $\mathbf{B}_j = \sup C$, with $\mathbf{B}_j \in C$, then by definition $\mathbf{B}_j \succeq \mathbf{B}_i$, for all $\mathbf{B}_i \in C$ with $i \neq j$. Let $\mathbf{G} = \mathbf{B}_j - \mathbf{B}_i$, thus $\mathbf{G} \in \mathbb{S}_+^q$ and $Tr\{\mathbf{G}\} > 0$. Hence, $Tr\{\mathbf{B}_j\} > Tr\{\mathbf{B}_i\}$, for all $\mathbf{B}_i \in C$ with $i \neq j$, but as we discussed at the end of Section II, a unique supremum or an infimum with respect to the Loewner partial ordering in the finite set C might not exist.

Here we address the computation of the supremum by finding a minimal-upper bound $\mathbf{B}_q \in \mathbb{S}_{++}^q$ to the set C such that $\mathbf{B}_q \succeq C$ and which is minimal in the sense that there is no smaller matrix $\mathbf{B}'_q \preceq \mathbf{B}_q$ such that $\mathbf{B}'_q \succeq C$. In [77], the authors implicitly introduced an algorithm for computing a minimal-upper bound to a finite set of positive definite matrices and redefined this element as the supremum of the set. Before discussing more details about it, we need to introduce the so-called penumbra $P(\mathbf{M})$ of a matrix $\mathbf{M} \in \mathbb{S}^q$ as the set $P(\mathbf{M}) := \{\mathbf{N} \in \mathbb{S}^q : \mathbf{N} \preceq \mathbf{M}\}$ [77, 78] and the following proposition:

Proposition 2 *Define* $\mathbf{M}$ *and* $\mathbf{N} \in \mathbb{S}^q$*, then* $\mathbf{M} \succeq \mathbf{N}$ *iff* $P(\mathbf{N}) \subseteq P(\mathbf{M})$.

Proof. *If* $P(\mathbf{N}) \subseteq P(\mathbf{M})$*, then* $\mathbf{N} \in P(\mathbf{M})$ *and then, by the definition of penumbra,* $\mathbf{M} \succeq \mathbf{N}$*. To prove the other implication, we define a matrix* $\mathbf{G} \in \mathbb{S}^q$ *such that* $\mathbf{N} \succeq \mathbf{G}$*. Then if* $\mathbf{M} \succeq \mathbf{N}$ *we have, by the transitivity property of the Loewner order,* $\mathbf{M} \succeq \mathbf{G}$*, namely,* $\mathbf{M} \succeq \mathbf{N} \succeq \mathbf{G}$*. Therefore, all the matrix elements in* $P(\mathbf{N})$ *are also in* $P(\mathbf{M})$*, thus,* $P(\mathbf{N}) \subseteq P(\mathbf{M})$*.* ■

The penumbra $P(\mathbf{M})$ is seen as an inverted cone of vertex $\mathbf{M}$ characterizing all matrices that are smaller than $\mathbf{M}$ [77, 78]. The authors in [77, 78] redefined the supremum of a set of matrices as the matrix associated to the vertex of the minimal penumbra covering the penumbras of all the matrices in the set. The minimal-penumbra vertex is a minimal-upper bound to the set with respect to the partially order set $(\mathbb{S}^q, \succeq)$. In [77], the minimal-penumbra vertex is computed by associating with each matrix $\mathbf{M} \in \mathbb{S}^q$ a ball in the subspace $\mathbb{S}_{\mathbf{A}} = \{\mathbf{A} : Tr\{\mathbf{A}\} = 0\}$, and the authors show that it is determined by the smallest ball enclosing the set of balls associated to each matrix in the set. The latter algorithm is implemented in an approximate manner, by solving instead the problem of finding the smallest enclosing ball of a set of points which correspond to samples from the boundaries of each ball. The success of this method to obtain a minimal-upper bound matrix depends on the samples chosen. For example, in the case of having two balls, it is easy to show that the smallest enclosing ball is tangent to each ball border at the two farthest points from the set of points defined by the intersection of a line passing through each ball center and each ball boundary. Therefore if the sampling procedure does not include this pair of points, then the resulting ball does not completely enclose both balls and, thus, the resulting matrix is not a minimal-upper bound. Moreover, when the dimension is larger than two, a simple analytical computation shows that this algorithm fails to obtain a minimal-upper bound matrix for the set formed by two diagonal matrices not comparable to each other according to Loewner order.

Here, instead, we propose a method for computing a suitable $\mathbf{B}_q$ for any dimension. First, we show that computing $\mathbf{B}_q$ is equivalent to finding the minimum-volume hyper-ellipsoid covering the set of hyper-ellipsoids associated to each matrix in the set C. And second, we show that this problem can be written as a convex objective

function with convex constraints which can be solved efficiently using semidefinite programming. An hyper-ellipsoid $\varepsilon \subset \mathbb{R}^q$ with non-empty interior and centered at the origin can be represented by the set $\varepsilon(\mathbf{F}) = \left\{\mathbf{x}^T\mathbf{F}^{-1}\mathbf{x} \leq \mathbf{1}\right\}$, where $\mathbf{F} \in \mathbb{S}^q_{++}$. Suppose $\varepsilon\left(\widetilde{\mathbf{F}}\right)$ is another hyper-ellipsoid similarly represented, where $\widetilde{\mathbf{F}} \in \mathbb{S}^q_{++}$. Then, the following statement holds:

Lemma 3 $\mathbf{F} \succeq \widetilde{\mathbf{F}}$ *iff* $\varepsilon(\mathbf{F}) \supseteq \varepsilon\left(\widetilde{\mathbf{F}}\right)$. ***Proof.*** *By the S-procedure [86], we have that* $\varepsilon\left(\widetilde{\mathbf{F}}\right) \subseteq \varepsilon(\mathbf{F})$ *if and only if there is a* $\lambda > 0$ *such that*

$$\begin{bmatrix} \mathbf{F}^{-1} & 0 \\ 0 & -1 \end{bmatrix} \preceq \lambda \begin{bmatrix} \widetilde{\mathbf{F}}^{-1} & 0 \\ 0 & -1 \end{bmatrix},$$

with equality when $\lambda = 1$, *implying the necessary condition* $\widetilde{\mathbf{F}} \preceq \mathbf{F}$. ■

Given a finite set of hyper-ellipsoids $C_\varepsilon := \{\varepsilon(\mathbf{F}_i) \mid \mathbf{F}_i \in \mathbb{S}^q_{++}, i = 1, \ldots, R\}$, we can always find a unique minimum volume hyper-ellipsoid, $\varepsilon(\mathbf{F}_{\text{jl}})$, containing the set C_ε, *i.e*, containing all $\varepsilon(\mathbf{F}_i)$ [86]. Since C_ε is convex, $\varepsilon(\mathbf{F}_{\text{jl}})$ is known as the Lowner-John ellipsoid of C_ε [86] and, as we show in the following statement, $\mathbf{F}_{\text{jl}}$ is a minimal-upper bound of the set $C_F := \{\mathbf{F}_i, i = 1, \ldots, R\}$ formed by all the matrices associated to the hyper-ellipsoids in C_ε.

Theorem 4 *The matrix* $\mathbf{F}_{\text{jl}}$, *associated to the Lowner-John ellipsoid of the set* C_ε, *is a minimal-upper bound of the set* C_F w.r.t *to the Loewner partial ordering.* ***Proof.*** *We will demonstrate this by contradiction. From the previous Lemma we have that* $\mathbf{F}_{\text{jl}} \succeq \mathbf{F}_i$, $i = 1, \ldots, R$. *Assume that there exists a matrix* $\mathbf{F}_{\text{o}} \notin C_F$ *such that* $\mathbf{F}_{\text{jl}} \succeq \mathbf{F}_{\text{o}} \succeq \mathbf{F}_i$, *therefore* $\varepsilon(\mathbf{F}_{\text{o}}) \supseteq \varepsilon(\mathbf{F}_i)$, *for* $i = 1, \ldots, R$, *and thus* $\varepsilon(\mathbf{F}_{\text{o}}) \supseteq \bigcup_{i=1}^{P} \varepsilon(\mathbf{F}_i)$. *Given that the volume of* $\varepsilon(\mathbf{F}_{\text{jl}})$ *is less than the volume of* $\varepsilon(\mathbf{F}_{\text{o}})$,

since it is the minimum volume hyper-ellipsoid enclosing all $\mathbf{F}_i$*, then* $|\mathbf{F}_{\text{jl}}| \leq |\mathbf{F}_{\text{o}}|$*, but by construction* $\mathbf{F}_{\text{jl}} \succeq \mathbf{F}_{\text{o}}$*, thus* $|\mathbf{F}_{\text{jl}}| \geq |\mathbf{F}_{\text{o}}|$*, which is a contradiction. Thus* $\mathbf{F}_{\text{o}} = \mathbf{F}_{\text{jl}}$*, and* $\mathbf{F}_{\text{jl}}$ *is a minimal-upper bound of the set* C_F*.* ■

Therefore, computing a minimal-upper bound matrix $\mathbf{B}_q$ of the set C, defined by

$$C := \{\mathbf{B}\mathbf{B}_q(\mathbf{t}, \mathbf{D}(\boldsymbol{\alpha})),\ \boldsymbol{\alpha} \in\ S\} \subset \mathbb{S}^q_{++}, \tag{4.36}$$

is equivalent to finding the Lowner-John ellipsoid of the set of hyper-ellipsoids associated to C. This is a particular case of a more general problem of computing the minimum volume hyper-ellipsoid $\varepsilon(\mathbf{B}) = \left\{\mathbf{x}^T\mathbf{B}^{-1}\mathbf{x} + 2\left(\mathbf{B}^{-1/2}\ \mathbf{b}\right)^T\mathbf{x} + \mathbf{b}^T\mathbf{b} \leq \mathbf{1}\right\}$ which covers the union of a set of non centered hyper-ellipsoids parameterized by the quadratic inequalities $\varepsilon_i(\mathbf{B}_i) = \left\{\mathbf{x}^T\mathbf{B}_i^{-1}\mathbf{x} + 2\mathbf{b}_i^T\mathbf{x} + c_i \leq 0\right\}$ for $i = 1, \ldots, m$. This problem can be posed as follows [86]:

$$\max_{\{\mathbf{B},\ \mathbf{b}\}} \left\{\log\left(\det\left(\mathbf{B}^{1/2}\right)\right)\right\} \tag{4.37}$$

subject to :

$$\tau_1 \geq 0,\ \tau_2 \geq 0, \ldots, \tau_m \geq 0,$$

$$\begin{bmatrix} \mathbf{B}^{-1} - \tau_i\mathbf{B}_i^{-1} & \mathbf{B}^{-1/2}\ \mathbf{b} - \tau_i\mathbf{b}_i \\ \left(\mathbf{B}^{-1/2}\ \mathbf{b} - \tau_i\mathbf{b}_i\right)^T & \mathbf{b}^T\mathbf{b} - 1 - \tau_i c_i \end{bmatrix} \preceq 0,$$

$$i = 1, \ldots, m.$$

The objective function and the set of constraints are convex, so it can be solved efficiently using semidefinite programming. In particular, we solve this problem using `CVX`, a package for specifying and solving convex programs [87, 88], for $\mathbf{B}_i = \mathbf{B}\mathbf{B}_q(\mathbf{t}, \mathbf{D}(\boldsymbol{\alpha}_i))$ for $\boldsymbol{\alpha}_i \in S$, $\mathbf{b}_i = \mathbf{b} = \mathbf{0}$, and $c_i = 1$. Therefore, the minimal-upper

bound $\mathbf{B}_q$ of the set C is given by

$$\mathbf{B}_q = \mathbf{B}_*, \tag{4.38}$$

where $\mathbf{B}_*$ is the optimal solution of (4.37.)

Using the following statement, we can even reduce the number of constraints in the above problem by considering only the set $C_{\mathrm{m}} \subseteq C$ formed by all the maximal elements of C.

Theorem 5 *Define C_{Fm} as the subset of C_F formed by all the maximal elements of C_F. Then, the Lowner-John ellipsoid $\varepsilon(\mathbf{F}_{\mathrm{jl}})$ of C_ε is also the Lowner-John ellipsoid of the set $C_{\varepsilon m}$ formed by the hyper-ellipsoids associated to the matrices in C_{Fm}.* ***Proof.*** *Since C_{Fm} is formed by all the maximal elements of C_F, then for $\mathbf{F}_i \in C_{Fm}$ and any $\mathbf{F}_j \in C_{Fc} = C_F - C_{Fm}$, we have that $\mathbf{F}_i \succeq \mathbf{F}_j$. From* Lemma 2, *$\varepsilon(\mathbf{F}_i) \supseteq \{\varepsilon(\mathbf{F}_j),$ for all $\mathbf{F}_j \in C_{Fc}\}$, which is true for all $\mathbf{F}_i \in C_{Fm}$, i.e., for all $\varepsilon(\mathbf{F}_i) \in C_{\varepsilon m}$, thus $C_{\varepsilon m} \supseteq \{\varepsilon(\mathbf{F}_j),$ for all $\mathbf{F}_j \in C_{Fc}\}$ and $C_\varepsilon = C_{\varepsilon m} \cup \{\varepsilon(\mathbf{F}_j),$ for all $\mathbf{F}_j \in C_{Fc}\} = C_{\varepsilon m}$. Therefore, $\varepsilon(\mathbf{F}_{\mathrm{jl}})$ is the Lowner-John ellipsoid for the set C_ε and $C_{\varepsilon m}$.* ■

In Figure 4.2 we illustrate the Lowner-John ellipsoid of a set formed by three ellipsoids in which two are maximal elements of the set.

Hence, using the above result we decrease the number of constraints in (4.37) by performing a pre-step which identifies the set C_{m}. Note that if C has a greatest element, it is the unique maximal element of C and therefore it is the supremum of the set and its associated hyper-ellipsoid is the Lowner-John ellipsoid of the set of hyper-ellipsoids associated to C. Therefore, there is no need to solve problem (4.37). Our algorithm searches and removes from the set of constraints the matrices whose

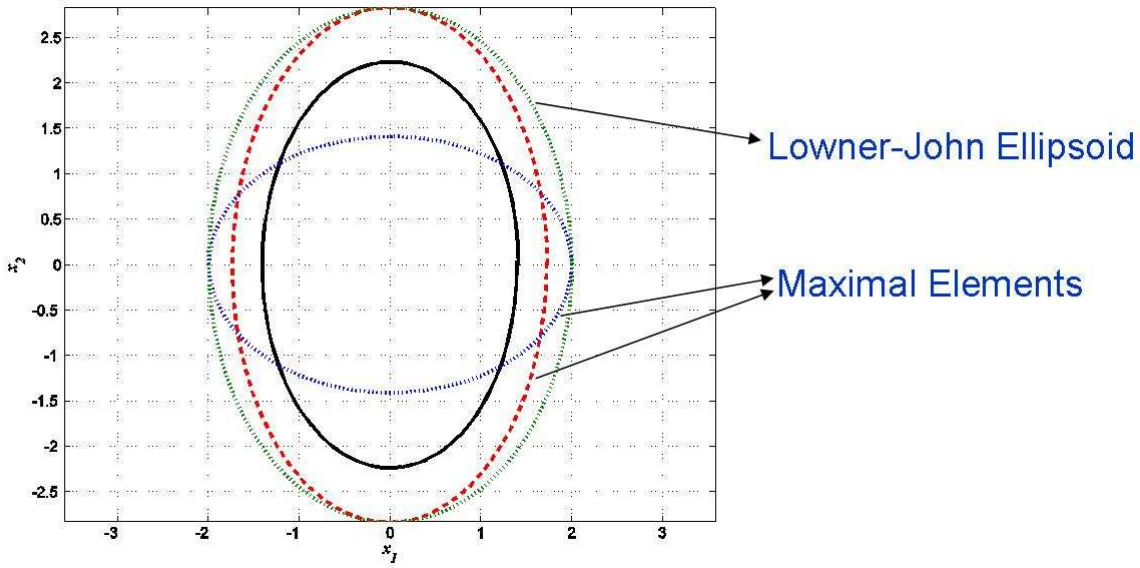


Figure 4.2: Illustration of Lowner-John ellipsoid of a set formed by three ellipsoids in which two are maximal elements of the set.

hyper-ellipsoid is fully enclosed by other hyper-ellipsoids. In particular, we evaluate in an iterative manner the membership in C_{m} of all elements in C. We define a membership indicator vector $\mathbf{i}_{C_{\mathrm{m}}}$ where $[\mathbf{i}_{C_{\mathrm{m}}}]_i = I_{C_{\mathrm{m}}}(\mathbf{F}_i)$ and the algorithm begins by assuming that all elements belong to C_{m}, namely, $\mathbf{i}_{C_{\mathrm{m}}} = \mathbf{1}_{R\times 1}$, where $R = |C|$. Then, all the values of the elements of $\mathbf{i}_{C_{\mathrm{m}}}$ are evaluated using the following iterative procedure:

- *Step 0*: Initialize $\mathbf{i}_{C_{\mathrm{m}}} = \mathbf{1}_{R\times 1}$ and set indexes $k = 1$, $l = 1$.
- *Step 1*: Evaluate membership of $\mathbf{F}_k$ to C_{m} (if $k > R$, terminate the algorithm):

$$\text{If } I_{C_{\mathrm{m}}}(\mathbf{F}_k) = \begin{cases} 0, \text{ set } k = k+1 \text{ and restart } \textit{Step 1}, \\ 1, \text{ set } l = l+1 \text{ and go to } \textit{Step 2}. \end{cases}$$

- *Step 2*: Evaluate membership of $\mathbf{F}_l$ to C_{m} (if $l > R$, set $k = k+1$, $l = 1$, and go to *Step 1*):

$$\text{If } I_{C_{\mathrm{m}}}(\mathbf{F}_l) = \begin{cases} 0, \text{ set } l = l+1 \text{ and restart } \textit{Step 2}, \\ 1, \text{ go to } \textit{Step 3}. \end{cases}$$

- *Step 3*: Compare $\mathbf{F}_k$ versus $\mathbf{F}_l$ *w.r.t.* the Loewner ordering:

$$\begin{cases} \text{if } \mathbf{F}_k \succeq \mathbf{F}_l, \text{ set } I_{C_{\mathrm{m}}}(\mathbf{F}_l) = 0,\ l = l+1, \text{ and go to } \textit{Step 2}, \\ \text{if } \mathbf{F}_l \succeq \mathbf{F}_k, \text{ set } I_{C_{\mathrm{m}}}(\mathbf{F}_k) = 0,\ k = k+1,\ l = 1, \text{ and go to } \textit{Step 1}, \\ \text{if } \textit{not comparable}, \text{ set } I_{C_{\mathrm{m}}}(\mathbf{F}_l) = 1,\ l = l+1, \text{ and go to } \textit{Step 2}. \end{cases}$$

Finally, once the algorithm terminates, the set C_{m} will be given by all elements such that $I_{C_{\mathrm{m}}}(\mathbf{F}_i) = 1$. To compare $\mathbf{F}_k$ versus $\mathbf{F}_l$, *w.r.t.* to the Loewner ordering, we apply the determinant test [89] to the matrix, $\mathbf{G} = \mathbf{F}_k - \mathbf{F}_l$. This test evaluates the principal minors of $\mathbf{G}$ and concludes on the matrix definiteness as follows: (i) $\mathbf{G}$ is positive definite, *i.e.*, $\mathbf{F}_k \succ \mathbf{F}_l$, if and only if all its leading principal minors are strictly positive, and it is negative definite, *i.e.*, $\mathbf{F}_l \succ \mathbf{F}_k$, if its k-th order leading principal minor is < 0 for k odd and > 0 for k even; (ii) $\mathbf{G}$ is positive semidefinite, *i.e.*, $\mathbf{F}_k \succeq \mathbf{F}_l$, if and only if all the principal minors are non-negative, and it is negative semidefinite, *i.e.*, $\mathbf{F}_l \succeq \mathbf{F}_k$, if all the k-th order principal minors are ≤ 0 for k odd and ≥ 0 for k even; (iii) $\mathbf{G}$ is indefinite, *i.e.*, $\mathbf{F}_k$ and $\mathbf{F}_l$ are not comparable, if none of the previous conditions are satisfied. Since all the matrices in the set C are block diagonal and the maximum size of one block is 2×2, then every matrix $\mathbf{G}$ is a symmetric tridiagonal matrix, of which the leading principal minors $\{f_{\mathbf{G}}(r),\ r = 1, \ldots, q\}$ can be computed iteratively as follows [90]:

$$f_{\mathbf{G}}(r) = \begin{cases} 1, \text{ for } r = 0, \\ [\mathbf{G}]_{11}, \text{ for } r = 1, \\ [\mathbf{G}]_{r\,r}\ f_{\mathbf{G}}(r-1) - \left([\mathbf{G}]_{r\,r-1}\right)^2 f_{\mathbf{G}}(r-2), \text{ for } 2 < r < q. \end{cases}$$

Note that the determinant of the tridiagonal matrix $\mathbf{G}$ is given by $|\mathbf{G}| = f_{\mathbf{G}}(q)$, and since all the principal minors of $\mathbf{G}$ are also tridiagonal matrices, then their values are computed efficiently using the above expression.

Following the ideas of [77], the issue of having a unique supremum of a set of positive definite matrices can be overcome by redefining the supremum as the matrix associated to the Lowner-John Ellipsoid of the set of hyperellipsoids associated to the maximal elements of the set C formed by the P-order BB matrices. This matrix $\mathbf{B}_q$ is unique in the sense that there is no other ellipsoid with minimal volume covering the hyper-ellipsoids associated to the set of maximal elements of C. It also has the properties of continuity, namely, it is positive definite.

In the following section we will derive the elements of the Barankin information matrix for the problem of changes in the parameters of Gaussian and Poisson distributions.

4.4 Change in parameters of Gaussian and Poisson distributions

In this section, we apply the proposed bound for two distributions generally encountered in signal processing. We analyze these two cases in a very general way, which means that the results presented here can be applied to a wide variety of estimation problems. Indeed, the parameters involved in the Gaussian distribution (mean and covariance) and in the Poisson distribution are assumed to be a function of the parameters $\boldsymbol{\eta}_j$, which generally represent physical parameters of interest in signal processing. An example of change of parameters in a Gaussian distribution in the radar context is direction-of-arrival (DOA) estimation. The varying cross-section fluctuations are

modeled with a Swerling 0 model [91], where the DOAs are hidden in the mean of the observations, leading, for example, to the so-called conditional MLE [92]. On the other hand, when the emitted signals are modeled with a Swerling 1-2, the DOAs are hidden in the covariance of the observations, leading, for example, to the so-called unconditional MLE [93]. In the context of particle detection, the Poisson distribution is generally used to model the particle counting process; *i.e.*, the observations and the parameter involved in the Poisson distribution become a function of the DOA [94].

4.4.1 Gaussian case

Let us assume that the vector of observations $\mathbf{x}_i \in \mathbb{R}^M$, for $i = 1, \ldots, N$, is modeled as follows:

$$\mathbf{x}_i = \mathbf{f}(\boldsymbol{\nu}_j) + \mathbf{n}_i, \tag{4.39}$$

where $\mathbf{f}(\cdot)$ is a vector of known functions; $\boldsymbol{\nu}_j \in \mathbb{R}^F$ is a known parameter vector; $\mathbf{n}_i$ is a zero-mean Gaussian random vector with covariance matrix $\mathbf{M}(\boldsymbol{\varphi}_j)$, with $\mathbf{M}(\cdot)$ a symmetric positive definite matrix of known functions; and $\boldsymbol{\varphi}_j \in \mathbb{R}^G$ is a known parameter vector. Then $\boldsymbol{\eta}_j = \left[\boldsymbol{\nu}_j^T,\ \boldsymbol{\varphi}_j^T\right]^T \in \mathbb{R}^L$, with $L = F + G$, and $\mathbf{x}_i$ are distributed as $\mathcal{N}\left(\mathbf{f}(\boldsymbol{\nu}_j), \mathbf{M}(\boldsymbol{\varphi}_j)\right)$. Here we are interested in deriving the elements of the Barankin information matrix for changes in the pdf parameters of $\mathbf{x}_i$, *i.e.*, the mean and covariance matrix. First, we analyze the general case of piecewise changes of mean and covariance. Second, we deduce two particular cases: i) piecewise changes of the mean and constant covariance matrix, *i.e.*, $\mathbf{M}(\boldsymbol{\varphi}_j) = \mathbf{M}(\boldsymbol{\varphi}) = \boldsymbol{\Sigma}$; and ii) piecewise changes of covariance and constant mean vector, *i.e.*, $\mathbf{f}(\boldsymbol{\nu}_j) = \mathbf{f}(\boldsymbol{\nu}) = \boldsymbol{\mu}$. Note that we restrict our analysis to the set of parameter vectors $\{\boldsymbol{\nu}_j\}$ and $\{\boldsymbol{\varphi}_j\}$ such that the functions in $\mathbf{f}(\boldsymbol{\nu}_j)$ and $\mathbf{M}(\boldsymbol{\varphi}_j)$ are injective. In other words, a change

in the values of $\boldsymbol{\nu}_j$ changes the values of $\mathbf{f}(\boldsymbol{\nu}_j)$, the mean of the distribution of $\mathbf{x}_i$. Similarly, a change in the values of $\boldsymbol{\varphi}_j$ implies a change in values of the covariance matrix $\mathbf{M}\left(\boldsymbol{\varphi}_j\right)$. Below, we compute the elements of the Barankin information matrix $\mathbf{BIM_t}$. Then, for each case, respectively, we derive closed-form expressions for the elements $\boldsymbol{\Phi} - \mathbf{1}_{q\times q}$ (see Appendix D.0.3 for details on their derivation) which are different from zero; namely, we evaluate

$$\begin{cases} [\boldsymbol{\Phi}]_{kk} \text{ for } \begin{cases} \alpha_k > 0, \\ \alpha_k < 0, \end{cases} \\ [\boldsymbol{\Phi}]_{kk+1} \text{ for } t_k + \alpha_k > t_{k+1} + \alpha_{k+1}. \end{cases} \tag{4.40}$$

Piecewise changes of mean and covariance matrix

For $\alpha_k > 0$, using Equation (4.15), we have that $[\boldsymbol{\Phi}]_{kk}$ is given by

$$[\boldsymbol{\Phi}]_{kk} = \begin{cases} \left(\frac{\left|\mathbf{M}\left(\boldsymbol{\varphi}_{k+1}\right)\right|^{1/2}\left|\mathbf{M}_k^{-1}\right|^{1/2}}{|\mathbf{M}(\boldsymbol{\varphi}_k)|}\right)^{\alpha_k} \\ \times \exp\left\{\frac{\alpha_k}{2}\mathbf{g}_k^T\mathbf{M}_k^{-1}\mathbf{g}_k\right\} \\ \times \exp\left\{-\alpha_k\ \mathbf{f}^T(\boldsymbol{\nu}_k)\left(\mathbf{M}\left(\boldsymbol{\varphi}_k\right)\right)^{-1}\mathbf{f}(\boldsymbol{\nu}_k)\right\} \\ \times \exp\left\{\frac{\alpha_k}{2}\mathbf{f}^T(\boldsymbol{\nu}_{k+1})\left(\mathbf{M}\left(\boldsymbol{\varphi}_{k+1}\right)\right)^{-1}\mathbf{f}(\boldsymbol{\nu}_{k+1})\right\}, \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{for } \mathbf{M}_k \in \mathbb{S}_{++}^M, \\ \infty, \text{ otherwise,} \end{cases} \tag{4.41}$$

where $\mathbf{M}_k = \left(2\left(\mathbf{M}\left(\boldsymbol{\varphi}_k\right)\right)^{-1} - \left(\mathbf{M}\left(\boldsymbol{\varphi}_{k+1}\right)\right)^{-1}\right)$
and $\mathbf{g}_k = 2\left(\mathbf{M}\left(\boldsymbol{\varphi}_k\right)\right)^{-1}\mathbf{f}(\boldsymbol{\nu}_k) - \left(\mathbf{M}\left(\boldsymbol{\varphi}_{k+1}\right)\right)^{-1}\mathbf{f}(\boldsymbol{\nu}_{k+1})$.

For $\alpha_k < 0$, using Equation (4.15), we have that $[\mathbf{\Phi}]_{kk}$ is given by

$$[\mathbf{\Phi}]_{kk} = \begin{cases} \left(\frac{|\mathbf{M}(\boldsymbol{\varphi}_k)|^{1/2} \left|\mathbf{M}_{k+1}^{-1}\right|^{1/2}}{\left|\mathbf{M}\left(\boldsymbol{\varphi}_{k+1}\right)\right|} \right)^{-\alpha_k} \\ \times \exp\left\{ \frac{-\alpha_k}{2} \mathbf{g}_{k+1}^T \mathbf{M}_{k+1}^{-1} \mathbf{g}_{k+1} \right\} \\ \times \exp\left\{ \alpha_k \ \mathbf{f}^T(\boldsymbol{\nu}_{k+1}) \left(\mathbf{M}\left(\boldsymbol{\varphi}_{k+1}\right)\right)^{-1} \mathbf{f}(\boldsymbol{\nu}_{k+1}) \right\} \\ \times \exp\left\{ \frac{-\alpha_k}{2} \mathbf{f}^T(\boldsymbol{\nu}_k) \left(\mathbf{M}\left(\boldsymbol{\varphi}_k\right)\right)^{-1} \mathbf{f}(\boldsymbol{\nu}_k) \right\}, \\ \qquad\qquad \text{for } \mathbf{M}_{k+1} \in \mathbb{S}_{++}^M, \\ \\ \infty, \text{ otherwise,} \end{cases} \tag{4.42}$$

where $\mathbf{M}_{k+1} = 2\left(\mathbf{M}\left(\boldsymbol{\varphi}_{k+1}\right)\right)^{-1} - \left(\mathbf{M}\left(\boldsymbol{\varphi}_k\right)\right)^{-1}$
and $\mathbf{g}_{k+1} = 2\left(\mathbf{M}\left(\boldsymbol{\varphi}_{k+1}\right)\right)^{-1} \mathbf{f}(\boldsymbol{\nu}_{k+1}) - \left(\mathbf{M}\left(\boldsymbol{\varphi}_k\right)\right)^{-1} \mathbf{f}(\boldsymbol{\nu}_k)$.

For $t_k + \alpha_k > t_{k+1} + \alpha_{k+1}$, using Equation (4.19), we have that $[\mathbf{\Phi}]_{kk+1}$is given as follows:

$$[\mathbf{\Phi}]_{kk+1} = \begin{cases} \left(\frac{\left|\mathbf{M}\left(\boldsymbol{\varphi}_{k+1}\right)\right|^{1/2} |\overline{\mathbf{M}}_k^{-1}|^{1/2}}{|\mathbf{M}(\boldsymbol{\varphi}_k)|^{1/2} \left|\mathbf{M}\left(\boldsymbol{\varphi}_{k+2}\right)\right|^{1/2}} \right)^{\beta_k} \\ \times \exp\left\{ \frac{\beta_k}{2} \overline{\mathbf{g}}_k^T \overline{\mathbf{M}}_k^{-1} \overline{\mathbf{g}}_k^T \right\} \\ \times \exp\left\{ -\frac{\beta_k}{2} \mathbf{f}^T(\boldsymbol{\nu}_k) \left(\mathbf{M}\left(\boldsymbol{\varphi}_k\right)\right)^{-1} \mathbf{f}(\boldsymbol{\nu}_k) \right\} \\ \times \exp\left\{ -\frac{\beta_k}{2} \mathbf{f}^T(\boldsymbol{\nu}_{k+2}) \left(\mathbf{M}\left(\boldsymbol{\varphi}_{k+2}\right)\right)^{-1} \mathbf{f}(\boldsymbol{\nu}_{k+2}) \right\} \\ \times \exp\left\{ \frac{\beta_k}{2} \mathbf{f}^T(\boldsymbol{\nu}_{k+1}) \left(\mathbf{M}\left(\boldsymbol{\varphi}_{k+1}\right)\right)^{-1} \mathbf{f}(\boldsymbol{\nu}_{k+1}) \right\}, \\ \qquad\qquad \text{for } \overline{\mathbf{M}}_k^{-1} \in \mathbb{S}_{++}^M, \\ \\ \infty, \text{ otherwise,} \end{cases} \tag{4.43}$$

where $\overline{\mathbf{M}}_k = \left(\mathbf{M}\left(\boldsymbol{\varphi}_k\right)\right)^{-1} + \left(\mathbf{M}\left(\boldsymbol{\varphi}_{k+2}\right)\right)^{-1} - \left(\mathbf{M}\left(\boldsymbol{\varphi}_{k+1}\right)\right)^{-1}$,
and $\overline{\mathbf{g}}_k = \left(\mathbf{M}\left(\boldsymbol{\varphi}_k\right)\right)^{-1} \mathbf{f}(\boldsymbol{\nu}_k) + \left(\mathbf{M}\left(\boldsymbol{\varphi}_{k+2}\right)\right)^{-1} \mathbf{f}(\boldsymbol{\nu}_{k+2}) - \left(\mathbf{M}\left(\boldsymbol{\varphi}_{k+1}\right)\right)^{-1} \mathbf{f}(\boldsymbol{\nu}_{k+1})$.

Piecewise changes of mean and constant covariance matrix

In this case $\mathbf{M}\left(\boldsymbol{\varphi}_j\right) = \mathbf{M}\left(\boldsymbol{\varphi}\right) = \boldsymbol{\Sigma}$, $\boldsymbol{\eta}_j = \left[\boldsymbol{\nu}_j^T,\ \boldsymbol{\varphi}^T\right]^T$, and $[\boldsymbol{\Phi}]_{kl}$ is given as follows:

For $\alpha_k > 0$, using Equation (4.41) and replacing $\mathbf{M}\left(\boldsymbol{\varphi}_k\right)$ and $\mathbf{M}\left(\boldsymbol{\varphi}_{k+1}\right)$ by $\boldsymbol{\Sigma}$, we have straightforwardly for $[\boldsymbol{\Phi}]_{kk}$:

$$[\boldsymbol{\Phi}]_{kk} \;=\; \exp\left\{\alpha_k \left(\mathbf{f}(\boldsymbol{\nu}_k) - \mathbf{f}(\boldsymbol{\nu}_{k+1})\right)^T \boldsymbol{\Sigma}^{-1} \left(\mathbf{f}(\boldsymbol{\nu}_k) - \mathbf{f}(\boldsymbol{\nu}_{k+1})\right)\right\}. \qquad (4.44)$$

For $\alpha_k < 0$, using Equation (4.42), $[\boldsymbol{\Phi}]_{kk}$ is given as follows:

$$[\boldsymbol{\Phi}]_{kk} \;=\; \exp\left\{-\alpha_k \left(\mathbf{f}(\boldsymbol{\nu}_{k+1}) - \mathbf{f}(\boldsymbol{\nu}_k)\right)^T \boldsymbol{\Sigma}^{-1} \left(\mathbf{f}(\boldsymbol{\nu}_{k+1}) - \mathbf{f}(\boldsymbol{\nu}_k)\right)\right\}. \qquad (4.45)$$

For $t_k + \alpha_k > t_{k+1} + \alpha_{k+1}$, using Equation (4.43), then$[\boldsymbol{\Phi}]_{kk+1}$is given as follows:

$$\begin{aligned}[\boldsymbol{\Phi}]_{kk+1} \;=\; \Bigg(\exp\Bigg\{\frac{\beta_k}{2}\Big(&\left(\mathbf{f}(\boldsymbol{\nu}_{k+1}) - \mathbf{f}(\boldsymbol{\nu}_k)\right)\boldsymbol{\Sigma}^{-1}\left(\mathbf{f}(\boldsymbol{\nu}_{k+1}) - \mathbf{f}(\boldsymbol{\nu}_k)\right)^T \\ &+ \left(\mathbf{f}(\boldsymbol{\nu}_{k+2}) - \mathbf{f}(\boldsymbol{\nu}_{k+1})\right)\boldsymbol{\Sigma}^{-1}\left(\mathbf{f}(\boldsymbol{\nu}_{k+2}) - \mathbf{f}(\boldsymbol{\nu}_{k+1})\right)^T \\ &- \left(\mathbf{f}(\boldsymbol{\nu}_k) - \mathbf{f}(\boldsymbol{\nu}_{k+2})\right)\boldsymbol{\Sigma}^{-1}\left(\mathbf{f}(\boldsymbol{\nu}_k) - \mathbf{f}(\boldsymbol{\nu}_{k+2})\right)^T\Big)\Bigg\}\Bigg). \qquad (4.46)\end{aligned}$$

Piecewise changes of covariance matrix and constant mean vector

In this case $\mathbf{f}(\boldsymbol{\nu}_j) = \mathbf{f}(\boldsymbol{\nu}) = \boldsymbol{\mu}$, $\boldsymbol{\eta}_j = \left[\boldsymbol{\nu}^T,\ \boldsymbol{\varphi}_j^T\right]^T$, and $[\boldsymbol{\Phi}]_{kl}$ is given as follows:

For $\alpha_k > 0$ using Equation (4.41) and replacing $\mathbf{f}(\boldsymbol{\nu}_k)$ and $\mathbf{f}(\boldsymbol{\nu}_{k+1})$ by $\boldsymbol{\mu}$, we have straightforwardly for $[\boldsymbol{\Phi}]_{kk}$:

$$[\boldsymbol{\Phi}]_{kk} = \begin{cases} \left(\frac{\left|\mathbf{M}\left(\boldsymbol{\varphi}_{k+1}\right)\right|^{1/2}}{|\mathbf{M}(\boldsymbol{\varphi}_k)||\mathbf{M}_k|^{1/2}} \right)^{\alpha_k}, \text{ for } \mathbf{M}_k \in \mathbb{S}_{++}^M, \\ \infty, \text{ otherwise,} \end{cases} \tag{4.47}$$

where $\mathbf{M}_k = 2\left(\mathbf{M}\left(\boldsymbol{\varphi}_k\right)\right)^{-1} - \left(\mathbf{M}\left(\boldsymbol{\varphi}_{k+1}\right)\right)^{-1}$.

For $\alpha_k < 0$, using Equation (4.42), $[\boldsymbol{\Phi}]_{kk}$ is given as follows:

$$[\boldsymbol{\Phi}]_{kk} = \begin{cases} \left(\frac{|\mathbf{M}(\boldsymbol{\varphi}_k)|^{1/2}}{\left|\mathbf{M}\left(\boldsymbol{\varphi}_{k+1}\right)\right||\mathbf{M}_{k+1}|^{1/2}} \right)^{-\alpha_k}, \text{ for } \mathbf{M}_{k+1} \in \mathbb{S}_{++}^M, \\ \infty, \text{ otherwise,} \end{cases} \tag{4.48}$$

where $\mathbf{M}_{k+1} = 2\left(\mathbf{M}\left(\boldsymbol{\varphi}_{k+1}\right)\right)^{-1} - \left(\mathbf{M}\left(\boldsymbol{\varphi}_k\right)\right)^{-1}$.

For $t_k + \alpha_k > t_{k+1} + \alpha_{k+1}$, using Equation (4.43), then$[\boldsymbol{\Phi}]_{kk+1}$is given as follows:

$$[\boldsymbol{\Phi}]_{kk+1} = \begin{cases} \left(\frac{\left|\mathbf{M}\left(\boldsymbol{\varphi}_{k+1}\right)\right|^{1/2}}{|\mathbf{M}(\boldsymbol{\varphi}_k)|^{1/2}\left|\mathbf{M}\left(\boldsymbol{\varphi}_{k+2}\right)\right|^{1/2}\left|\overline{\mathbf{M}}_k\right|^{1/2}} \right)^{\beta_k}, \\ \qquad\qquad \text{for } \overline{\mathbf{M}}_k \in \mathbb{S}_{++}^M, \\ \infty, \text{ otherwise,} \end{cases} \tag{4.49}$$

where $\overline{\mathbf{M}}_k = \left(\mathbf{M}\left(\boldsymbol{\varphi}_k\right)\right)^{-1} + \left(\mathbf{M}\left(\boldsymbol{\varphi}_{k+2}\right)\right)^{-1} - \left(\mathbf{M}\left(\boldsymbol{\varphi}_{k+1}\right)\right)^{-1}$.

The elements of the Barankin bound for each case are obtained by using Equation (4.35), recalling that $A_k = [\boldsymbol{\Phi}]_{kk} - 1$ and $B_k = [\boldsymbol{\Phi}]_{kk+1} - 1$, from Equations (4.21) and (4.22), respectively.

4.4.2 Poisson case

Assume that the measurements $x_i \in \mathbb{N}+\{0\}$, for $i = 1, \ldots, N$, are distributed as a Poisson distribution with parameter $f(\boldsymbol{\eta}_j)$, where $f(\cdot)$ is a known function and $\boldsymbol{\eta}_j \in \mathbb{R}^L$ is a known parameter vector. Similarly to the Gaussian case, we restrict our analysis to the set of parameter vectors $\{\boldsymbol{\eta}_j\}$ such that the function $f(\boldsymbol{\eta}_j)$ is injective. Therefore, we derive closed-form expressions for the elements of the matrix $\boldsymbol{\Phi} - \mathbf{1}_{q\times q}$ for piecewise changes of the parameter $\boldsymbol{\eta}_j$. Below, we evaluate $[\boldsymbol{\Phi}]_{kk}$ for $\alpha_k > 0$ and $\alpha_k < 0$, and $[\boldsymbol{\Phi}]_{kk+1}$ for $t_k + \alpha_k > t_{k+1} + \alpha_{k+1}$. Note that since $x_i \in \mathbb{N}$ we replace the integral operator by the summation operator.

For $\alpha_k > 0$, $[\boldsymbol{\Phi}]_{kk}$ becomes

$$[\boldsymbol{\Phi}]_{kk} = \exp\left\{\frac{\alpha_k\left(f(\boldsymbol{\eta}_{k+1})-f(\boldsymbol{\eta}_k)\right)^2}{f(\boldsymbol{\eta}_{k+1})}\right\}, \tag{4.50}$$

For $\alpha_k < 0$, $[\boldsymbol{\Phi}]_{kk}$ becomes

$$[\boldsymbol{\Phi}]_{kk} = \exp\left\{\frac{-\alpha_k\left(f(\boldsymbol{\eta}_k)-f(\boldsymbol{\eta}_{k+1})\right)^2}{f(\boldsymbol{\eta}_k)}\right\}. \tag{4.51}$$

For $t_k + \alpha_k > t_{k+1} + \alpha_{k+1}$, $[\boldsymbol{\Phi}]_{kk+1}$is given as follows:

$$\begin{aligned}[\boldsymbol{\Phi}]_{kk+1} &= \exp\left\{\beta_k\frac{\left(f(\boldsymbol{\eta}_{k+1})-f(\boldsymbol{\eta}_k)\right)^2}{2f(\boldsymbol{\eta}_{k+1})}\right\} \\ &\quad\times\exp\left\{\beta_k\frac{\left(f(\boldsymbol{\eta}_{k+2})-f(\boldsymbol{\eta}_{k+1})\right)^2}{2f(\boldsymbol{\eta}_{k+1})}\right\} \\ &\quad\times\exp\left\{-\beta_k\frac{\left(f(\boldsymbol{\eta}_k)-f(\boldsymbol{\eta}_{k+2})\right)^2}{2f(\boldsymbol{\eta}_{k+1})}\right\}.\end{aligned} \tag{4.52}$$

Similarly, as in the Gaussian case, the elements of the Barankin bound for each case are obtained by using Equation (4.35) with $A_k = [\mathbf{\Phi}]_{kk} - 1$ and $B_k = [\mathbf{\Phi}]_{kk+1} - 1$.

4.5 Numerical examples

In this section, as an illustration, we compare the MSE between the true values of the change-point locations and their maximum likelihood estimations with our bounds. In particular, we first introduce the MLE of change-point locations assuming the total number of changes is known. Then we analyze the cases of multiple changes in (i) the mean of a Gaussian distribution with fixed variance, (ii) the variance of a Gaussian assuming a fixed mean, and (iii) the mean rate of a Poisson distribution.

4.5.1 Maximum likelihood estimation

The MLE of $\mathbf{t}$ is the solution to the following problem:

$$\hat{\mathbf{t}}_{\mathrm{ML}} = \arg\max_{\mathbf{t}} \sum_{i=1}^{q+1} \ln p_i(\mathbf{x}_{t_{i-1}+1,}, \ldots, \mathbf{x}_{t_i}; \boldsymbol{\eta}_i), \tag{4.53}$$

where $t_0 = 0$ and $t_{q+1} = N$ by definition. There is no known closed-form expression for $\hat{\mathbf{t}}_{\mathrm{ML}}$ so it has to be estimated via numerical computations. To solve this multidimensional optimization problem efficiently, we apply dynamic programming (DP), explained in detail in [95], in our context of change-point estimation. The main advantage of the DP approach is that it does not need to evaluate all the possible combinations of values for $\mathbf{t}$ in (4.53). The details on how to implement DP to solve (4.53) can be found in [95], Chapter 12. In all our examples below, we illustrate the

average MSE performance of the MLE for 1000 Monte Carlo experiments. We studied the performance as a function of the signal-to-noise ratio (SNR), which is defined accordingly in each example, and as a function of the distance between change points. Here we chose $q = 3$ and the number of samples $N = 80$. In each example below, we set $t_2 = 40$, and $t_3 = 60$, and we analyze two scenarios for change point t_1: In the first one, we set $t_1 = 20$ such that each segment has the same number of samples, and in the second scenario, $t_1 \in [2, 38]$. Note that the unbiasedness properties of the MLE have been studied in [96] for a single change-point and for multiple change-points in [97]. The asymptotic results derived in [96] and [97] are applicable only for the case of a Gaussian distribution with changes in the mean. However, in the case of having a finite interval the MLE is expected to be biased independently of the distribution. On the other hand, it seems reasonable to assume that for large SNR values the MLE is unbiased for a subset of the parameter space, i.e, subintervals, and specially for change-points located equidistant from their neighboring change-points or the interval limits. For example, in all the examples below, the bias of the MLE for $\mathbf{t} = [20\ 40\ 60]$ is approximately zero for all the SNR ranges considered in each scenario.

4.5.2 Changes in the mean of a Gaussian distribution

We consider the scenario of a time series with three change points in the mean values of a Gaussian distribution with common variance. We recall the closed-form expressions obtained for computing $[\mathbf{\Phi}]_{kk}$, namely, Equations (4.44) and (4.45), and define the SNR for the k^{th} change point as follows:

$$SNR_k = \left(\mathbf{f}(\boldsymbol{\nu}_{k+1}) - \mathbf{f}(\boldsymbol{\nu}_k)\right)^T \mathbf{\Sigma}^{-1} \left(\mathbf{f}(\boldsymbol{\nu}_{k+1}) - \mathbf{f}(\boldsymbol{\nu}_k)\right), \tag{4.54}$$

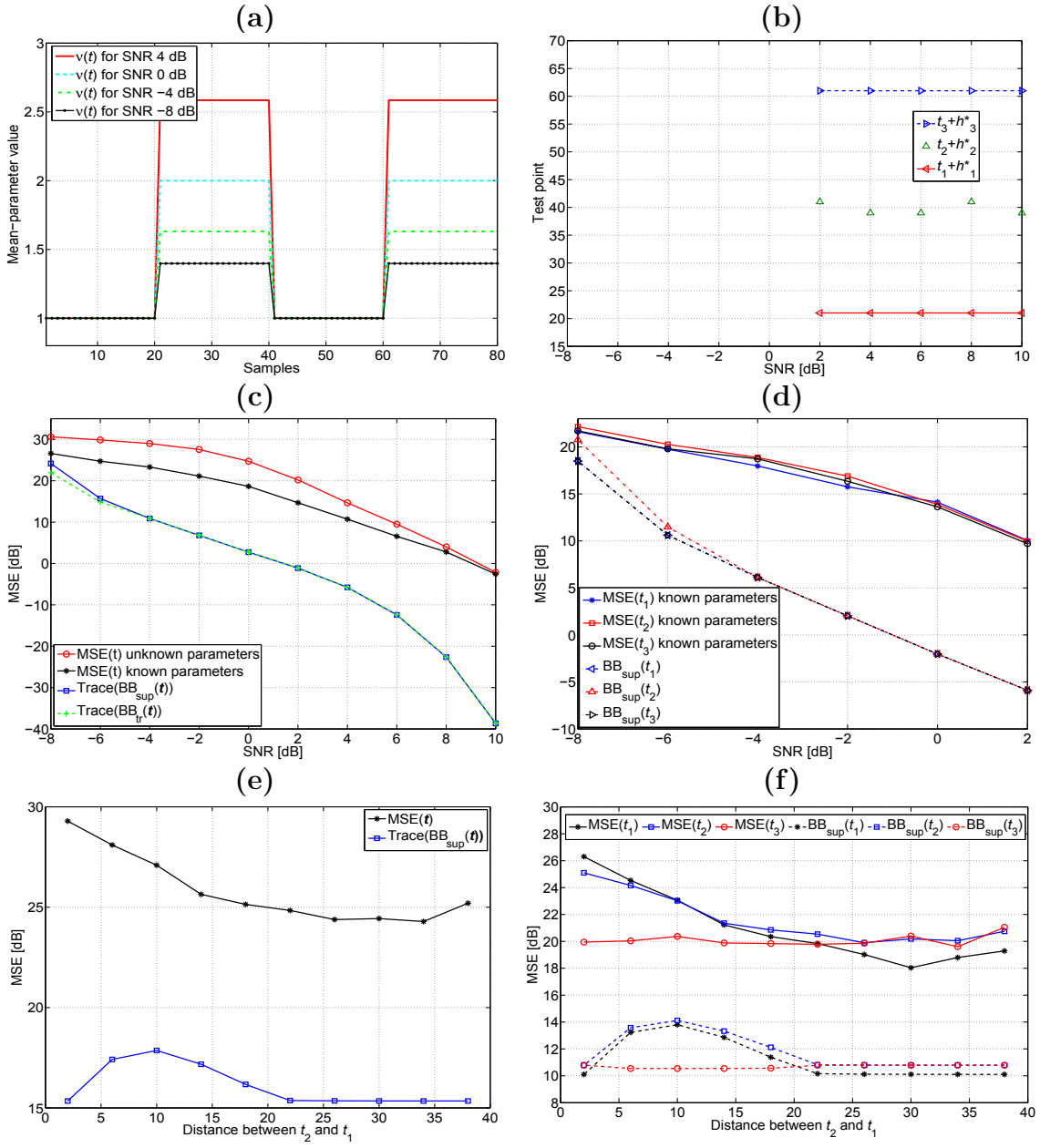


Figure 4.3: Performance analysis for estimating change-points of the mean in a Gaussian distribution: (a) Mean values as a function of sample time for different SNR values; (b) Test points associated with the BB given by the minimal-upper bound of C, $\mathbf{BB}_{\text{sup}}$, as a function of SNR; (c) MSE of the change-point vector using the MLE of $\mathbf{t}$ and its Barankin bound given by $\mathbf{BB}_{\text{sup}}$, and by the matrix with maximum trace in C, $\mathbf{BB}_{\text{tr}}$; (d) MSE of each change-point as a function of SNR using the MLE of t_1, t_2, and t_3 and their corresponding Barankin bound $\mathbf{BB}_{\text{sup}}(t_i)$, $i = 1, \ldots, 3$; (e) MSE of change-point vector using the MLE of $\mathbf{t}$ and its Barankin bound, $\mathbf{BB}_{\text{sup}}(\mathbf{t})$, as a function of the distance between t_2 and t_1 for SNR $= -6$ [dB]; (f) MSE of each change-point and their respective $\mathbf{BB}_{\text{sup}}$ as a function of the distance between t_2 and t_1 for SNR $= -6$ [dB]. ©[2010] IEEE.

where $\mathbf{f}(\boldsymbol{\nu}_k) \in \mathbb{R}^M$ is the mean vector of the k^{th} segment and $\boldsymbol{\Sigma} \in \mathbb{R}^{M\times M}$ is the common covariance matrix. In our example, $M = 1$ and, without loss of generality, we choose $f(\boldsymbol{\nu}_k) = \nu_k$ and $\Sigma = \sigma^2 = 1$, thus $\boldsymbol{\eta}_k = \left[\boldsymbol{\nu}_k^T,\ 1\right]^T$. Here, we set $\nu_1 = 1$, and ν_2, ν_3, and ν_4 are set such that $SNR_1 = SNR_2 = SNR_3 = SNR$. In particular, $\nu_k = \nu_{k-1} + (-1)^k \sqrt{\sigma^2 SNR}$ for $k = 2, 3, 4$. Figure 4.3(a) illustrates the mean values as a function of sample time for different SNR values. In Figure 4.3(c), we illustrate the MSE performance of the MLE for the change-point vector, and the BB as a function of the SNR. In particular, $\text{MSE}_{\text{known}}$ is the MSE performance of the MLE for the change-point vector, assuming knowledge of the means and variance. $\text{MSE}_{\text{unknown}}$ is the MSE performance of the MLE for a more realistic case when no knowledge of the distribution parameters is available. The $\mathbf{BB}_{\text{sup}}$ is given by the minimal-upper bound matrix $\mathbf{B}_q$ of the set C computed using the algorithm presented in Section III.D, and $\mathbf{BB}_{\text{tr}}$ is the matrix in C that has the maximum trace. We illustrate the trace of $\mathbf{BB}_{\text{sup}}$ and $\mathbf{BB}_{\text{tr}}$ since we are comparing the MSE performance for the change-point vector estimates. Note that, in view of the discussion presented in Section III.D, we compute $\mathbf{BB}_{\text{tr}}$ only in this example to show that $\mathbf{BB}_{\text{tr}}$ does not necessarily coincide with supremum of the set unless $\mathbf{BB}_{\text{sup}} \in C$. In this particular scenario, we found that $\mathbf{BB}_{\text{sup}}$ belongs to the set C for SNR values equal to and larger than 2 dB. Therefore, we have the optimal test points $\{\alpha_1^*, \alpha_2^*, \alpha_3^*\}$ associated to the matrix $\mathbf{BB}_{\text{sup}}$ defining the Lowner-John Ellipsoid , which are presented in Figure 4.3(b). For SNR values above 2 dB no change point is overlapped, therefore, each bound depends only on its corresponding diagonal element $[\boldsymbol{\Phi}]_{ii}$, which is equivalent to the resulting analysis of considering one change point located at $t = 20$, assuming $N = 40$. Moreover, it is important to mention that in this example, $[\boldsymbol{\Phi}]_{ii}$ is symmetric with respect to α_i, and since all segments have the same length, then both α_i and $-\alpha_i$ are optimal solutions for the bound on t_i. In Figure 4.3(b) we illustrate only one optimal solution.

When the SNR > 2 dB, we found the set C had several maximal elements that are not mutually comparable, thus, $\mathbf{BB}_{\text{sup}} \notin C$ and did not show up in Figure 4.3(b). Finally, it can be seen that the test point approached the true change point values as the SNR increases; *i.e.*, α_2 tends to -1 as SNR increases.

In Figure 4.3(d), we illustrate the $\text{MSE}_{\text{known}}$ and $\mathbf{BB}_{\text{sup}}$ for change-point t_i, $i = 1, 2, 3$ as a function of SNR. It is noteworthy to mention that we did not illustrate the performance for higher SNR range in this example, since we found that for SNR values larger than 10 dB the bound tends quickly to zero. On the other hand, computing MSE values in these examples for larger SNR requires a large number of Monte Carlo simulations, since the higher the SNR, the smaller the probability of an error. For example, a single realization with an error of only 1 unit in one of the change-points, among 1000 realizations in the Monte Carlo simulation, amounts to an MSE of -30dB. Similar observations hold for the example of changes in the mean rate of a Poisson distribution.

We also analyze the MSE performance as a function of the distance between change points for a fixed SNR value. In Figure 4.3(e), for SNR $= -6$ dB, we illustrate the diagonal elements of $\mathbf{BB}_{\text{sup}}$ and the MSE of the MLE for the change-point vector $\mathbf{t}$, assuming knowledge of the distribution parameters, as a function of the distance between change points t_1 and t_2. In Figure 4.3(f) we illustrate the BB and the MSE of the MLE for each change-point. We observe that the MSE of the MLE for t_1 and t_2 increases as the distance between change points t_1 and t_2 decreases. Similarly, their respective BB predict the same behavior for distances between t_1 and t_2 equal to and larger than 10 time-units; however, for distances smaller than 10 time-units their respective bounds decrease to the same value, as the did for distances larger than 22 time-units. This bound behavior is expected to take place as our Barankin-type lower

bound approximation considers only one change-point per parameter. Therefore, in our problem the test-point values are lower and upper bounded by the adjacent change-point parameters, which does not allow for evaluating errors, in estimating each change-point, beyond these limits. Thus, as the change-points get closer, the test-point domains become limited, and the bound cannot take into account either estimated errors given by estimates of t_1 which are larger than the true value of t_2, or estimated errors given by estimates of t_2 which are lower than the true value of t_1.

4.5.3 Changes in the variance of a Gaussian distribution

We consider the same scenarios as above, but with a time series with three change points in the variance of a Gaussian distribution and a common mean. We recall the closed-form expressions obtained for computing $[\mathbf{\Phi}]_{kk}$, namely, Equations (4.47) and (4.48), and define the SNR for the k^{th} change point as follows:

$$SNR_k = \frac{\left|\mathbf{M}\left(\boldsymbol{\varphi}_{k+1}\right)\right|}{\left|\mathbf{M}\left(\boldsymbol{\varphi}_k\right)\right|}, \tag{4.55}$$

where $\mathbf{M}\left(\boldsymbol{\varphi}_k\right) \in \mathbb{R}^{M\times M}$ is the covariance matrix of the k^{th} segment. In our example, $M = 1$, and, without loss of generality, we choose $\mathbf{M}\left(\boldsymbol{\varphi}_k\right) = \varphi_k$, and the mean to be equal to zero since the BIM does not depend on the mean, thus $\boldsymbol{\eta}_k = [0,\ \varphi_k]^T$. Here, we set $\varphi_1 = 1$, and variances φ_2, φ_3, and φ_4 are set such that $SNR_1 = SNR_2 = SNR_3 = SNR$. In practice, $\varphi_k = \varphi_{k-1}SNR$. In Figure 4.4(a), we illustrate sigma-parameter values as a function of sample time for different SNR values. In Figure 4.4(c), we illustrate the MSE performance of the MLE for the change-point vector as a function of the SNR and its respective Barankin bound, $\mathbf{BB}_{\text{sup}}$. In particular, we illustrate the $\text{MSE}_{\text{unknown}}$ and $\text{MSE}_{\text{known}}$ of $\mathbf{t}$ for SNR ranging from 1 to 30 dB.

In Figure 4.4(d) we focus on SNR ranging between 1 to 10 dB, and we illustrate the MSE for change-point estimates of t_1, t_2, and t_3, using the MLE and their respective bounds given by the diagonal elements of $\mathbf{BB}_{\text{sup}}$. In this scenario $\mathbf{BB}_{\text{sup}}$ belongs to set C for SNR values larger than 4 dB. It can be seen that the MSE of the MLE slowly approaches the BB as the SNR increases. In this example, the BB is the same for all change-points for SNR values above 2 dB. It can be seen that for all the SNR ranges illustrated, the maximum differences between the BB and both the $\text{MSE}_{\text{known}}$ and $\text{MSE}_{\text{unknown}}$ are approximately 7 dB and 17 dB, respectively. For SNR values lower than 2 dB the $\mathbf{BB}_{\text{sup}}$ is greater than the MSE of the MLE because the Barankin bound derivation does not consider the set of admissible values of the estimator. In our example, the MLE computation restricts the search to the range between 1 and N, and thus the MLE variance has an upper limit, which the BB computation does not consider. Moreover, the BB assumes that the estimator is unbiased at the test-points, thus for low SNR the comparison against the MLE's MSE is inappropriate because the optimal test-points tend to go to the extreme of the intervals associated to each change-point causing some bias. Also, we illustrate in Figure 4.4(b), the optimal test points $[\alpha_1^*, \alpha_2^*, \alpha_3^*]^T$ associated to the matrix $\mathbf{BB}_{\text{sup}}$. It can be seen that for all the SNR range there are no overlaps between test points and, as in the previous example, all test points approach to 1 or -1, namely, they are close to the true change-point values as SNR increases. Therefore, for large SNR values $[\mathbf{BB}_{\text{sup}}]_{kk} = \frac{\sqrt{2\ SNR_k - 1}}{SNR_k}$, which tends to 0 as $SNR_k \to \infty$.

In Figures 4.4(e) and (f), for SNR = 4 dB, we illustrate the BB and the MSE of the MLE for t_1, t_2, and t_3, assuming knowledge of the distribution parameters, as a function of the distance between change points t_1 and t_2. Above 10 units, the BB for all the change-points remains the same for distances between change-points t_1 and

t_2. The BB for t_1 increases as the distance between change-points t_1 and t_2 increases from zero to 10 units. As in the previous example, the bound in this range is overly optimistic since the test-point domains become limited.

4.5.4 Changes in the mean rate of a Poisson distribution

Now we consider a time series with three change points in the mean rate of a Poisson distribution. As in the previous examples, we recall the closed-form expressions for $[\mathbf{\Phi}]_{kk}$, *i.e.*, Equations (4.50) and (4.51). Then we define the SNR for the k^{th} change point detector as follows,

$$SNR_k = \frac{\left(f(\boldsymbol{\eta}_k) - f(\boldsymbol{\eta}_{k+1})\right)^2}{f(\boldsymbol{\eta}_k)^2}, \tag{4.56}$$

where $f(\boldsymbol{\eta}_k)$ is the mean rate of the k^{th} segment. Here, without loss of generality, we set $f(\boldsymbol{\eta}_k) = \eta_k$. The mean rate is set to $\eta_1 = 1$ and the mean rates η_2, $\eta_{3,}$ and, η_4 are set such that $SNR_1 = SNR_2 = SNR_3 = SNR$. In practice, $\eta_k = \eta_{k-1}\left(1 + \sqrt{SNR}\right)$. In Figure 4.5(a), we illustrate the mean-rate-values as a function of sample time for different SNR values

Figures 4.5(c) and (d), illustrate the $\text{MSE}_{\text{unknown}}$ and $\text{MSE}_{\text{known}}$ performance for the change-point vector and each change points t_1, t_2, and t_3, respectively. In this case, the MSE values, as well as the bounds for t_1, t_2, and t_3, are not the same for the same SNR values. In fact, it can be seen that the MSE values for t_3 are lower than the MSE values for t_2, and these last are lower than the MSE values for t_1. This difference in performance is due to the fact that in our example the difference between the means of contiguous segments are not the same, which is a direct consequence

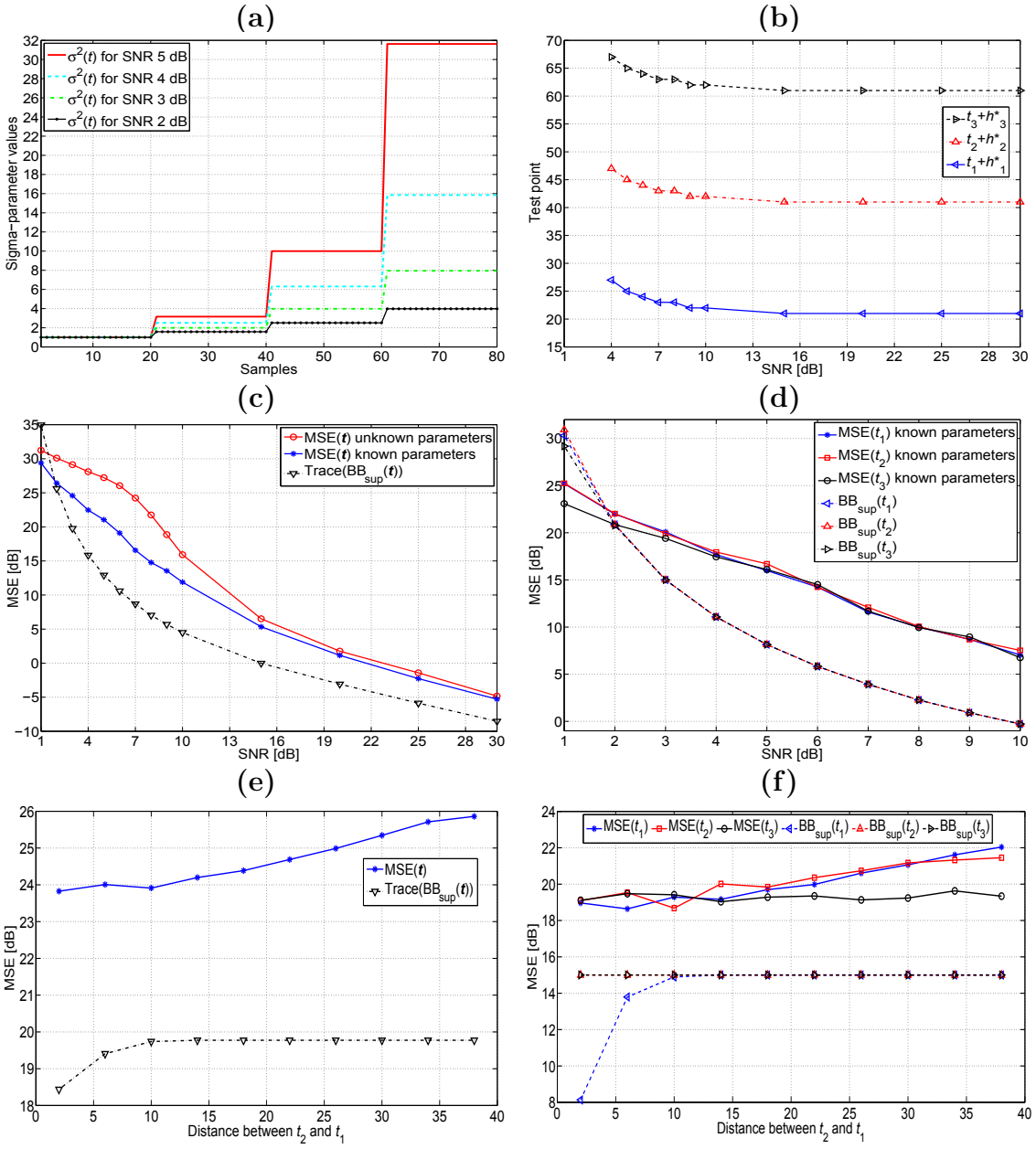


Figure 4.4: Performance analysis for estimating change-points of the variance in a Gaussian distribution: (a) Sigma-parameter values as a function of sample time for different SNR values; (b) Test points associated with the BB given by the minimal-upper bound of C, $\mathbf{BB}_{\text{sup}}$, as a function of SNR; (c) MSE of the change-point vector using the MLE of $\mathbf{t}$ and its Barankin bound given by $\mathbf{BB}_{\text{sup}}$; (d) MSE of each change-point as a function of SNR using the MLE of t_1, t_2, and t_3 and their corresponding Barankin bound $\mathbf{BB}_{\text{sup}}(t_i)$, $i = 1, \ldots, 3$; (e) MSE of change-point vector using the MLE of $\mathbf{t}$ and its Barankin bound, $\mathbf{BB}_{\text{sup}}(\mathbf{t})$, as a function of the distance between t_2 and t_1 for SNR $= \mathbf{4}$ [dB]; (f) MSE of each change-point and their respective $\mathbf{BB}_{\text{sup}}$ as a function of the distance between t_2 and t_1 for SNR $= 4$ [dB]. ©[2010] IEEE.

of the definition used for SNR. In practice, for any SNR, the differences between the means for segments $[t_3 + 1,\ N]$ and $[t_2 + 1,\ t_3]$ is larger than the difference between the means for segments $[t_2 + 1,\ t_3]$ and $[t_1 + 1,\ t_2]$. In Figure 4.5(b) we illustrate the test points associated to the matrix $\mathbf{BB}_{\text{sup}}$. As in the previous examples, the test points tend to the true change-point values as the SNR increases. Finally, in Figures 4.5(e) and (f), we illustrate the MSE performance, assuming known mean rates, as a function of the distance between change points for SNR $= -6$ dB. The bounds for change-point t_2 and t_3 are constant in all the illustrated range, though, the MSE of the MLE for t_2 slightly varies as t_1 approaches t_2. As we discussed in the previous examples, the bound for t_1 is overly optimistic for small distances between t_2 and t_1, due to the constrained test-point domain.

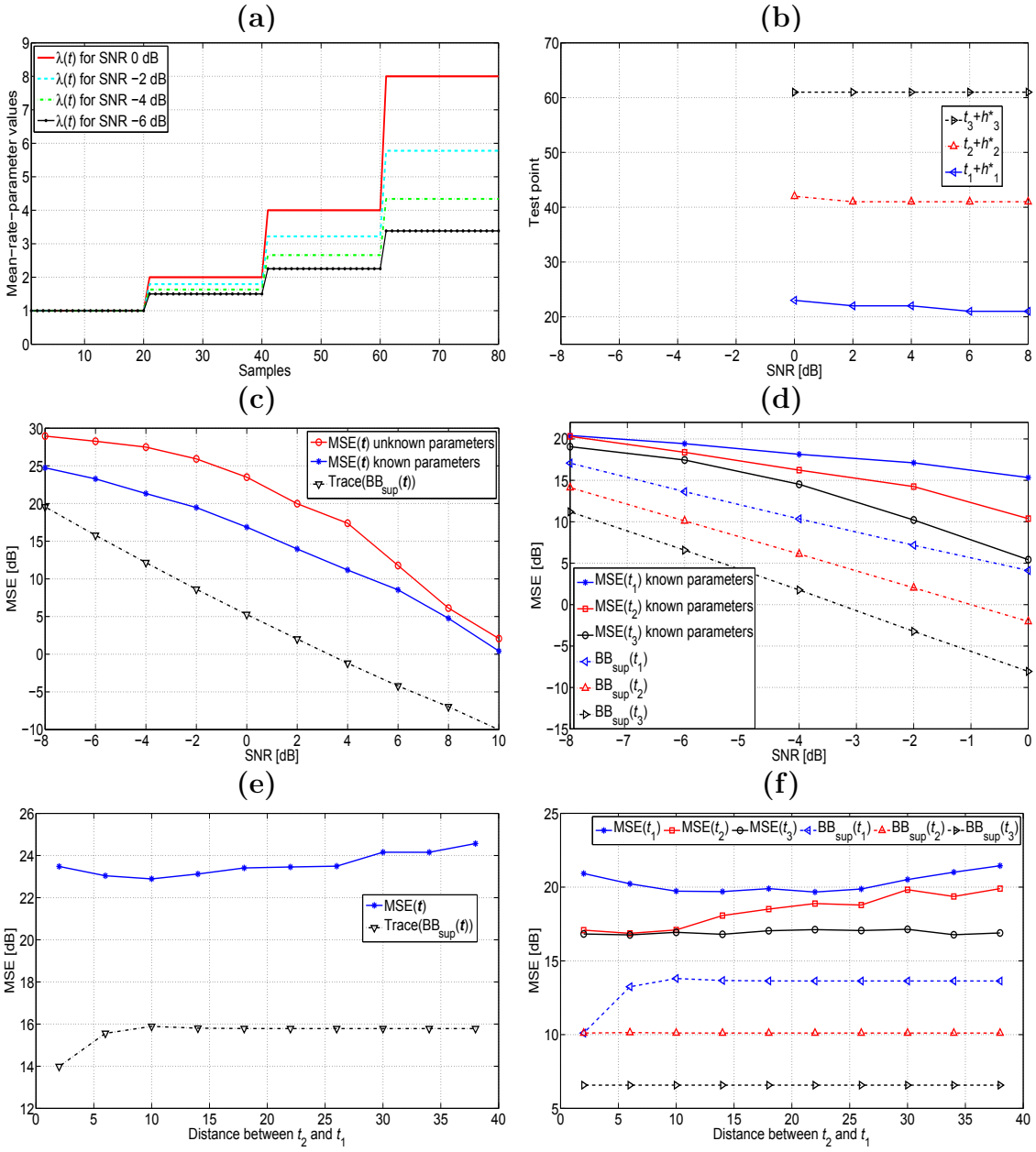


Figure 4.5: Performance analysis for estimating change-points in the mean rate of a Poisson distribution distribution: (a) Mean-rate-values as a function of sample time for different SNR values; (b) Test points associated with the BB given by the minimal-upper bound of C, $\mathbf{BB}_{\text{sup}}$, as a function of SNR; (c) MSE of the change-point vector using the MLE of $\mathbf{t}$ and its Barankin bound given by $\mathbf{BB}_{\text{sup}}$; (d) MSE of each change-point as a function of SNR using the MLE of t_1, t_2, and t_3 and their corresponding Barankin bound $\mathbf{BB}_{\text{sup}}(t_i)$, $i = 1, \ldots, 3$; (e) MSE of change-point vector using the MLE of $\mathbf{t}$ and its Barankin bound, $\mathbf{BB}_{\text{sup}}(\mathbf{t})$, as a function of the distance between t_2 and t_1 for SNR $= -6$ [dB]; (f) MSE of each change-point and their respective $\mathbf{BB}_{\text{sup}}$ as a function of the distance between t_2 and t_1 for SNR $= -6$ [dB]. ©[2010] IEEE.

4.6 Summary

We investigated a simplified version of the Barankin bound on multiple change-point estimation. The approximate Barankin information matrix was spelled, revealing an interesting tri-diagonal structure, meaning that the estimation of one change point is naturally perturbed by its two neighbors. Moreover, the Barankin information matrix can be reduced to a block diagonal structure leading to closed-form for the elements of its inverse. The main limitation posed by this HCR approximation is a reduced search space for the BIM that leads to a loose Barankin bound. We also discussed the existence and computation of the supremum with respect to the Loewner partial ordering, on the finite set of candidate BB solutions. To overcome this problem, we computed a suitable minimal-upper bound to this set given by the matrix associated with the Lowner-John Ellipsoid of the set of hyper-ellipsoids associated to each maximal element of the set of candidate bound matrices. Two important distributions in signal and image processing were investigated, the Gaussian case and the Poisson case, for which we obtained closed-form expressions for all the elements of the Barankin information matrix. Finally, we illustrated our analysis by presenting various simulation results, including our problem of estimating changes in the variance of a Gaussian distribution ($d = 0$).

Chapter 5

Conclusions and Future Work

In this chapter, we summarize the key contributions of our work in Section 5.1 and then provide future work in Section 5.2.

5.1 Conclusions

In this dissertation, we contributed with a physical model and with statistical algorithms for analyzing uterine contractions using MMG. In particular, we proposed a forward electromagnetic model of human myometrial contractions. Our modeling approach takes into account electrophysiological and anatomical knowledge jointly at the cellular, tissue, and organ level. We applied a bidomain approach for modeling the propagation of the myometrium transmembrane potential v_{m} on the uterus and used this approach to compute the action potential ϕ and the magnetic field $\boldsymbol{B}$ at the abdominal surface. We introduced a modified version of the FitzHugh-Nagumo equation for modeling the ionic currents in each cell. Though our ionic current model does not explicitly consider C_a^{2+} dynamics, the simplicity of the FitzHugh-Nagumo allowed for capturing the nuances of the uterine myocyte response, but it can be used to model

the propagating action potential under well defined conditions, as shown in this work. We also incorporated the anisotropic nature of the uterus by considering conductivity tensors, and we proposed a general approach to design conductivity-tensor orientation for any uterine shape. Using Archie's law and an analytical expression for the transmembrane potential propagation speed as function of the model parameters, we estimated conductivity-tensor values in the extracellular and intracellular domains. We introduced a four-compartment volume conductor geometry for the problem and proposed a discretized model solution using finite element methods. We illustrated our approach through a numerical example of a uterine contraction at term. Considering a spherical uterus and one pacemaker located in the fundus, we obtained a travelling transmembrane potential depolarizing from -56mV to -16 mV and an average potential in the plateau area of -25 mV with a duration, before hyperpolarization, of 35 [s], which is a good approximation with respect to the average recorded transmembrane potentials reported in the technical literature. Similarly, the percentage of myometrial cells contracting as a function of time had the same symmetric properties and length as the intrauterine pressure waveforms of a pregnant human myometrium at term.

We also developed a general analysis for the detection of uterine MMG contractions. In particular, we have proposed a two-stage statistical time-segment discriminator using a single channel of MMG measurements. We assumed that the preprocessed channels are modeled by a piecewise time-varying AR model of a certain order, with an input given by a white Gaussian noise with time-varying variance. Therefore, we first designed a statistical model-based segmentation algorithm based on the SIC to estimate the time-instants of changes in the parameters. To discriminate time segments that contain a contraction, we evaluated features such as the time segment

power (RMS) and the dominant frequency component (FOZC). Then, we applied the non-supervised cluster algorithm K-means to classify the RMS values, obtaining then a discrete-time binary decision signal indicating the presence of a contraction. Since each single channel detector provided local information regarding the presence of a burst of activities, we also analyzed the fusion of the decision signals from all the sensors, as in the parallel distributed detection approach. We proposed a distributed processing approach providing estimates of spatial-temporal propagation of uterine activities and RMS values, the time evolution of a percentage of active sensors and the time evolution of total RMS values.

We applied our detection algorithm to real MMG records obtained from 10 patients with gestational ages between 31 and 40 weeks, who were admitted to the hospital for contractions. We found that the RMS values discriminated the presence of time segments with contractions. However, that result was not obtained in the case of the FOZC values. We evaluated the performance of our detection algorithm by computing the DR, FAR, and CORR, respectively, using as a reference the patient's feedback. We observed that the maximum average DR and average CORR were achieved consistently in all groups of channels when d equaled to 0 and 1, in the frequency range 0.2-0.4 Hz. Also, in the same frequency range for $d = 0, 1$, the lowest average FAR were obtained in the group of channels G1. We found that in the frequency range 0.1-0.4 Hz, the maximum averages for DR and CORR were obtained for $d = 0, 3$. Thus, on average, a variance based algorithm ($d = 0$) is suitable to detect contractions using the RMS values. Our distributed processing approach, applied to real data, proved helpful in understanding uterine MMG contraction activity both spatially and temporally. For example, our approach detected uterine activity much earlier than the patient began to sense it. It also enabled visualizing the relative

location of the origin of uterine contractions and quantifying the amount of energy delivered during a contraction.

Our statistical model-based segmentation algorithm is based on the detection of multiple change points, using a binary search algorithm; therefore, the problem is simplified by testing the hypothesis for a single change point. Using the asymptotic distribution of the test statistics based on SIC, we obtained optimal thresholds in the sense of the Neyman-Pearson criterion; therefore, we controlled the probability of false alarm and maximized the probability of change-point detection in each stage of the binary search algorithm. We also proposed an estimate of the model order d. However, in practice, this approach performed well only in a group of channels from the same patient, suggesting that a different model for segmentation should be attempted. For example, to avoid over segmentation, time segments with piecewise constant time-varying model orders should be considered.

We also investigated in detail the performance of multiple change-point estimates as a function of the MSE for an independent vector sequence, including our aforementioned problem of estimating changes in the variance of a Gaussian distribution ($d = 0$). We studied the global performance for the class of unbiased estimator of change points in a sequence assuming known the number of changes. In particular, we studied the Hammersley-Chapman-Robbins (HCR) bound, a simplified version of the Barankin bound, on multiple change-point estimation. The approximate Barankin information matrix was derived, revealing an interesting tri-diagonal structure, meaning that the estimation of one change point is naturally perturbed by its two neighbors. Moreover, the Barankin information matrix can be reduced to a block diagonal structure leading to closed-form for the elements of its inverse. The main limitation posed by this HCR approximation is a reduced search space for the BIM that leads to a looser Barankin

bound. With respect to the Loewner partial ordering, we also discussed the existence and computation of the supremum on the finite set of candidate BB solutions. To overcome this problem, we computed a suitable minimal-upper bound to this set, given by the matrix associated with the Lowner-John Ellipsoid of the set of hyper-ellipsoids associated to each maximal element of the set of candidate bound matrices. Two important distributions in signal and image processing were investigated, the Gaussian case and the Poisson case, for which we obtained closed-form expressions for all the elements of the Barankin information matrix. Finally, we analyzed the case of piecewise changes of variance($d = 0$) and evaluated the MSE performance as a function of the variances and distances between the change point.

5.2 Future work

In future work, we will consider a more realistic model for the geometry of the uterus, fetus, and abdominal shape. We will also include more realistic ionic current models as in [33–35], and will consider spatial variations of the fiber orientation.

In further studies using MMG measurements, we will address the optimization of the detection by evaluating additional features of the preprocessed measurements. We will also develop a biomagnetically compatible pressure measurement device for a better recording of the intrauterine pressure and provide a more precise approach to validate the performance of our method with real measurements, as well as to allow for comparison with other methods.

We will further study Barankin-type lower bounds, considering distribution parameters in addition to the multiple change-point localizations for the time-dependent case.

Appendix A

Designing Uterine Anisotropy

Denote C as the curve of symmetry and define it using the following parametric representation as a function of a single parameter t:

$$C : t \longmapsto \boldsymbol{r}_C(t),\ t_1 \leq t \leq t_2, \tag{A.1}$$

where $\boldsymbol{r}_C(t)$ is a point defined with respect to the global coordinate system, and $\boldsymbol{r}_C(t_1)$ and $\boldsymbol{r}_C(t_2)$ represent the extreme points of the curve. For example, $\boldsymbol{r}_C(t) = (x(t), y(t), z(t))$ with respect to the Cartesian system. Define $\widehat{k}(r) = \left.\frac{d\boldsymbol{r}_C(t)}{dt}\right|_{t_0} \Big/ \left\| \left.\frac{d\boldsymbol{r}_C(t)}{dt}\right|_{t_0} \right\|$ as the unitary vector field with direction given by the tangent vector of $\boldsymbol{r}_C(t)$ at t_0, where $\boldsymbol{r}_C(t_0)$ is the closest point to r such that $\left\langle \left.\frac{d\boldsymbol{r}_C(t)}{dt}\right|_{t_0}, \overrightarrow{\boldsymbol{r}_C(t_0)\ r} \right\rangle = 0$. Then, we define $\widehat{\boldsymbol{t}}_1(\boldsymbol{r})$ to be contained in the plane formed by $\widehat{k}(\boldsymbol{r})$ and $\widehat{\boldsymbol{n}}(\boldsymbol{r})$ as follows:

$$\widehat{\boldsymbol{t}}_1(\boldsymbol{r}) = \beta\widehat{k}(\boldsymbol{r}) + \gamma\widehat{\boldsymbol{n}}(\boldsymbol{r}), \tag{A.2}$$

subject to the following conditions:

$$\left\langle \widehat{\boldsymbol{t}}_1(\boldsymbol{r}), \widehat{\boldsymbol{n}}(\boldsymbol{r}) \right\rangle = 0, \text{ and} \tag{A.3}$$

$$\left\langle \widehat{\boldsymbol{t}}_1(\boldsymbol{r}), \widehat{\boldsymbol{t}}_1(\boldsymbol{r}) \right\rangle^2 = 1. \tag{A.4}$$

Therefore, replacing (A.2) in (A.3) and (A.4) we obtain the following system of equations:

$$\beta \left\langle \widehat{k}(\boldsymbol{r}), \widehat{\boldsymbol{n}}(\boldsymbol{r}) \right\rangle + \gamma = 0, \text{ and} \tag{A.5}$$

$$\beta^2 + \gamma^2 + 2\ \beta\ \gamma \left\langle \widehat{k}(\boldsymbol{r}), \widehat{\boldsymbol{n}}(\boldsymbol{r}) \right\rangle = 1. \tag{A.6}$$

Solving for all $\boldsymbol{r}$, such that $\left\langle \widehat{k}(\boldsymbol{r}), \widehat{\boldsymbol{n}}(\boldsymbol{r}) \right\rangle \neq 1$, we obtain that $\beta = \pm \frac{1}{\sqrt{1-\left\langle \widehat{k}(\boldsymbol{r}), \widehat{\boldsymbol{n}}(\boldsymbol{r}) \right\rangle}}$ and $\gamma = \mp \frac{\left\langle \widehat{k}(\boldsymbol{r}), \widehat{\boldsymbol{n}}(\boldsymbol{r}) \right\rangle}{\sqrt{1-\left\langle \widehat{k}(\boldsymbol{r}), \widehat{\boldsymbol{n}}(\boldsymbol{r}) \right\rangle}}$. Then $\widehat{\boldsymbol{t}}_2(\boldsymbol{r}) = \widehat{\boldsymbol{t}}_1(\boldsymbol{r}) \times \widehat{\boldsymbol{n}}(\boldsymbol{r}) = \beta \widehat{k}(\boldsymbol{r}) \times \widehat{\boldsymbol{n}}(\boldsymbol{r})$, since by definition it is mutually orthogonal to $\widehat{\boldsymbol{t}}_1(\boldsymbol{r})$. Hence, given $\widehat{\boldsymbol{t}}_1(\boldsymbol{r})$ and $\widehat{\boldsymbol{t}}_2(\boldsymbol{r})$, we define $\boldsymbol{a}_3(\boldsymbol{r})$ as follows:

$$\boldsymbol{a}_3(\boldsymbol{r}) = \widehat{\boldsymbol{t}}_1(\boldsymbol{r}) \cos(\alpha) + \widehat{\boldsymbol{t}}_2(\boldsymbol{r}) \sin(\alpha), \tag{A.7}$$

where α is the fiber orientation angle with respect to $\widehat{\boldsymbol{t}}_1(\boldsymbol{r})$. In order to take into account complex fiber orientations α can be modeled as a spatial function defined over the domain of interest. Given our uterine volume assumptions, the points $\boldsymbol{r}_C(t_1)$ and $\boldsymbol{r}_C(t_2)$ are the only points that satisfy the condition $\left\langle \widehat{k}(\boldsymbol{r}), \widehat{\boldsymbol{n}}(\boldsymbol{r}) \right\rangle = 1$. Since at $\boldsymbol{r}_C(t_1)$ and $\boldsymbol{r}_C(t_2)$ we cannot define $\widehat{\boldsymbol{t}}_1(\boldsymbol{r})$ and $\widehat{\boldsymbol{t}}_2(\boldsymbol{r})$ using the curve C as a global reference, we set $\boldsymbol{a}_3(\boldsymbol{r}) = 0$, defining a point of isotropic conductivity. In Fig. A.1 we represent the fiber orientation $\boldsymbol{a}_3(\boldsymbol{r})$ with respect to the local coordinate axes given by $\{\widehat{\boldsymbol{n}}(\boldsymbol{r}), \widehat{\boldsymbol{t}}_1(\boldsymbol{r}), \widehat{\boldsymbol{t}}_2(\boldsymbol{r})\}$.

For the case of a uterine volume such that C is parallel to the z axis, then $\boldsymbol{a}_3(\boldsymbol{r})$ can be written as a function of $\widehat{\boldsymbol{n}}(\boldsymbol{r})$ as follows:

$$\boldsymbol{a}_3(\boldsymbol{r}) = (P \cos(\alpha) + F \sin(\alpha))\, \widehat{\boldsymbol{n}}(\boldsymbol{r}), \tag{A.8}$$

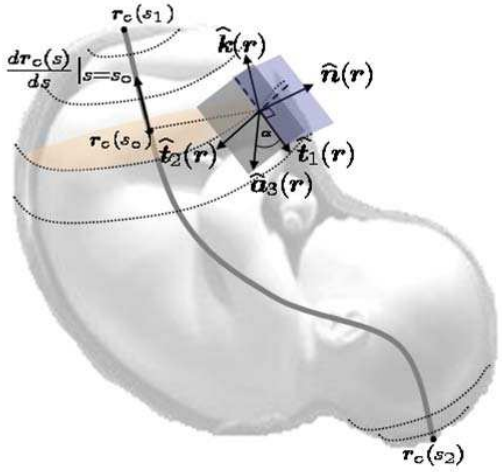


Figure A.1: Simplified illustrations of $\boldsymbol{a}_3(\boldsymbol{r})$ with respect to the local coordinates axis given by $\{\widehat{\boldsymbol{n}}(\boldsymbol{r}), \widehat{\boldsymbol{t}}_1(\boldsymbol{r}), \widehat{\boldsymbol{t}}_2(\boldsymbol{r})\}$. The blue plane contains the vectors $\widehat{\boldsymbol{n}}(\boldsymbol{r})$, $\widehat{\boldsymbol{k}}(\boldsymbol{r})$, and $\widehat{\boldsymbol{t}}_1(\boldsymbol{r})$, and it is perpendicular to the gray plane formed by vectors $\widehat{\boldsymbol{t}}_1(\boldsymbol{r})$, $\widehat{\boldsymbol{t}}_1(\boldsymbol{r})$ and $\boldsymbol{a}_3(\boldsymbol{r})$. The orange plane is the cross section of the uterus perpendicular to the vector $\frac{d\boldsymbol{r}_C(t)}{dt}\Big|_{t_0}$. The gray curve is the curve of symmetry $\boldsymbol{r}_C(t)$ with $\boldsymbol{r}_C(t_1)$ and $\boldsymbol{r}_C(t_2)$ extreme points of the curve.

where

$$P = \begin{bmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & -1/a \end{bmatrix}, \; F = \begin{bmatrix} 0 & b & 0 \\ -b & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \tag{A.9}$$

$$a = \frac{\nabla_z f(\boldsymbol{r})}{\sqrt{\left(\nabla_x f(\boldsymbol{r})\right)^2 + \left(\nabla_y f(\boldsymbol{r})\right)^2}}, \text{ and } b = \frac{\|\nabla f(\boldsymbol{r})\|}{\sqrt{\left(\nabla_x f(\boldsymbol{r})\right)^2 + \left(\nabla_y f(\boldsymbol{r})\right)^2}}, \tag{A.10}$$

with ∇_j the j-th component of the gradient. In the case of a spherical myometrium, $\boldsymbol{a}_3(\boldsymbol{r})$ is given as by

$$\boldsymbol{a}_3(\boldsymbol{r}) = \begin{bmatrix} \frac{zx\cos\alpha}{\sqrt{x^2+y^2}R} + \frac{y\sin\alpha}{2\sqrt{x^2+y^2}} \\ \frac{zy\cos\alpha}{\sqrt{x^2+y^2}R} - \frac{x\sin\alpha}{\sqrt{x^2+y^2}} \\ -\frac{\sqrt{x^2+y^2}}{R}\cos\alpha \end{bmatrix}, \tag{A.11}$$

where $R = \sqrt{x^2 + y^2 + z^2}$. Note that for $\alpha = 0$ the main axis of the fibers runs vertically from the fundus to the cervix.

Appendix B

Determination of C_α for the SIC Change-point Detector Based on an AR-model

Let $\zeta_n = \max_\tau\{\zeta_n(\tau) : 2d+1 < \tau \leq n-d-3\}$, where $\zeta_n(\tau) = (n-d)\,\ln\hat{\sigma_0^2} - (\tau - d)\,\ln\hat{\sigma_1^2}(\tau) - (n-\tau)\,\ln\hat{\sigma_2^2}(\tau)$. [98] showed that the asymptotic distribution of ζ_n is given by

$$\Pr[\frac{\zeta_n - b_n(d+2)}{a_n(d+2)} \leq x] \xrightarrow[n\to\infty]{\mathcal{D}} \exp\{-2\,\exp\{-x/2\}\}, \tag{B.1}$$

where

$$\begin{aligned} b_n(x) &= \frac{\{2\ln\ln(n) + x/2\ln\ln\ln(n) - \ln(\Gamma(x/2))\}^2}{2\ln\ln(n)} \\ a_n(x) &= \sqrt{\frac{b_n(x)}{2\ln\ln(n)}}. \end{aligned}$$

Let $\lambda = \max_{2d+1<\tau\leq n-d-3}\{SIC(n) - SIC(\tau)\}$. Then, $\zeta_n = \lambda + (d+2)\,\ln(n)$. Thus

$$\begin{aligned}
1-\alpha &= \Pr(\mathrm{SIC}(n) < \min_{2\,d+1\leq\tau\leq n-d-3} \mathrm{SIC}(\tau) + C_\alpha|\mathrm{H}_0) \\
&= \Pr(\max_{2\,d+1\leq\tau\leq n-d-2} \mathrm{SIC}(n) - \mathrm{SIC}(\tau) < C_\alpha|\mathrm{H}_0) \\
&= \Pr(\lambda + (d+2)\,\ln(n) < C_\alpha + (d+2)\,\ln(n)|\mathrm{H}_0) \\
&= \Pr(\zeta_n - b_n(d+2) < C_\alpha + (d+2)\,\ln(n) - b_n(d+2)|\mathrm{H}_0) \\
&= \Pr(\frac{\zeta_n - b_n(d+2)}{a_n(d+2)} < \frac{C_\alpha + (d+2)\,\ln(n) - b_n(d+2)}{a_n(d+2)}|\mathrm{H}_0). \qquad \text{(B.2)}
\end{aligned}$$

(B.3)

Hence, using (B.1)

$$1-\alpha \;\cong\; \exp\left(-2\,\exp-\left(\frac{C_\alpha + (d+2)\,\ln(n) - b_n(d+2)}{2\,a_n(d+2)}\right)\right) \qquad \text{(B.4)}$$

Then, solving for C_α

$$C_\alpha = b_n(d+2) - 2\,a_n(d+2)\left\{\ln\ln((1-\alpha)^{-1}) - \ln(2)\right\} - (d+2)\ln(n). \qquad \text{(B.5)}$$

Appendix C

Proof of Lemma 1

Proof. We need to proof that for all $\mathbf{y} \in \mathbb{R}^q$ with $\mathbf{y} \neq \mathbf{0}$, $\mathbf{y}^T (\mathbf{A} - \mathbf{B}) \mathbf{y} > 0$ if $\lambda_1 \leq 1$. Since $\mathbf{A}$ is pd and $\mathbf{B}$ is psd, there exist a non-singular matrix $\mathbf{F}$ such that

$$\mathbf{F}^T \mathbf{B} \mathbf{F} = \mathbf{diag} (\lambda_1, \ldots, \lambda_m, \lambda_{m+1}, \ldots, \lambda_q) = \Lambda \text{ and} \tag{C.1}$$

$$\mathbf{F}^T \mathbf{A} \mathbf{F} = \mathbf{I} \tag{C.2}$$

Thus, $\mathbf{B} = \left(\mathbf{F}^T\right)^{-1} \Lambda \left(\mathbf{F}\right)^{-1}$ and $\mathbf{A} = \left(\mathbf{F}^T\right)^{-1} \mathbf{I} \left(\mathbf{F}\right)^{-1}$ and $\mathbf{y}^T (\mathbf{A} - \mathbf{B}) \mathbf{y} = \mathbf{y}^T \left(\mathbf{F}^T\right)^{-1} (\mathbf{I} - \Lambda) \left(\mathbf{F}\right)^{-1} \mathbf{y}$. Let $\mathbf{z} = \left(\mathbf{F}\right)^{-1} \mathbf{y}$, because $\mathbf{F}$ is not singular $\left(\mathbf{F}\right)^{-1} \mathbf{y} = \mathbf{0}$ for $\mathbf{y} = \mathbf{0}$, therefore our problem is equivalent to analyze the positiveness of $r = \mathbf{z}^T (\mathbf{I} - \Lambda) \mathbf{z}$, for $\mathbf{z} \neq \mathbf{0}$. Since $\lambda_{m+1} = \cdots = \lambda_q = 0$, $r = \sum_{i=1}^{m} (1 - \lambda_i) z_i^2 + \sum_{i=1}^{m} z_i^2$. Hence, if $\lambda_1 \leq 1$, then $(1 - \lambda_i) \geq 0$, for $i = 1, \ldots, m$ thus $r > 0$ and $\mathbf{A} \succ \mathbf{B}$. On the other hand, if $\lambda_1 > 1$, we can always find a $\mathbf{z}$ vector such that $r \leq 0$ or $r > 0$, thus $\mathbf{A}$ and $\mathbf{B}$ are not mutually comparable. ■

Appendix D

Computing Elements of the Barankin Information Matrix

D.0.1 Computing diagonal elements of $\mathbf{\Phi}$

For $\alpha_k > 0$, Equation (4.14) becomes

$$
\begin{aligned}
[\mathbf{\Phi}]_{kk} &= \int_{\Omega} \frac{\prod_{i=1}^{t_1} p_1^2(\mathbf{x}_i;\boldsymbol{\eta}_1)\cdots\prod_{i=t_{k-1}+1}^{t_k+\alpha_k} p_k^2(\mathbf{x}_i;\boldsymbol{\eta}_k)}{\prod_{i=1}^{t_1} p_1(\mathbf{x}_i;\boldsymbol{\eta}_1)\cdots\prod_{i=t_q+1}^{N} p_{q+1}(\mathbf{x}_i;\boldsymbol{\eta}_{q+1})} \\
&\quad \times \prod_{i=t_k+\alpha_k+1}^{t_{k+1}} p_{k+1}^2(\mathbf{x}_i;\boldsymbol{\eta}_{k+1})\cdots\prod_{i=t_q+1}^{N} p_{q+1}^2(\mathbf{x}_i;\boldsymbol{\eta}_{q+1})d\mathbf{X}
\end{aligned}
$$

$$
\begin{aligned}
[\mathbf{\Phi}]_{kk} &= \int_{\Omega} \prod_{i=1}^{t_1} p_1(\mathbf{x}_i;\boldsymbol{\eta}_1)\cdots\prod_{i=t_{k-1}+1}^{t_k} p_k(\mathbf{x}_i;\boldsymbol{\eta}_k) \\
&\quad \times \frac{\prod_{i=t_k+1}^{t_k+\alpha_k} p_k^2(\mathbf{x}_i;\boldsymbol{\eta}_k)}{\prod_{i=t_k+1}^{t_k+\alpha_k} p_{k+1}(\mathbf{x}_i;\boldsymbol{\eta}_{k+1})} \\
&\quad \times \prod_{i=t_k+\alpha_k+1}^{t_{k+1}} p_{k+1}(\mathbf{x}_i;\boldsymbol{\eta}_{k+1})\cdots\prod_{i=t_q+1}^{N} p_{q+1}(\mathbf{x}_i;\boldsymbol{\eta}_{q+1})d\mathbf{X}.
\end{aligned}
$$

After some straightforward simplifications, we have that

$$
\begin{aligned}
[\mathbf{\Phi}]_{kk} &= \int_{\Omega} \prod_{i=t_k+1}^{t_k+\alpha_k} \frac{p_k^2(\mathbf{x}_i;\boldsymbol{\eta}_k)}{p_{k+1}(\mathbf{x}_i;\boldsymbol{\eta}_{k+1})} d\mathbf{x}_{t_k+1} \cdots d\mathbf{x}_{t_k+\alpha_k} \\
&= \left(\int_{\Omega} \frac{p_k^2(\mathbf{x};\boldsymbol{\eta}_k)}{p_{k+1}(\mathbf{x};\boldsymbol{\eta}_{k+1})} d\mathbf{x} \right)^{\alpha_k}.
\end{aligned}
$$

Similar analysis can be applied to solve for $\alpha_k < 0$.

D.0.2 Computing non-diagonal elements of $\mathbf{\Phi}$

For $\alpha_k > 0$ and $\alpha_l > 0$, Equation (4.16) becomes

$$
\begin{aligned}
[\mathbf{\Phi}]_{kl} &= \int_{\Omega} \frac{\prod_{i=t_{k-1}+1}^{t_k+\alpha_k} p_k(\mathbf{x}_i;\boldsymbol{\eta}_k) \prod_{i=t_k+\alpha_k+1}^{t_{k+1}} p_{k+1}(\mathbf{x}_i;\boldsymbol{\eta}_{k+1})}{\prod_{i=t_{k-1}+1}^{t_k} p_k(\mathbf{x}_i;\boldsymbol{\eta}_k) \prod_{i=t_k+1}^{t_{k+1}} p_{k+1}(\mathbf{x}_i;\boldsymbol{\eta}_{k+1})} \\
&\qquad \times p\left(\mathbf{X};\boldsymbol{\theta}_0 + \mathbf{h}_l\right) d\mathbf{X}
\end{aligned}
$$

$$
\begin{aligned}
[\mathbf{\Phi}]_{kl} &= \int_{\Omega} \frac{\prod_{i=t_k+1}^{t_k+\alpha_k} p_k(\mathbf{x}_i;\boldsymbol{\eta}_k)}{\prod_{i=t_k+1}^{t_k+\alpha_k} p_{k+1}(\mathbf{x}_i;\boldsymbol{\eta}_{k+1})} \prod_{i=1}^{t_1} p_1(\mathbf{x}_i;\boldsymbol{\eta}_1) \\
&\quad \cdots \prod_{i=t_k+1}^{t_k+\alpha_k} p_{k+1}(\mathbf{x}_i;\boldsymbol{\eta}_{k+1}) \prod_{i=t_k+\alpha_k+1}^{t_{k+1}} p_{k+1}(\mathbf{x}_i;\boldsymbol{\eta}_{k+1}) \\
&\quad \cdots \prod_{i=t_{l-1}+1}^{t_l+\alpha_l} p_l(\mathbf{x}_i;\boldsymbol{\eta}_l) \cdots \prod_{i=t_q+1}^{N} p_{q+1}(\mathbf{x}_i;\boldsymbol{\eta}_{q+1}) d\mathbf{X} \\
&= 1.
\end{aligned}
$$

The cases ($\alpha_k < 0,\ \alpha_l < 0$), ($\alpha_k < 0,\ \alpha_l > 0$), and $t_k + \alpha_k < t_l + \alpha_l$ are solved using same approach as above. For the overlapping case, i.e., $t_k + \alpha_k > t_l + \alpha_l$, is more difficult. Replacing $l = k+1$ and keeping in mind that $\alpha_k > 0$ and $\alpha_{k+1} < 0$, Equation (4.16) becomes

$$[\mathbf{\Phi}]_{kk+1} = \int_{\Omega} \prod_{i=1}^{t_1} p_1(\mathbf{x}_i;\boldsymbol{\eta}_1)\cdots \times \frac{\prod_{i=t_{k-1}+1}^{t_k+\alpha_k} p_k(\mathbf{x}_i;\boldsymbol{\eta}_k) \prod_{i=t_k+\alpha_k+1}^{t_{k+1}} p_{k+1}(\mathbf{x}_i;\boldsymbol{\eta}_{k+1})}{\prod_{i=t_{k-1}+1}^{t_k} p_k(\mathbf{x}_i;\boldsymbol{\eta}_k) \prod_{i=t_k+1}^{t_{k+1}} p_{k+1}(\mathbf{x}_i;\boldsymbol{\eta}_{k+1})} \times \frac{\prod_{i=t_k+1}^{t_{k+1}+\alpha_{k+1}} p_{k+1}(\mathbf{x}_i;\boldsymbol{\eta}_{k+1}) \prod_{i=t_{k+1}+\alpha_{k+1}+1}^{t_{k+2}} p_{k+2}(\mathbf{x}_i;\boldsymbol{\eta}_{k+2})}{\prod_{i=t_{k+1}+1}^{t_{k+2}} p_{k+2}(\mathbf{x}_i;\boldsymbol{\eta}_{k+2})} \cdots \times \prod_{i=t_q+1}^{N} p_{q+1}(\mathbf{x}_i;\boldsymbol{\eta}_{q+1}) d\mathbf{X}$$

$$\begin{aligned}
[\mathbf{\Phi}]_{kk+1} &= \int_{\Omega} \prod_{i=1}^{t_1} p_1(\mathbf{x}_i;\boldsymbol{\eta}_1) \prod_{i=t_{k+1}+\alpha_{k+1}+1}^{t_k+\alpha_k} \frac{p_k(\mathbf{x}_i;\boldsymbol{\eta}_k) p_{k+2}(\mathbf{x}_i;\boldsymbol{\eta}_{k+2})}{p_{k+1}(\mathbf{x}_i;\boldsymbol{\eta}_{k+1})} \\
&\qquad \cdots \prod_{i=t_q+1}^{N} p_{q+1}(\mathbf{x}_i;\boldsymbol{\eta}_{q+1}) d\mathbf{X} \\
&= \left(\int_{\Omega} \frac{p_k(\mathbf{x};\boldsymbol{\eta}_k) p_{k+2}(\mathbf{x};\boldsymbol{\eta}_{k+2})}{p_{k+1}(\mathbf{x};\boldsymbol{\eta}_{k+1})} d\mathbf{x} \right)^{\beta_k},
\end{aligned}$$

where $\beta_k = (t_k + \alpha_k) - (t_{k+1} + \alpha_{k+1})$.

D.0.3 Computing the elements of Φ for changes in mean and covariance matrix of Gaussian distribution

In this case $\boldsymbol{\eta}_j = \left[\boldsymbol{\nu}_j^T,\ \boldsymbol{\varphi}_j^T\right]^T$, and the data likelihood is given as follows,

$$p(\mathbf{X};\mathbf{t}) = \frac{1}{(2\pi)^{N\ M/2} \prod_{j=1}^{q+1} \left|\mathbf{M}\left(\boldsymbol{\varphi}_j\right)\right|^{(t_j - t_{j-1})/2}} \times \exp\left\{ -\frac{1}{2} Tr\left\{ \sum_{j=1}^{q+1} \mathbf{M}\left(\boldsymbol{\varphi}_j\right)^{-1} \times \left(\sum_{i=t_{j-1}+1}^{t_j} \left(\mathbf{x}_i - \mathbf{f}(\boldsymbol{\nu}_j)\right)\left(\mathbf{x}_i - \mathbf{f}(\boldsymbol{\nu}_j)\right)^T \right) \right\} \right\}.$$

For $\alpha_k > 0$, using Equation (4.15), we have that $[\mathbf{\Phi}]_{kk}$ is given as follows:

$$
\begin{aligned}
[\mathbf{\Phi}]_{kk} &= \left(\int_{\mathbb{R}^M} \frac{(2\pi)^{M/2} \left|\mathbf{M}\left(\boldsymbol{\varphi}_{k+1}\right)\right|^{1/2}}{\left((2\pi)^{M/2} \left|\mathbf{M}\left(\boldsymbol{\varphi}_{k}\right)\right|^{1/2}\right)^2} \right. \\
&\quad \times \exp\left(- \left(\mathbf{x}_i - \mathbf{f}(\boldsymbol{\nu}_k)\right)^T \left(\mathbf{M}\left(\boldsymbol{\varphi}_k\right)\right)^{-1} \left(\mathbf{x}_i - \mathbf{f}(\boldsymbol{\nu}_k)\right)\right) \\
&\quad \left. \times \exp\left\{ + \frac{1}{2} \left(\mathbf{x}_i - \mathbf{f}(\boldsymbol{\nu}_{k+1})\right)^T \left(\mathbf{M}\left(\boldsymbol{\varphi}_{k+1}\right)\right)^{-1} \left(\mathbf{x}_i - \mathbf{f}(\boldsymbol{\nu}_{k+1})\right)\right\} \, d\mathbf{x}_i \right)^{\alpha_k}
\end{aligned}
$$

$$
\begin{aligned}
[\mathbf{\Phi}_{kk} &= \left(\frac{\left|\mathbf{M}\left(\boldsymbol{\varphi}_{k+1}\right)\right|^{1/2}}{(2\pi)^{M/2} \left|\mathbf{M}\left(\boldsymbol{\varphi}_{k}\right)\right|} \right)^{\alpha_k} \\
&\quad \left(\int_{\mathbb{R}^M} \exp\left\{ -\frac{1}{2} \left(\mathbf{x}_i^T \mathbf{M}_k \mathbf{x}_i - 2\mathbf{g}_k^T \mathbf{x}_i\right)\right\} d\mathbf{x}_i \right)^{\alpha_k} \\
&\quad \times \exp\left\{ -\frac{\alpha_k}{2} \mathbf{f}^T(\boldsymbol{\nu}_k) 2 \left(\mathbf{M}\left(\boldsymbol{\varphi}_k\right)\right)^{-1} \mathbf{f}(\boldsymbol{\nu}_k)\right\} \\
&\quad \times \exp\left\{ \frac{\alpha_k}{2} \mathbf{f}^T(\boldsymbol{\nu}_{k+1}) \left(\mathbf{M}\left(\boldsymbol{\varphi}_{k+1}\right)\right)^{-1} \mathbf{f}(\boldsymbol{\nu}_{k+1})\right\},
\end{aligned}
$$

where $\mathbf{M}_k = \left(2\left(\mathbf{M}\left(\boldsymbol{\varphi}_k\right)\right)^{-1} - \left(\mathbf{M}\left(\boldsymbol{\varphi}_{k+1}\right)\right)^{-1}\right)$ and
$\mathbf{g}_k = 2\left(\mathbf{M}\left(\boldsymbol{\varphi}_k\right)\right)^{-1} \mathbf{f}(\boldsymbol{\nu}_k) - \left(\mathbf{M}\left(\boldsymbol{\varphi}_{k+1}\right)\right)^{-1} \mathbf{f}(\boldsymbol{\nu}_{k+1})$. The integral above has a finite value for $\mathbf{M}_k$ positive definite (pd). Hence, and after some straightforward algebraic derivations, we obtain the expression in (4.41). The case $\alpha_k < 0$ is obtained proceeding similarly as above.

For $t_k + \alpha_k > t_{k+1} + \alpha_{k+1}$, using Equation (4.19), we have that $[\mathbf{\Phi}]_{kk+1}$is given as follows:

$$
\begin{aligned}
&[\mathbf{\Phi}_{kk+1} = \\
&\left(\int_{\mathbb{R}^M} \frac{\left|\mathbf{M}\left(\boldsymbol{\varphi}_{k+1}\right)\right|^{1/2}}{(2\pi)^{M/2}\left|\mathbf{M}\left(\boldsymbol{\varphi}_{k}\right)\right|^{1/2}\left|\mathbf{M}\left(\boldsymbol{\varphi}_{k+2}\right)\right|^{1/2}} \right. \\
&\times \exp\left\{-\frac{1}{2}\left(\mathbf{x}_i - \mathbf{f}(\boldsymbol{\nu}_k)\right)^T \left(\mathbf{M}\left(\boldsymbol{\varphi}_{k}\right)\right)^{-1}\left(\mathbf{x}_i - \mathbf{f}(\boldsymbol{\nu}_k)\right)\right\} \\
&\times \exp\left\{-\frac{1}{2}\left(\mathbf{x}_i - \mathbf{f}(\boldsymbol{\nu}_{k+2})\right)^T \left(\mathbf{M}\left(\boldsymbol{\varphi}_{k+2}\right)\right)^{-1}\left(\mathbf{x}_i - \mathbf{f}(\boldsymbol{\nu}_{k+2})\right)\right\} \\
&\times \exp\left\{\frac{1}{2}\left(\mathbf{x}_i - \mathbf{f}(\boldsymbol{\nu}_{k+1})\right)^T \left(\mathbf{M}\left(\boldsymbol{\varphi}_{k+1}\right)\right)^{-1}\left(\mathbf{x}_i - \mathbf{f}(\boldsymbol{\nu}_{k+1})\right)\right\} \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad d\mathbf{x}_i)^{\beta_k}
\end{aligned}
$$

$$
\begin{aligned}
[\mathbf{\Phi}_{kk+1} \quad = \quad & \left(\frac{\left|\mathbf{M}\left(\boldsymbol{\varphi}_{k+1}\right)\right|^{1/2}}{(2\pi)^{M/2}\left|\mathbf{M}\left(\boldsymbol{\varphi}_{k}\right)\right|^{1/2}\left|\mathbf{M}\left(\boldsymbol{\varphi}_{k+2}\right)\right|^{1/2}}\right)^{\beta_k} \\
& \left(\int_{\mathbb{R}^M} \exp\left\{-\frac{1}{2}\left(\mathbf{x}_i^T \overline{\mathbf{M}}_k \mathbf{x}_i - 2\overline{\mathbf{g}}_k^T \mathbf{x}_i\right)\right\} d\mathbf{x}\right)^{\beta_k} \\
& \times \exp\left\{-\frac{\beta_k}{2}\mathbf{f}^T(\boldsymbol{\nu}_k)\left(\mathbf{M}\left(\boldsymbol{\varphi}_{k}\right)\right)^{-1}\mathbf{f}(\boldsymbol{\nu}_k)\right\} \\
& \times \exp\left\{-\frac{\beta_k}{2}\mathbf{f}^T(\boldsymbol{\nu}_{k+2})\left(\mathbf{M}\left(\boldsymbol{\varphi}_{k+2}\right)\right)^{-1}\mathbf{f}(\boldsymbol{\nu}_{k+2})\right\} \\
& \times \exp\left\{\frac{\beta_k}{2}\mathbf{f}^T(\boldsymbol{\nu}_{k+1})\left(\mathbf{M}\left(\boldsymbol{\varphi}_{k+1}\right)\right)^{-1}\mathbf{f}(\boldsymbol{\nu}_{k+1})\right\}
\end{aligned}
$$

where

$$
\overline{\mathbf{M}}_k = \left(\mathbf{M}\left(\boldsymbol{\varphi}_{k}\right)\right)^{-1} + \left(\mathbf{M}\left(\boldsymbol{\varphi}_{k+2}\right)\right)^{-1} - \left(\mathbf{M}\left(\boldsymbol{\varphi}_{k+1}\right)\right)^{-1},
$$

and

$$
\begin{aligned}
\overline{\mathbf{g}}_k \quad = \quad & \left(\mathbf{M}\left(\boldsymbol{\varphi}_{k}\right)\right)^{-1}\mathbf{f}(\boldsymbol{\nu}_k) + \left(\mathbf{M}\left(\boldsymbol{\varphi}_{k+2}\right)\right)^{-1}\mathbf{f}(\boldsymbol{\nu}_{k+2}) \\
& - \left(\mathbf{M}\left(\boldsymbol{\varphi}_{k+1}\right)\right)^{-1}\mathbf{f}(\boldsymbol{\nu}_{k+1}).
\end{aligned}
$$

Hence, and after some straightforward algebraic derivations, we obtain the expression in (4.43)

References

[1] R. Young and R. O. Hession, "Three-dimensional structure of the smooth muscle in the term-pregnant human uterus," *Obstetrics and Gynecology*, vol. 93, pp. 94–99, January 1999.

[2] *McGraw-Hill Concise Dictionary of Modern Medicine*. The McGraw-Hill Companies, Inc., 2002.

[3] R. Lamont, "Looking to the future," *BJOG: An International Journal of Obstetrics and Gynaecology*, vol. 110, pp. 131–135, April 2003.

[4] J. D. Iams, F. F. Johnson, and C. Hamer, "Uterine activity and symptoms as predictors of preterm labor," *Obstetrics and Gynecology*, vol. 76, no. 1, pp. 42S–46S, 1990.

[5] H. Eswaran, H. Preissl, J. D. Wilson, P. Murphy, S. Robinson, and C. Lowery, "First magnetomyographic recordings of uterine activity with spatial-temporal information with 151-channel sensor array," *American Journal of Obstetrics and Gynecology*, vol. 187, pp. 145–151, 2002.

[6] J. Gondry, J. Duchene, and C. Marque, "First results on uterine EMG monitoring during pregnancy," *IEEE*, pp. 2609–2610, 1992.

[7] D. Devedeux, C. Marque, D. Mansour, S. Germain, and J. Duchene, "Uterine electromyography: a critical review," *American Journal of Obstetrics and Gynecology*, vol. 169, pp. 1636–1653, 1993.

[8] J. Duchene, D. Devedeux, D. Mansour, and C. Marque, "Analyzing uterine EMG: tracking instantaneous burst frequency," *IEEE Engineering in Medicine and Biology*, vol. 14, pp. 125–132, 1993.

[9] N. Radhakrishnan, J. D. Wilson, C. Lowery, H. Eswaran, and P. Murphy, "A fast algorithm for detecting contractions in uterine electromyography," *IEEE Engineering in Medicine and Biology*, vol. 19, pp. 89–94, March - April 2000.

[10] K. Horoba, J. Jezewski, J. Wrobel, and S. Graczyk, "Algorithm for detection of uterine contractions from electrohysterogram," in *Proc.23rd Int. Conference of IEEE EMBS*, pp. 461–464, 2001.

[11] J. Jezewski, K. Horoba, A. Matonia, and J. Wrobel, "Quantitative analysis of contraction patterns in electrical activity signal of pregnant uterus as an alternative to mechanical approach," *Physiological Measurement*, vol. 26, pp. 753–767, July 2005.

[12] P. S. La Rosa, H. Eswaran, C. Lowery, H. Preissl, and A. Nehorai, "Forward modeling of uterine EMG and MMG contractions," in *IFMBE Proceedings 11th World Congress on Medical Physics and Biomedical Engineering* (O. Dssel and W. C. Schlegel, eds.), vol. 25/4, (Munich, Germany), pp. 1600–1603, Springer Berlin Heidelberg, Sept. 2009.

[13] P. S. La Rosa, H. Eswaran, C. Lowery, H. Preissl, and A. Nehorai, "Forward electromagnetic model of uterine contractions during pregnancy," *submitted to Physics in Medicine and Biology*, 2010.

[14] P. S. La Rosa, A. Nehorai, H. Eswaran, C. Lowery, and H. Preissl, "Detecting uterine MMG contractions using a multiple change point detector and the K-means cluster algorithm," in *International Congress Series* (Cheyne, B. Ross, G. Stroink, and H. Weinberg, eds.), vol. 1300, pp. 745–748, June 2007.

[15] P. S. La Rosa, A. Nehorai, H. Eswaran, C. Lowery, and H. Preissl, "Detection of uterine MMG contractions using a multiple change point estimator and the K-means cluster algorithm," *Transactions of Biomedical Engineering*, vol. 55, pp. 453–467, Feb. 2008.

[16] P. S. La Rosa, A. Renaux, and A. Nehorai, "Barankin bounds for multiple change points estimation," in *Second IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, (St. Thomas, U.S. Virgin Islands.), pp. 37–40, IEEE, Dec. 2007.

[17] P. S. La Rosa, A. Renaux, C. Muravchik, and A. Nehorai, "Barankin-type lower bound on multiple change-point estimation," *To appear in Transactions on Signal Processing*, 2010.

[18] L. Tung, *A bidomain model for describing ischemic myocardial D-C potentials (Ph. D. thesis).* M.I.T. Cambridge, Massachusets., 1978.

[19] W. Miller and D. Geselowitz, "Simulation studies of the electrocardiogram, i. the normal heart," *Circulation Research*, vol. 43, no. 2, pp. 301–315, 1978.

[20] R. FitzHugh, "Impulses and physiological states in theoretical models of nerve membrane," *Biophysical J.*, vol. 1, pp. 445–466, 1961.

[21] J. Nagumo, S. Arimoto, and S. Yoshizawa, "An active pulse transmission line simulating nerve axon," in *Proc. IRE*, vol. 50, 1962.

[22] R. FitzHugh, "Mathematical models of excitation and propagation in nerve," *H.P. Schwan, ed. Biological Engineering*, vol. Chapter 1, pp. 1–85, 1969.

[23] A. Benveniste and M. Basseville, *Detection of Abrupt Changes in Signals and Dynamical Systems*. Springer-Verlag, Berlin, 1986.

[24] M. Basseville and I. V. Nikiforov, *Detection of Abrupt Changes, Theory and Application*. Englewood Cliffs, NJ: Prentice-Hall, Apr. 1993.

[25] F. Gustafsson, *Adaptive Filtering and Change Detection*. Wiley, 1 ed., Oct. 2000.

[26] H. Cramér, *Mathematical Methods of Statistics*, vol. 9 of *Princeton Mathematics*. New-York: Princeton University Press, Sept. 1946.

[27] E. W. Barankin, "Locally best unbiased estimates," *The Annals of Mathematical Statistics*, vol. 20, pp. 477–501, Dec. 1949.

[28] T. Chard and J. G. Grudzinskas, *The Uterus*. Cambridge, University Press, first ed., 1995.

[29] S. Weiss, T. Jaermann, P. Schmid, P. Staempli, P. Niederer, R. Caduff, and M. Bajka, "Three-dimensional fiber architecture of the nonpregnant human uterus determined ex vivo using magnetic resonance diffusion tensor imaging," *Anatomical Record. Part A, Discoveries in Molecular, Cellular, and Evolutionary Biology*, vol. 288, no. 1, pp. 84–90, 2006.

[30] R. C. Young, "Myocytes, myometrium, and uterine contractions.," *Annals of the New York Academy of Sciences*, vol. 1101, no. 1, pp. 72–84, 2007.

[31] H. F. Andersen and M. L. Barclay, "A computer model of uterine contractions based on dicrete contractile elements," *Obstetrics and Gynecology*, vol. 86, pp. 108–111, 1995.

[32] R. Young, "A computer model of uterine contractions based on action potential propagation and intercellular calcium waves," *Obstetrics and Gynecology*, vol. 89, pp. 604–608, 1997.

[33] L. Bursztyn, O. Eytan, A. J. Jaffa, and D. Elad, "Modeling myometrial smooth muscle contraction," *Annals of the New York Academy of Sciences*, vol. 1101, no. 1, pp. 110–138, 2007.

[34] L. Bursztyn, O. Eytan, A. J. Jaffa, and D. Elad, "Mathematical model of excitation-contraction in a uterine smooth muscle cell," *American Journal of Physiology - Cell Physiology*, vol. 292, no. 5, pp. C1816–C1829, 2007.

[35] S. Rihana and C. Marque, "Modelling the electrical activity of a uterine cell, a mathematical model apporach," in *Proc. The 3rd Eurpean Medical and Biological Engineering Conference, Prague*, 2005.

[36] R. E. Garfield and W. L. Maner, "Physiology and electrical activity of uterine contractions," *Seminars in Cell and Developmental Biology*, vol. 18, pp. 289–295, Jun. 2007.

[37] M. J. Peters, J. G. Stinstra, and M. Hendriks, "Estimation of the electrical conductivity of human tissue," *Electromagnetics*, vol. 21, pp. 545–557, 2001.

[38] R. Plonsey and D. B. Heppner, "Considerations of quasistationarity in electrophysiological systems," *Bull. Math. Biophys.*, vol. 29, no. 4, pp. 657–664, 1967.

[39] T. Kawarabayashi, T. Kishikawa, and H. Sugimori, "Effect of oxytocin on spontaneous electrical and mechanical activities in pregnant human myometrium," *American Journal of Obstetric and Gynecology*, vol. 155, pp. 671–676, Sept. 1986.

[40] H. C. Parkington, M. A. Tonta, S. P. Brennecke, and H. A. Coleman, "Contractile activity, membrane potential, and cytoplasmic calcium in human uterine smooth muscle in the third trimester of pregnancy and during labor," *American Journal of Obstetrics and Gynecology*, vol. 181, pp. 1445–1451, December 1999.

[41] M. Wikland and B. Lindblom, "Relationship between electrical and mechanical activity of the isolated term-pregnant human myometrium," *European Journal of Obstetrics and Gynecology and Reproductive Biology*, vol. 20, pp. 337–346, December 1985.

[42] P. C. Franzone, L. Guerri, and S. Tentoni, "Mathematical modeling of the excitation process in myocardial tissue: Influence of fiber rotation on wavefront propagation and otential field," *Mathematical Bioscienses*, vol. 101, pp. 155–235, 1990.

[43] M. J. Peters, J. G. Stinstra, S. Uzunbajakau, and N. Srinivasan, "Fetal magnetocardiography," *In Advances in Electromagnetic Fields in Living Systems*, vol. 4, pp. 1–40, 2005.

[44] R. Weber and F. Dickstein, "On the influence of a volume conductor on the orientation of currents in a thin cardiac issue," *Lecture Notes in Computer Science*, vol. 2674, pp. 111–121, December 2003.

[45] H. Eswaran, H. Preissl, J. D. Wilson, P. Murphy, and C. Lowery, "Prediction of labor in term and preterm pregnancies using non-invasive magnetomyographic recordings of uterine contractions," *The Journal of Maternal-Fetal and Neonatal Medicine*, vol. 190, pp. 1598–1603, 2004.

[46] M. Khalil and J. Duchene, "Uterine EMG analysis: a dynamic approach for change detection and classification," *IEEE Transactions on Biomedical Engineering*, vol. 47, pp. 748–756, June 2000.

[47] C. Ramon, H. Preissl, P. Murphy, J. D. Wilson, C. Lowery, and H. Eswaran, "Synchronization analysis of the uterine magnetic activity during contractions," *BioMedical Engineering Online*, vol. 4:55, 2005.

[48] J. Ramondt, C. van Kooten, A. Verhoeff, and H. Wallenburg, "Computer analysis of mechanical and electrical uterine activity," *Medicine Biology Engineering Computation*, vol. 24, pp. 351–355, 1986.

[49] S. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation.," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 674–693, June 1989.

[50] J. Duchene and F. Goubel, "Surface electromyogram during voluntary contraction: processing tools and relation to physiological events," *Critical review biomedical engineering*, vol. 21, pp. 313–397, June 1993.

[51] G. Inbar, J. Allin, O. Paiss, and H. Kranz, "Monitoring surface EMG spectral changes by the zero crossing rate," *Medicine Biology Engineering Computation*, vol. 31, p. 597, 1984.

[52] L. J. Vostrikova, "Detecting 'disorder' in multidimensional random processes," *Soviet Mathematics Doklady*, vol. 24, pp. 55–59, 1981.

[53] J. Chen and A. K. Gupta, "Testing and locating variance changepoints with applications to stock," *Journal of the American Statistical Association*, vol. 92, pp. 739–747, June 1997.

[54] H. Akaike, "Information theory and an extension of the maximum likelihood principle," in *Proceedings of the 2nd international symposium of information theory* (B. N. Petrov and B. A. K. E. Csaki, eds.), pp. 267–281, 1973.

[55] G. Schwarz, "Estimating the dimension of a model," *The Annals of Statistics*, vol. 6, pp. 461–464, March 1978.

[56] B. Kedem, "Spectral analysis and discrimination by zero-crossings," *Proceedings of the IEEE*, vol. 74, pp. 282–304, November 1986.

[57] K. Horoba, J. Jezewski, J. Wrobel, and S. Graczyk, "Some methods for classification and analysis of multivariate observations," in *Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability*, no. 1, pp. 281–297, Berkeley, University of California Press, 1967.

[58] C. M. Bishop, *Neural Networks for Pattern Recognition.* Oxford, England: Oxford University Press, 1995.

[59] R. Nagarajan, H. Eswaran, J. D. Wilson, P. Murphy, C. Lowery, and H. Preissl, "Analysis of uterine contractions: a dynamical approach," *The Journal of Maternal-Fetal and Neonatal Medicine*, vol. 14, pp. 8–21, March 2003.

[60] B. Brodskya and B. Darkhovskyb, *Non-Parametric Statistical Diagnosis. Problems and Methods.* Kluwer Academic Publishers Dordrecht/Boston/London, 2000.

[61] B. Brodskya and B. Darkhovskyb, "Asymptotically optimal methods of change-point detection for composite hypotheses," *Journal of Statistical Planning and Inference*, vol. 133, pp. 123–138, 2005.

[62] R. Kakarala and A. O. Hero, "On achievable accuracy in edge localization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, pp. 777–781, July 1992.

[63] A. Bartov and H. Messer, "Analysis of inherent limitations in localizing step-like singularities in a continuous signal," in *Proc. of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, (Paris, FR), pp. 21–24, June 1996.

[64] A. M. Reza and M. Doroodchi, "Cramér-Rao lower bound on locations of sudden changes in a steplike signal," *IEEE Transactions on Signal Processing*, vol. 44, pp. 2551–2556, Oct. 1996.

[65] J.-Y. Tourneret, M. Chabert, and M. Ghogho, "Detection and estimation of multiplicative jumps," in *Proc. of the 8th IEEE Signal Processing Workshop on Statistical Signal and Array Processing*, pp. 20–23, June 1996.

[66] A. Bartov and H. Messer, "Lower bound on the achievable DSP performance for localizing step-like continuous signals in noise," *IEEE Transactions on Signal Processing*, vol. 46, pp. 2195–2201, Aug. 1998.

[67] A. Swami and B. Sadler, "Analysis of multiscale products for step detection and estimation," *IEEE Transactions on Information Theory*, vol. 44, pp. 1043–1051, Apr. 1999.

[68] S. Zacks, *The Theory of Statistical Inference.* J. Wiley & Sons, N. York, 1971.

[69] A. Swami and B. Sadler, "Cramér-Rao bounds for step-change localization in additive and multiplicative noise," in *Proceedings IEEE-SP Workshop on Statistical Signal and Array Proc., Portland, OR*, pp. 403–406, Sept. 1998.

[70] J. V. Braun and H. G. Muller, "Statistical methods for DNA sequence segmentation," *Statistical Science*, vol. 13, no. 2, pp. 142–162, 1998.

[71] C. H. Wu and C. H. Hsieh, "Multiple change-point audio segmentation and classification using an MDL-based Gaussian model," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 14, pp. 647–657, Mar. 2006.

[72] N. Dobigeon, J.-Y. Tourneret, and J. D. Scargle, "Change-point detection in astronomical data by using a hierarchical model and a Bayesian sampling approach," in *IEEE/SP 13th Workshop on Statistical Signal Processing*, pp. 369–374, July 2005.

[73] D. G. Chapman and H. Robbins, "Minimum variance estimation without regularity assumptions," *The Annals of Mathematical Statistics*, vol. 22, pp. 581–586, Dec. 1951.

[74] L. Knockaert, "The Barankin bound and threshold behaviour in frequency estimation," *IEEE Transactions on Signal Processing*, vol. 45, pp. 2398–2401, Sept. 1997.

[75] A. Ferrari and J.-Y. Tourneret, "Barankin lower bound for change-points in independent sequences," in *Proc. of IEEE Workshop on Statistical Signal Processing (SSP)*, (St. Louis, MO, USA), pp. 557–560, Sept. 2003.

[76] J. M. Hammersley, "On estimating restricted parameters," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 12, no. 2, pp. 192–240, 1950.

[77] B. Burgeth, A. Bruhna, N. Papenberga, M. Welka, and J. Weickert, "Mathematical morphology for matrix fields induced by the Loewner ordering in higher dimensions," *Signal Processing*, vol. 87, pp. 277–290, Feb. 2007.

[78] F. Pukelsheim, *Optimal Design of Experiments*. John Wiley & Sons, INC., 1st ed., 1993.

[79] J. Borwein and A. Lewis, *Convex Analysis and Nonlinear Optmization: Theory and Examples*. Springer-Verlag, 1st ed., 2000.

[80] J. D. Gorman and A. O. Hero, "Lower bounds for parametric estimation with constraints," *IEEE Transactions on Information Theory*, vol. 26, Nov. 1990.

[81] I. Reuven and H. Messer, "A Barankin-type lower bound on the estimation error of a hybrid parameter vector," *IEEE Transactions on Information Theory*, vol. 43, pp. 1084–1093, May 1997.

[82] R. J. McAulay and L. P. Seidman, "A useful form of the Barankin lower bound and its application to PPM threshold analysis," *IEEE Transactions on Information Theory*, vol. 15, pp. 273–279, Mar. 1969.

[83] R. J. McAulay and E. M. Hofstetter, "Barankin bounds on parameter estimation," *IEEE Transactions on Information Theory*, vol. 17, pp. 669–676, Nov. 1971.

[84] J. S. Abel, "A bound on mean square estimate error," *IEEE Transactions on Information Theory*, vol. 39, pp. 1675–1680, Sept. 1993.

[85] J. Tabrikian and J. Krolik, "Barankin bounds for source localization in an uncertain ocean environment," *IEEE Transactions on Signal Processing*, vol. 47, pp. 2917–2927, Nov. 1999.

[86] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[87] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming (web page and software)," *http://stanford.edu/ boyd/cvx*, December 2008.

[88] M. Grant and S. Boyd, "Graph implementations for nonsmooth convex programs," *Recent Advances in Learning and Control (a tribute to M. Vidyasagar)*, December 2008.

[89] G. Strang, *Introduction to Linear Algebra*. Wellesley, MA : Wellesley-Cambridge Press, 2005.

[90] T. Muir, *A treatise on the theory of determinants*. Dover Publications, 1960.

[91] P. Swerling, "Probability of detection for fluctuating targets," *IEEE Transactions on Information Theory*, vol. 6, pp. 269–308, Apr. 1960.

[92] J. F. Böhme, "Estimation of source parameters by maximum likelihood and non linear regression," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 731–734, 1984.

[93] J. F. Böhme, "Estimation of spectral parameters of correlated signals in wavefields," *Signal Processing*, vol. 10, pp. 329–337, 1986.

[94] Z. Liu, A. Nehorai, and E. Paldi, "Statistical analysis of a generalized compound eye detector array," in *Proc. 4th IEEE Sensor Conference*, Oct.-Nov. 2005.

[95] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume 2: Detection Theory*. Prentice Hall PTR, 1st ed., Jan. 1998.

[96] D. V. Hinkley, "Inference about the change-point in a sequence of random variables," *Biometrika*, vol. 57, no. 1, pp. 1–17, 1970.

[97] Y.-C. Yao and S. T. Au, "Least-squares estimation of a step function," *Sankhya: The Indian Journal of Statistics, Series A*, vol. 51, pp. 370–381, October 1989.

[98] R. A. Davis, D. Huang, and Y. Yao, "Testing for a change in the parameter values and order of an autoregressive model," *The Annals of Statistics*, vol. 23, pp. 282–304, February 1995.

Vita

Patricio Salvatore La Rosa Araya
patricio.larosa@go.wustl.edu

Campus Box 8005, 660 South Euclid Avenue
Division of General Medical Sciences
Department of Internal Medicine
Washington University School of Medicine
St. Louis, MO 63112

Summary: Ph.D. in Electrical Engineering from the Electrical and Systems Engineering Department at Washington University in St. Louis (WUSTL) USA, August 2010. Ph.D. work involves modeling and detection of uterine contractions using magnetomyography and derivations of lower bounds on the mean-squared error for 1-D segmentation algorithms, under the guidance of Professor Arye Nehorai. Recipient of student travel-support award, following presentation of article at IEEE CAMSAP, 2007. Co-advisor of the undergraduate research project awarded first-place at the St. Louis Area Undergraduate Research Symposium (STLAURS), 2009. Departmental representative responsible for accompanying six ESE undergraduate students during their study abroad program on Medical Imaging at the University of Tbingen MEG-Center, and the Max Planck Institute for Biological Cybernetics, Germany, 2009. M. Sc. degree in Electrical Engineering from University of Chile with maximum distinction, 2003.

Research Interests: Statistical Array Signal Processing and Applications to Medical Imaging and Biosignal Processing; Physical Modeling, including Bioelectromagnetism, Optical Mapping, and Non-linear Dynamics.

Education:

- *2*003 to 2010: Ph.D. in Electrical Engineering, Washington University in St. Louis; Adviser: Prof. Arye Nehorai; Dissertation: *Modeling and Detection of Uterine Contractions using Magnetomyography*;
- *2*010: M.Sc. in Electrical Engineering, Washington University in St. Louis; Adviser: Prof. Arye Nehorai
- *2*001 to 2003: M.Sc. in Electrical Engineering, University of Chile, Santiago, Chile ; Adviser: Prof. Guillermo Gonzalez R.; Specialization: System Identification and Control; Thesis: *Recognition of handwritten numbers in the principal component space and its orthogonal extensions*
- *1*995 to 2000: B.Sc. in Engineering at the Pontifical Catholic University of Chile, Santiago, Chile; Major: Civil Industrial and Electrical Engineering

Research Experience:

- *A*ugust 2010-: Postdoctoral Research Associate in the General Medical Sciences Division of the Department of Internal Medicine at the Washington University School of Medicine; Advisor: Dr. William Shannon
- *J*an. 2006 to July 2010: Research Assistant at the Center for Sensor Signal and Information Processing of the Electrical and Systems Engineering Department at Washington University in St. Louis; Adviser: Prof. Arye Nehorai; Research on detection and modeling of uterine contractions using magnetomyography
- *A*ug. 2003 to Dec. 2005: Research Assistant at the Signal and Image Research Laboratory of the Electrical and Computer Engineering Department at the University of Illinois at Chicago; Adviser: Prof. Arye Nehorai; Research on robust mean-squared error beamforming for narrowband and wideband signals
- *M*arch 2002 to June 2003: Research Assistant at the Electrical Engineering Department, University of Chile; Adviser: Prof. Guillermo Gonzalez; Research on fault detection and diagnosis in time-variable nonlinear dynamic systems using wavelet transform; Project funded by the National Agency of Science and Technology of Chile: FONDECYT 1020741
- *A*ug. 2000 to March 2002: Research Assistant at the Electrical Engineering Department, University of Chile; Adviser: Prof. Guillermo Gonzalez; Research on

system identification theory to pattern recognition of handwritten numbers using statistical quadratics discriminators and neural networks; Project funded by the National Agency of Science and Technology of Chile: FONDECYT 1000977

Awards:

- *J*an. 2008: Recipient of student travel-support, following evaluation and presentation of conference paper, at the Second IEEE International Workshop on Computational advances in Multi-Sensor Adaptive Processing (CAMSAP)
- *M*ay 2003: Graduated with maximum distinction from the Master in Sciences Program in Electrical Engineering at University of Chile
- *1*995-2000: Recipient of the John Paul II Foundation scholarship for pursuing undergraduate studies in Engineering Sciences at the Pontifical Catholic University of Chile

Activities and Professional Affiliation:

- Reviewer for: *IEEE Transactions on Signal Processing* and *IEEE Transactions on Biomedical Engineering*
- Student Member of IEEE Signal Processing Society since 2003

Dissertation and Thesis:

- P. S. La Rosa, "Modeling and Detection of Uterine Contractions using Magnetomyography," Dissertation to obtain the degree of Ph.D. in Electrical Engineering, Washington University in St. Louis, MO, USA. August, 2010
- P. S. La Rosa, "Filtering and Recognition of Manuscript and Printed Digits in the Principal Component Space and its Orthogonal Extensions", (in Spanish) Thesis to obtain the degree of M. Sc. In Electrical Engineering, University of Chile, Santiago, Chile, May, 2003

Journal Papers:

1. P. S. La Rosa, A. Renaux, C. H. Muravchik, and A. Nehorai, "Barankin-type lower bound for multiple change-point estimation," *t*o appear in IEEE Trans. on Signal Processing, 2010

2. P. S. La Rosa, A. Nehorai, H. Eswaran, C. Lowery, and H. Preissl, "Detection of uterine MMG contractions using a multiple change-point estimator and the K-means cluster algorithm," *I*EEE Trans. on Biomedical Engineering, Vol. 55 (1), No. 2, pp. 453-467, Feb. 2008

3. Y. Eldar, A. Nehorai, and P. S. La Rosa, "A competitive mean-squared error approach to beamforming," *I*EEE Trans. on Signal Processing, Vol. 55, pp. 5143 - 5154, Nov. 2007

4. Y. Eldar, A. Nehorai, and P. S. La Rosa, "An expected least-squares beamforming approach to signal estimation with steering vector uncertainties," *I*EEE Signal Processing Letters, Vol. 13, No. 5, pp. 288-291, May 2006

5. M. Duarte and P. La Rosa, "Designing SISO observers with unknown input," *J*ournal of Mathematical Control and Information (IMA), Vol. 20, No. 4, pp. 387-391, December 2003

6. G. D. Gonzalez, C. Perez, and P. La Rosa, "Enfoque global de la identificacin de sistemas mediante modelos de caja negra,(Global approach for system identification by black box models)," *Anales del Instituto de Ingenieros de Chile (Review of Chilean Institute of Engineers)*, Vol. 114, No. 2, pp. 47-61, August 2002

Journal Papers in Preparation:

1. P. S. La Rosa, H. Eswaran, H. Preissl, and A. Nehorai, "Forward electromagnetic model of uterine contractions during pregnancy" *s*ubmitted to Physics in Medicine and Biology, Jun. 2010

Conference Papers:

1. P. S. La Rosa, H. Eswaran, C. Lowery, H. Preissl, and A. Nehorai, "Forward modeling of uterine EMG and MMG contractions," *i*n IFMBE Proceedings 11th

World Congress on Medical Physics and Biomedical Engineering, O. Dssel and W. C. Schlegel, Eds., Vol. 25/4, Munich, Germany, pp. 1600-1603, Springer Berlin Heidelberg, Sept. 2009

2. P. S. La Rosa, A. Renaux, A. Nehorai, "Barankin bounds for multiple change points estimation," *i*n Proceedings of the Second IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), St. Thomas, U.S. Virgin Islands, pp. 35-40, Dec. 2007 (*Invited paper for special session on minimal bounds in estimation; Recipient of student travel-support award following presentation at CAMSAP*)

3. P. S. La Rosa, A. Nehorai, H. Eswaran, C. Lowery, and H. Preissl, "Detection of uterine MMG contractions using a multiple change-point estimator and K-means cluster algorithm," *I*nternational Congress Series, Vol. 1300, pp. 745-748, Jun. 2007

4. G. D. Gonzalez, G. Ceballos, R. Paut, D. Miranda, P. La Rosa, "Fault detection and identification through variance of wavelet transform of system outputs," *i*n Recent Advances in Intelligent Systems and Signal Processing, N. E. Mastorakis, C. Manikoupoulos, G. E. Antoniou, V. M. Mladenov, and I. F. Gonos, Eds. Athens, Greece: WSEAS Press, pp. 47-53, 2003

5. R. Salazar, P. Rodrguez, M. A. Duarte-Mermoud, P. S. La Rosa, M. A. Bustos and R. I. Gonzlez, "Control of a levitation system by adaptive feedback passivity," *P*roceedings of the 10 Latin-American Congress on Automatic Control, Guadalajara, Mexico, December 3-6, 2002. Proceedings in CD, Track No. 156, 6 pages

6. M. Duarte and P. La Rosa, "MRAC using observers with unknown inputs," *I*n Recent Advances in Circuits, Systems and Signal Processing, E. Mastorakis and G. Antoniou Eds. Electrical and Computer Engineering Series, WSEAS Press, pp. 238-243, July 2002 (ISBN: 960-8052-64-5)

7. G. D. Gonzalez, C. Perez, P. Rodriguez and P. La Rosa, "Linear regression applied to digit recognition," *P*roceedings of the Third International Conference on Intelligent Processing and Manufacturing of Materials (IPMM2001), Vancouver, Canada, July 2001 (CD ROM Proceedings)

Talks & Poster Presentations:

1. "Detection and modeling of uterine contractions using magnetomyography," invited talk at the Department of Radiology, WUSTL, May 6 th, 2010

2. "Detection and modeling of uterine MMG," invited talk at the MEG-Center, U. Tubingen, Germany, May 14 th, 2009

3. "Detection of uterine contractions using MMG," invited talk at the Statistics Seminar of the Mathematics Department, WUSTL, Jan. 20 th, 2009

4. "Detection of uterine contractions using MMG," invited talk at the Electrical Engineering Department, Pontifical Catholic University of Chile, Santiago, Chile, Jan. 16 th, 2009

5. "Forward modeling of uterine contractions using EMG and MMG," poster presentation at the Imaging Science Pathway Retreat: 21st-century Imaging Sciences: Pathways to Discovery, St. Louis, MO, April 18 th, 2008

6. "Barankin bounds for multiple change-point estimation," presentation at the Special Session on Minimal Bounds in Estimation, Second International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), St. Thomas, U.S. Virgin Islands, Dec. 12 th, 2007

Teaching Experience:

- *A*ug. 2007 to May 2010: Instructor of ESE 497 Undergraduate Research at the Department of Electrical and Systems Engineering, Washington University in St. Louis; Responsible for teaching and guiding undergraduate students in the execution of research projects in the areas of microphone array signal processing and robotic sensing

- *M*ay 2007 to Dec. 2009: Instructor of study groups on Detection and Estimation Theory, Introduction to Mathematical Statistics, Linear Statistical Models, Multivariate Analysis, and Linear Algebra at the Center for Signal and Sensing Information Processing, Department of Electrical and Systems Engineering, Washington University in St. Louis; Responsible for teaching as well as assigning, grading, and correcting homework for junior graduate students; Responsible for preparing lecture notes on Introduction to Mathematical Statistics, Linear Statistical Models, and Multivariate Analysis.

- *A*ug. 2001 to July 2003: Lecturer of MMS620 Microprocessors and Microntrollers, Programs of Applied Electrical Engineering and Industrial Process and Automatization Engineering, Professional Institute DUOC U.C. Santiago, Chile; Responsible for teaching as well as assigning, grading, and correcting homework for undergraduate students

- *A*ug. 2001 to July 2003: Lecturer of EIS411 Power Electronics, Programs of Applied Electrical Engineering and Industrial Process and Automatization Engineering, Professional Institute DUOC U.C. Santiago, Chile; Responsible for teaching as well as assigning, grading, and correcting homework for undergraduate students

- *M*ar. 2002 to July 2002: Teaching Assistant of IEE3633 System Design in Automatic Control at the Department of Electrical Engineering, Pontifical Catholic University of Chile, Santiago, Chile; Responsible for teaching helping sessions as well as grading homework

- *M*ar. 2002 to July 2002: Teaching Assistant of EM717 System Identification for Control Systems at the Department of Electrical Engineering, University of Chile, Santiago, Chile; Responsible for teaching helping sessions as well as grading homework

- *A*ug. 2001 to Dec. 2001: Teaching Assistant of EL63E Optimization for Automatic Control at the Department of Electrical Engineering, University of Chile, Santiago, Chile; Responsible for teaching helping sessions as well as grading homework

- *M*ar. 2001 to July 2001: Teaching Assistant of EL42D Control Systems at the Department of Electrical Engineering, University of Chile, Santiago, Chile; Responsible for teaching helping sessions as well as grading homework

- *M*ar. 2001 to July 2001: Teaching Assistant of IEE2512 Telecommunications at the Department of Electrical Engineering, Pontifical Catholic University of Chile, Santiago, Chile; Responsible for grading homework

Advising Experience:

- *J*une 2009 to Dec. 2009: Undergraduate research co-advisor of the former ME student Chase La-Font; Project title: "Robotic Microphone Sensing: Optimizing Source Estimation and Algorithms for Adaptive Control," Electrical and Systems Engineering Department, WUSTL

- *J*une 2009 to Dec. 2009: Undergraduate research co-advisor of the ESE students Evan Nixon and Ian Beil; Project title: "Source Localization: Acoustic Vector-sensor Array Processing," Electrical and Systems Engineering Department, WUSTL

- *J*une 2009 to Dec. 2009: Undergraduate research co-advisor of the ESE students Raphael Schwartz and Zachary Knudsen; Project title: "Robotic Microphone Sensing: Data Processing Architectures for Real-Time, Acoustic Source Position Estimation," Electrical and Systems Engineering Department, WUSTL

- *J*une 2009 to Dec. 2009: Undergraduate research co-advisor of the ESE student Joy Chiang; Project title: "Characterizing Odors Using Electronic Nose Sensors," Electrical and Systems Engineering Department, WUSTL

- *M*ay 2009: Departmental representative responsible for accompanying six ESE undergraduate students during their study abroad program on Medical Imaging at the University of Tbingen MEG-Center, and the Max Planck Institute for Biological Cybernetics, Germany

- *J*an. 2009 to May 2009: Undergraduate research co-advisor of the former EE student Joshua York; Project title: "3D Acoustic Source Location Using Acoustic Vector Sensors," Electrical and Systems Engineering Department, WUSTL

- *J*une 2008 to Dec. 2009: Undergraduate research co-advisor of the former EE student Joshua York; Project title: "Acoustic Source Localization Using Microphone Array and Cross-Correlation Methods," Electrical and Systems Engineering Department, WUSTL. *Winner of first place award for best overall presentation at St. Louis Area Undergraduate Research Symposium (STLAURS), April 2009*

- *2*002 to 2003: Undergraduate thesis advisor of the former EE students Claudio Montagno and Miguel Dussaillant; Project title: "Methods for command recognition based on acoustic and articulation techniques." Thesis to obtain the degree in Applied Engineering, Program of Industrial Process and Automatization Engineering, DUOC U.C. Chile

August 2010

Biomagnetic Analysis of Uterine Activity, La Rosa Araya, Ph.D. 2010