Bayesian Active Learning for Personalization and Uncertainty Quantification in Cardiac Electrophysiological Model

Jwala Dhamala
jd1336@rit.edu

Follow this and additional works at: https://scholarworks.rit.edu/theses

Recommended Citation
Dhamala, Jwala, "Bayesian Active Learning for Personalization and Uncertainty Quantification in Cardiac Electrophysiological Model" (2020). Thesis. Rochester Institute of Technology. Accessed from

This Dissertation is brought to you for free and open access by RIT Scholar Works. It has been accepted for inclusion in Theses by an authorized administrator of RIT Scholar Works. For more information, please contact ritscholarworks@rit.edu.
Bayesian Active Learning for
Personalization and Uncertainty Quantification in
Cardiac Electrophysiological Model

by

Jwala Dhamala

A dissertation submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
in Computing and Information Sciences

B. Thomas Golisano College of Computing and
Information Sciences

Rochester Institute of Technology
Rochester, New York
February 3, 2020
Bayesian Active Learning for Personalization and Uncertainty Quantification in Cardiac Electrophysiological Model

by

Jwala Dhamala

Committee Approval:
We, the undersigned committee members, certify that we have advised and/or supervised the candidate on the work described in this dissertation. We further certify that we have reviewed the dissertation manuscript and approve it in partial fulfillment of the requirements of the degree of Doctor of Philosophy in Computing and Information Sciences.

Dr. Linwei Wang
Dissertation Advisor
Date

Dr. Daniel Ornt
External Chair
Date

Dr. Christopher Kanan
Committee Member
Date

Dr. Natalia Trayanova
Committee Member
Date

Dr. Nathan D. Cahill
Committee Member
Date

Dr. Pengcheng Shi
Committee Member
Date

Certified by:

Dr. Pengcheng Shi
Director, Computing and Information Sciences
Date
Bayesian Active Learning for Personalization and Uncertainty Quantification in Cardiac Electrophysiological Model

by

Jwala Dhamala

Abstract

Cardiacvascular disease is the top death causing disease worldwide. In recent years, high-fidelity personalized models of the heart have shown an increasing capability to supplement clinical cardiology for improved patient-specific diagnosis, prediction, and treatment planning. In addition, they have shown promise to improve scientific understanding of a variety of disease mechanisms.

However, model personalization by estimating the patient-specific tissue properties that are in the form of parameters of a physiological model is challenging. This is because tissue properties, in general, cannot be directly measured and they need to be estimated from measurements that are indirectly related to them through a physiological model. Moreover, these unknown tissue properties are heterogeneous and spatially varying throughout the heart volume presenting a difficulty of high-dimensional (HD) estimation from indirect and limited measurement data. The challenge in model personalization, therefore, summarizes to solving an ill-posed inverse problem where the unknown parameters are HD and the forward model is complex with a non-linear and computationally expensive physiological model.

In this dissertation, we address the above challenge with following contributions. First, to address the concern of a complex forward model, we propose the surrogate modeling of the complex target function containing the forward model – an objective function in deterministic estimation or a posterior probability density function in probabilistic estimation – by actively selecting a set of training samples and a Bayesian update of the prior over the target function. The efficient and accurate surrogate of the expensive target function obtained in this manner is then utilized to accelerate either deterministic or probabilistic parameter estimation. Next, within the framework of Bayesian active
learning we enable active surrogate learning over a HD parameter space with two novel approaches: 1) a multi-scale optimization that can adaptively allocate higher resolution to heterogeneous tissue regions and lower resolution to homogeneous tissue regions; and 2) a generative model from low-dimensional (LD) latent code to HD tissue properties. Both of these approaches are independently developed and tested within a parameter optimization framework. Furthermore, we devise a novel method that utilizes the surrogate pdf learned on an estimated LD parameter space to improve the proposal distribution of Metropolis-Hastings for an accelerated sampling of the exact posterior pdf. We evaluate the presented methods on estimating local tissue excitability of a cardiac electrophysiological model in both synthetic data experiments and real data experiments. Results demonstrate that the presented methods are able to improve the accuracy and efficiency in patient-specific model parameter estimation in comparison to the existing approaches used for model personalization.
Acknowledgments

I would like to thank my advisor Linwei Wang for providing me the freedom to explore topics of my interest, shaping the direction of my dissertation and teaching me to be a better researcher. I am grateful for her guidance and insightful feedbacks. I would like to thank Natalia Trayanova, Pengcheng Shi, Nathan Cahill, Christopher Kanan and Daniel Ornt for serving in my dissertation committee and setting time aside to discuss with me. I am especially thankful to Christopher Kanan who motivated and discussed with me during the latter part of my PhD. I am very grateful for the opportunity to collaborate with Natalia Trayanova and Dana Brooks. I am thankful to all my collaborators and coauthors.

I would like to thank my parents who supported me in all my academic pursuit and taught me the value of patience and hard work. I want to thank my big sister who served as a role model in my early years and taught me various academic and non-academic skills. I especially want to thank my husband Kushal Kafle who supported and encouraged me throughout the PhD beginning with the application to PhD programs upto defending my PhD dissertation. Thank you for engaging with me on scientific discussions while being a constant source of love and happiness. Without Kushal this journey would have been much more difficult and this dissertation would not have taken the current direction.

I would like to thank all the undergraduate and graduate students that I worked with during my PhD. I have learned a lot about research and life from them. Finally, I have been very fortunate to have my friends especially Emily Myers, Colin Wielga, Ryan Thomas Ford, Robik Shrestha, Kishan KC, Jared Van Cor, Roland Sanford, Nuthan Munaiah, Xuan Guo and Azar Rahimi around me.
To my husband Kushal Kafle
## Contents

1 Introduction ............................................. 1  
  1.1 Problem Definition ................................. 2  
  1.2 Objectives and Contributions ..................... 4  
  1.3 Outline ............................................. 5  

2 Background and Related Works .......................... 7  
  2.1 Models of Cardiac Electrophysiology ............... 7  
     2.1.1 Anatomical Model ............................ 8  
     2.1.2 Electrophysiology Model ..................... 8  
     2.1.3 Electrocardiography Model ................... 10  
  2.2 Personalization of Model Parameters ............... 13  
     2.2.1 Parameter Estimation of Complex Physiological Models 13  
     2.2.2 High-dimensional Parameter Estimation .......... 14  
     2.2.3 Bayesian Active Approximation of Expensive Functions 14  

3 Spatially Adaptive Multi-Scale Optimization .......... 17  
  3.1 Introduction ....................................... 18  
  3.2 Spatially-Adaptive Multi-Scale Optimization ....... 19  
     3.2.1 Multi-Scale Hierarchy ....................... 21  
     3.2.2 Coarse-to-Fine Optimization ............... 21  
     3.2.3 Adaptive Spatial Resolution Adjustment ...... 24  
  3.3 Evaluation on Synthetic Data ....................... 25  
     3.3.1 Performance Based on the Size and Location of the Infarct 27  
     3.3.2 Comparison with Optimization using Uniform Resolution 29  
     3.3.3 Identifiability Based on the Size of the Infarct .... 30  
     3.3.4 Identifiability Based on the Amount of Measurement Data 31
3.4 Evaluation on a Blinded EP Model and In-vivo MRI Scar 33
3.5 Evaluation on 120-lead ECG and Catheter Data 35
3.6 Discussion 36
   3.6.1 Threshold for Adaptive Coarsening 36
   3.6.2 Effect of Refining Multiple Nodes Versus a Single Node 37
   3.6.3 Parameters of the GPO 38
   3.6.4 Performance in the Presence of Two Infarcts 38
3.7 Conclusion 38

4 High-dimensional Bayesian Optimization via an Embedded Generative Model 40
   4.1 Introduction 41
   4.2 HD Parameter Estimation 42
      4.2.1 LD-to-HD Parameter Generation via VAE 43
      4.2.2 Bayesian Optimization with Embedded Generative Model 44
   4.3 Evaluation on Synthetic Data 45
      4.3.1 VAE Architecture and Training 45
      4.3.2 Comparison with Existing Methods 46
      4.3.3 The Effect of VAE-encoded Knowledge About the LD Manifold 47
   4.4 Evaluation on 120-lead ECG and Catheter Data 48
   4.5 Conclusion 49

5 Bayesian Optimization on Large Graphs via a Graph Convolutional Generative Model 50
   5.1 Introduction 51
   5.2 Personalizing HD Parameters on Unstructured Meshes 52
      5.2.1 Graph Convolutional VAE 52
      5.2.2 Bayesian Optimization on Large Graphs 54
   5.3 Synthetic Experiments 55
      5.3.1 gVAE as a Generative Model 56
      5.3.2 gVAE-based Parameter Optimization 57
      5.3.3 Feature Sharing Across Geometries 58
   5.4 Real Data Experiments 58
   5.5 Conclusion and Future Work 59
6 Gaussian Process-based Markov Chain Monte Carlo 60
   6.1 Introduction ................................................. 61
   6.2 Probabilistic Parameter Estimation ........................... 63
       6.2.1 Active Construction of the GP Surrogate .......... 64
       6.2.2 GP Surrogate Accelerated Metropolis-Hastings ...... 67
   6.3 Evaluation on Synthetic Data ................................. 71
       6.3.1 Validation of the Accuracy and Efficiency .......... 72
       6.3.2 Comparison with Directly Sampling the GP Surrogate . 72
       6.3.3 Analysis of the Sampled Posterior Distributions .... 74
   6.4 Evaluation on a Blinded EP Model and In-vivo MRI Scar .... 79
   6.5 Evaluation on 120-lead ECG and Catheter Data ................ 81
   6.6 Discussion .................................................... 84
       6.6.1 Effect of the Gaussian Proposal Distribution ...... 84
       6.6.2 Quality of the Surrogate vs. Acceptance Rate .... 85
       6.6.3 Related Works ............................................ 86
   6.7 Conclusion .................................................... 88

7 Conclusion and Future Work 89
   7.1 Limitations ................................................... 90
   7.2 Future Work .................................................. 91
List of Figures

2.1 A pipeline showing patient-specific anatomy generation. .................. 8
2.2 Result of one-factor-at-a-time sensitivity analysis of the AP model. Mean (solid black) and standard deviation (std) (shaded blue) throughout the action potential duration resulting from the variation in each model parameter. .................................................. 10
2.3 A schematic diagram of normal sinus rhythm for a human heart as seen on ECG [77]. ................................................................. 12

3.1 A work-flow diagram of the spatially adaptive multi-scale optimization framework. ................................................................. 20
3.2 Multi-Scale hierarchy with 20, 15, 10, 5, and 1 final clusters (left to right); color is repeated for unconnected clusters. ................. 20
3.3 Schematics of 26 pre-defined ventricular segments for optimization with uniform resolution. ............................................... 25
3.4 Synthetic experiments. Examples of the progression of the presented spatially-adaptive multi-scale optimization. ..................... 26
3.5 Synthetic experiments. Examples of tissue excitability estimated using: 1) uniform BOBYQA, 2) adaptive BOBYQA, and 3) adaptive GPO in the presence of infarcts of varying sizes and locations. .... 28
3.6 Synthetic experiments. Comparison of uniform BOBYQA, adaptive BOBYQA, and adaptive GPO in terms of DC and RMSE based on 35 different infarct cases of varying sizes and locations (bar: mean, line: standard deviation). ................................................. 28
3.7 Synthetic experiments. Comparison of the final dimensions of unknowns obtained by using presented methods based on 35 infarct cases of various sizes (bar: mean, line: standard deviation). ........ 29
3.8 Synthetic experiments. Left: tree narrows down along the infarct region showing that the infarct is identified. Right: tree takes first a few steps along the infarct region, but later splits along homogeneous healthy region showing that the infarct is not identified. 

3.9 Synthetic experiments. Performance of the presented method using 12-lead ECG in comparison to that using 120-lead ECG in terms of DC and RMSE. Comparison is based on 35 different infarct cases of varying sizes and locations (bar: mean, line: standard deviation).

3.10 Synthetic experiments. Estimation using 120-lead ECG shows good accuracy with splits along the infarct region, whereas using 12-lead ECG is unable to identify the infarct.

3.11 Results of local parameter estimation using the presented method, where measurement data is generated from epicardial potentials simulated by a multi-scale ionic electrophysiological model blinded to the presented estimation method. Infarct from 3D-MRI is used as the ground-truth.

3.12 Real-data experiments. Estimated tissue excitability using the presented adaptive methods and uniform BOBYQA. In-vivo voltage maps are used as reference data.

3.13 Real-data experiments. Comparison of the real ECG data and the simulated ECG with estimated tissue excitability.

3.14 Performance analysis of the presented framework: when threshold for spatial coarsening changes from 0.1% to 20% of the maximum gain in global optimum (left), and when two nodes vs. one node is split for spatial refinement (right).

3.15 Left: comparison of the performance of the presented method when empirical kernel parameters are used vs. when optimized kernel parameters are used. Right: examples of estimation results using the presented method in the presence of two infarcted tissue regions.

4.1 Outline of the Bayesian Optimization with an embedded Generative Model, with the dimension of each VAE layer labeled.

4.2 Comparison of BO-VAE EI Post-1 (blue bar) with: 1) FH and FS (green bars); and 2) BO-VAE using standard EI, EI Isotropic, and EI Post-K (yellow bars) in terms of DC, RMSE, and number of model evaluations (from left to right).
LIST OF FIGURES

4.3 Left: Examples of estimated parameters with BO-VAE, FH, and FS. Right: Progression of FH on the multi-scale hierarchy for parameter estimation of (a) (green leaf: homogeneous tissues; red leaf: heterogeneous tissues). .......................... 46

4.4 Comparison of training points selected by EI (a) and EI Post-1 (b), and examples of the estimated parameters by the two acquisition functions (c-d). .................................................. 47

4.5 (a-b): Examples of estimated parameters using five vs. two dimensional latent codes. (c-d): Latent code manifold based on (c) infarct location, and (d) infarct size. ................................. 48

4.6 Model parameter estimated with BO-VAE, FH, and FS on real-data study. ................................................................. 48

5.1 Outline of the presented method, with dimensions labeled within the VAE. ................................................................. 53

5.2 (a) Comparison of reconstruction accuracy using gVAE with a 2d manifold vs. PCA and fVAE [20] with various-dimensional manifolds. (b)-(c): Plots of 2d latent codes from gVAE colored by infarct location (b) and infarct size (c). .................................................. 56

5.3 (a) Accuracy and (b) computational cost of gVAE (green) versus fVAE [20] (yellow), FH [17] (purple), and FS [82] (red). (c)-(e): Examples of estimated parameters. .................................................. 57

5.4 (a) Reconstruction accuracy and (b) convergence of test losses from fine-tuned gVAE vs. gVAE trained from scratch (termed fully-trained) with varying training data size. (c) Examples of estimated parameters using gVAE fine-tuned with 7360 data. .................................................. 58

5.5 Estimated parameters with gVAE, fVAE, FH, and FS on real-data studies. ................................................................. 59

6.1 Work-flow diagram of the presented framework. .................... 64

6.2 Examples of exact bivariate marginal pdfs superimposed with the training points collected during GP construction, which are spread in the parameter space but are most concentrated in regions of high probability density. .................................................. 67

6.3 Examples of the exact posterior pdfs (red) vs. those obtained by the presented method (green) and sampling the GP surrogate pdf (blue). 73
6.4 Comparison of the efficiency between the presented method (GP-accelerated MH) and the standard MH method (Direct MH). Left: computational cost. Right: acceptance rate. 73
6.5 Examples of parameter estimation results when using clusters from the adaptive method vs. segments from the uniform method as LD representation of the spatial parameter space. 75
6.6 Example of high uncertainty associated with over-parameterization when the adaptive method assigned several small clusters to represent the heterogeneous regions around the infarct (c-e). 77
6.7 (a) Regions of the heart obtained by the adaptive method. (b) Estimated univariate and bivariate marginal pdf plots. (c) Trace plot for parameters of regions 6 and 7 (top) shows a negative correlation with a switching behavior. 78
6.8 Mean, mode, and standard deviation of posterior pdfs estimated from epicardial potentials simulated by a multi-scale ionic EP model blinded to the presented estimation method. 80
6.9 Mean, mode, and standard deviation of the posterior parameter pdfs estimated from 120-lead ECG data. 81
6.10 Real-data experiments for case 1 from Fig. 6.9: regions of the heart to be parameterized and the corresponding marginal probability density plots. 82
6.11 Real-data experiments for case 2 from Fig. 6.9: regions of the heart to be parameterized and the corresponding marginal probability density plots. 83
6.12 Real-data experiments for case 3 from Fig. 6.9. (a) Regions of the heart to be parameterized. (b) Univariate and bivariate marginal pdf plots. 84
6.13 Acceptance rate of the GP-accelerated MH method: (a) decreases as the standard deviation of the Gaussian proposal increases, (b) decreases as the Kullback-Leibler divergence from the GP surrogate to the exact posterior pdf increases. 85
List of Tables

5.1 Comparison of reconstruction accuracy with the presented gVAE, PCA, and IVAE with various depths [20] in five different geometries. 56

6.1 Mean absolute errors in the estimated mean, mode, and standard deviation against directly sampling the exact posterior pdf: the presented method (GP-accelerated MH) vs. sampling the GP surrogate pdf only (MH on GP). 74
Chapter 1

Introduction

In the past few decades, heart disease has continued to remain the leading cause of mortality and morbidity worldwide. A recent study published by the American Heart Association (AHA) in 2017 shows that in the United States alone cardiovascular disease accounted for nearly 800,000 deaths (about one in every three deaths) [6].

One of the concerns in the treatment of cardiovascular diseases is that the treatment and therapy provided to one patient may not work similarly for another patient with a similar cardiac condition. This can lead to failure of the treatment method and re-occurrence of the disease. Patients differ in the anatomy and physiology, and therefore, “one-size-fits-all” treatment strategies that do not take into account inter-patient variabilities may be less effective. In this context, computational models of the heart that have the ability to integrate both generic physics based multi-scale knowledge about the structure and function of a heart and patient-specific variations in physiology and anatomy have emerged as an important patient-specific treatment tool. These personalized computational models of the heart have been shown to facilitate the planning of surgery such as the coronary bypass graft surgery [65] or therapy such as the re-synchronization therapy [42, 68], the design and experimentation of medical devices [48], and the study of the mechanism of various cardiac diseases [4, 56, 76, 81, 86].
1.1 Problem Definition

Personalization of a cardiac electrophysiological (EP) model has two major ingredients: personalization of the anatomy and personalization of the physiology. Advances in medical imaging techniques over the past few decades have enabled reconstruction of image-based anatomical models at an unprecedented level of detail. However, the estimation of subject-specific organ tissue properties that are in the form of parameters of a physiological model still suffers from several unresolved critical challenges. First, tissue properties of an organ typically cannot be directly measured but needs to be estimated from sparse measurements that are indirectly related to the unknown tissue properties through a complex nonlinear physiological process. Measurements obtained with invasive procedure such as optical mapping and catheter mapping can be used. However, in addition to suffering from lack of resolution and precision, these procedures are risky to the patients. Alternately, non-invasive data measured on the body surface can be utilized. While it is safer to acquire data non-invasively from the patients, these measurements are more remote from the unknown tissue properties increasing the challenge in tissue properties estimation. Additionally, these measurements also contain errors and uncertainties. Another critical challenge in tissue properties estimation is that tissue properties are spatially varying throughout the 3D organ where local abnormality and heterogeneity often hold important therapeutical implications. The personalization of organ tissue properties, therefore, translates to an ill-posed inverse problem in which the forward model contains a computationally expensive non-linear physiological simulation model and the unknowns are represented by high-dimensional (HD) model parameters at the resolution of the discrete organ mesh. For clarity below we summarize the major challenges associated with physiological model personalization and uncertainty quantification into three categories.

- **High-dimensional parameter space:** Tissue properties are heterogeneous and spatially varying throughout the cardiac mesh. Through discretization of the cardiac mesh, this continuous space of tissue properties can be expressed with model parameters at the dimensionality of the resolution of the cardiac mesh (thousands or millions). Direct estimation of these HD model parameters from indirect measurements suffers from two major issues. First, an increasing dimension of unknown pa-
rameters increases the ill-posedness of the inverse problem given limited measurement data, resulting in the un-identifiability of model parameters. Second, the exploration of the HD parameter space in the presence of a computationally expensive forward model becomes computationally prohibitive.

- **High computational cost:** Significant progress has been made in the multi-scale computational modeling of cardiac electrophysiology resulting in sophisticated simulation models with cell-to-organ level details. However, with an increase in the model fidelity comes an increase in the computational cost needed to evaluate these models. Since model parameters are indirectly related to the measurements through this computationally expensive simulation, existing optimization and inference methods will require repeated evaluation of this simulation model for parameter personalization. Therefore, there is a challenge of high computational cost associated with the need to repeatedly evaluate the simulation model.

- **Sources of uncertainty:** The personalization of model parameters is subject to various sources of uncertainty. For example, model parameters at the resolution of the cardiac mesh cannot be estimated directly and they have to be approximated with a low-dimensional (LD) representation [46, 68, 82, 83, 87]. Additionally, measurement data are noisy and sparse. This uncertainty when propagated to the output of the personalized model hinders its reliable application in clinical cardiology. To rigorously understand and quantify the variability of the predictions made by a personalized model, it is important to quantify the uncertainty associated with the estimated model parameters. The estimated uncertainty in the prediction provides a measure of the quality of the prediction which is useful in determining whether the prediction is within an acceptable limit. Additionally, the estimated uncertainty can provide various other valuable knowledge to the clinicians or the scientists, for example, the non-identifiability of certain tissue properties, the correlation between various tissue properties, and the uncertain modeling elements. This knowledge can be useful in determining what additional data needs to be gathered to improve the model’s reliability.
1.2 Objectives and Contributions

Personalized cardiac models have found extensive use in patient-specific treatment and therapy planning, the design and experimentation of medical devices, and the study of disease progression [74]. One current challenge with personalized models in these applications is that there are various free parameters related to the physiology (e.g., tissue properties, tissue distribution, etc) or geometry (e.g., fiber directions, segmentation threshold, etc) that need to be estimated from external measurements [44, 65, 74]. The estimation of parameters from noisy measurements can introduce uncertainty in the estimated parameter values. Another important challenge is that these models employ various implicit and explicit assumptions adding stochasticity to the model’s output. Given the various sources of input variabilities, a point-estimate of the model’s parameters is not sufficient for its reliable use in clinical and scientific study. First, it is important to quantify the quality and confidence on the predicted values to allow informed-decision making. Second, it is important to understand various factors that are contributing to the variations in model’s output to allow clinicians and scientists to either control these factors by soliciting additional measurements or to simply take into account these factors when making decisions. In this context, the main objective of this dissertation is to provide novel computational methods for uncertainty quantification and parameter estimation in cardiac models.

The driving application here is the estimation of electrophysiological model parameters using non-invasive ECG data and the quantification of uncertainty in these model parameters. Because the measurements are indirectly related to the model parameters through a non-linear physiological model, in addition to the challenge of uncertainty, this application also presents a challenge of solving an ill-posed inverse problem with high-dimensional (HD) unknowns as described in Section 1.1. To handle the above challenges in personalized modeling, this dissertation presents following major contributions:

- **Bayesian Active Surrogate Modeling:** Model personalization requires the handing of an objective function or a posterior probability density function that contains the complex and computationally expensive simulation model. Our central idea to enable an accurate model personalization within feasible computation is to learn an approximation (surrogate) of this computationally expensive target function and
use it to guide the exploration of the parameter space during optimization or inference. In this thesis, we introduce the concept of Bayesian active learning for surrogate learning. In particular, a set of training data points are sequentially and actively selected to learn an accurate and efficient surrogate model of the target function. This surrogate is then used for either accelerated optimization or inference.

• **Active surrogate modeling in HD parameter space:** Surrogate modeling in HD space is difficult because the number of samples necessary to ensure an accurate surrogate is exponential in the number of dimensions. To enable surrogate modeling of an expensive target function defined over a HD space, we present two different approaches. First, we present a geometry-based approach that learns a surrogate of the objective function progressively over gradually increasing dimensions of the parameter space as defined by the multi-scale hierarchy of the cardiac anatomy and further refined by a spatially-adaptive decision criterion. Second, we present a data-driven approach that embeds the learning of the surrogate model into a low-dimensional (LD) generative space of tissue properties that is obtained by learning a generative model from the LD latent space to the HD space of tissue properties.

• **Surrogate accelerated Metropolis-Hastings (MH) Sampling:** The surrogate of a computationally expensive objective function can be used to facilitate a sample efficient optimization. In this thesis, we extend this idea to probabilistic estimation. In specific, we present a novel approach to incorporate the surrogate model of the target probability distribution function into the proposal distribution of the Metropolis-Hastings (MH) sampling that will enable sampling with a reduced computation cost and a higher acceptance rate.

1.3 Outline

This dissertation is organized as follows. Chapter 2 reviews the background and related works on cardiac model personalization and Bayesian active surrogate learning. Chapter 3 presents a spatially adaptive multi-scale surrogate-based optimization to enable the estimation of HD model parameters [17, 22]. Chapter 4 presents a HD Bayesian active optimization that is enabled through
an embedded generative model [19, 20]. Chapter 5 presents a novel approach to encode the anatomical knowledge into a generative model during parameter estimation [21]. By incorporating the anatomical knowledge into the generative model, the approach presented in this chapter bridges the gap between the geometry-based approach presented in Chapter 3 and the data-driven approach presented in Chapter 4. Chapter 6 details how a surrogate of the probability density function learned with Bayesian active learning can be incorporated to accelerate MH sampling [18, 23]. Finally, Chapter 7 provides some of the future research directions stemming from this dissertation.
Chapter 2

Background and Related Works

2.1 Models of Cardiac Electrophysiology

The heart is an electro-mechanically coupled organ whose function is to pump blood through the blood vessels to regulate the flow of oxygen, nutrients, and waste products in the body. The rhythmic contraction and expansion of the heart muscles for blood flow is controlled by the electrical function of the heart. Specifically a node in the upper left atrium called sinoatrial node first releases an electrical signal that will induce a pattern of electrical propagation throughout the heart muscles. Any abnormality in the electrical property of heart tissues can result in an abnormal pattern of electric propagation which, in turn, might result in a lethal breakdown of the mechanical function of the heart.

A 3D computational model of the cardiac electrophysiology can be obtained by integrating the equations describing a cell scale cardiac electrophysiology model onto a 3D anatomical model of the ventricles. To personalize these 3D cardiac electrophysiology models, the measurement data from patients is utilized. In the following sections, we introduce the anatomical modeling of the heart and the torso, cardiac electrophysiology model, and measurement model in brief. For more detail, we refer the readers to [13,77].
2.1.1 Anatomical Model

The anatomical model of a patient-specific heart and torso can be obtained from the Computed Tomography (CT) scan images. A CT scan produces a set of images that correspond to slices of the body along the torso. As outlined in Fig. 2.1, a 3D bi-ventricular anatomical model is built from CT images by taking the following steps. First, a set of images along the cross-section of the ventricles (parasternal short axis) from the apex to the base of the heart is extracted. On these images, the epicardium, left endocardium, and right endocardium are segmented using an in-house semi-automatic routine. The binary images thus obtained are then utilized to obtain a triangular mesh. A tetrahedral mesh is generated from the triangular mesh by utilizing the iso2mesh package [27]. Finally, on the tetrahedral mesh a cloud of unstructured points (meshfree nodes) are added to represent the 3D myocardial volume. The fiber structures on the meshfree nodes are mapped from the standard ventricular fibrous structure mathematical model established in the literature as described in [75, 77].

The generation of the torso anatomy follows the same line as the generation of the ventricles anatomy. Specifically first the torso is segmented on the CT scan images using an in-house semi-automatic routine. These segmented binary images are then used to build a triangular mesh of the torso.

2.1.2 Electrophysiology Model

The cardiac electrophysiology can be simulated in a patient-specific anatomical model by integrating a cell scale model to tissue and organ scale [13, 14]. We describe the cell-scale electrophysiology models and how they are integrated to the organ scale in this section.

The electrical activity in the cardiac muscle cells occurs due the flow of
different ions through ion channels and transporters in the cell membrane. The resulting membrane current is known as Transmembrane potential or action potential. Over the past few decades, a wide range of mathematical models of cardiac electrophysiology varying in the level of detail and complexity have been developed [13, 14]. These models can be broadly classified into three types: biophysical, Eikonal, and phenomenological. Biophysical models capture the microscopic level ionic interactions within the cardiac cell and through the cell membrane. These models are biophysically detailed and contain a large number of parameters to describe the function of different ion channel, pump, and exchanger. The Eikonal models, on the other hand, focus only on the propagation of the electrical wave-front and cannot accurately describe the action potential. Phenomenological models offer a level of detail that lies in between these two types of models. They capture the key macroscopic dynamical properties such as the inward and outward current. Because these models have a small number of parameter and a faster execution time compared to the biophysical models, they offer a practical balance between model plausibility and computational feasibility for personalization of model parameters [14]. Therefore, these models have found widespread use in model personalization [32, 49, 52, 63, 68].

In this dissertation, we consider local parameter estimation for the two-variable phenomenological Aliev-Panfilov (AP) model [2]:

\[
\begin{align*}
\frac{\partial u}{\partial t} &= \nabla (D \nabla u) - cu(u - \theta)(u - 1) - uv, \\
\frac{\partial v}{\partial t} &= e(u, v)(-v - cu(u - \theta - 1)).
\end{align*}
\] (2.1)

where \( u \) is the normalized transmembrane action potential, \( v \) is the recovery current, and \( e = e_0 + (e_1 v)/(u + e_2) \) controls the coupling between \( u \) and \( v \). Parameter \( D \) is the diffusion tensor that describes conductivity anisotropy and is determined by the local fiber structure: the ratio of conductivities in the longitudinal and transverse fiber directions is set to 4 : 1 according to literature [29, 75]. Parameter \( c \) controls the repolarization and \( \theta \) controls the excitability of the cell.

To select the parameter to be estimated in the AP model, we conduct a one-factor-at-a-time sensitivity analysis of the AP model with baseline values of the parameters as documented in the literature [2]: \( \theta = 0.15, c = 8, e_0 = 0.002, e_1 = 0.2, \) and \( e_2 = 0.3 \). The shaded blue region in Fig. 2.2 shows the
standard deviation of $u$ throughout the action potential duration resulting from the variations in each model parameter. As shown, the average of these standard deviations is the highest as a result from the variations in parameter $\theta$ (tissue excitability). Furthermore, tissue excitability is documented to be closely associated with the ischemic severity of the myocardial tissue [14, 64, 77]. Therefore, in this dissertation, we will consider the estimation of parameter $\theta$ of the AP model. The physiological bounds on the value of parameter $\theta$ is $[0, 0.5]$ in which $\theta = 0.15$ represents normal excitability and an increase in the value represents an increase in the severity of tissue infarct until $\theta = 0.5$ represents necrotic tissue. For the rest of the parameters, their values are fixed to standard values as specified above.

The meshfree method is utilized to spatially discretize the AP model on the patient-specific 3D bi-ventricular model and generate Transmembrane potential activity model [77].

\section*{2.1.3 Electrocardiography Model}

Generally in clinical cardiology, there are two ways to measure cardiac electrophysiology: 1) invasive methods such as a point-to-point catheter mapping on the heart surface [61, 68], and 2) non-invasive method such as recording potentials on the body surface also known as electrocardiogram (ECG) [59, 77, 84]. In the following section, we briefly describe the major measurement models used in this study.
CHAPTER 2. BACKGROUND AND RELATED WORKS

Measurement Model
The transmembrane potential \( u(\theta) \) can be, in general, observed by two means. First, it generates extracellular potential \( Y_b \) in the thorax following the quasi-static approximation of the electromagnetic theory [57]. This relationship can be modeled by a Poisson’s equation within the heart and a set of Laplace’s equations, one for each organ, external to the heart and within the body torso. Given a discrete subject-specific mesh of the thorax, these equations are solved as described in [79] to obtain a linear relationship between extracellular potential \( Y_b \) and transmembrane potential \( u(\theta) \) through a subject-specific transfer matrix \( H_b \):

\[
Y_b = H_b u(\theta).
\]  

(2.2)

When measured on the heart surface, \( Y_b \) is known as electrograms; when measured at the body surface, \( Y_b \) is known as ECG—the latter is non-invasively available in practice but more challenging for parameter estimation, because it represents a more remote and sparse observation of \( u(\theta) \). This is the main measurement model that we utilize for parameter estimation.

Alternatively, transmembrane potential \( u(\theta) \) may be locally recorded on the epicardium through optical mapping [62, 78]:

\[
Y_e = H_e u(\theta),
\]  

(2.3)

where \( Y_e \) represents coarsely distributed subset of transmembrane potential \( u(\theta) \) measured at local regions of the epicardium, indexed by a sparse binary matrix \( H_e \). This is a second measurement model that we will use in some of the parameter estimation experiments. For clarity, throughout the remainder of this dissertation, we use \( Y \) to denote either \( Y_e \) or \( Y_b \), and \( H \) to denote either \( H_e \) or \( H_b \).

Electrocardiography
Electrocardiography is a non-invasive method that involves measuring potentials on various locations on the body surface over time. A schematic of a normal ECG signal is shown in Fig. 2.3. It consists of following major segments: P wave, PR segment, QRS complex, ST segment, and T wave. Different segment of the ECG corresponds to different stages of electrical activity in the heart which can be utilized to trace specific abnormal electrical function of
the heart. In particular, the P wave corresponds to atrial depolarization, PR segment corresponds to the propagation of the activation through the Atrioventricular node and the Purkinje fiber, QRS complex corresponds to the depolarization of the ventricles, ST segment corresponds to the stage when all myocytes remain at the plateau and all regions in the ventricles are at depolarized state, and T wave corresponds to the re-polarization of the ventricles.

Classical ECG consisted of a recording system consisting of three electrodes positioned on the left arm, right arm, and left leg of the patient from which three limb voltages $V_I$, $V_{II}$, and $V_{III}$ were calculated [26]. An extension and the most commonly used 12-lead ECG standard consists of six additional electrodes that provide six additional measurements $V_{1-6}$.

Often the standard 12-lead ECG is not adequate for clinical diagnosis and treatment planning of heart diseases and a more elaborate Body Surface Potential Map (BSPM) is utilized. A BSPM consists of potentials recorded on the body surface with tens to hundreds of electrodes [51]. In this dissertation, we primarily use the 120-lead BSPM (referred to as 120-lead ECG throughout the remainder of this dissertation).
2.2 Personalization of Model Parameters

2.2.1 Parameter Estimation of Complex Physiological Models

Cardiac tissue properties typically cannot be directly measured but needs to be estimated from sparse and indirect measurements. The estimation of cardiac model parameters, therefore, consists of an ill-posed inverse problem in which the forward model contains a non-linear and computationally expensive simulation model. This forward model when used in the objective function for a deterministic parameter estimation or in the probability distribution function for a probabilistic parameter estimation results in a challenge of handling an intractable function for parameter estimation. In particular, this target function of interest does not have a closed form expression, its derivatives are difficult to obtain, its convexity property is unknown, and it is computationally expensive to evaluate.

In the past few decades, significant progress has been made in deterministic approaches that handle complex objective function containing the simulation model in parameter estimation. In particular, various methods of derivative-free optimization were presented to handle the analytically-intractable objective function, such as the use of the subplex method [83], Bound Optimization BY Quadratic Approximation (BOBYQA) [82], and New Unconstrained Optimization Algorithm (NEWUOA) [68]. More recently, effective methods of multi-scale derivative-free optimization were developed [10,11]. Essentially, all these methods involve a non-intrusive, repeated evaluation of the physiological model. Hence, the computational cost in model personalization is mostly dependent on the number of evaluations of the objective function required by these methods.

Deterministic approaches focus on obtaining a single value of the model parameters that best fits the available data (under given optimization criteria); they do not provide an uncertainty measure associated with the estimated patient-specific parameter values. Limited works have been reported on probabilistic approaches to estimate the distribution of patient-specific parameters of a cardiac model. Recent studies handle the complex posterior probability density function (pdf) of model parameters by building an efficient surrogate of the simulation model using methods such as kriging [67] and polynomial chaos [44]. This approximate but efficient simulation is then used to replace the exact simulation in the posterior pdf for a substantially faster Markov
Chain Monte Carlo (MCMC) sampling of the posterior pdf [44, 67].

2.2.2 High-dimensional Parameter Estimation

While various deterministic and probabilistic approaches to estimate model parameters have been proposed over the past few decades, none of the existing methods can be directly applied to estimate the high-dimensional (HD) model parameters corresponding to the spatially-varying tissue properties. This is because the increased dimension of unknowns: 1) increases the ill-posedness of the problem, where estimation of parameters at the resolution of the discrete cardiac mesh is likely to be unidentifiable given limited indirect measurements, and 2) leads to an exponential increase in the number of model evaluations, where each model evaluation itself is computationally intensive.

As a result, most previous works resort to the estimation of parameters at a reduced spatial dimension. Many works focus on the estimation of a global parameter [49, 53] by assuming uniform tissue property throughout the myocardium. Although such estimation allows a fast model calibration, it does not capture the locally varying tissue properties. Another approach is to reduce the dimension of the unknowns by partitioning the cardiac mesh into pre-defined segments and to assume uniform parameter value within each segment. This reduces the spatial field of tissue properties to a low-resolution representation (in the range of 3-26 segments) [32, 44, 46, 68, 82, 83, 87]. As a result of this drastically decreased resolution, the solution has a limited ability to reflect different levels of tissue heterogeneity, or the existence of local abnormal tissue with various sizes and distributions. Additionally, as the number of pre-defined segments increases, a good initialization is shown to become critical to the accuracy of parameter estimation [83]. This unfortunately relies on the availability of measurement data that can reveal diseased regions a priori. Therefore, a critical gap remains in personalized modeling between the needs for a high resolution estimation of spatially-varying parameters and the difficulty to accommodate such high dimensionality by existing estimation approaches.

2.2.3 Bayesian Active Approximation of Expensive Functions

In domains such as the computational astrophysics [15] and geothermal research [16] one often needs to estimate a probability density of the simulation
model parameters that neither has a tractable analytical form nor is cheap to evaluate. Similarly, in applications like optimal drug design [80] and optimal user preferences selection [1] the objective function to be optimized does not have a closed-form expression and is costly to evaluate. In these applications, MCMC sampling methods that consist of repeated sampling and evaluation of the unknown function for learning its approximation are used [39]. However, MCMC methods are not designed to actively choose sample points on which to evaluate the expensive function. Therefore, these function approximation technique require a large number of samples making them computationally expensive.

On the other hand, recent Bayesian active learning methods are designed to actively select sample points with an aim to achieve an accurate result with as few evaluations of the expensive function as possible [7, 39, 40]. Therefore, Bayesian active learning methods such as Bayesian optimization [7, 38] and Bayesian quadrature [41, 54] have been used in various tasks that require modeling an intractable and expensive unknown function. These class of methods have two major components. First, a surrogate model of the expensive function that is learned from the samples collected so far. Second, an acquisition or a utility function that uses the surrogate of the expensive function in guiding the selection of future samples. The acquisition function has a capability to balance between exploration of the support of the unknown function and exploitation of the important regions in the support.

For surrogate modeling a Gaussian process (GP) is often utilized [?, 7, 46, 47, 70]. A GP is a random process with an infinite number of random variables, any finite subset of which has a joint Gaussian distribution [60]. It thus provides a distribution over functions \( f(\cdot) \sim \mathcal{GP}(\mu(\cdot), \kappa(\cdot, \cdot)) \) that is fully characterized by a mean function \( \mu(\cdot) \) and a covariance function \( \kappa(\cdot, \cdot) \) [60]. GP has been successfully utilized in many non-linear Bayesian regression tasks. In a typical regression setting, given a set of training input-output pairs \( \Theta = [\theta_1, \theta_2, \ldots, \theta_n]^T \) and \( G = [g_1, g_2, \ldots, g_n]^T \), one is interested in making a prediction \( g_* \) for any given \( \theta_* \). In standard regression with GP, this is achieved by estimating an unobserved latent function that is responsible for generating \( g \) from \( \theta: g = f(\theta) + \mathcal{N}(0, \sigma^2) \). A GP is used to define prior distribution over the latent function \( f(\cdot) \sim \mathcal{GP}(0, \kappa(\cdot, \cdot)) \). Then, using the properties of Gaussian distribution, the predictive mean \( \mu(\theta_*) \) and variance \( \sigma^2(\theta_*) \) can be obtained
as:

\[ \mu(\theta_s) = k^T(K + \xi^2I)^{-1}G, \]
\[ \sigma^2(\theta_s) = \kappa(\theta_s, \theta_s) - k^T(K + \xi^2I)^{-1}k, \]

where \( k = [\kappa(\theta_s, \theta_1), \kappa(\theta_s, \theta_2), \ldots, \kappa(\theta_s, \theta_n)]^T \) and \( K \) is the positive definite co-variance matrix with \( K_{i,j} = \kappa(\theta_i, \theta_j) \). In active learning setting, these training input-output pairs are gradually collected by using an acquisition function. The acquisition function takes the predictive GP mean and variance for sample selection and trades off between a large GP posterior mean given the current GP belief of the unknown function (i.e., exploitation) and a large GP posterior variance to reduce the uncertainty of the GP belief of the unknown function (i.e., exploration).

While Bayesian active approximation approaches are widely used for approximating and optimizing expensive functions, they do not work well when the function is defined over a HD (dimension > 15) space [39,47,70]. This is because the number of training data required for nonparametric regression with GP grows exponentially with the dimension on with the function is defined. In addition, the active sample point selection involves global optimization of a HD acquisition function which is an inherently difficult [39,47,70].

This thesis builds-upon Bayesian active approximation of expensive functions defined over a HD space for both deterministic and probabilistic estimation. In the following chapters, we describe various methods for HD Bayesian active surrogate modeling and approaches to accommodate these surrogate models in optimization or density estimation.
Chapter 3

Spatially Adaptive Multi-Scale Optimization

The estimation of spatially varying parameter values of a 3D cardiac EP model is important for revealing abnormal tissues with altered material properties and for building patient-specific models. As explained in Chapter 2, to reduce the dimension of unknown parameters existing works in local parameter estimation typically represent the heart with a small number of pre-defined segments. Such low-resolution approaches have limited ability to estimate tissues with varying sizes, locations, and distributions.

In this chapter, we present a novel framework that goes beyond a uniform low-resolution approach and achieves a higher resolution estimation of tissue properties represented by spatially non-uniform resolution. This is achieved by two central elements: a multi-scale coarse-to-fine surrogate-based optimization and a spatially adaptive decision criterion that retains lower resolution in homogeneous tissue regions and allows higher resolution in heterogeneous tissue regions. The surrogate model-based optimization allows to incorporate and approximate complex objective function in optimization and the spatially-adaptive multi-scale approach allows to handle a high-dimensional unknown.
CHAPTER 3. SPATIALLY ADAPTIVE MULTI-SCALE OPTIMIZATION

3.1 Introduction

i. A surrogate-based multi-scale optimization framework that progressively optimizes from lower resolution to higher resolution, where a lower resolution solution is used to facilitate the higher resolution optimization.

ii. An adaptive decision criterion that allocates higher resolution to regions with heterogeneous tissue properties, allowing a higher resolution parameter estimation without using a large number of unknowns.

iii. Development of the adaptive decision criterion with two optimization methods: 1) Gaussian process (GP) surrogate based optimization (GPO) [7], and 2) Bound Optimization BY Quadratic Approximation (BOBYQA) [58]; and comparison with standard BOBYQA carried out on a pre-defined division of cardiac mesh into 26 segments.

iv. Evaluation of the presented framework in the estimation of local tissue excitability in a 3D cardiac electrophysiological (EP) model, using several measurement data including 120-lead ECG, 12-lead ECG, and epicardial action potentials.

We test the performance of the framework presented in this chapter on three categories of experiments. First, on synthetic experiments, we evaluate the presented method in estimating local tissue excitability: 1) in the presence of infarcted tissues of varying sizes and locations, and 2) using 120-lead ECG vs. using 12-lead ECG. Second, we apply the presented method to estimate local tissue excitability on seven patients with prior myocardial infarction, where a subset of epicardial potentials simulated from a high-resolution ionic electrophysiological model blinded to the parameter estimation framework is used as measurement data. Finally, we conduct two pilot studies on patients who underwent catheter ablation of ventricular tachycardia due to prior tissue infarction by using non-invasive 120-lead ECG as measurement data. Results show that the presented framework is able to characterize local abnormal infarct tissues without using any knowledge of the abnormal tissue derived from other data modalities.
3.2 Spatially-Adaptive Multi-Scale Optimization

Estimation of the three dimensionally distributed tissue excitability in a cardiac mesh with \( m \) meshfree nodes can be formulated as an optimization problem:

\[
\hat{\theta} = \arg \max_{\theta \in \Omega} J(\theta),
\]

with bounds constraint: \( \Omega = \{ \theta \in \mathbb{R}^m | lb \leq \theta \leq ub \} \). The objective function \( J \) could be any measure of similarity between the model output \( \Phi \) and the measured ECG \( Y \), such as a measure of the similarity in signal magnitude by squared error, or a measure of the similarity in signal morphology by correlation coefficient. Because the measurement ECG data tend to be noisy and there is often a scale difference between the measured and simulated data, we use an objective function consisting of correlation coefficient and sum of squared error.

\[
J(\theta) = \sum_{i=1}^{l} \left\{ \frac{1}{t} \sum_{j=1}^{t} (y_{ij} - \bar{y}_i)(\Phi_{ij} - \bar{\Phi}_i) - \lambda \sum_{j=1}^{t} (\Phi_{ij} - y_{ij})^2 \right\},
\]

where \( l \) is the number of body surface leads, \( t \) is the number of discrete time samples, and \( \bar{\Phi}_i \) and \( \bar{y}_i \) are respectively the mean of simulated and measured ECG on the \( i^{th} \) body surface lead. The correlation coefficient takes value in the range \([-1, 1]\) and the value of the squared error depends on the unit of measurement data. Parameter \( \lambda \) is empirically adjusted using the knowledge of the two.

The direct estimation of \( \theta \) involves a high-dimensional (HD) optimization that is infeasible due to high ill-posedness and prohibitively large computation. To maximize the resolution of local parameter estimation using a limited dimensionality of unknowns, we present a spatially-adaptive multi-scale scheme that can be used in conjunction with any optimization method. This framework consists of three key components: 1) a multi-scale hierarchical representation of the spatial domain, 2) a coarse-to-fine optimization, and 3) a criterion for adaptive spatial resolution adjustment that favors higher resolution in heterogeneous regions. The workflow chart in Fig. 3.1 shows how each of the key components work together in the presented framework. Given a discrete cardiac mesh, we first construct the multi-scale representation of the
spatial domain. Parameter estimation is then carried out as an iterative procedure of coarse-to-fine optimization with an adaptive adjustment of spatial resolution, until further refinement does not produce major improvement in the global optimum. Individual components of the framework are detailed in the following sections.
3.2.1 Multi-Scale Hierarchy

To facilitate a coarse-to-fine optimization, a multi-scale representation of the cardiac mesh is needed. By exploiting the spatial smoothness of tissue properties (i.e., neighboring tissues have similar material properties), we adopt a bottom-up agglomerative hierarchical clustering [34]. It begins with every node in the cardiac mesh as separate clusters. As one moves up the hierarchy, a pair of the closest clusters is merged until all the nodes belong to a single cluster. The distance between two clusters is measured by the average of the pairwise Euclidean distance between every node in the two clusters. Fig. 3.2 shows the clusters of this multi-scale hierarchy at several scales, illustrating (from left to right) how two clusters at a lower scale is merged into a cluster at a higher scale. The finest level in the multi-scale hierarchy corresponds to the beginning stage of the clustering. At this stage parameter estimation would correspond to a direct HD local parameter estimation at the resolution of the cardiac mesh. The coarsest level in the multi-scale hierarchy corresponds to the end stage of the clustering where every nodes belong to the same cluster. At this stage, parameter estimation would correspond to a global parameter estimation.

3.2.2 Coarse-to-Fine Optimization

In optimizing a complex objective function, a coarse-to-fine optimization may: 1) facilitate faster convergence by exploiting coarser-scale information in finer-scale optimization, and 2) decrease the risk of being stuck in a bad local minimum by optimizing a cascade of increasingly complex objective functions over a gradually increasing resolution.

The optimization starts from the coarsest scale of the multi-scale hierarchy. It then progressively optimizes the parameters at a higher resolution that is determined by the adaptive spatial resolution adjustment criterion (Section 3.2.3). For clarity, in this study, we represent the change in resolution during the coarse-to-fine optimizations using a tree, in which the leaf nodes represent the clusters in the present resolution. At the $k^{th}$ iteration, we consider the optimization of $\hat{\theta}^{(k)}$ corresponding to the parameter values on the leaf nodes:

$$\hat{\theta}^{(k)} = \arg \max_{\theta^{(k)} \in \Omega} J(\theta^{(k)}).$$

(3.3)
In the presented framework, any optimization method that can handle the objective function (3.2) can be used. As an example, below we present the presented framework with GPO [70]. The use of the presented framework with BOBYQA [58] follows the same line and will be discussed in the experiment section.

A GP is first used to approximate the objective function (3.3). The GP assumes a smoothness property of the objective function which is mainly determined by its co-variance function. Because the most commonly used squared exponential kernel is known to enforce a large smoothness property, we take the “Mátern 5/2” kernel that enables less smooth function [60]:

\[ \kappa(\theta_1^{(k)}, \theta_2^{(k)}) = \alpha^2 (\gamma + \sqrt{5d/\gamma} + 5d^2/3\gamma^2) \exp(-\sqrt{5d/\gamma}), \]

where \(d = ||\theta_1^{(k)} - \theta_2^{(k)}||\), and \(\gamma\) and \(\alpha\) are kernel parameters. Because no prior knowledge about the objective function (3.3) is available, we use a zero mean function for simplicity.

A typical surrogate based optimization consists of two major ingredients: 1) using the current surrogate of the objective function, find out points in the solution space that can improve the approximation of the objective function well especially in the region of global maximum, and 2) evaluate the objective function at the point determined in the previous step and update the surrogate. Because the initial knowledge of the GP plays an important role in determining how fast a good GP can be obtained in the subsequent updates, here we introduce an additional step to initialize the GP using the coarser-scale optimum. Below we describe each of these major steps:

**Initialize Using the Coarser-Scale Optimum:** Depending on the optimization method used, different strategies can be used. If the optimization starts with a single point, such as the BOBYQA used here, the lower resolution optimum can be used as initialization. If the optimization can start with multiple points, such as the GPO used here, a set of higher resolution points derived from the lower resolution optimum can be used. For two sibling leaf nodes obtained by recent refinement, a set of two values whose average equals to the value on the parent node (lower resolution optimum) are generated. This is done by first uniformly sampling a set of values from the parameter bound [\(lb, ub\)] for one of the leaf nodes. Then, corresponding set of values for the second leaf node are generated by subtracting previous samples from twice the value on the parent node. For leaf nodes that did not undergo resolution change,
their parameter values are equal to their current optimum. These points are used to obtain an initial GP surrogate.

**Determine the Next Query Point:** In the context of optimization, a good surrogate should both approximate the objective function well globally, and have higher accuracy in the region of global maximum. For the former, the query point is determined to explore the solution space where the predictive uncertainty $\sigma(\theta^{(k)})$ of the current GP is high. For the latter, the query point is determined to exploit the current GP at the solution space where the predictive mean $\mu(\theta^{(k)})$ is high. Overall, this $n^{th}$ query point is obtained by maximizing the upper confidence bound (UCB) of the GP using BOBYQA.

$$\theta_n^{(k)} = \arg \max_{\theta^{(k)} \in \Omega} \left( \mu(\theta^{(k)}) + \beta^{1/2} \sigma(\theta^{(k)}) \right),$$  \hspace{1cm} (3.4)

where the parameter $\beta = 2 \log(\pi^2 n^2 / 6 \eta)$, $\eta \in (0, 1)$ balances between exploitation and exploration. The predictive mean and uncertainty in (3.4) is evaluated by using the Sherman-Morrison-Woodbury formula \[60\]:

$$\mu(\theta_n^{(k)}) = \kappa^T (\kappa + \varsigma^2 I)^{-1} J_{1:n-1},$$  \hspace{1cm} (3.5)

$$\sigma(\theta_n^{(k)}) = \kappa(\theta_n^{(k)}, \theta_n^{(k)}) - \kappa^T (\kappa + \varsigma^2 I)^{-1} \kappa,$$  \hspace{1cm} (3.6)

where $\kappa = \begin{bmatrix} \kappa(\theta_n^{(k)}, \theta_1^{(k)}) & \kappa(\theta_n^{(k)}, \theta_2^{(k)}) & \cdots & \kappa(\theta_n^{(k)}, \theta_{n-1}^{(k)}) \end{bmatrix}$ and $\kappa$ is the positive definite co-variance matrix $\kappa_{i,j} = \kappa(\theta_i^{(k)}, \theta_j^{(k)})$. A small noise $\varsigma$ is added for numerical stability during matrix inversion.

**Update the GP:** Once a new query point is obtained, the objective function (3.3) is evaluated at this point. GP is updated at the new point-objective function value pair $\{\theta_n^{(k)}, J_n\}$.

$$\begin{bmatrix} J_{1:n-1} \\ J_n \end{bmatrix} \sim \mathcal{N} \left( \mu(\theta_n^{(k)}), \begin{bmatrix} \kappa(\theta_n^{(k)}, \theta_1^{(k)}) & \kappa(\theta_n^{(k)}, \theta_2^{(k)}) & \cdots & \kappa(\theta_n^{(k)}, \theta_{n-1}^{(k)}) \\ \kappa(\theta_1^{(k)}, \theta_n^{(k)}) & \kappa(\theta_2^{(k)}, \theta_n^{(k)}) & \cdots & \kappa(\theta_{n-1}^{(k)}, \theta_n^{(k)}) \end{bmatrix} \begin{bmatrix} \kappa(\theta_n^{(k)}, \theta_1^{(k)}) & \kappa(\theta_n^{(k)}, \theta_2^{(k)}) & \cdots & \kappa(\theta_n^{(k)}, \theta_{n-1}^{(k)}) \end{bmatrix} \right).$$  \hspace{1cm} (3.7)

This provides an updated GP. In this study, we set the values of the kernel hyperparameters empirically. These steps are run in iteration until the query points do not change significantly over several iterations. In this way, a GP surrogate of the objective function along with the optimum at a given resolution is obtained.
3.2.3 Adaptive Spatial Resolution Adjustment

If coarse-to-fine optimization is done on the full multi-scale hierarchy, the number of unknowns will again quickly become too high in dimension to be effectively optimized. To overcome this, we aim for a non-uniform resolution with higher resolution in heterogeneous regions and lower resolution in homogeneous regions. To achieve this, instead of refinement of all the leaf nodes after optimization at each scale, we selectively refine the leaf nodes that are heterogeneous and coarsen those that are homogeneous.

To identify the clusters that are heterogeneous versus homogeneous in tissue properties, we propose a greedy criterion based on gain in the global optimum. Intuitively, if a cluster is homogeneous, its refinement is expected to yield children clusters with similar parameter values; as a result, there will be a minimal gain in the objective function in representing this region with higher resolution. The contrary is true for heterogeneous clusters.

For each pair of sibling leaf nodes \( \hat{\theta}_{p,i}^{(k-1)}, \hat{\theta}_{p,i+1}^{(k-1)} \) that share the common parent node \( \hat{\theta}_p^{(k-1)} \), we evaluate the gain resulting from the refinement as the change in the objective function value on \( \hat{\theta}_p^{(k)} \) (current optimum) versus replacing the values of the children nodes \( \hat{\theta}_{p,i}^{(k)}, \hat{\theta}_{p,i+1}^{(k)} \) with the value of their parent node \( \hat{\theta}_p^{(k-1)} \):

\[
r_{i}^{(k)} = J(\hat{\theta}_p^{(k)}) - J(s^{(k)}),
\]

where \( s^k = (\hat{\theta}_p^{(k)} \setminus (\hat{\theta}_{p,i}^{(k)}, \hat{\theta}_{p,i+1}^{(k)}), \hat{\theta}_p^{(k-1)}) \) is the parameter vector obtained by replacing \( \hat{\theta}_{p,i}^{(k)} \) and \( \hat{\theta}_{p,i+1}^{(k)} \) with \( \hat{\theta}_p^{(k-1)} \) in the current optimum \( \hat{\theta}_p^{(k)} \). For leaf nodes that do not have a sibling due to previous coarsening, no resolution change has occurred but their value may have changed as a result of resolution-change elsewhere. For such nodes \( \hat{\theta}_q^{(k)} \), the gain \( r_q^{(k)} \) equals the change in the objective function due to the change in \( \hat{\theta}_q^{(k)} \) during the coarsening and after the optimization.

Based on \( r_i^{(k)} \), two actions are taken. For the leaf node or the pair of leaf nodes that has the maximum gain \( r_i^{(k)} \), we hypothesize that they represent highly heterogeneous regions and allow a higher resolution representation (i.e., refinement). For the pair of leaf nodes that bring negligible or negative gain \( (r_i^{(k)} < \epsilon) \), we assume that the refinement suggested in the previous scale is not
beneficial and thus retract the refinement (i.e., coarsening). The refinement and coarsening is done by using the multi-scale hierarchy (Section 3.2.1) as a reference.

3.3 Evaluation on Synthetic Data

In synthetic experiments, we evaluate the performance of the presented framework in combination with the GPO and the BOBYQA (termed as adaptive GPO and adaptive BOBYQA respectively for the remainder of this chapter) in the presence of infarcts of varying sizes and locations. We also compare the presented framework with optimization using 26 pre-defined segments of ventricles, with 17 segments on the left ventricle (LV) based on the American Heart Association (AHA) standard and 9 segments on the right ventricle (RV) as shown in Fig. 3.3. Finally, we study the identifiability of local tissue excitability with respect to: 1) the size of the infarct, and 2) the amount of measurement data.

In total, we consider 35 synthetic cases with different settings of infarcts on three realistic CT-derived human heart-torso geometrical models. We set infarcts of sizes varying from 1% to 65% of the ventricles at different LV and RV locations by using various combinations of the pre-defined segments, or random locations when the size is smaller than one pre-defined segment. To represent healthy and infarcted tissues, the parameter $\theta$ of the AP model (2.1) is uniformly sampled from $[0.149, 0.151]$ and $[0.399, 0.501]$ respectively. For
each infarct case, the action potentials are first simulated using the AP model. 120-lead ECG is then generated using the forward model and corrupted with 20dB Gaussian noise. The admissible range of tissue excitability parameter for the optimization is set to $[0, 0.52]$. To evaluate the accuracy of the estimated local parameters, we use two metrics: 1) root mean squared error (RMSE) $= ||\hat{\theta} - \theta||^2$ between the true and estimated parameter values, and 2) dice coefficient (DC) $= \frac{2|S_1 \cap S_2|}{|S_1| + |S_2|}$, where $S_1$ and $S_2$ are the sets of nodes in the true and estimated regions of the infarct; these regions are determined for each infarct case by calculating a threshold value that minimizes the intra-region variance [55] on the estimated parameter values. Both metrics are evaluated at the resolution of the cardiac mesh.
3.3.1 Performance Based on the Size and Location of the Infarct

We first summarize the performance of the presented framework in the presence of infarcts of varying sizes using three representative examples. Fig. 3.4(a) shows an example of estimation on a small infarct (3.6%). The tree shows that once a homogeneous healthy region is found in stage 1, the optimization keeps the entire region at a very low resolution and refines only along the heterogeneous region that contains the infarct. It continues narrowing down the infarct with higher resolution, generating a narrow yet deep tree. The final result shown in stage 3 of Fig. 3.4(a) is achieved with only a dimension of 10 unknowns. In comparison, if a uniform resolution is used, a dimension of 128 would be needed to achieve an estimation at the same resolution. Fig. 3.4(b) shows another example with a medium sized infarct (28.7%). Since the infarct spans a larger number of clusters, it is not until stage 2 before the tree can be split along one major branch. In addition, because both normal and infarcted tissues are large enough to be represented by low-resolution homogeneous clusters, an overall lower resolution solution is obtained with a wider yet shallower tree. While the presented method converges at a dimension of 7, a uniform resolution would require a dimension of 16 for an estimation at similar resolution. Finally, Fig. 3.4(c) shows an example with a larger sized infarct (62.4%). A larger sized infarct can be represented by big clusters at low resolution and has a boundary that needs to be represented by a larger number of smaller clusters. Therefore, until stage 2 the large homogeneous regions (both normal and infarcted) are split into major clusters. In the following stages, the border region is split into multiple branches increasing the resolution along the border yielding a wider tree in the first a few steps and narrow branches in later stages. In this case, while the presented method converges at a dimension of 13, a uniform resolution estimation would require a dimension of 64.

Additional examples of the estimated tissue excitability for different infarcts are provided in Fig. 3.5. Fig. 3.6 summarizes the accuracy in estimation of the presented adaptive BOBYQA (magenta bar) and adaptive GPO (blue bar) with respect to the size and location of the infarct. Specifically as shown in Fig. 3.6, while both the presented methods are in general able to identify infarcted tissues at various locations and of various sizes, the accuracy in the presence of septal infarcts show a noticeable decrease. As shown in an ex-
Figure 3.5: Synthetic experiments. Examples of tissue excitability estimated using: 1) uniform BOBYQA, 2) adaptive BOBYQA, and 3) adaptive GPO in the presence of infarcts of varying sizes and locations (red: infarcted tissue, blue: healthy tissue). Infarct locations and sizes (in percentage ventricles covered) from top to bottom: anterior (3.17%), inferior (9.84%), and septal (24.31%).

Figure 3.6: Synthetic experiments. Comparison of uniform BOBYQA, adaptive BOBYQA, and adaptive GPO in terms of DC and RMSE based on 35 different infarct cases of varying sizes and locations (bar: mean, line: standard deviation).

AMPLE OF SEPTAL INFARCT IN FIG. 3.5(C), THIS DROP IN ACCURACY IS ASSOCIATED WITH THE PRESENCE OF FALSE POSITIVES AT THE LATERAL VENTRICULAR WALLS. THIS IMPLIES THE EXISTENCE OF MULTIPLE PARAMETER CONFIGURATIONS THAT FIT THE MEASUREMENT DATA WELL. THE CHALLENGE IN DEALING WITH SEPTAL INFARCTS IS CONSISTENT WITH THAT REPORTED IN LITERATURE [79]. INTERESTINGLY, IN THE PRESENCE OF SEPTAL INFARCT, ADAPTIVE GPO SHOWS HIGHER ACCURACY THAN THE ADAPTIVE BOBYQA (FIG. 3.6). THIS COULD BE BECAUSE ADAPTIVE GPO WAS ABLE TO AVOID LOCAL MINIMA Owing TO THE INITIALIZATION WITH MULTIPLE POINTS. THE ADAPTIVE BOBYQA, HOWEVER, IN GENERAL HAS HIGHER EFFICIENCY, ESPECIALLY IN THE ESTIMATION OF HIGHER DIMENSION
of unknowns. Because both presented methods show similar performance, in the remaining sections of this paper, we consider performance analysis of the presented framework using adaptive GPO only unless explicitly stated otherwise.

### 3.3.2 Comparison with Optimization using Uniform Resolution

We compare the performance of the presented adaptive methods with the state-of-the-art BOBYQA using 26 pre-defined segments (termed as uniform BOBYQA for the remainder of this paper). We do not include comparison with GPO using 26 pre-defined segments because it performs poorly for a dimension higher than 20 both as reported in literature [70] and as observed in our experiments. Experimental setup as described in Section 3.3 is used. Because uniform BOBYQA is sensitive to initialization, each experiment case is run twice using random initialization and the better result is picked for comparison. The summary statistics of accuracy as shown in Fig. 3.6 demonstrates that the presented adaptive methods are more robust to infarcts of various sizes and locations. This improvement is statistically significant in both DC and RMSE (paired-\(t\) tests on 35 synthetic experiments, \(p < 0.0025\)). In particular, using uniform BOBYQA, it is difficult to estimate an infarct of size equal to or less than a single segment. In comparison, the presented methods are able to identify infarcts smaller than a single segment with a small number of unknowns (10-14). Furthermore, as shown in Fig. 3.5, uniform BOBYQA tends to show false-positives across multiple segments, failing to accurately reveal the spatial distribution of the infarcted tissues.

The major computational cost in local parameter estimation comes from
the repeated evaluation of the CPU-intensive simulation model. We compare the computational cost of the presented adaptive methods and uniform BOBYQA in terms of the computation time and the number of model evaluations used. This is based on 35 synthetic experiments conducted on a computer with Xeon E5 2.20GHZ processor and 128 GB RAM. In summary, uniform BOBYQA takes on average $3.33 \pm 0.83$ hours for convergence, while the presented methods take on average $6.07 \pm 3.02$ hours for convergence. Because, in the presented methods, the tree varies with varying infarct sizes leading to a significant difference in the number of model evaluations, we compare the number of model evaluations into two categories. For smaller and larger infarcts, the trees are deeper involving multiple coarse-to-fine optimizations and an average of 10-12 final dimensions of unknowns (Fig. 3.7). Such cases in general require a larger number of model evaluations. Our experiments show that the presented methods take at most twice as many model evaluations. However, it should be noted that uniform BOBYQA for such cases as shown in Fig 3.5 and 3.6 suffer from limited accuracy in estimation. For average sized infarcts, the trees are shallower and end up with an average of 7-8 unknown dimensions (Fig. 3.7), resulting in a fewer number of coarse-to-fine optimizations of a smaller number of unknowns. Compared to uniform BOBYQA, the presented methods require a similar number of model evaluations. In the presented methods, the adaptive spatial resolution adjustment step also necessitates model evaluations. Supposing that at a stage there are $2^n$ leaf nodes, a minimum of $n$ model evaluations (when all nodes have sibling) and a maximum of $2n$ model evaluations (when no nodes have sibling) are needed. At the current stage where the number of leaf nodes are typically $< 20$, the number of model evaluations needed for spatial resolution adjustment is insignificant in comparison to that needed for optimization (at a range thousands).

3.3.3 Identifiability Based on the Size of the Infarct

To investigate the limit of the presented method in terms of the smallest size of the infarct that can be identified, we conduct a set of experiments by gradually decreasing the size of an infarct until a size that cannot be identified is found. For such a size of infarct, six different experiments are repeated. In summary, an infarct of size $\geq 3\%$ can be estimated with good accuracy, $\sim 2\%$ can be identified with lower accuracy, and $\sim 1\%$ cannot be identified. Fig. 3.8 shows
Figure 3.8: Synthetic experiments. Left: tree narrows down along the infarct region showing that the infarct is identified. Right: tree takes first a few steps along the infarct region, but later splits along homogeneous healthy region showing that the infarct is not identified.

two examples of estimation for infarcts of different sizes. For an infarct of size 3.6%, the tree progressively branches by allocating higher resolution along the infarct and lower resolution along the homogeneous normal regions. However, for an infarct of size 2.0%, the tree branches along the infarct only for first a few steps (until H-3) and later branches along the homogeneous normal tissue. This shows that no significant gain in objective function was obtained by dividing the heterogeneous regions containing the infarct. Rather, the gain in objective function by refining the homogeneous regions (fine tuning of parameter values) is comparable to that by refining the heterogeneous region. Hence, in this case the infarct was not observable given the current data and the objective function used.

3.3.4 Identifiability Based on the Amount of Measurement Data

We investigate the change in the performance of the presented method when the amount of measurement data is decreased from 120-lead ECG to 12-lead ECG on 35 synthetic experiments. Experimental setup as described in Section 3.3 is used. In addition, 12-lead ECG is extracted from the simulated and noisy 120-lead ECG. As summarized in Fig. 3.9, an overall decrease in the accuracy across both metrics is observed. This is mainly due to larger false positive regions and sometimes non-unique solutions. The accuracy in estimation when infarcts are located at inferior LV and RV regions shows major decline. Fig 3.10 shows an example of estimation when an infarct is present in the inferior LV. Using 120-lead ECG, the presented method is able to progressively narrow down to the infarct region. Using 12-lead ECG, all
Figure 3.9: Synthetic experiments. Performance of the presented method using 12-lead ECG in comparison to that using 120-lead ECG in terms of DC and RMSE. Comparison is based on 35 different infarct cases of varying sizes and locations (bar: mean, line: standard deviation).

Figure 3.10: Synthetic experiments. Estimation using 120-lead ECG shows good accuracy with splits along the infarct region, whereas using 12-lead ECG is unable to identify the infarct.

The nodes in the tree take the value of 0.15 showing that the infarct is not identified. The drop in accuracy, especially in the RV and inferior LV, could be explained by the limited presence of leads in the inferior and RV side in the 12-lead ECG set.
3.4 Evaluation on a Blinded EP Model and \textit{In-vivo} MRI Scar

We next test the performance of the presented method in estimating tissue property of 3D myocardial infarct delineated from high resolution \textit{in-vivo} magnetic resonance (MR) images. Compared to the settings in synthetic experiments, these MRI-derived 3D infarct has the following characteristics that can be expected to increase the difficulty of local parameter estimation: 1) heterogeneous tissues due to the presence of both dense scar core and gray zone, 2) the presence of a single or multiple scars with complex spatial distribution and irregular boundary, and 3) the presence of both transmural and non-transmural scars. Due to the lack of \textit{in-vivo} electrical mapping data, here measurement data for parameter estimation are generated by a high resolution (average resolution of 350\,µm), multi-scale (sub-cellular to organ scale) \textit{in-silico} ionic electrophysiological model on the MRI-derived patient-specific ventricular models as detailed in [4]. Data are extracted from 300-400 epicardial sites, temporarily down-sampled to a 5\,ms resolution, and corrupted with 20\,dB Gaussian noise. Note that although no \textit{in-vivo} electrical data were available for parameter estimation, the experiments are designed to mimic a real-data scenario because: 1) the 3D electrophysiological model used to generate the measurement data is known to be capable of generating high-fidelity electrophysiological simulation of patient-specific heart [4], 2) this model is unknown to the framework of parameter estimation and thus no “inverse crime” is involved, and 3) only a subset of epicardial surface data corrupted with noise is used as measurement data. Fig. 3.11 shows the local tissue excitability estimated using the presented method on 7 cases. Based on the fundamental assumption of the presented method (higher resolution in heterogeneous regions), we expect that its performance will be closely tied to the underlying heterogeneity of the tissue property. Therefore, we analyze the performance of the presented method with respect to the following two factors that affect the heterogeneity of the scar:

1) \textit{Scar Transmurality}: Because scar non-transmurality increases tissue heterogeneity across the myocardial wall, it is expected that the estimation accuracy in the presence of dense transmural scars would be higher than that in the presence of non-transmural scars. Specifically cases 4-7 have some dense transmural scars surrounded by gray zone. In these cases, the location of the
dense transmural scar is accurately identified with the estimated parameter value close to the documented parameter value for infarct core (i.e., 0.5) [14]. In comparison, the non-transmural scars in case 1 anterior-septal region and case 4 apical region are missed. The tissue property for the non-transmural scar in case 3 apical region and case 5 inferior region, although identified, are not accurately estimated. Overall, in the presence of non-transmural scars, the estimation either misses the scar, or produces regions of abnormal tissues larger than the scar with parameter value in between the values for healthy and infarct core.

2) Dense vs. Gray Zone: The presence of gray zone is also expected to increase the difficulty in parameter estimation. For example, tissue property in dense scar in cases 4, 6 apical region and case 2 anterior region are correctly estimated, whereas the gray zone in case 1 septal region is missed. Additionally, as shown in Fig. 3.11 case 5 view 1, the presence of gray zone within the dense scar decreases the estimation accuracy. In these cases, the location of scar is identified, but the value of the parameter does not accurately reflect the heterogeneous presence of the dense scar and the gray zone.
3.5 Evaluation on 120-lead ECG and Catheter Data

We conduct real-data study on two patients who underwent catheter ablation of ventricular tachycardia due to prior tissue infarction as described in [66]. The patient-specific heart-torso geometry is constructed from 3D CT images. The tissue excitability is estimated from 120-lead ECG using adaptive GPO, adaptive BOBYQA, and uniform BOBYQA. Because uniform BOBYQA is sensitive to initialization, we use the result of the global parameter estimation to initialize this optimization. For validation of the result, we consider: 1) the relation between the estimated tissue excitability and in-vivo epicardial bipolar voltage data obtained from catheter mapping, and 2) the similarity between the real ECG data and those simulated with the estimated tissue excitability. Note that voltage data can be used only as a reference and not the gold standard for the estimated tissue excitability.

Fig. 3.12 shows a comparison between the estimated tissue excitability and the bipolar voltage data, in which the first column shows the original catheter maps (red: dense scar, pink: healthy tissue), the second column shows the catheter mapping registered to the CT-derived cardiac mesh (red: dense scar $\leq 0.5$ mV, green: scar border $= 0.5 - 1.5$ mV, blue: normal $> 1.5$ mV). The catheter maps from patient case 1 reveal a dense scar on the basal lateral region of the LV. As shown in Fig. 3.12, tissue excitability estimated using the presented adaptive methods reveal abnormal tissues in the same region. In this case, tissue excitability estimated using the uniform BOBYQA does not align with the catheter maps. Similarly, the catheter maps from patient case 2 reveal a dense infarct distributed across the lateral and inferior LV regions. As shown in Fig. 3.12, the tissue excitability estimated from the presented methods and uniform BOBYQA successfully reveal the dense abnormal tissues in this region. However, the estimation results from uniform BOBYQA also shows some dense abnormal tissues in the apical region.

Fig. 3.13 shows a comparison of the real ECG data and those simulated with the estimated tissue excitability on a few example leads. In patient case 1, the simulated ECG fits the real ECG data on majority of the leads, although the fit is poor on a few leads (see an example of a good and a poor fit in Fig. 3.13(a)). The average RMSE between the simulated and real ECG across all the leads is 0.2133 mV. As shown in Fig. 3.13(b), in patient case 2, the simulated ECG fits the real ECG data with lower accuracy: RMSE 0.4388 mV.
3.6 Discussion

3.6.1 Threshold for Adaptive Coarsening

In the presented framework, if the refinement of a parent node into two leaf nodes does not yield a significant gain in the objective function, such refinements are retracted. To do this, a threshold value in the gain of the objective function must be pre-determined. We study the influence of this threshold value using synthetic experiments on five different infarct cases. On each case parameter estimation is done with the following values of the threshold: 0.1%, 0.5%, 1%, 5%, 10% and 20% of the maximum gain in global optimum at the present resolution. The following observations are made: When a small value of the threshold is used (0.1%), the refinements that contribute small gain
to the global optimum are retained, resulting in a large number of unknowns in the higher levels of the multi-scale hierarchy and a decreased ability to go deeper into the multi-scale hierarchy. As a result, wider but shallower trees are obtained. Alternatively, when a large value of the threshold is used (20%), only those refinements that contribute a large gain in the global optimum are retained, resulting in an increased ability to go deeper in the multi-scale hierarchy but at the cost of accuracy. As a result, the final tree is narrower and deeper. In overall, as shown in Fig. 3.14 left, a threshold value of 5% resulted in good accuracy consistently.

3.6.2 Effect of Refining Multiple Nodes Versus a Single Node

Because an optimization method can accurately estimate only a limited dimensions of unknowns in local parameter estimation, we choose to create spatial partitions in the cardiac mesh sparingly. Thus, we refine a single node at each scale. Alternatively, the refinement of more than one node could result in a different final resolution of the cardiac mesh, and consequently different parameter values. On eight synthetic experiments, we test the performance of the presented method when one node is refined vs. when two nodes are refined. As shown in Fig. 3.14 right, because the refinement of two nodes results in a need to optimize a larger number of unknowns at each scale, there is a decrease in the accuracy. The ability of the presented framework to go deeper into the multi-scale hierarchy is also impacted.
3.6.3 Parameters of the GPO

The GP is dependent on three kernel parameters: 1) the length scale ‘γ’, 2) the covariance amplitude ‘α’, and 3) the observation noise ‘ς’ as described in Section 3.2.2. The optimization of these kernel parameters along with each update of the GP could result in higher accuracy. The most commonly used approach is to maximize the marginal likelihood under the current GP, 

\[ p(J|\theta_{1:n},\alpha,\gamma,\varsigma) = \mathcal{N}(J|\mu_1, K_{\alpha,\gamma} + \varsigma I) \].

As shown in Fig. 3.15 left, adaptive GPO with kernel parameter optimization does achieve higher accuracy compared to that without. Therefore, we will include kernel parameter optimization in the future development of the presented method.

3.6.4 Performance in the Presence of Two Infarcts

On six synthetic experiments, we test the performance of the presented framework in the presence of two infarcts. As shown in Fig. 3.15 right, both the infarcts are revealed. However, when two infarcts are spatially close some false positives might be obtained at the narrow healthy tissue region separating the two infarcts (Fig. 3.15 right case (c)). A large number of studies will be carried out in future to investigate the performance of the presented framework in the presence of multiple infarcts.

3.7 Conclusion

This chapter presented a novel framework that is able to estimate spatially-varying parameters using a small number of unknowns, achieved through a
coarse-to-fine optimization and a spatially-adaptive resolution that is higher at heterogeneous regions. It is demonstrated on the estimation of local tissue excitability in a cardiac EP model on both synthetic and real-data experiments. Through comparison studies with optimization using pre-defined segments, we show the ability of the presented framework in revealing heterogeneous tissue properties with higher accuracy. One improvement to this framework is to improve its ability to go deeper and wider into the multi-scale hierarchy. Another future step is to integrate the presented framework with a probabilistic estimation to quantify the uncertainties in local parameters.
Chapter 4

High-dimensional Bayesian Optimization via an Embedded Generative Model

A common solution to reduce the dimension of unknown parameters is to partition the anatomical mesh into a fixed small number of segments [83, 87]. In Chapter 3 we presented an alternative approach in which the partitioning of the cardiac mesh into segments is obtained through a coarse-to-fine optimization along a pre-defined multi-scale hierarchy of the cardiac mesh. Both strategies are based on anatomy-based explicit partitioning of the cardiac mesh into a set of segments. Without a priori knowledge of the distribution of the tissues on the cardiac mesh, the low-dimensional (LD) parameter space thus obtained could result in solutions that are either too low in resolution to reflect tissue heterogeneity or too high in dimension to be reliably estimated.

In this chapter, we present a novel concept that embeds a generative variational auto-encoder (VAE) into the objective function of Bayesian optimization, providing an implicit LD search space that represents the generative code of the high-dimensional (HD) spatially-varying tissue properties. In addition, the VAE-encoded knowledge about the generative code is further used to guide the exploration of the search space. By embedding a generative process into the objective function of a surrogate-based Bayesian optimization we are able to embed both HD surrogate learning and active search of sample points into a LD latent space enabling a HD surrogate-based Bayesian optimization.
4.1 Introduction

Numerous works have been presented in estimating the spatially-varying tissue properties in the form of high-dimensional (HD) model parameters. Most existing methods choose to represent spatially-varying tissue properties via a low-dimensional (LD) partitioning of the underlying geometrical model, either as pre-defined segments [82], or iteratively optimized in a coarse-to-fine fashion [17, 68]. This LD-to-HD definition directly exploits the spatial proximity and hierarchical composition of the underlying geometry. However, it is of such limited expressiveness that the number of partitioning is either too low to faithfully represent high-resolution tissue properties, or too high to allow effective optimization.

Here, we present a drastic alternative that replaces the explicitly defined LD or multi-scale representation of the parameter space with an implicit LD latent encoding. It is achieved by embedding within the optimization a stochastic LD-to-HD generative model that describes the generation of the HD spatially-varying tissue properties from a LD manifold. This generative model is obtained with a variational auto-encoder (VAE), trained from a large set of spatially-varying tissue properties reflecting regional tissue abnormality with various locations, sizes, and distributions. Once trained, the VAE is integrated with a surrogate-based Bayesian optimization [7] in two novel ways. First, the generative model (the VAE decoder) is embedded within the objective function to provide an implicit LD search space for the optimization of HD parameters. Second, the posterior distribution of the LD latent code as learned from the VAE encoder is used as prior knowledge within the BO for an efficient exploration of the LD manifold. To the best of our knowledge, this is the first work that utilizes a probabilistic generative process within an optimization framework for estimating HD patient-specific model parameters.

The presented method is applied to the estimation of local tissue excitability of a cardiac electrophysiological model using non-invasive electrocardiogram (ECG) data. On both synthetic and real data experiments, the presented method is compared against existing methods based on explicitly-defined LD [83] or multi-scale representation of the parameter space [22]. Experiments demonstrate that the presented method can achieve a drastic reduction in computational cost while improving the accuracy of the estimated parameters. Beyond the specific model considered in this study, the presented
method provides an efficient and reliable solution to a wider range of HD model parameter estimation problems.

4.2 HD Parameter Estimation

The HD parameter $\theta$ in the electrophysiological model (2.1) can be estimated by minimizing the sum of the squared errors between the measured physiological signals $Y$ and signals simulated by the composite of the electrophysiological model and the measurement model: $M(\theta) = H u(\theta)$. This can be achieved by solving the following maximization objective:

$$\hat{\theta} = \arg \max_{\theta} L(\theta) = \arg \max_{\theta} \{-||Y - M(\theta)||^2\}. \quad (4.1)$$

To enable the estimation of $\theta$ at the resolution of the cardiac mesh, the presented method embeds within the Bayesian optimization framework a stochastic generative model that generates $\theta$ from a LD manifold. It includes two major components as outlined in Fig. 4.1: 1) the construction of a generative model of HD spatially-varying tissue properties at the resolution of the cardiac mesh, and 2) a novel Bayesian optimization method utilizing the embedded generative model.

![Figure 4.1: Outline of the Bayesian Optimization with an embedded Generative Model, with the dimension of each VAE layer labeled.](image-url)
4.2.1 LD-to-HD Parameter Generation via VAE

Generative VAE model: We assume that the spatially varying tissue properties at the resolution of a cardiac mesh \( \theta \) is generated by a small number of unobserved continuous random variables \( z \) in a LD manifold. To obtain the generative process from \( z \) to \( \theta \), the VAE consists of two modules: a probabilistic deep encoder network with parameters \( \alpha \) that approximates the intractable true posterior density as \( q_\alpha(z|\theta) \); and a probabilistic deep decoder network with parameters \( \beta \) that can probabilistically reconstruct \( \theta \) given \( z \) as \( p_\beta(\theta|z) \). Both networks consist of three fully-connected layers as shown in Fig. 4.1.

To train the VAE, we generate \( \Theta = \{\theta^{(i)}\}_{i=1}^N \) consisting of \( N \) configurations of heterogeneous tissue properties in a patient-specific cardiac mesh. The training involves optimizing the variational lower bound on the marginal likelihood of each training data \( \theta^{(i)} \) with respect to network parameters \( \alpha \) and \( \beta \):

\[
\mathcal{L}(\alpha; \beta; \theta^{(i)}) = -D_{KL}(q_\alpha(z|\theta^{(i)})||p_\beta(z)) + E_{q_\alpha(z|\theta^{(i)})} \log p_\beta(\theta^{(i)}|z), 
\]

where we model \( p_\beta(\theta|z) \) with a Bernoulli distribution. To optimize Eq. (4.2), stochastic gradient descent with standard backpropagation can be utilized. Assuming the approximate posterior \( q_\alpha(z|\theta^{(i)}) \) as a Gaussian density and the prior \( p_\beta(z) \sim \mathcal{N}(0, 1) \), their KL divergence can be derived analytically as:

\[
D_{KL}(q_\alpha(z|\theta^{(i)})||p_\beta(z)) = -\frac{1}{2} \sum_j (1 + \log(\sigma_j^2) - \mu_j^2 - \sigma_j^2),
\]

where \( j \) is along the dimensions of \( z \), and \( \mu \) and \( \sigma^2 \) are mean and variance from \( q_\alpha(z|\theta^{(i)}) \). Because stochastic latent variables are utilized, the gradient of the expected negative reconstruction term during backpropagation cannot be directly obtained. The popular re-parameterization trick is utilized to express \( z \) as a deterministic variable as \( z^{(i)} = \mu^{(i)} + \sigma^{(i)} \epsilon \), where \( \epsilon \sim \mathcal{N}(0, I) \) is noise [43].

Probabilistic modeling of the latent code: The trained encoder provides an approximated posterior density of the LD latent code \( q_\alpha(z|\theta) \). This represents valuable knowledge about the probabilistic distribution of \( z \) learned from a large training dataset. To utilize this in the subsequent optimization, we integrate \( q_\alpha(z|\theta) \) over the training data \( \Theta \) to obtain the density \( q_\alpha(z) \) as a mixture of Gaussians \( 1/N \sum_i^N \mathcal{N}(\mu^{(i)}, \Sigma^{(i)}) \), where \( \mu^{(i)} \) and \( \Sigma^{(i)} \) are mean and covariance from \( q_\alpha(z|\theta^{(i)}) \). Because the number of mixture components in \( q_\alpha(z) \)
scales linearly with the number of training data, we approximate \( q_\alpha(z) \) with a single Gaussian density as \( \mathcal{N}(1/N \sum_i^N \mu(i), 1/N \sum_i^N (\Sigma(i) + \mu(i) \mu(i)^T) - \mu \mu^T) \). Alternatively, we approximate \( q_\alpha(z) \) with a mixture of Gaussians with \( K \ll N \) components, where k-means clustering with the Bregman divergence [30] as a similarity metric is used to reduce the number of mixture components.

In this way, we obtain a generative model \( p_\beta(\theta|z) \) of HD tissue properties from an implicit LD manifold, and prior knowledge of the LD manifold \( q_\alpha(z) \) from the probabilistic encoder. Both will be embedded into Bayesian optimization to enable efficient and accurate HD parameter estimation.

4.2.2 Bayesian Optimization with Embedded Generative Model

Representing \( \theta \) with the expectation of the trained decoder \( p_\beta(\theta|z) \), we obtain:

\[
\hat{z} = \arg \max_z L(z) = \arg \max_z \|Y - M(E[p_\beta(\theta|z)])\|^2,
\]

which allow us to optimize the HD parameter \( \theta \) in an implicit LD manifold of \( z \). For Bayesian optimization, we assume a zero mean Gaussian process (GP) with an anisotropic Matérn 5/2 kernel as a prior over the objective function (4.4). The optimization then consists of two iterative steps: 1) select point in the LD manifold that allows the GP to globally approximate Eq. (4.4) (exploration) while locally refining the area of optimum (exploitation); and 2) update the GP.

**VAE-informed Acquisition function:** To select points on LD manifold, we adopt the expected improvement (EI) function that picks a point with maximum expectation of improvement over the current best objective function value \( f_m \) [7]. For a GP posterior \( \sim \mathcal{N}(\mu(.), \sigma(.)) \), it can be obtained as:

\[
\text{EI}(z) = (\mu(z) - f_m)\Phi\left(\frac{\mu(z) - f_m}{\sigma(z)}\right) + \sigma(z)\phi\left(\frac{\mu(z) - f_m}{\sigma(z)}\right),
\]

where \( \Phi \) and \( \phi \) are the cumulative distribution function and density function of the standard normal distribution. Here, the first term controls exploitation (through high \( \mu \)) and the second term controls exploration (through high \( \sigma \)). Because using only \( f_m \) can lead to excessive exploitation, it is common to augment \( f_m \) with a constant trade-off parameter \( \varepsilon \) as: \( f_m + \varepsilon \) for increased
CHAPTER 4. HD BAYESIAN OPTIMIZATION VIA AN EMBEDDED VAE

exploration [7]. Here, we utilize the VAE-encoded knowledge about the LD manifold \( q_a(z) \) to enforce higher exploration in the areas of high probability density for \( z \), and lower elsewhere. Specifically we define \( \varepsilon(z) = -\frac{1}{m} \sum_i w_i (z - \mu_i) \Sigma_i^{-1} (z - \mu_i) \), where \( w, \mu, \) and \( \Sigma \) are the weight, mean, and variance of the \( K \) Gaussian mixture components in \( q_a(z) \).

**GP Update**: After a new point \( z^{(n)} \) is selected by maximizing the modified EI, the objective function (4.4) is evaluated at the HD parameter given by the mean of the generative model \( p_g(\theta | z^{(n)}) \). The GP is then updated by adding the new pair of \( z^{(n)} \) and objective function value, and maximizing the log marginal likelihood with respect to kernel parameters: length scales and kernel amplitude.

### 4.3 Evaluation on Synthetic Data

We include 27 synthetic experiments on three CT-derived human heart-torso models. In each case, an infarct sized \( 2\% - 40\% \) of the heart was placed at differing locations using various combinations of the American Heart Association (AHA) segments [9]. The value of the parameter \( \theta \) in the infarcted and the healthy region is set to \( 0.5 \) and \( 0.15 \), respectively. 120-lead ECG is simulated and corrupted with 20dB Gaussian noise as measurement data. We evaluate the accuracy in estimated parameters with two metrics: 1) root mean square error (RMSE) between the true and estimated parameters; and 2) dice coefficient (DC) = \( \frac{2(|S_1 \cap S_2|)}{|S_1| + |S_2|} \), where \( S_1 \) and \( S_2 \) are the sets of nodes in the true and estimated regions of infarct; these regions are determined by Otsu’s thresholding method [55].

### 4.3.1 VAE Architecture and Training

For each heart, we generate a training dataset of tissue properties with various heterogeneous infarcts. Each infarct is generated by random region growing in which, starting with one infarct node, one out of the five closest neighbors of the present infarct is randomly added to the infarct until an infarct of desired size is obtained. It is then added to training data. Because infarcts thus generated tend to be very irregular, we also include infarcts generated by growing the infarct with the node closest to its center. For each heart,
CHAPTER 4. HD BAYESIAN OPTIMIZATION VIA AN EMBEDDED VAE

4.3.2 Comparison with Existing Methods

The presented method (termed as BO-VAE) is compared against two common approaches based on explicit LD representation: 1) optimization over fixed 18 segments (fixed-segment (FS) method); and 2) coarse-to-fine optimization along a fixed multi-scale hierarchy (fixed-hierarchy (FH) method). As summarized in Figs. 4.2(a)-(b), BO-VAE (blue bar) is more accurate than the other two methods (green bars) in both DC and RMSE (paired t-tests, p < 0.012). This is achieved at a reduction of the computational cost by: 87.57% for the FS method and 98.73% for the FH method (Fig. 4.2(c)).

As expected, the FS method shows the lowest accuracy: as illustrated in Fig. 4.3, the estimated parameters either miss the infarct or include large
false positives. The FH method overcomes this issue, although only to a limited extent. As shown in the final multi-scale representation obtained by the FH method in the right panel of Fig. 4.3: several dimensions are wasted at representing homogeneous healthy regions (green nodes) across different scales, which limits its ability to optimize deeper along the tree. In contrast, BO-VAE is not limited by such explicitly-imposed anatomy-based structure, allowing it to attain higher accuracy with only two latent dimensions and 1-10% of the computation time.

4.3.3 The Effect of VAE-encoded Knowledge About the LD Manifold

To study the effect of incorporating the VAE-encoded $q_\alpha(z)$ in the EI function, we compare the standard EI with EI augmented with three types of distributions on $z$: 1) $p_\beta(z) \sim N(0, 1) \text{ (EI Isotropic)}$, 2) approximated $q_\alpha(z)$ with a single Gaussian density (EI Post-1); and 3) approximated $q_\alpha(z)$ with a mixture of 10 Gaussian densities (EI Post-K). As shown in Fig. 4.2, the estimation accuracy using all three distributions is higher than that without using any, among which EI Post-1 has the highest accuracy. Fig. 4.4 illustrates that, when $q_\alpha(z)$ is utilized, the exploration gradually proceeds from the region of high probability density to the region of low probability density (Fig. 4.4(b)). In comparison, with standard EI, the exploration is spread in an attempt to reduce overall variance (Fig. 4.4(a)); this could result in incorrect (Fig. 4.4(c)) or suboptimal (Fig. 4.4(d)) solutions.

We also experimented with HD latent code $z$. As shown in Fig. 4.5(a)-
Figure 4.5: (a-b): Examples of estimated parameters using five vs. two dimensional latent codes. (c-d): Latent code manifold based on (c) infarct location, and (d) infarct size.

Figure 4.6: Model parameter estimated with BO-VAE, FH, and FS on real-data study.

(b), there was only a marginal improvement in accuracy with a five vs. a two dimensional (2d) latent code. It suggests that, given the focus of the training data on local infarcts, a 2d latent code may be sufficient to capture the necessary generative factors. The plot of these 2d latent codes in Fig. 4.5(c)-(d) show that they cluster by infarct location and their radial direction accounts for the infarct size.

4.4 Evaluation on 120-lead ECG and Catheter Data

Real-data studies are conducted on two patients with previous myocardial infarction. Patient-specific heart and torso meshes are constructed from axial CT images. Tissue excitability is estimated from 120-lead ECG data. The results are evaluated by in-vivo bipolar voltage data which, although not a direct measure of tissue excitability, provides a reasonable reference about the
region of infarcts. The first two columns of Fig. 4.6 show the original voltage data (red: dense infarct; purple: healthy tissue; green: infarct border) and the same data registered to cardiac meshes.

The voltage map in case 1 (Fig. 4.6(a)) shows a highly heterogeneous infarct spread over a large region in the lateral LV. The estimated parameters by all methods capture this region of infarct. For this accuracy, the FH and FS methods required 4056 and 1058 model evaluations, whereas BO-VAE required only 105 model evaluations. By contrast, as shown in Fig. 4.6(b), case 2 has a smaller region of dense scar in the lateral LV. The estimated parameters by BO-VAE and FH correctly reveal this region of scar, whereas the FS method is less accurate. In this case, BO-VAE required 105 model evaluations in comparison to the FH and FS methods that required 5798 and 1501 model evaluations, respectively.

4.5 Conclusion

In this chapter, we presented a novel approach to estimating HD model parameters by optimizing their LD generative code, achieved by embedding within the Bayesian optimization a generative model of HD tissue properties from a LD manifold. Experiments show a gain in accuracy with drastically reduced computation. Future improvements include two direction: 1) to incorporate more realistic training data from high resolution 3D imaging for a more expressive generative model and potentially improved estimation of highly heterogeneous tissues; and 2) to investigate alternative means to incorporate the knowledge of latent manifold to more efficiently guide the selection of points during optimization.
Chapter 5

Bayesian Optimization on Large Graphs via a Graph Convolutional Generative Model

Most previous works on parameter estimation as described in Chapters 2 and 3 rely on a geometry-based partitioning of the anatomical model to obtain a low-dimensional (LD) representation of the spatially varying unknown tissue properties in optimization. While these methods exploit the anatomical information of the heart, the pre-defined set of partitions have limited ability to represent the complex tissue distribution and heterogeneity at a high-resolution. Chapter 4 presented an alternative to utilize a variational auto-encoder as a more expressive low to high-dimensional (HD) generative model that allowed to embed the HD optimization into the LD latent space. Its Euclidean nature, however, neglects the rich anatomical information in the heart.

In this chapter, we present a novel graph convolutional generative model to allow generative modeling of non-Euclidean data, and utilize it to embed Bayesian optimization of large graphs into a small latent space. This approach bridges the gap of previous works by introducing an expressive generative model that is able to incorporate the knowledge of spatial proximity and hierarchical compositionality of the underlying geometry.
5.1 Introduction

Numerous works have been presented in estimating spatially-varying tissue properties in the form of high-dimensional (HD) model parameters for model personalization. Most existing methods choose to represent spatially-varying tissue properties via a low-dimensional (LD) partitioning of the underlying geometrical model, either as pre-defined segments [32, 49, 63, 82, 83, 87], or iteratively optimized in a coarse-to-fine fashion [10, 17, 68]. This LD-to-HD definition directly exploits the spatial proximity and hierarchical composition of the underlying geometry. However, it is of such limited expressiveness that the number of partitioning is either too low to faithfully represent high-resolution tissue properties, or too high to allow effective optimization.

By contrast the work in Chapter 3 presented the use of a data-driven generative model of HD tissue properties, via a variational auto-encoder (VAE), to embed the optimization into a LD latent space [20]. Being more expressive, this VAE-based generative model is able to represent high-resolution tissue properties with a latent code sufficiently small for effective optimization. However, as the regular VAE is defined over Euclidean data, it does not take into account the valuable geometry information in the data, nor does it allow transferring among different geometry without first establishing point-by-point correspondence.

If we view organ tissue properties over a 3D geometrical model as an image, convolutional neural networks (CNNs) are a natural choice to incorporate knowledge of the spatial proximity and hierarchical composition of the image [8]. However, standard CNNs have been most successful on data with an underlying Euclidean structure (i.e., image grids). Generalizing CNNs to non-Euclidean domains is an emerging area of research [8], where significant efforts have been presented on addressing the challenges of defining convolution [28], pooling, and up-sampling operations [85]. However, most developments to non-Euclidean CNNs are focused on supervised discriminative networks. To date, very limited work have been presented to enable generative modeling of non-Euclidean data.

In this chapter, we present a novel VAE architecture that allows generative modeling of data over non-Euclidean domains, and utilize this generative model to embed Bayesian optimization of large graphs into a LD manifold. The presented approach bridges the gap of previous works by introducing
an expressive generative model that is able to represent high-resolution tissue properties with a small latent code, while incorporating the geometrical knowledge in the data and being transferable across geometries. We evaluate the presented method in synthetic and real-data experiments of estimating tissue excitability in a cardiac electrophysiological model, where we compare the expressiveness of the generative model and the accuracy of the subsequent parameter optimization with those obtained by using a linear reconstruction model based on principal component analysis (PCA) and a regular fully-connected VAE [20]. We further demonstrate the feasibility of transferring the presented non-Euclidean VAE across patients. To our knowledge, this is the first introduction of a graph convolutional VAE and its use to enable Bayesian optimization over large graphs.

5.2 Personalizing HD Parameters on Unstructured Meshes

We seek parameter $\theta$ that minimizes the sum of squared errors between model output $M(\theta) = H(\theta)$ and patient’s measurements $Y$ as:

$$\hat{\theta} = \arg \max_\theta \{-||Y - M(\theta)||^2\}. \quad (5.1)$$

Directly solving (5.1) for the spatially-distributed $\theta$ is difficult [17]. Below we describe how we learn a LD-to-HD generation of $\theta$ that accounts for the underlying geometry, and embed the HD optimization into the expressive LD manifold.

5.2.1 Graph Convolutional VAE

We model the generation of spatially-distributed $\theta$ with a VAE [43]. A VAE consists a probabilistic encoder network with parameters $\alpha$ that approximates the intractable true posterior density as $q_{\alpha}(z|\theta)$; and a probabilistic decoder network with parameters $\beta$ that represents the likelihood as $p_{\beta}(\theta|z)$. For data defined on a Euclidean grid, structural information is incorporated in VAE through CNNs. Here, we present a novel VAE architecture that enables convolution, pooling, and unpooling over non-Euclidean geometry of the heart.
Local Connectivity & Graph Convolution: We model the cardiac mesh as a graph: \( G = (\mathcal{V}, \mathcal{E}, \mathcal{U}) \), where vertices \( \mathcal{V} \) consist of all \( N \) meshfree nodes and edges \( \mathcal{E} \) exist between each meshfree node and its \( k \) nearest neighbors. \( \mathcal{U} \in [0, 1]^{N \times N \times 3} \) consists of edge attributes \( \mathbf{u}(i,j) \), calculated as the normalized \((x_1 - x_2, y_1 - y_2, z_1 - z_2)\) if an edge \((i,j) \in \mathcal{E}\) exists between vertices at \((x_1, y_1, z_1)\) and \((x_2, y_2, z_2)\) and 0 otherwise. On this graph, we use a convolution operator based on spatial continuous convolution kernels because it was shown to allow better generalization to similar graphs [28]. Specifically given the graph \( G \) and \( M \)-dimensional input features \( \mathbf{f}(i) | i \in \mathcal{V} \), the \( l \)-th convolution kernel is:

\[
g_l(\mathbf{v}) = \sum_{\mathbf{p} \in \mathcal{P}} w_{\mathbf{p},l} \prod_{i=1}^{d} N_{m, \mathcal{P}_{i,p}}(v_i),
\]

where \((N_{m,1}^{(1)}(\mathcal{P}_{1,p}), \ldots, N_{m,1}^{(d)}(\mathcal{P}_{d,p}))\) denotes \( d \) open B-spline bases of degree \( m \) based on equidistant knot vectors with \( d \)-dimensional kernel size of \( \mathbf{k} = (k_1, \ldots, k_d) \), \( \mathcal{P} \) is the Cartesian product of the B-spline bases, and \( w_{\mathbf{p},l} \) are the trainable parameters. Given kernel functions \( \mathbf{g} = (g_1, \ldots, g_M) \) and input features \( \mathbf{f} \in \mathcal{R}^M \), the spatial convolution operator for each vertex \( i \in \mathcal{V} \) with a neighborhood \( \mathcal{N}(i) \) based on its edge connectivity is then defined as:

\[
(f \ast \mathbf{g})(i) = \frac{1}{|\mathcal{N}(i)|} \sum_{l=1}^{M} \sum_{j \in \mathcal{N}(i)} f_l(j) g_l(\mathbf{v}(i,j)).
\]

Hierarchical Composition & Pooling: To define pooling and unpooling operations necessary for the encoding-decoding architecture, a hierarchical representation of the graph is needed. We obtain this by an efficient
multilevel graph clustering method based on minimizing the normalized cuts (Graclus) [24], which reduces the graph size by half the number of vertices at each coarsening. We store hierarchical graph representation in matrices which reduces pooling/unpooling operations to efficient matrix multiplications. Specifically if $G$ is a graph with $N_1$ vertices and $G_c$ is its coarsened graph with $N_2 < N_1$ vertices, we populate a binary matrix $P_{N_1 \times N_2}$, where $P_{ij} = 1$ if the $i^{th}$ vertex in $G$ was grouped to the $j^{th}$ vertex in $G_c$ and $P_{ij} = 0$ otherwise. Given $M$ feature maps $F \in \mathbb{R}^{N_1 \times M}$ over the vertices of graph $G$ and $F_c \in \mathbb{R}^{N_2 \times M}$ over graph $G_c$, the average pooling in the encoder can be obtained by $F_c = P_n^{T} F$ and unpooling in the decoder by $F = P F_c$, where $P_n$ is column normalized from $P$.

Graph Convolutional VAE: Using these building blocks we construct a VAE architecture as shown in Fig. 5.1. It is trained by optimizing the variational lower bound on the marginal likelihood of the training data $L(\alpha; \beta; \theta(i)) = -D_{KL}(q_\theta(z|\theta(i)) || p(z)) + E_{q_\theta(z|\theta(i))}[\log p_\theta(z|\theta(i))|z])$. We set $q_\alpha(z|\theta)$ and $p_\beta(z|\theta)$ to be Gaussian parameterized by the graph convolutional networks. The prior $p(z) \sim \mathcal{N}(0,1)$ is set to be an isotropic Gaussian, producing an analytical form for the KL divergence. Using the reparameterization trick [43], standard stochastic gradient methods can be used to optimize $L(\alpha; \beta; \theta(i))$.

5.2.2 Bayesian Optimization on Large Graphs

Bayesian optimization is a popular choice in optimizing complex objective functions such as (5.1) [7]. It begins by defining a surrogate over the objective function. The optimization then consists of two iterative steps: 1) actively find a point that optimizes a utility function based on the surrogate, and 2) update the surrogate with the newly-selected point. Direct Bayesian optimization over HD space is difficult and its use over large graphs has not been reported [7]. To enable this, we reformulate the original objective function in (5.1) as follows:

$$\hat{z} = \arg \max_z \{-||Y - M(E[p_\beta(\theta|z)])||^2\}. \quad (5.4)$$

This allows us to embed surrogate construction and active selection of training points in a LD manifold. We initialize the Gaussian process (GP) surrogate of (5.4) with a zero mean function and an anisotropic Matérn 5/2 kernel.
Active Selection of Training Points: To select a training point, we maximize the expected improvement (EI) utility function that favors a point with the highest expected improvement over the current optimum $f^+$ [7]:

$$
\text{EI}(z) = (\mu(z) - f^+)\Phi\left(\frac{\mu(z) - f^+}{\sigma(z)}\right) + \sigma(z)\phi\left(\frac{\mu(z) - f^+}{\sigma(z)}\right), \tag{5.5}
$$

where $\mu(z)$ and $\sigma(z)$ are the predictive mean and standard deviation of the GP, $\Phi$ is the cumulative normal distribution, and $\phi$ is the normal density function. The first term promotes exploitation, while the second term promotes exploration.

GP Update: After picking a new point $z^{(i)}$, the value of the optimization objective (5.4) is evaluated at $z^{(i)}$ as $J^{(i)}$. The GP is then updated by including the new input-output pair of $(z^{(i)}, J^{(i)})$ and maximizing the log marginal likelihood for kernel hyperparameters: length-scales and co-variance amplitude.

5.3 Synthetic Experiments

We evaluate the presented graph convolutional VAE (termed as gVAE) by: 1) its reconstruction accuracy, and 2) optimization accuracy of the gVAE-based Bayesian optimization, both in comparison to existing methods. Accuracy is evaluated by the sum of squared error (SSE) in $\theta$, and the dice coefficient (DC) of the abnormal region obtained by thresholding $\theta$ with Otsu’s method.

We synthetically generate data of heterogeneous tissue excitability via random region growing in each cardiac model. Beginning with a single meshfree node as abnormal, we randomly grow the abnormal region by adding one of the nearest neighbors of the abnormal nodes, until the abnormal region reaches a desired size (2% to 40% of the heart volume). On average, we generated 78,208 ± 12,541 data for training and 13,545 ± 7654 data each for validation and testing. The various layers and sizes of feature maps in each layer of the presented gVAE architecture are detailed in Fig. 5.1. We use B-spline basis degree of $m = 1$ with kernel size of $k_1 = k_2 = k_3 = 5$ in all graph convolution layers. All models are trained with a learning rate of 0.001 with Adam optimizer.
CHAPTER 5. BAYESIAN OPTIMIZATION VIA A GRAPH CONVOLUTIONAL VAE

5.3.1 gVAE as a Generative Model

On five different human heart models constructed from CT images, we first evaluate the ability of the presented gVAE model to reconstruct tissue excitability in comparison to: 1) PCA-based linear reconstruction, and 2) fully-connected VAE (termed as fVAE) [20] with three to five hidden layers (termed as fVAE-3h, -4h, and -5h, respectively). As summarized in Table 5.1, PCA being a linear model has the lowest accuracy. Compared to fVAE, the reconstruction accuracy of gVAE is consistently higher in both DC and SSE, even when its number of trainable parameters is similar to or lower than fVAE. However, gVAE is more expensive to train: 37.21 hrs vs. 7.51 mins for fVAE-3h in TITAN Xp GPU. Nevertheless, note that the number of trainable parameters for gVAE does not increase for larger meshes.

We further compare gVAE with two-dimensional (2d) latent codes vs. fVAE and PCA with various latent dimensions. As shown in Fig. 5.2(a), to achieve the reconstruction accuracy of gVAE with 2d latent codes, at least 13 principles components are required with PCA: this increase in dimension

<table>
<thead>
<tr>
<th>Anatomy</th>
<th>SSE 1</th>
<th>SSE 2</th>
<th>SSE 3</th>
<th>SSE 4</th>
<th>SSE 5</th>
<th>DC 1</th>
<th>DC 2</th>
<th>DC 3</th>
<th>DC 4</th>
<th>DC 5</th>
<th>Trainable parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCA</td>
<td>12.73</td>
<td>14.18</td>
<td>12.35</td>
<td>23.85</td>
<td>23.26</td>
<td>30.80</td>
<td>39.87</td>
<td>41.31</td>
<td>49.17</td>
<td>54.42</td>
<td>NA</td>
</tr>
<tr>
<td>fVAE-3h [20]</td>
<td>8.03</td>
<td>8.45</td>
<td>8.44</td>
<td>13.07</td>
<td>13.70</td>
<td>61.77</td>
<td>66.20</td>
<td>60.45</td>
<td>70.30</td>
<td>70.72</td>
<td>2,087,822</td>
</tr>
<tr>
<td>fVAE-4h</td>
<td>7.97</td>
<td>8.29</td>
<td>8.33</td>
<td>12.47</td>
<td>13.99</td>
<td>61.76</td>
<td>64.60</td>
<td>66.58</td>
<td>71.51</td>
<td>69.84</td>
<td>2,613,134</td>
</tr>
<tr>
<td>fVAE-5h</td>
<td>7.42</td>
<td>8.01</td>
<td>8.19</td>
<td>13.96</td>
<td>12.41</td>
<td>64.59</td>
<td>65.72</td>
<td>62.21</td>
<td>68.04</td>
<td>74.50</td>
<td>3,138,446</td>
</tr>
<tr>
<td>gVAE</td>
<td>6.66</td>
<td>6.89</td>
<td>6.79</td>
<td>11.28</td>
<td>11.43</td>
<td>68.43</td>
<td>70.92</td>
<td>70.70</td>
<td>75.10</td>
<td>76.86</td>
<td>2,778,069</td>
</tr>
</tbody>
</table>

Figure 5.2: (a) Comparison of reconstruction accuracy using gVAE with a 2d manifold vs. PCA and fVAE [20] with various-dimensional manifolds. (b)-(c): Plots of 2d latent codes from gVAE colored by infarct location (b) and infarct size (c).
will make the subsequent optimization difficult. With fVAE, a similar SSE is attained with four latent dimensions, while a similar DC could not be attained with even 50 latent dimensions. This may be because fVAE does not consider the geometry underlying the spatial distribution of tissue excitability. Fig. 5.2(b) and (c) show that the latent code learned with gVAE are clustered by the location of the abnormal tissue and its radial direction encodes the size of the abnormal tissue.

5.3.2 gVAE-based Parameter Optimization

On 40 synthetic cases on two different heart geometries, we conduct experiments on estimating unknown tissue excitability. In each case, we set an abnormal region by using various combinations of AHA segments which are very different from the training set. Measurement data was simulated, sub-sampled and corrupted with 20dB Gaussian noise. We compare the accuracy of gVAE-based Bayesian optimization with three previous approaches: 1) optimization on 17 fixed segments (termed as FS) [68, 82], 2) coarse-to-fine optimization along a fixed multi-scale mesh hierarchy (termed as FH) [17], and 3) optimization on a LD manifold obtained with fVAE-3h [20]. Result in Fig. 5.3(a) shows that gVAE-based Bayesian optimization is more accurate than all other approaches in both DC and SSE (paired t-test, $p < 0.05$). Fig. 5.3(c)-(f) shows visual comparison on a few examples. The computational cost of gVAE-based optimization is much lower than that of FH ($> 22x$) and FS ($> 7.5x$), and similar to that of fVAE-based optimization.
5.3.3 Feature Sharing Across Geometries

To demonstrate the feasibility of transferring the presented gVAE across different geometries, we take a pre-trained gVAE, fix the learned features in the encoder’s graph convolution layers, and fine-tune the remaining layers for a different anatomy. We compare this training strategy to training a gVAE from scratch. Results in Fig. 5.4(a) show that a pre-trained model can be fine-tuned with as small as 1088 new examples. In comparison, gVAE could not be trained from scratch with $\leq 7360$ samples, shown both by the low reconstruction accuracy (Fig. 5.4(a): DC = 0.24; SSE = 17.68) and a flat test loss plot (Fig. 5.4(b)). Test loss plots in Fig. 5.4(b) also show that a pre-trained model starts with a lower loss and a larger size of training data leads to faster convergence. Parameter optimization via a gVAE fine-tuned with 7360 data on 20 cases achieved an average DC and SSE of 53.10 and 11.01 respectively. Fig. 5.4(c) shows some examples of the estimated parameters.

5.4 Real Data Experiments

We conduct real-data studies on two patients with chronic myocardial infarction. Patient-specific heart-thorax models are obtained from axial CT images. Using 120-lead ECG as measurements, we evaluate the presented gVAE in estimating tissue excitability in comparison to the fVAE [20], FH [17], and FS [82] methods. Training dataset and network architectures are as described in Section 5.3. We qualitatively evaluate the results with in-vivo catheter mapping data which, as shown in Fig. 5.5, provides a reference for the location of
the abnormal (red, voltage $\leq 0.5\text{mV}$) and healthy (purple, voltage $> 1.5\text{mV}$) regions.

Case 1 has a large heterogeneous abnormal region in the lateral LV region (Fig. 5.5(a)). All methods are able to localize this region, but with varying degree of heterogeneity. Optimizations based on gVAE and fVAE are much faster requiring only 100 model evaluations, in comparison to FH and FS that required 4056 and 1058 model evaluations, respectively. By contrast, case 2 has a smaller but dense abnormal region in the lateral LV (Fig. 5.5(b)). While all methods identify the general location of this abnormality, gVAE more accurately differentiates the region of dense core and border. In comparison, fVAE and FH estimate a larger border region; and the abnormal region revealed by FS is less accurate. Again, in this case, gVAE and fVAE required only 100 model evaluations, whereas FH and FS required 5798 and 1501 model evaluations, respectively.

5.5 Conclusion and Future Work

We presented a novel graph convolutional VAE model that allows generative modeling of data defined over non-Euclidean structures and integrated it with Bayesian optimization to enable Bayesian optimization on data defined over large graphs. Experiments showed higher accuracy in both reconstructing the tissue excitability and estimating them from indirect measurements in comparison to existing baselines. A future direction is to incorporate realistic data from high resolution 3D images and investigate transfer learning of models trained on these data for accurate and efficient model personalization.
Chapter 6

Gaussian Process-based Markov Chain Monte Carlo

In previous chapters we described several deterministic approaches to estimate the high-dimensional (HD) patient-specific tissue properties in the form of model parameters. However, significant uncertainty can be associated with the estimated values of the model parameters which, if left unquantified, will lead to unknown variability in the output of the physiological model that will hinder their reliable clinical adoption. Probabilistic estimation of model parameters, however, remains an unresolved challenge.

In this chapter, we present a novel approach on probabilistic estimation of model parameters by integrating the surrogate modeling of the posterior pdf of model parameters into accelerating the Metropolis-Hastings (MH) sampling of the exact posterior pdf. It is achieved by two main components: 1) construction of a Gaussian process (GP) surrogate of the exact posterior pdf by actively selecting training points that allow for a good global approximation accuracy with a focus on the regions of high posterior probability density; and 2) use of the GP surrogate to improve the proposal distribution in MH sampling, in order to improve the acceptance rate.
6.1 Introduction

The application of patient-specific modeling in clinical practice faces a critical barrier with respect to the variability in the simulation output. This output variability arises from different sources of uncertainty inside the physiological model when built from data. Primary sources of uncertainty include the anatomy of the model (e.g., shape of the heart), tissue properties (e.g., excitability and contractility of the heart muscle), and boundary conditions. Given the continued progress in high-resolution 3D imaging techniques, highly accurate patient-specific models of the heart are now possible [4, 68]. By contrast, the personalization of tissue properties faces several critical challenges that can contribute to the uncertainty in the obtained patient-specific values. First, cardiac tissue properties typically cannot be directly measured; they must be estimated from sparse, noisy, and indirect data. This results in correlation and non-identifiability of tissue properties in different regions of the heart. Second, it is impossible to estimate the tissue properties at the resolution level of the discrete cardiac mesh and a representation of the parameter space at a reduced dimension is necessary. The choice of different low-dimensional (LD) representations will contribute to different uncertainties in the resulting tissue properties of the patient-specific model. Hence, to rigorously understand and quantify the variability and reliability of the predictions made by a patient-specific model, it is important to properly quantify the uncertainty associated with the model parameters that represent the estimated tissue properties for each specific patient.

In the past few decades, significant progress has been made in deterministic approaches to parameter estimation. However, because these methods focus on obtaining a single value of the model parameters that best fits the available data (under given optimization criteria), they do not provide an uncertainty measure associated with the estimated patient-specific parameter values. By contrast, limited progress has been made in probabilistic approaches to estimating the patient-specific parameters of a cardiac model, where the uncertainty in these parameters can be described by their posterior probability density function (pdf) given the measurement data. This posterior pdf of model parameters, in theory, could be estimated using standard Markov Chain Monte Carlo (MCMC) methods that involve repeated non-intrusive evaluations of the posterior pdf. Unfortunately, in this case, the posterior
pdf consists of an analytically-intractable simulation model, each evaluation of which evokes a computationally expensive simulation that could take hours or even days to complete. As a result, it is infeasible to use standard MCMC methods to obtain a posterior pdf of the parameters for expensive simulation models [25, 33, 44].

In this chapter, we present a novel Gaussian process (GP)-accelerated Metropolis-Hastings (MH) sampling framework to overcome the current challenges associated with the probabilistic estimation of patient-specific model parameters. In this framework, a GP surrogate is built to approximate the posterior pdf of the model parameters, which is then used to accelerate the MH sampling of the exact posterior pdf. The key contributions of this work include the following:

(i) We present a strategy of active GP construction that, rather than randomly exploring the parameter space, actively selects training points to approximate the posterior pdf with higher accuracy in regions of high posterior density.

(ii) We present a mechanism that utilizes the efficient GP surrogate to modify and improve the proposal distribution of the MH sampling. Specifically, the GP surrogate is utilized to initially test the acceptance for each proposed candidate, and only those that are initially accepted will be evaluated by the exact posterior pdf for final acceptance, eliminating the need to evoke the expensive simulation model at highly unlikely candidates. This improves the acceptance rate of MH sampling without compromising its accuracy.

(iii) We apply the presented framework to the probabilistic estimation of local tissue excitability in a 3D cardiac electrophysiological (EP) model. Using input data from simulated 120-lead electrocardiographic (ECG) data and validation on synthetic infarct settings, we validate the accuracy and establish its computational cost against direct MH sampling of the exact posterior pdf. Further, we compare its performance with that of directly sampling the surrogate posterior pdf as done in existing works [67]. Our approach is also noteworthy in that limited work has been reported on estimating tissue properties using non-invasive ECG data.

(iv) We further evaluate the presented method in estimating tissue excitabil-
ity in a variety of experimental settings with different input data and validation data. This includes using: 1) input data from a subset of epicardial action potentials generated from an EP model blinded to the presented estimation framework, with validation data of myocardial scar from \textit{in-vivo} magnetic resonance images (MRI); and 2) input data from \textit{in-vivo} 120-lead ECG data from post-infarction patients, with validation data from \textit{in-vivo} voltage mapping.

(v) We evaluate the presented estimation framework on two different LD representations of the parameter space. We analyze the estimated posterior pdf and demonstrate how the uncertainty of the obtained solution is associated with the underlying dimensionality reduction method of choice. This highlights the importance of quantifying the uncertainty in estimated parameter values in patient-specific modeling.

(vi) We provide additional analyses of the estimated posterior pdf in relation to other factors that may contribute to the uncertainty of the estimation solution, including tissue heterogeneity, parameter coupling, and model over-parameterization.

### 6.2 Probabilistic Parameter Estimation

A stochastic relationship between measurement data $Y$ and model parameter $\theta$ can be expressed as:

\[
Y = M(\theta) + \epsilon, \quad (6.1)
\]

where $M$ consists of the whole-heart electrophysiological model and the measurement model as described in Section 2.1.2 and 2.1.3. $\epsilon$ is the noise term that accounts for measurement error and modeling error other than that arising from the value of the parameter $\theta$. Using Bayes’ rule, the unnormalized posterior density of the model parameter $\theta$ has the form:

\[
\pi(\theta|Y) \propto \pi(Y|\theta)\pi(\theta). \quad (6.2)
\]

Assuming uncorrelated Gaussian noise $\epsilon \sim \mathcal{N}(0, \sigma^2_{\epsilon}\mathbf{I})$, the likelihood $\pi(Y|\theta)$ can be written as:

\[
\pi(Y|\theta) \propto \exp\left(-\frac{1}{2\sigma^2_{\epsilon}}\|Y - M(\theta)\|^2\right). \quad (6.3)
\]
where $||.||$ is the Frobenius norm. The prior distribution $\pi(\theta)$ quantifies \textit{a priori} knowledge about the parameters. Here, a uniform distribution bounded within $[0, 0.52]$ is used. So, $\pi(\theta)$ is constant on this interval and 0 off it.

A direct MCMC sampling of the posterior pdf in Eq. (6.2) is infeasible because its slow convergence requires a significant number of evaluations of the expensive electrophysiological model (at an order of $10^5$). Below we present an accelerated Metropolis-Hastings method that consists of two major ingredients. The first ingredient involves the rapid construction of a computationally-efficient surrogate of the expensive posterior pdf (6.4) via active GP construction. The second ingredient involves the use of the efficient GP surrogate to modify the proposal distribution in MH sampling in order to improve its acceptance and convergence rate. Fig. 6.1 shows a high-level work-flow of the presented framework. In the following sections, we describe each component in detail.

### 6.2.1 Active Construction of the GP Surrogate

Because of the analytic properties and the ability to provide probabilistic prediction estimates, recently GP has found widespread use in active learn-
that is concerned with gathering the most informative training data in cases in which collecting a large number of input-output pairs \( \{\Theta, G\} \) is prohibitively expensive [69]. In the context of this study, generating training pairs \( \{\Theta, G\} \) for building the GP surrogate requires expensive evaluation of the exact posterior pdf at each input \( \theta \). Therefore, a method for actively selecting the training points from the parameter space is important. Below, we describe the method that actively selects the training points to obtain an approximation of the posterior pdf model that has higher accuracy in the regions of high posterior pdf.

A GP fitted to the log-posterior pdf is typically better than a GP fitted to the posterior pdf because, in general, the former has longer length scales and lower dynamic range than the latter. Therefore, we fit a GP model for the log of the un-normalized posterior pdf obtained by replacing \( \pi(Y|\theta) \) in Eq. (6.2) with Eq. (6.3) as given below:

\[
g(\theta) = -\frac{1}{2\sigma^2}||Y - M(\theta)||^2 + \log(\pi(\theta)).
\]  

We first define a GP prior over the unknown function (6.4). Through the co-variance function of a GP, assumptions about properties of the function being modeled such as its smoothness and periodicity can be specified. Here, we take an anisotropic Matérn 5/2 co-variance function [60] that enforces an assumption of twice differentiable function:

\[
\kappa(\theta_i, \theta_j) = \alpha^2 \exp\left(-\sqrt{5\delta(\theta_i, \theta_j)}\right) \left(1 + \sqrt{5\delta(\theta_i, \theta_j)} + 5/3d^2(\theta_i, \theta_j)\right),
\]  

where \( d^2(\theta_i, \theta_j) = (\theta_i - \theta_j)^\top \Lambda (\theta_i - \theta_j), \Lambda \) is a diagonal matrix with the square of the characteristics length scales along each dimension of \( \theta \) as the diagonal elements, and \( \alpha^2 \) is the co-variance amplitude. Because no prior knowledge about the posterior pdf is available, here for simplicity we take a zero mean function, which is a commonly and effectively used mean function in GP modeling [60]. The active construction of GP consists of an iteration of two major steps: 1) find a point in the sample space that improves the approximation of Eq. (6.4), especially in the regions of high posterior pdf, and 2) update the GP at this point.

1. **Find optimal training points in the parameter space to update the GP:** Here, we assume that the optimal training points are those that will: 1) allow the GP to globally approximate the Eq. (6.4) well, and 2) identify the regions
of high posterior probability. For the former, points are chosen where the predictive uncertainty $\sigma(\theta)$ of current GP is high (to facilitate exploration of uncertain space). For the latter, points are chosen where the predictive mean $\mu(\theta)$ of the current GP is high (to exploit the current knowledge about the space of high posterior probability). This is done by finding the point that maximizes the upper confidence bound of the GP [71]:

$$\theta_{n+1} = \arg \max_\theta \{ \mu(\theta) + \beta^{1/2} \sigma(\theta) \}.$$  \hfill (6.6)

The parameter $\beta = 2 \log(\pi^2 n^2 / 6 \eta)$, $\eta \in (0, 1)$ balances between exploitation and exploration of the parameter space [7, 71]. Eq. (6.6) is optimized using a bound constrained derivative-free optimization method known as Bound Optimization BY Quadratic Approximation (BOBYQA) [58]. The predictive mean and uncertainty in Eq. (6.6) are evaluated using Eq. (2.4).

2. Updating the GP surrogate at selected training points: Once a new training point is obtained, Eq. (6.4) is evaluated at this point and the GP is updated at the newly obtained $\{\theta_{n+1}, g_{n+1}\}$ pair. This GP captures the updated belief over Eq. (6.4) after having observed the new training point.

$$\begin{bmatrix} G \\ g_{n+1} \end{bmatrix} \sim \mathcal{N} \left( 0, \begin{bmatrix} K + \zeta^2 I & k \\ k^T & \kappa(\theta_{n+1}, \theta_{n+1}) \end{bmatrix} \right).$$ \hfill (6.7)

where $k = [\kappa(\theta_{n+1}, \theta_1), \kappa(\theta_{n+1}, \theta_2), \ldots, \kappa(\theta_{n+1}, \theta_n)]^T$, $K$ is the co-variance matrix, and $\zeta = 0.001$ is a small noise term that is added for numerical stability. After every several updates of the GP, we optimize the hyperparameters (length scales $\Lambda$ and covariance amplitude $\alpha$) by maximizing the marginal likelihood $\log p(\mathbf{g}_{1:n+1} | \mathbf{g}_{1:n+1}, \alpha, \Lambda)$. These two steps iterate until the training point selected by optimizing the upper confidence bound (6.6) changes little over a few iterations ($\leq 0.005$, 15 iterations). The hyperparameters length scales $\Lambda$ and covariance amplitude $\alpha$ are once again optimized by maximizing the marginal likelihood $\log p(\mathbf{g}_{1:n+1} | \theta_{1:n+1}, \alpha, \Lambda)$. Fig. 6.2 gives examples of the training points selected during the GP construction, which reveals that these points are spread in the sample space but are more concentrated in the regions of high posterior probability. We take the predictive mean function of the resulting GP as the surrogate of Eq. (6.4). In this way, we can obtain a GP-based surrogate of the
Figure 6.2: Examples of exact bivariate marginal pdfs superimposed with the training points collected during GP construction, which are spread in the parameter space but are most concentrated in regions of high probability density.

exact posterior pdf $\pi(\theta|Y)$, denoted by $\pi^*(\theta|Y)$ for the remainder of this chapter, that is cheap to evaluate and is most accurate in regions of high posterior probability.

### 6.2.2 GP Surrogate Accelerated Metropolis-Hastings

MH is the most widely used MCMC method. It begins from an arbitrary sample $\theta_n$ and generates a Markov chain of samples that come from an invariant distribution. Specifically at each step in the MH algorithm, a candidate sample $\theta_c$ is proposed using a proposal distribution $q(\theta_c|\theta_n)$. This candidate is accepted with a probability given by:

$$
\rho_{mh}(\theta_n, \theta_c) = \min\left(1, \frac{q(\theta_n|\theta_c)\pi(\theta_c|Y)}{q(\theta_c|\theta_n)\pi(\theta_n|Y)}\right).
$$

If accepted, $\theta_{n+1} = \theta_c$. If rejected, $\theta_{n+1} = \theta_n$. This is repeated until the samples converge to the target distribution.

The success of the MH largely relies on the choice of the proposal distribution. If the proposal distribution is much narrower than the target distribution, the MH will spend too much time exploring the sampling space, resulting in bad mixing. Conversely, if the proposal distribution is much wider than the target distribution, the MH will make wide jumps in the sampling space, resulting in a large number of rejections. Ideally, a proposal distribution similar
to the target distribution is desired for a higher acceptance rate with a good mixing. However, to obtain such a proposal distribution is notoriously difficult and a Gaussian distribution is the most commonly used proposal distribution in practice. Meanwhile, some previous works have proposed the modification of this generic proposal distribution using various approximate target distributions [12, 25]. Below, we describe how a GP surrogate of the posterior pdf is utilized to modify the Gaussian proposal distribution in MH to accelerate sampling with good mixing and a higher acceptance rate.

Specifically, we present a two-step test of acceptance. In the first step, a candidate \( c_1 \) proposed by a standard Gaussian proposal distribution \( q(\theta|\theta_{n}) \) is tested for acceptance by the GP surrogate of the posterior pdf \( \pi^*(\theta|Y) \) with acceptance probability given by:

\[
\rho_1(\theta_n, \theta_{c1}) = \min \left( 1, \frac{q(\theta_n|\theta_{c1})\pi^*(\theta_{c1}|Y)}{q(\theta_{c1}|\theta_n)\pi^*(\theta_n|Y)} \right). \tag{6.9}
\]

The candidate for the second test of acceptance is determined by the outcome of the previous step, i.e., \( \theta_c = \theta_{c1} \) if accepted and \( \theta_c = \theta_n \) if rejected. In other words, the candidates for the second step are effectively generated from the transition probability in the first step, which defines the effective proposal distribution for the second step:

\[
q^*(\theta_c|\theta_n) = \rho_1(\theta_n, \theta_c)q(\theta_c|\theta_n) + r(\theta_n)\delta_{\theta_n}(\theta_c), \tag{6.10}
\]

where \( r(\theta_n) = 1 - \int \rho_1(\theta_n, \theta_c)q(\theta_c|\theta_n)d\theta_c \) is the probability that the chain remains at \( \theta_n \) and \( \delta_{\theta_n}(.) \) denotes the Dirac mass at \( \theta_n \). Using this modified proposal distribution, the proposed candidate sample is accepted by the exact posterior pdf with a probability given by:

\[
\rho_2(\theta_n, \theta_c) = \min \left( 1, \frac{q^*(\theta_n|\theta_c)\pi(\theta_c|Y)}{q^*(\theta_c|\theta_n)\pi(\theta_n|Y)} \right). \tag{6.11}
\]

Depending on whether the candidate was accepted or rejected in the first step, the acceptance rate in Eq. (6.11) can be calculated. When a candidate is accepted in the first step, i.e., \( \theta_c = \theta_{c1} \), we obtain \( q^*(\theta_c|\theta_n) = \)
$\rho_1(\theta_n, \theta_c)q(\theta_c|\theta_n)$, which can be further simplified using Eq. (6.9) as follows:

$$q^*(\theta_c|\theta_n) = \min \left( q(\theta_c|\theta_n)\pi^*(\theta_n|Y), q(\theta_n|\theta_c)\pi^*(\theta_c|Y) \right) \frac{1}{\pi^*(\theta_n|Y)}$$

$$= \min \left( \frac{q(\theta_c|\theta_n)\pi^*(\theta_n|Y)}{q(\theta_n|\theta_c)\pi^*(\theta_c|Y)}, 1 \right) q(\theta_n|\theta_c)\pi^*(\theta_c|Y) \frac{\pi^*(\theta_n|Y)}{\pi^*(\theta_c|Y)}$$

$$= \rho_1(\theta_c, \theta_n)q(\theta_n|\theta_c)\pi^*(\theta_c|Y) \frac{\pi^*(\theta_n|Y)}{\pi^*(\theta_n|Y)} \frac{\pi^*(\theta_n|Y)}{\pi^*(\theta_n|Y)}.$$  \hspace{1cm} (6.12)

Substituting Eq. (6.12) into Eq. (6.11), the acceptance probability in the second step can be simplified to:

$$\rho_2(\theta_n, \theta_c) = \min \left( 1, \frac{\pi(\theta_c|Y)\pi^*(\theta_n|Y)}{\pi(\theta_n|Y)\pi^*(\theta_c|Y)} \right).$$  \hspace{1cm} (6.13)

When a candidate is rejected in the first step, then $\theta_c = \theta_n$ and $\rho_2(\theta_n, \theta_c) = 1$. In other words, samples that are rejected in the first step do not need to be evaluated by the exact posterior pdf. This improves the proposal distribution for the MH method and reduces the need for evaluating the expensive posterior pdf at candidates that are highly unlikely to be accepted.

**Convergence of the GP-accelerated MH Sampling:** The convergence of the MH with modified proposal distribution follows the same line as that of the standard MH [3, 12, 25, 33]. Given any initial sample, the Markov chain generated by the MH converges to an invariant distribution if the transition probability meets the following properties: 1) irreducibility, and 2) aperiodicity. A sufficient, but not necessary, condition to ensure convergence to an invariant distribution is reversibility (detailed balance). The presented GP-accelerated MH satisfies these criteria as follows:

1) **Reversibility:** Similar to the standard MH, through the inclusion of the acceptance rate, the transition probability is designed to meet the criteria of detailed balance by construction. Specifically, the detailed balance condition is given by:

$$\pi(\theta_n|Y)T(\theta_c|\theta_n) = \pi(\theta_c|Y)T(\theta_n|\theta_c).$$  \hspace{1cm} (6.14)
where $T$ denotes the transition probability of the presented GP-accelerated MH method that is defined by:

$$T(\theta_c|\theta_n) = \rho_2(\theta_n, \theta_c)q^*(\theta_c|\theta_n) + \left(1 - \int \rho_2(\theta_n, \theta_c)q^*(\theta_c|\theta_n)d\theta_c\right)\delta_{\theta_n}(\theta_c).$$

(6.15)

Proof. When $\theta_c = \theta_n$, Eq. (6.14) is automatically satisfied. When $\theta_c \neq \theta_n$, Eq. (6.14) can be simplified using Eq. (6.15) and Eq. (6.11) as:

$$\pi(\theta_n|Y)T(\theta_c|\theta_n) = \pi(\theta_n|Y)\rho_2(\theta_n, \theta_c)q^*(\theta_c|\theta_n)$$

$$= \min\left(q^*(\theta_c|\theta_n)\pi(\theta_n|Y), q^*(\theta_n|\theta_c)\pi(\theta_c|Y)\right)$$

$$= \pi(\theta_c|Y)\min\left(1, \frac{q^*(\theta_c|\theta_n)\pi(\theta_n|Y)}{q^*(\theta_n|\theta_c)\pi(\theta_c|Y)}\right)q^*(\theta_n|\theta_c)$$

(6.16)

2) Aperiodicity: Because the acceptance criteria in the presented method always allow for rejection of the samples as in the standard MH, the presented GP-accelerated MH is also aperiodic.

3) Irreducibility: If $\pi(\theta_c|Y) > 0$, $\forall \theta_c \in \Omega$ implies $\pi^*(\theta_c|Y) > 0$, $\forall \theta_c \in \Omega$, where $\Omega$ is the support of the exact posterior pdf $\pi(\theta_c|Y)$, then the Markov chain generated by the presented method is $\pi$-irreducible.

Proof. To ensure the condition of irreducibility in the standard MH, the proposal distribution is chosen to satisfy $q(\theta_c|\theta_n) > 0$, $\forall \theta_c, \theta_n \in \Omega$. As a result, the transition probability in the standard MH, $T_{mh}(\Omega|\theta_n) > 0$, $\forall \theta_n \in \Omega$. Using the condition that the GP surrogate of the posterior distribution $\pi^*(\theta_c|Y) > 0$, $\forall \theta_c \in \Omega$, we can obtain similar results for the transition probability in the first step (alternatively the proposal distribution for the second step) of the presented method, i.e., $q^*(\theta_c|\theta_n) > 0$, $\forall \theta_c, \theta_n \in \Omega$. Without the loss of generality, assuming that $\theta_c \neq \theta_n$, we obtain $q^*(\theta_c|\theta_n) = \rho_1(\theta_n, \theta_c)q(\theta_c|\theta_n) > 0$, $\forall \theta_c, \theta_n \in \Omega$ from Eq. (6.10). This implies that $\rho_2(\theta_n, \theta_c)q^*(\theta_c|\theta_n) > 0$, $\forall \theta_c, \theta_n \in \Omega$. Therefore, the effective transition probability Eq. (6.15) of the presented method $T(\theta_c|\theta_n) > 0$, $\forall \theta_c, \theta_n \in \Omega$. Hence, under the given condition that $\pi^*(\theta_c|Y) > 0$, $\forall \theta_c \in \Omega$, the presented
method is π-irreducible. Here, the GP surrogate of the posterior distribution
\( \pi^*(\theta_c | Y) \) is an exponential function. Therefore, \( \pi^*(\theta_c | Y) > 0, \quad \forall \theta_c \in \mathbb{R}^d \).

In practice, while the theoretical convergence is guaranteed, an inaccurate surrogate model could lead to a biased sampling with a low acceptance rate in the presented method.

### 6.3 Evaluation on Synthetic Data

In experiments with synthetic data, we first evaluate the accuracy and efficiency of the presented method (GP-accelerated MH) against: 1) the baseline of directly sampling the exact posterior pdf using the standard MH (direct MH), and 2) the previously reported approach of directly sampling the surrogate posterior pdf using the standard MH (MH on GP) [67]. We then analyze and interpret the obtained posterior pdfs in relation to different factors contributing to parameter uncertainty, primarily under the setting of parameter estimation using two different LD representations of the spatial parameter space as described in Chapter 3.

In total, we consider 14 synthetic cases with seven different settings of infarcts, each of which is estimated on two different LD representations of the spatial parameter space. To represent healthy and infarcted tissues in each case, the parameter \( a \) of the AP model (2.1) is set to 0.15 and 0.50, respectively. Measurement data for parameter estimation is generated in two steps. First, action potentials on the cardiac mesh are simulated using the AP model (2.1). Then, 120-lead ECG are generated using the forward model (2.2) and corrupted with 20 dB Gaussian noise.

All MCMC sampling runs on four parallel MCMC chains of length 20,000 with a common Gaussian proposal distribution and four different initial points. The variance of the Gaussian proposal distribution is tuned by rapidly sampling the GP surrogate pdf until obtaining an acceptance rate of \( \sim 0.22 \), which is documented to enable good mixing and faster convergence in higher dimensional problems [3, 33]. The four initial points are obtained by conducting a rapid sampling of the GP surrogate pdf, constructing four clusters of the samples using a Gaussian mixture model, and using the mean of each as the starting points for each chain. After discarding initial burn-in samples and selecting alternate samples to avoid auto-correlation in each chain, the samples
from four chains are combined. The convergence of all the MCMC chains is tested using trace plots, Geweke statistics, and Gelman-Rubin statistics [3,33].

To differentiate the infarcted and healthy regions from the estimated tissue properties, we calculate a threshold value that minimizes the intra-region variance on the estimated parameter values [55].

### 6.3.1 Validation of the Accuracy and Efficiency

We first validate the accuracy of the presented method against directly sampling the exact posterior pdf using the standard MH method. Fig. 6.3 presents four examples of posterior pdfs obtained from different synthetic data cases. As shown, the presented sampling strategy (green curve) closely reproduces the true posterior pdf (red curve) obtained from direct MH.

Next, we compare the computational cost of the presented method with that of the direct MH in terms of the number of model evaluations needed and actual computation times. The comparison is based on 14 synthetic cases run on a computer with a Xeon E5 2.20 GHz processor and 128 GB RAM. The presented method reduces the number of model evaluations by an average of 64.47% despite the overhead of constructing the GP surrogate which, as highlighted in the purple bar in Fig. 6.4 left, is very small compared with the number of model evaluations required for sampling. The computation time is reduced from 41.073±2.028 hours with direct MH to 7.961±2.028 hours with the presented method.

The efficiency of the sampling method is also measured in terms of its acceptance rate. Here, acceptance rate refers to the fraction of the accepted candidates out of all those proposed to the exact posterior pdf. As shown in the right panel of Fig. 6.4, the presented method improves the acceptance rate from 0.2653±0.0500 of the direct MH to 0.3988±0.0788. This means that the presented method is able to improve the proposal distribution by filtering out a large portion of candidate samples that would eventually be rejected by the exact posterior pdf, thereby avoiding evoking expensive simulation models on these candidates.

### 6.3.2 Comparison with Directly Sampling the GP Surrogate

Directly sampling the GP surrogate pdf instead of the exact pdf, as commonly done in existing methods [67], requires significantly less computation because
model simulation is not needed. However, the sampling accuracy also becomes critically reliant on the accuracy of the surrogate model. As illustrated in the examples in Figs. 6.3(c) and 6.3(d), sampling the GP surrogate (blue curve) produces a distribution that is different from the exact pdf not only in general shape but also in locations of the mode. In comparison, while the accuracy in the GP surrogate affects the efficiency of the presented method, the estimated pdf converges to the exact pdf as obtained by the direct MH. Using the mean, mode, and standard deviation of the exact pdf as the baseline, Table 6.1 shows that sampling errors of the presented method are significantly lower than those
Table 6.1: Mean absolute errors in the estimated mean, mode, and standard deviation against directly sampling the exact posterior pdf: the presented method (GP-accelerated MH) vs. sampling the GP surrogate pdf only (MH on GP).

<table>
<thead>
<tr>
<th>Method</th>
<th>Mean</th>
<th>Mode</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>GP-accelerated MH</td>
<td>0.012</td>
<td>0.030</td>
<td>0.006</td>
</tr>
<tr>
<td>MH on GP</td>
<td>0.039</td>
<td>0.058</td>
<td>0.015</td>
</tr>
</tbody>
</table>

from sampling the surrogate only (paired $t$-test on 139 estimated parameters, $p < 0.001$).

6.3.3 Analysis of the Sampled Posterior Distributions

Several factors can contribute to the uncertainty in the estimated model parameters, including but not limited to the sparse measurement data, parameter correlation, model over-parameterization, and limited spatial resolution in comparison with the underlying tissue heterogeneity (which we will refer to as “model under-parameterization” for the remainder of the chapter). Some of these factors vary with the method used to represent the parameter space. For example, different LD representations of the parameter space may result in different correlations between each dimension of the parameter, as well as different over- and/or under-parameterization of the model. Therefore, in this section, we consider the presented approach on two types of LD representations of the parameter space: a uniform division of the cardiac mesh into 10 segments using the AHA standard, and a non-uniform division of the cardiac mesh into 7-15 clusters using a method that aims to adaptively group homogeneous nodes of the cardiac mesh together [22]. Below, we analyze and interpret the estimated posterior pdfs in relation to the aforementioned contributing factors to uncertainty. We elaborate on three examples with a varying degree of “parameter heterogeneity”, a term we use to denote a state in which a dimension of the parameter representation is too low in resolution to reflect the underlying heterogeneity.

In the first example shown in Fig. 6.5, the infarct in the septum region is better represented by the uniform method than the adaptive method with respect to parameter heterogeneity. Specifically, the uniform method generates three heterogeneous segments that are 37.32% of the heart volume; in com-
Figure 6.5: Examples of parameter estimation results when using clusters from the adaptive method vs. segments from the uniform method as LD representation of the spatial parameter space. Case 1, case 2, and case 3 respectively show an example in which the infarct is represented better by the segments, equally by both segments and clusters, and better by the clusters in terms of parameter heterogeneity.

parison, the adaptive method generates four heterogeneous clusters that are 53.09% of the heart volume. In this case, the mode estimated using the uniform method accurately captures the infarct region (Dice coefficient: 0.5896), whereas the adaptive method captures the infarct with false positives at the right ventricle (Dice coefficient: 0.3447). Correspondingly, while the uniform method shows high confidence in the solution obtained in the region of true infarct, the adaptive method shows high uncertainty in the region of true infarct. In addition, it is interesting to note that both LD representations result in high parameter uncertainty at the right ventricle region and the anterior region adjacent to the region of infarct. This uncertainty, independent of the LD representation of choice, may indicate difficulty in estimating parameters in this region of the heart due to non-identifiability and limited measurement data.

In the second example that has an infarct in the anterior region as shown in Fig. 6.5, both the uniform and adaptive methods show similar parameter heterogeneity in representing the infarct. Specifically, with the uniform method, the infarct lies completely within a segment such that only one segment that
is 10.42% of the heart volume contains heterogeneous tissue; in comparison, the adaptive method generates three heterogeneous clusters that is 10.16% of the heart volume. As shown, using the uniform method, the estimated mode reveals higher parameter value in the segment that contains the infarct (Dice coefficient: 0.1622) with high confidence. However, the standard deviation plot shows high uncertainty in the estimated parameters throughout the cardiac mesh, possibly to compensate for the fact that the true infarct is much smaller than the segment representing the infarct. In comparison, using the adaptive method, the estimated mode has high parameter value in a narrower region representing the compact infarct with higher accuracy (Dice coefficient: 0.5763). Likewise, high parameter uncertainty is obtained in a more concentrated region that overlaps the heterogeneous clusters. Additionally, similar to the first example, both LD representations show high uncertainty in the region that is adjacent on the left to the region of true infarct. This indicates difficulty in accurately estimating parameters in those regions, possibly again due to non-identifiability and limited data.

Finally, Fig. 6.5 case 3 shows an example in which an infarct is better represented by the adaptive method than the uniform method with respect to parameter heterogeneity. Specifically, uniform method generates five segments with heterogeneous tissue, totaling 45.01% of the heart volume; in comparison, the adaptive method also generates five clusters with heterogeneous tissue, but totaling only 30.59% of the heart volume. In this case, the mode obtained using the adaptive method captured the region of infarct with higher accuracy (Dice coefficient: 0.6220) than the mode obtained using the uniform method (Dice coefficient: 0.3269). A closer look at the standard deviation plots reveals that, in general, high uncertainty is obtained in the regions of heterogeneous tissue when using either method. Overall, because a higher proportion of the cardiac mesh is associated with heterogeneous representations when using the uniform method (model under-parameterization), an overall larger region of high parameter uncertainty is obtained with the uniform method compared with the adaptive method.

While the issue of model under-parameterization is more evident with the uniform method as explained above, the issue of model over-parameterization is more evident with the adaptive method. This is because the adaptive method as described in [17] could result in a large number of small clusters at heterogeneous regions around the infarct. Consequently, parameter values at
Figure 6.6: Example of high uncertainty associated with over-parameterization when the adaptive method assigned several small clusters to represent the heterogeneous regions around the infarct (c-e). Top row: clusters represented by one dimension of the parameter space. Middle row: univariate marginal density plot of the estimated parameter. Bottom row: estimated mean, mode and standard deviation of the parameter.

these regions are associated with a high uncertainty as seen in the standard deviation plots obtained by using the adaptive method in Fig. 6.5, likely due to the issue of non-identifiability given limited and indirect measurement data. An example case, in which this issue is more pronounced, is shown in Fig. 6.6. Here, several small clusters (c-e) are formed by the adaptive method at the region of the inferior-lateral infarct and its border. For parameter values at these clusters, the MH sampling has difficulty converging to a distribution with distinct modes, which is reflected as high uncertainty in the resulting parameter values. These examples show that probabilistic parameter estimation can help reveal the issue of identifiability, which cannot be observed when a single point estimate is being sought.

Finally, different LD representations of the parameter space can also result in different correlations among the parameters in each dimension. In Fig. 6.7, we show an example with an infarct localized in the septal region, in which the parameter of one of the regions shows both a positive correlation and a negative correlation with parameters in other regions. As shown at the top
Figure 6.7: (a) Regions of the heart obtained by the adaptive method (red: infarct, green: non-infarct/mixed). (b) Estimated univariate and bivariate marginal pdf plots. (c) Trace plot for parameters of regions 6 and 7 (top) shows a negative correlation with a switching behavior, whereas that of regions 5 and 7 (bottom) shows a positive correlation.

of Fig. 6.7(c), the parameter in region 7 and region 6 exhibits a negative correlation with a switching behavior (i.e., when the parameter in region 7 is estimated in a healthy range, the parameter in region 6 is estimated in an unhealthy range). At the same time, the parameter in region 7 and region 5 exhibit a positive correlation (i.e., when the parameter in region 7 is estimated in a healthy range, the parameter in region 6 also is estimated in a healthy range; Fig. 6.7(c), bottom). This is reflected as higher parameter uncertainty in all three regions.
6.4 Evaluation on a Blinded EP Model and *In-vivo* MRI Scar

In this section, we study the presented method in quantifying the uncertainty in model parameters for post-infarction human hearts, where validation data for the 3D myocardial infarct is available from *in-vivo* magnetic resonance imaging. Compared with the infarct settings in synthetic data experiments, these MRI-derived 3D infarcts have the following characteristics that increase the heterogeneity in tissue properties: 1) the presence of both dense scar core and gray zone, 2) the presence of a single or multiple scars with complex spatial distribution and irregular boundaries, and 3) the presence of both transmural and non-transmural scars. The resolution to which such heterogeneity can be captured is largely limited by the method of dimensionality reduction. Because previous work has shown that an adaptive non-uniform LD representation may be able to better represent tissue heterogeneity [17], the experiments below are conducted using only the adaptive LD representation of the parameter space. Because *in-vivo* electrical mapping data were unavailable, here measurement data for probabilistic parameter estimation are generated by a high-resolution (average resolution: 350 µm) multi-scale (sub-cellular to organ scale) *in-silico* ionic electrophysiological model on the MRI-derived patient-specific ventricular models as detailed in [4]. Data used for parameter estimation are extracted from 300-400 epicardial sites, temporarily down-sampled to a 5 ms resolution, and corrupted with 20 dB Gaussian noise. Note that although no *in-vivo* electrical data were available, the experiments are designed to mimic a real-data scenario because: 1) the 3D EP model used to generate the measurement data is known to be capable of generating high-fidelity EP simulation of patient-specific hearts [4], 2) this model is unknown to the framework of GP-accelerated MH and thus no “inverse crime” is involved, and 3) only a small subset of epicardial data corrupted with noise is used as measurement data.

For clarity, below we analyze the performance of the presented method with respect to two contributing factors to the heterogeneity of the scar: 1) scar transmurality, and 2) gray zone. For ground truth, these regions were determined from the 3D infarcts mapped to the high resolution cardiac mesh from MRI [4].

1) *Scar transmurality*: In examples case 1, case 2, and case 3 shown in
Fig. 6.8: Mean, mode, and standard deviation of posterior pdfs estimated from epicardial potentials simulated by a multi-scale ionic EP model blinded to the presented estimation method. Purple circles denote areas of non-transmural scars (cases 1, 2, and 3) or gray zones (cases 4 and 5).

Fig. 6.8, a portion of the scar is non-transmural. From left to right, the non-transmural portion of the scar lies respectively on the lateral wall, the anterior-septal region, and the anterior-lateral wall of the left ventricle as denoted by a purple circle. In all three cases, the estimated mode misses these regions of non-transmural scar; the estimated mean exhibits higher parameter values deviating from that for healthy tissue, yet the value is not as high as that for scar tissue. In all cases, parameter values in these regions are associated with high uncertainty. This provides a useful confidence measure for the estimated mode, suggesting that these regions do not consist entirely of healthy tissue as reflected by the mode.

2) **Gray zone:** Case 4 and case 5 in Fig. 6.8 show examples in which a transmural dense scar is surrounded by gray zone. In both cases, the mean and mode estimates obtain high parameter values in the region of dense scar and gray zone. The parameter values in the regions of dense scar are higher than those in gray zones. In addition, the mean estimates reveal the gray zone to be wider than the estimates from MRI data, whereas the mode estimates do not. Parameters for these border regions are associated with high uncertainty.
as shown in the standard deviation plots, reflecting the underlying tissue heterogeneity in these regions and the possible model under-parametrization as a result of the LD representation of the parameter space.

### 6.5 Evaluation on 120-lead ECG and Catheter Data

We conduct real-data studies on three patients who underwent catheter ablation of ventricular tachycardia due to prior myocardial infarction [66]. The patient-specific heart-torso geometrical models are constructed from axial computerized tomography images. The uncertainty of tissue excitability in the AP model (2.1) is estimated from 120-lead ECG. Similarly, all experiments are conducted using the adaptive LD representation of the parameter space [17]. For validation of the results, we consider the relation between the estimated tissue excitability and the in-vivo epicardial bipolar voltage data obtained from catheter mapping. It should, however, be noted that voltage data can be used only as a reference, not as the gold standard for the measure of tissue excitability. Below we focus our analysis on how the obtained parameter uncertainty is associated with the heterogeneity of the underlying tissue.

**Case 1:** The voltage data for case 1 (Fig. 6.9(a)) shows a dense infarct at inferolateral left ventricle (LV) with a heterogeneous region extending to
CHAPTER 6. GAUSSIAN PROCESS-BASED MCMC

Figure 6.10: Real-data experiments for case 1 from Fig. 6.9: regions of the heart to be parameterized and the corresponding marginal probability density plots.

Lateral LV. Adaptive dimensionality reduction as described in [17] generates eight regions of the heart to be parameterized, five of which are listed in Fig. 6.10 along with the estimated posterior marginal pdfs for their parameters. As shown, the parameter for the region of infarct core (a) is correctly estimated with high confidence. The parameter for the region of immediate border to the infarct (e) has a mean/mode value that is in between the healthy and infarct core, correctly indicating a border zone, whereas other regions in the infarct border are estimated as either healthy (c-d) or infarcted (b). For all these regions around the heterogeneous infarct border (b-e), uncertainties of the estimation are higher. This produces an estimation with correct posterior mode/mean with high confidence at the infarct core, and high uncertainty at the heterogeneous infarct border.

Case 2: The voltage data for case 2 (Fig. 6.9(b)) shows a massive yet quite heterogeneous infarct at lateral LV. The adaptive dimensionality reduction method generates 12 regions of the heart to be parameterized, five of which are listed in Fig. 6.11 along with the estimated marginal pdfs for their parameters. As shown, for remote healthy regions (a-b), their parameters are correctly estimated with high confidence. For heterogeneous border regions close to the infarct (c-d), their parameters are estimated in the healthy range but with lower confidence. For the region that corresponds to the infarct (e), its abnormal parameter is correctly captured but with a high uncertainty – likely reflecting the heterogeneous nature of tissue properties in this region. As summarized in Fig. 6.9(b), while the estimation correctly reveals the region
of infarct as in case 1, it is also associated with a higher uncertainty compared with the less heterogeneous infarct in case 1.

**Case 3:** The voltage data for case 3 (Fig. 6.9(c)) shows low voltage at lateral LV and RV, although it was not certain whether the low voltage on lateral RV was due to the presence of an infarct or fat layer. After dimensionality reduction with the adaptive method, there are seven regions of the heart that remain to be parameterized (Fig. 6.12(a)). The infarct region in lateral LV (region 1) is estimated with a distribution that has medium uncertainty and a mode of 0.257. This could indicate the presence of infarcted tissue along with some healthy tissue (heterogeneity). In contrast, the marginal distribution for the healthy apical region (region 2) is estimated with a very narrow uni-modal distribution with a mode of 0.142. Interestingly, several regions in the lateral RV (region 4, 5, and 7) show very high uncertainty with a distribution of the parameter value extending from healthy to infarct range. Overall, results show an estimate of healthy tissue in the apical region with high confidence, an estimate of heterogeneous tissue with infarct at lateral LV with medium confidence (Fig. 6.9(c)), and an estimate of the ambiguous region in lateral RV as healthy with high uncertainty.
6.6 Discussion

6.6.1 Effect of the Gaussian Proposal Distribution

The effect of the Gaussian proposal distribution on the presented GP-accelerated MH is similar to that on the standard MH. A narrow proposal distribution takes small steps in the sampling space, resulting in a high acceptance rate but slower mixing. By contrast, a wide proposal distribution takes wide steps in the sampling space, resulting in a low acceptance rate but fast mixing. As shown in Fig. 6.13(a), as the width of the Gaussian proposal distribution increases, the acceptance rate of the presented framework decreases.

When using the standard MH for higher dimensional posterior pdfs, a proposal distribution that results in an acceptance rate of $\sim 0.22$ is documented to enable good mixing and faster convergence [3,33]. However, tuning the pro-
CHAPTER 6. GAUSSIAN PROCESS-BASED MCMC

Figure 6.13: Acceptance rate of the GP-accelerated MH method: (a) decreases as the width (standard deviation) of the Gaussian proposal increases, (b) decreases as the Kullback-Leibler divergence from the GP surrogate to the exact posterior pdf increases. Data is taken from the sampled distributions in synthetic data experiments.

Proposal distribution to obtain a good acceptance rate by repeated trial-and-error sampling of the exact posterior pdf is impossible here due to high computational cost. In this study, assuming that the GP surrogate pdf captures the shape of the exact posterior pdf, we leverage the efficiency of the surrogate pdf and tune the proposal distribution by rapidly sampling the GP surrogate pdf. Similarly, as described earlier, the starting point for each of the parallel MC chains is also initialized from fast sampling of the GP surrogate pdf. In other words, the GP surrogate not only accelerates the MH through the two-stage sampling scheme, but also allows for fast tuning of the proposal distribution and selection of initial samples. For a fair evaluation of the efficacy of the presented GP-accelerated MH method, the tuned proposal distribution and initial samples were used with both the direct MH and GP-accelerated MH in the presented experiments.

6.6.2 Quality of the Surrogate vs. Acceptance Rate

To understand how the acceptance rate of the presented GP-accelerated MH method is related to the accuracy of the GP surrogate, we measure the quality of the GP surrogates built in the synthetic data experiments by their Kullback-Leibler (KL) divergence to the exact posterior pdfs. The KL-divergence from
the GP surrogate pdf $\pi^*(\mathbf{\theta}|Y)$ to the exact posterior pdf $\pi(\mathbf{\theta}|Y)$ is defined by:

$$D_{KL}(\pi(\mathbf{\theta}|Y)||\pi^*(\mathbf{\theta}|Y)) = \int_{-\infty}^{\infty} \pi(\mathbf{\theta}|Y) \log \left( \frac{\pi(\mathbf{\theta}|Y)}{\pi^*(\mathbf{\theta}|Y)} \right) d\mathbf{\theta}. \quad (6.17)$$

Because above KL divergence cannot be obtained analytically, we estimate it using Monte Carlo simulation [35] as follows:

$$D_{KL}(\pi(\mathbf{\theta}|Y)||\pi^*(\mathbf{\theta}|Y)) = \frac{1}{N} \sum_{i=1}^{N} \log \left( \frac{\pi(\mathbf{\theta}_i|Y)}{\pi^*(\mathbf{\theta}_i|Y)} \right), \quad (6.18)$$

where $\mathbf{\theta}_i$ are N random samples of $\pi(\mathbf{\theta}|Y)$ obtained from direct MH method. As shown in Fig. 6.13(b), in most cases, the KL divergence of the surrogate pdf is low, indicating a high accuracy of the surrogate pdf. However, some surrogate pdfs had a high KL divergence from the exact pdf, indicating limited accuracy which is most likely related to the increased complexity in the shape of the unnormalized log exact posterior pdf. As expected, a negative correlation between the quality of the GP surrogate and the acceptance rate of the GP-accelerated MH method (correlation coefficient = -0.777) can be observed in Fig. 6.13(b). In other words, a more accurate GP surrogate will result in a higher efficiency of the presented sampling method, whereas a less accurate GP surrogate would be less effective in accelerating the MH sampling. However, a more accurate GP surrogate would also be more expensive to construct, whereas a less accurate one would be faster to construct. How to balance between these two steps, as well as how to construct an accurate surrogate without evoking a large number of model evaluations, are to be investigated in future works.

### 6.6.3 Related Works

Related works on uncertainty quantification in personalized models can be broadly categorized into two types: 1) forward uncertainty quantification, and 2) inverse uncertainty quantification. Forward uncertainty quantification focuses on the uncertainty in model output as a result of variations in different model parameters. To overcome the challenge of repeatedly evaluating the expensive simulation model, methods such as generalized polynomial chaos or stochastic collocation are commonly used. In the domain of electrocardiography, this includes examples such as the study of sensitivity of model output to
conductivity parameters [31], the study of sensitivity of measured ECG signals on heart motion [73], and the study of sensitivity of ECG signal components as a result of variations in sub-endocardial ischemia [37]. These works are fundamentally different from the presented work that focuses on the inverse uncertainty quantification.

The inverse uncertainty quantification focuses on the uncertainty within the model (such as the estimated model parameters) as a result of different uncertain factors involved in the personalization of the model. Among existing works, the approach presented in [46], although applied to a brain tumor growth model, is most common in its spirit with the presented framework. Specifically, the work presented in [46] also utilizes a two-stage sampling method (GPHMC) in which first a GP surrogate of the unnormalized negative log posterior pdf is learned with HMC and then its gradient is utilized in HMC for efficient sampling. To compare the presented method and the GPHMC method, we conducted experiments on six synthetic cases. Implementation of the GPHMC method as detailed in [46] is utilized. We take 50 initialization points and 3000 exploratory points during the construction of the GP surrogate. The comparison is presented with respect to the two major elements of these methods: 1) GP surrogate construction, and 2) posterior pdf sampling. For the constructed GP surrogates, their mean KL divergences to the exact posterior pdf are respectively 6.58 and 38.57 for the presented method and the GPHMC method. This gain in accuracy by the presented method could be because of the utilization of a Matérn 5/2 kernel in GP, and an active scheme with a derivative-free deterministic optimization to select training points. An increase in the number of training points in the GPHMC method may increase the accuracy of the GP surrogate. For the posterior pdf sampling, surprisingly we observed an acceptance rate of $0.1$ with GPHMC although the accuracy of the GP surrogate was comparable between the two methods. We speculate that given the non-smooth and complex shape of the negative log posterior pdf and its first derivative in this study, a GP – especially one with a squared exponential kernel that assumes an infinitely differentiable prior over the negative log posterior pdf – could not accurately approximate its local derivatives. This inaccuracy in the approximated local derivatives may then lead to poor candidate samples proposed by the HMC, resulting in low acceptance rate. In contrast, because the presented method only depends on the approximate global shape of the log posterior pdf without
utilizing its derivative information in the sampling, a smoother approximation such as a GP with Matérn 5/2 kernel could increase the acceptance rate.

The selection of points for the construction of a GP surrogate shares common intuition with active learning [40], Bayesian optimization [7], and multi-armed bandits problems [71] in which based on a history of actions and rewards a decision needs to be made on the next best point to query from the solution space. More recently, there has been an interest in utilizing these methods to approximate intractable pdfs [39]. In contrast to these works that focus on obtaining a surrogate model that can directly replace the exact pdf, the presented framework focuses on utilizing this surrogate to accelerate the sampling without a compromise in accuracy.

6.7 Conclusion

In this study, we presented a novel framework to efficiently yet accurately sample the posterior distribution of parameters in patient-specific cardiac electrophysiological models. This is achieved by first an active construction of an efficient GP surrogate of the posterior pdf, followed by the use of this surrogate to improve the proposal distribution of the standard MH. The presented method is evaluated on both synthetic and real data experiments. The future work will investigate methods to further improve the accuracy of the surrogate model, without requiring a large number of model evaluations.
Chapter 7

Conclusion and Future Work

While computational modeling and medical imaging technologies have made tremendous progress in the past few decades there are still many unresolved hurdles to estimating the model parameters along with their uncertainties [74]. The Bayesian active surrogate learning based approaches presented in this dissertation aims at taking us a step forward in this direction by addressing the associated challenges of high-computational cost and high input-dimensionality in model personalization and uncertainty quantification. In clinical and scientific study of cardiac diseases and disease mechanisms, uncertainty measures of modeling components or the model output are very important. The uncertainty in model prediction provides a measure of the quality of the prediction and helps to quantify the reliability of the model. The uncertainty also provide various knowledge about the modeling components such as unidentifiable parameters and interactions/correlations between the various modeling components. Finally, the uncertainty measure can also provide knowledge to the scientists and clinician on what additional data needs to be collected or what modeling component needs to be improved to improve the reliability of the model.

This dissertation focuses on Bayesian active learning approaches in application to cardiac model personalization and uncertainty quantification. To handle the existing challenge of learning a surrogate of a complex function defined over a high-dimensional (HD) unknown space of model parameters we presented two fundamental contributions in Bayesian active learning: a multi-scale surrogate-based optimization with adaptive spatial refinement de-
cision criteria and a generative variational auto-encoder (VAE) enabled high-dimensional (HD) Bayesian optimization. The actively learned surrogate of the expensive objective function was used for efficient parameter optimization in Chapters 3-5. Finally, in Chapter 6 a novel approach to utilize the surrogate of an unknown posterior density function in accelerating its Metropolis-Hastings sampling was presented. Future research directions are listed below:

7.1 Limitations

Recent developments in multi-scale cardiac simulation models, medical imaging technologies and high-performance computing resources have enabled enormous progress in personalized cardiac simulation models. These models have shown increasing capability in surgical planning, design and experimentation on medical devices, and understanding various cardiac disease mechanism [65, 68, 74]. However, there are still various limitations of personalized model, especially in relation to their clinical adoptions. Below we detail some limitations in specific to the personalized models presented in this dissertation work.

One major limitation of the approaches presented here is that they were tested on a simple two-variable phenomenological model called the Aliev-Panfilov model. While this model offers a faster execution time and the key macroscopic insights of cardiac activation, it does not offer detailed insights of microscopic ionic activity and cellular kinetics. Furthermore, the anatomical model lacks various details present in a human heart. For example, the fibers are mapped from a simple canine heart model, his bundle branches and purkinje systems are absent, and the temporal change in anatomy is ignored. In summary, the physiological model used in this paper lacks details in order to faithfully explain complicated disease mechanisms or ion channel interactions.

Another important limitation of personalizing physiological models and quantifying uncertainty in these models is the huge computational-cost associated with them. A single evaluation of the physiological model can take minutes or hours to run in a high performance computing system. Personalization and uncertainty quantification requires a large number of runs (in a range of thousands or millions) of these models making them computationally very expensive. This also restricts their application in clinical setting where a quick decision is needed. This thesis presented various approaches to re-
roduce the computational cost of uncertainty quantification and personalization primarily through building and using an efficient surrogate model. However, there is much room for further improvement in this regard.

Currently there are various difficulties in the clinical adoption of the personalized cardiac models [74]. One important challenge in terms of their use in clinical cardiology is that these models can produce variability in outcomes across the temporal and spatial scales. The main goal of the Bayesian active surrogate learning approaches developed in these paper is to be able to capture these uncertainties. Therefore, as a preliminary investigation of the feasibility of these methods, these methods have been extensively evaluated on synthetic and real data. However, they have not been studied in clinical setting with clinician. These methods have not been validated for clinical use. A future extension is to investigate these approaches with a multi-scale model and validate them with clinical and experimental data. This naturally means that the models and these methods need to be deployed with high-performance computing resources.

7.2 Future Work

Quantification of the uncertainty in model parameters that are estimated by using a generative VAE model for low-to-high dimensional mapping: One immediate future direction is to explore the integration of the HD Bayesian optimization via an embedded VAE method presented in Chapter 5 with the uncertainty quantification framework presented in Chapter 6. The goal is to quantify the uncertainty in estimated model parameters arising from both measurement errors and the low-dimensional representation of model parameters obtained with a generative model.

Bayesian active learning beyond Gaussian processes: A standard Gaussian process is useful in modeling a stationary process. In future, one important direction is to investigate various alternatives to Gaussian process in Bayesian active learning such as a deep Gaussian process or a non-linear neural net. These alternatives should be able to capture functions with complex and non-stationary properties.
CHAPTER 7. CONCLUSION AND FUTURE WORK

Improvement of the generalization ability of VAE: Another future direction is to improve the generalization ability of the VAE to generalize to realistic conditions where tissue abnormality is more complex in terms of the shape, transmurality, and heterogeneity. To this end, the first step is to investigate the limitation on the accuracy of the estimated parameters contributed by the generative model vs. that contributed by indirect measurements. Next steps are to use realistic training data from high resolution 3D imaging to train a more expressive generative model and to investigate further strategies to improve VAE generalization.

Transfer learning and life-long learning of the generative model: Chapter 5 presented an initial study that showed the feasibility of transfer learning in a graph convolutional VAE. Because medical data comprising of spatially varying tissue properties are not readily available a future work is to investigate ways to enable accurate transfer learning and life long learning in a graph convolutional VAE.

Generalization beyond the application of cardiac model personalization: Complex and costly objective functions are common in various domains [1, 15, 16, 80]. A future direction is to validate the presented approaches in various applications that involve simulation models beyond a cardiac electrophysiological model.

Uncertainty quantification in forward models: The forward model consisting of the complex and computationally expensive physiological model itself has many important applications. For example, it has been used for surgical planning, medical device design, and understanding biomechanics of cardiovascular disease progression [4, 5, 36, 50, 74]. This dissertation revolves around the surrogate modeling of a complex function that can be utilized in uncertainty quantification. Therefore, a future direction is to extend the presented Bayesian active surrogate learning approaches in the uncertainty quantification of the forward models. The uncertainty obtained in this manner can help scientists to understand the reliability of the model’s prediction, to quantify the risk associated with model’s prediction, to determine what additional data needs to be collected, and to determine what modelling element needs to be improved.
Uncertainty quantification of image-based personalized model: Image-based personalized simulation models have shown promising results in important clinical problems [4, 68, 74]. While these approaches differ from the presented approaches by using the patient’s medical images instead of the physiological signals for model personalization, they share the challenge of having a complex simulation model in the forward model. Therefore, a future direction is to extend the Bayesian active surrogate learning of a complex function discussed in this dissertation to these image-based personalization strategies. One direct extension is to investigate surrogate based uncertainty quantification to quantify the uncertainties arising from scar segmentation and thresholds used to assign tissue types in imaged-based personalized models. The quantification of this uncertainty can provide scientists and clinicians a measure of quality associated with the model’s prediction and helps them in informed decision-making.
Bibliography


[46] Matthieu Lé, Hervé Delingette, Jayashree Kalpathy-Cramer, Elizabeth R Gerstner, Tracy Batchelor, Jan Unkelbach, and Nicholas Ayache. MRI


BIBLIOGRAPHY


