

IMPERIAL COLLEGE LONDON

DEPARTMENT OF COMPUTING

Physics-Inspired Neural Networks For Cardiac Electro-Physiology Characterization

Author:
Alon Oved

Supervisor:
Dr Marta Varela

Submitted in partial fulfillment of the requirements for the MSc degree in
Computing Specialism (AI & ML) of Imperial College London

September 2021

Abstract

Atrial Fibrillation (AF) is the most common sustained arrhythmia. AF is correlated with increased morbidity (particularly stroke and heart failure) and mortality. Catheter ablation is the dominant AF treatment, although its long-term success rates are disappointing, especially for persistent AF patients (43%).

Cardiac electrophysiology (EP) models are fundamental tools for characterizing the creation and propagation of cardiac electrical signals. Individual parameters within these models represent several EP properties that can be inferred from action potential measurements. Inferring such properties would be clinically valuable in the diagnosis and treatment of AF.

To accurately reproduce action potentials and estimate EP parameters, we utilize EP - Physics-Inspired Neural Networks (EP-PINNs) and train them with sparse *in silico* data of membrane potentials in 1D and 2D spatial geometry. We demonstrate EP-PINNs' capability in simultaneously reconstructing the propagation of cardiac action potentials and approximating EP parameters related to action potential duration, tissue excitability, and conduction.

Furthermore, we show EP-PINNs' potential in detecting local pathologies associated with arrhythmia by identifying heterogeneities in EP properties. A leading example of a pathology linked with AF is fibrosis, which is generally characterized by local heterogeneity in conduction velocity, which we have successfully estimated.

Finally, we utilize experimental optical mapping data to establish EP-PINNs' capability to be deployed as a clinical tool. EP-PINNs successfully detected the effect of anti-arrhythmic drugs on action potential duration in cardiac tissue.

Acknowledgments

I wish to thank my supervisor, Dr. Marta Varela, for her committed guidance and assistance throughout the project. She was always available for queries and meetings, day and night, all in an effort to guarantee my project's progress and my well-being. She helped me improve my research skills, provided valuable feedback to my work, and challenged me to achieve project objectives we didn't think we will reach.

I would also like to thank Clara Herrero-Martin for being an amazing colleague for this project. It was a pleasure to work with you.

Special thanks to my close friends and family for your support, love, encouragement, and trust throughout this journey. It would not have been this memorable without you.

Contents

1	Introduction	1
1.1	Atrial Fibrillation	1
1.2	Aims	2
1.3	Thesis Outline	2
2	Background	4
2.1	Cardiac Electrophysiology Models	4
2.2	Parameter Estimation in EP Models	7
2.3	Physics-Inspired Neural Networks	8
2.4	Cardiovascular Applications of PINNs	11
3	Methods	12
3.1	Datasets	12
3.1.1	Generating Ground Truth EP Data	12
3.1.2	Experimental Optical Mapping Data	13
3.2	Design of EP-PINNs	14
3.3	Training and Evaluation	17
3.4	EP-PINNs for Monodomain Forward Mode	20
3.4.1	1D	20
3.4.2	2D	21
3.5	EP-PINNs for Parameter Estimation	22
3.5.1	Monodomain Model	22
3.5.2	Canine Atrial Model	23
3.5.3	Experimental Data	23
4	Results	25
4.1	EP-PINNs for Monodomain Forward Mode	25
4.1.1	1D	25
4.1.2	2D	26
4.2	EP-PINNs for Parameter Estimation	28
4.2.1	Monodomain Model	28
4.2.2	Canine Atrial Model	31
4.2.3	Experimental Data	32

5 Discussion	33
5.1 Forward Solution	33
5.2 Parameter Inference	35
5.3 Limitations and Future Modifications	36
A Software Design	43
B Inverse Mode - 1D	46
C Inverse Mode - 2D	48

Chapter 1

Introduction

1.1 Atrial Fibrillation

Atrial Fibrillation (AF) is the most common clinically diagnosed cardiac arrhythmia and is associated with a markedly increased risk of several health issues such as stroke, heart failure, development of dementia, ischemia, systemic embolism, and even death [1]. In 2010, AF prevalence was calculated at 33 million people globally, and although it varies among different regions, AF is estimated to affect 1-4% of adults in Australia, Europe, and the USA. Between 1990 and 2010 mortality rate associated with AF was doubled. Furthermore, the financial burden on healthcare services related to AF is growing and estimated at 1% of NHS budget in the UK and 16-26\$ billion in the US annually. Worldwide prevalence is projected to rise due to global factors such as aging population and increased detection rates [2, 3].

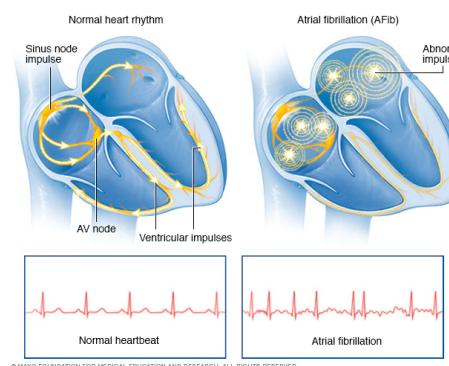


Figure 1.1: A healthy normal heart rhythm (left) as opposed to Atrial Fibrillation (right) [4].

AF is characterized by irregular or rapid heart rate that is caused by a chaotic discharge of electric signals in the atria, the two upper chambers of the heart. The biological and physical mechanisms of AF have been studied for decades. Although there is no consensus over what triggers it, several hypotheses suggested, such as multiple re-entrant wavelets of AP, rapidly discharging automatic foci, and a single re-entrant circuit with fibrillatory conduction [1].

Symptomatic patients may receive an anti-arrhythmic drug (AAD) treatment, although cases of persistent AF usually require an invasive treatment as catheter ablation, mainly the pulmonary vein isolation (PVI) type [5]. Catheter ablation is a complex procedure that is executed by inserting tubes through veins. Ablation is designed to restore local cardiac rhythm by creating lesions in the atrium, typically around the pulmonary veins, consequently eliminating targeted electrical pathways. These selected pathways cause rhythm abnormalities, thus blocking them destroys cardiac muscle that triggered the arrhythmia [6]. One-year freedom from AF is an accepted metric to measure treatment success for persistent AF cases. A single catheter ablation treatment's success rate is unsatisfactory and stands at only 43% of the cases. It increases to 69% by repeating the ablation procedure several times with or without AADs [5].

There exist two main procedures to acquire a patient's cardiac activity information and record their electrical potentials:

- Using electrodes placed on the chest skin (electrocardiogram, ECG);
- Using a catheter to contact the myocardium directly (contact electrogram, EGM). These invasive measurements are only usually available when the patient is undergoing an ablation procedure.

To clinically diagnose arrhythmias and other cardiovascular diseases, experts analyze these signals. Although diagnosis is achieved with great accuracy, sparse measurements of the electrical activity offer little direct information about electrophysiology properties [7].

1.2 Aims

Current catheter ablation strategies are based on general anatomic approaches [8]. We hypothesize that if we improve the characterization of the properties of cardiac tissue, we may be able to find additional places in which to ablate and thus improve treatment efficacy. If successful, the improved treatment to AF will be patient-specific as oppose to a standard ablation scheme.

Our approach involves utilizing physics-inspired neural networks (PINNs) to identify relevant electrophysiological (EP) properties. The designed PINNs incorporate a selected computation cardiac EP model. By setting a set of parameters, the model would represent either a healthy cardiac activity or an arrhythmic condition. These PINNs are designed to predict parameters related to EP properties such as the conduction velocity, the excitation threshold, and action potential duration.

Our main aim is to estimate these EP parameters from both *in silico* (synthetically generated) data and experimental optical mapping data of transmembrane potentials.

1.3 Thesis Outline

This thesis is organized as follows:

- Chapter 2 describes the background topics relevant to this project, such as the cardiac EP model, current parameter estimation strategies, physics-inspired neural networks, and other related works;
- Chapter 3 provides details on the techniques and methods used to achieve the objectives we have set. These include data acquisition, software and PINNs design, training procedures, and the various tasks executed to test our hypothesis;
- Chapter 4 specifies the results for the implemented tests;
- Chapter 5 is a discussion on the project's results, limits, and suggestions for future works.

Contributions

Several parties contributed to this project. Dr. Marta Varela, besides supervision, delivered the Matlab data files to generate the synthetic datasets. Clara Herrero-Martin assisted with code development as well as providing the programming solution for the 2D forward heterogeneity task.

Chapter 2

Background

2.1 Cardiac Electrophysiology Models

Electrical depolarization of the cell membrane prompts a mechanical contraction within the cardiac tissue. Subsequent repolarization (restitution) follows as the potential across the cell membrane decreases back to its resting value. That sequence of supra-threshold depolarization and repolarization is called action potential (AP). Cardiac muscle consists of several cell types and the most prevalent are excitable myocytes and fibroblasts. Above a voltage threshold, voltage-gated channels permeable to positively charged ions open, leading to membrane depolarization. This initiation of an AP generates additional local depolarization within the tissue. This current propagates to neighboring tissues and throughout the heart (see Figure 2.1) [9].

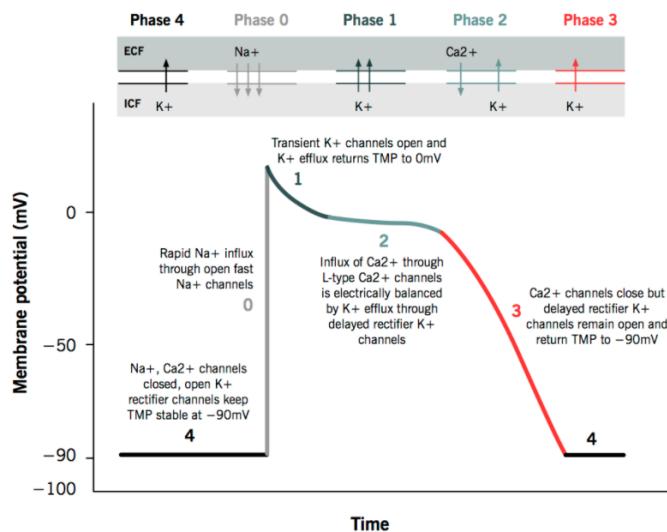


Figure 2.1: Typical AP of cardiac muscle [10]

Cardiac electrophysiology (EP) models are essential tools for studying the influence of cardiac tissue properties on the generation and propagation of cardiac electrical

signals [9]. Since the 1960s, researchers have introduced various EP models designed to emulate different EP properties and cardiac regions for humans or other animals (see [11]).

Each model carries a different level of complexity based on a different set of assumptions. Mathematically, most EP models are formulated as a reaction-diffusion system, consisting of a partial differential equation (PDE) that defines the changes in electrical potential with respect to time and space. The diffusion PDE is typically coupled with a single ordinary differential equation (ODE) or more, which describes the relationship between the flux of ions across the cell membrane and local state variables such as membrane potential [9].

The monodomain diffusion equation of membrane potential, noted V_m in equation (2.1), assumes that the conductivity in the extracellular space is proportional to the intracellular conductivity [9]. It is the base EP model on which we build this project.

$$\frac{\partial V_m}{\partial t} = \nabla(D\nabla V_m) - \frac{i_{ion}}{C_m} \quad (2.1)$$

C_m is the cell membrane capacitance, i_{ion} is the current of ions across the membrane and D is the diffusion tensor. It is worth noting that V_m is a function of both time and space, and its units are mV [9].

Aliev Panfilov model [12] described mathematically in the coupled equations (2.2) and (2.3) is our core model. It can be coupled to the monodomain equation (2.1) for the diffusion of V . The model outlines the AP as a traveling excitation wave followed by a non-excitatory region and models the ionic current, outlined by the interaction between V and W , using 6 parameters. The ODE, equation (2.3), relates to the non-observable, non-diffusible variable W , which controls the refractoriness and restitution attributes of the model.

$$\frac{\partial V}{\partial t} = \vec{\nabla}(D\vec{\nabla}V) - kV(V - a)(V - 1) - VW \quad (2.2)$$

$$\frac{\partial W}{\partial t} = (\epsilon + \frac{\mu_1 W}{V + \mu_2})(-W - kV(V - b - 1)) \quad (2.3)$$

The diffusion term, $\vec{\nabla}(D\vec{\nabla}V)$, can be simplified to $D\nabla^2V$ under the assumption of homogeneous and isotropic conduction (i.e., the diffusion tensor D is a constant scalar across space and time). The value of D is mainly influenced by the myocardium conductivity and is a vital factor of the AP propagation velocity [9].

The parameter a is related to the excitation threshold (i.e., the minimum potential V required to trigger an action potential), and b determines the action potential duration (APD) and refractoriness [12].

The rest of the parameters are less interpretable and were kept as constants throughout the project. To reproduce observed electrical signals, we adopted the values for the model parameters as empirically analyzed by Aliev Panfilov [12] (see Table 2.1). The required boundary conditions (BC) for this computational model are usually the Neumann BC, also called the no flux BC: $\frac{\partial V}{\partial \vec{n}} = 0$ on the boundary of the cardiac

Table 2.1: Default values for model parameters.

Parameter	value
D	0.1
a	0.01
b	0.15
k	8
ϵ	0.002
μ_1	0.2
μ_2	0.3

tissue. This condition prevents membrane potential leakage to external regions [9]. For completeness, note that V (non-dimensional, typically in the range of $[0, 1]$) is a surrogate of the cardiac tissue membrane potential V_m , and t is a scaled unit for time. V can be converted into membrane voltage in mV units and temporal variable t to time in ms units by the following affine maps:

$$V_m[mV] = -80 + 100v \quad t[ms] = 12.9t \quad (2.4)$$

Each set of parameters and initial conditions (IC) to the EP model can represent either a healthy cardiac activity or an arrhythmic condition. Healthy cardiac activity is commonly displayed as unidirectional smooth wave propagation from a single origin (Planar wave by large sources and centrifugal waves by point-like stimulation) - see Figure 2.2 a-b. Arrhythmias are usually modeled as one or more re-entrant waves: spiral waves in 2D - Figure 2.2c and scroll waves in 3D. Furthermore, heterogeneities in some model parameters can model localized pathology such as fibrosis, scar, or ischemia (Figure 2.2 d-e). For example, a limited regional decrease of D emulates a slower AP propagation in these regions [8].

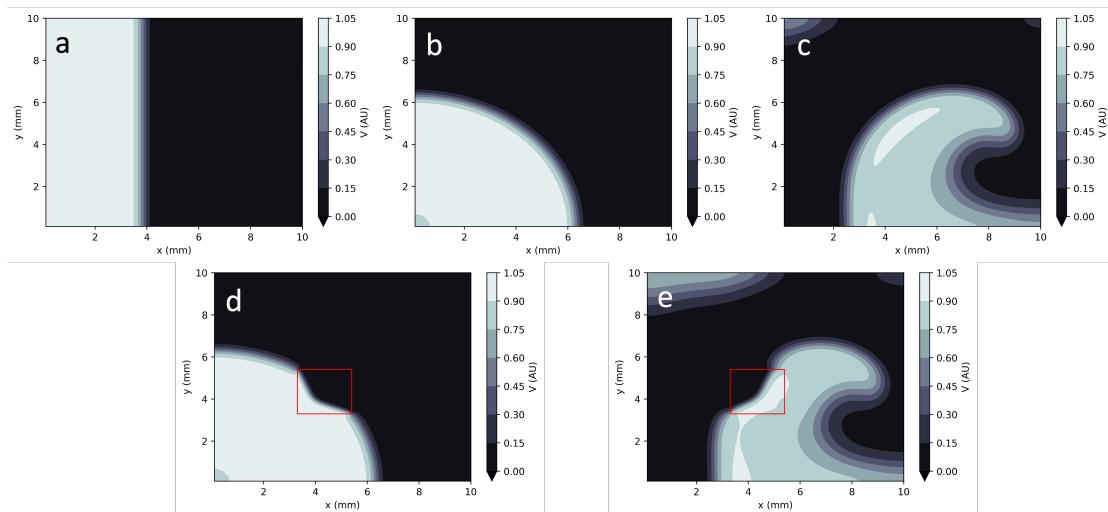


Figure 2.2: a) Planar wave; b) Centrifugal wave; c) Spiral wave; d) Centrifugal wave in the presence of heterogeneity in D ; e) Spiral wave in the presence of heterogeneity in D .

Other EP models can express more biological attributes of the AP and its propagation, such as individual membrane ionic currents. Those models offer more flexibility but are inherently more complex. Such a model is the 14-current 30-variable 100s-parameter (as opposed to our 1, 2, and 6, respectively) Canine Atrial model [13] that combines several degrees of EP remodeling caused by AF. The canine atrial model is used in this project, as discussed in the Methods chapter 3.5.2.

2.2 Parameter Estimation in EP Models

Computational cardiac EP models could be used to characterize accurately and efficiently an individual's underlying condition and tailor patient-specific treatment. The main idea is to estimate a subset of the EP model parameters using the relevant measurements of the patient. In cases where the clinical data for an individual is not sufficient, it is possible to include population-based techniques [14]. Estimating EP parameters from the measured cardiac electrical signals would assist in improving personal treatment strategies for cardiovascular arrhythmias (such as AF). Recent studies suggest that the irregular electrical activity that characterizes AF can be associated with heterogeneities in cardiac tissue, such as fibrosis, which can be retrieved by parameters related to conduction velocity [8].

Parameter estimation problems, commonly referred to as inverse problems, are prevalent in science [15]. Many methods exist for solving inverse problems: regression (linear and non-linear), iterative models, Fourier techniques, Bayesian methods, and more [16]. However, algorithms can become excessively complex and computationally intensive when the described physical system is defined by non-linear differential equations (partial and/or ordinary), as the Aliev Panfilov model [17].

Dokos and Lovell's study [18] is an example for estimating parameters with cardiac ionic models. They use the Beeler-Reuter EP model [19], which is composed of several ODEs and parametrized with 63 parameters. They have made several assumptions in order to reduce the number of parameters. Subsequently, the problem's dimensionality and the number of optimum points are limited. The method they used was the non-linear least-squares and could fit and estimate parameters only to one AP. Therefore, when trained with data of more than one AP, the model performed poorly. Furthermore, to fulfill its potential, the model requires large datasets and intensive computational resources.

Another common method in parameter estimation is Bayesian optimization [15, 20]. Bayesian methods usually make assumptions about the a priori statistical distributions of the data, the model, and the parameters involved in order to obtain the posterior distribution of the target parameters. These methods offer advantages in results interpretability and uncertainty analysis. However, their current challenges remain the ability to recover parameters with small, sparse, and noisy clinical data as well as the identifiability of model parameters.

2.3 Physics-Inspired Neural Networks

In recent years, artificial neural networks (NN) have been proven to be a powerful tool for solving data-driven scientific problems. Utilizing NNs, some fields have seen tremendous performance growth, such as computer vision, neural language processing, computational biology, genomics, and more [21–23].

Several requirements are necessary for NNs to converge and perform well: large high-quality datasets, suitable algorithms (architectures), and advanced computing resources. A leading challenge in many physical and biological problems is data acquisition which can be expensive, technologically limited, and even invasive. Although practically NNs may lack robustness or even fail to converge in small-size dataset conditions, several methods exist to cope with data shortage (e.g., data augmentation). Physics-inspired neural networks (PINNs) are an exciting and relatively new approach for such scenarios [24].

PINNs utilize the system's theoretical physics laws in the NNs, through the architecture and the loss function. This information is not exploited in standard machine learning algorithms. Incorporating the system's known dynamics and properties into the NN adds valuable signals that can boost the training process and enable convergence. Moreover, that added scientific knowledge may function as regularization and assist in overcoming data shortage [24].

Solving time-dependent PDEs and ordinary differential equations (ODEs) using PINNs is a new approach compared with conventional techniques, mostly numerical methods, such as finite-difference, finite-element, and finite-volume methods [25]. Each problem is characterized by specific system dynamics, initial and boundary conditions, symmetries, invariances, equivariances, and physics rules that PINNs can inherently enforce. PINNs use the automatic differentiation tools that already exist to compute terms in the relevant PDEs to create a minimization function based on the residuals [24, 26].

The setup for a general PDE can be written as the following system:

$$\frac{\partial u}{\partial t} = F(u(x, t); \lambda), \quad x \in \Omega, \quad t \in [0, T] \quad (2.5)$$

Where $u(x, t)$ denotes the latent solution, $F(u(x, t); \lambda)$ is a non-linear operator parametrized by $\lambda \in \mathbf{R}^P$ (P is the number of parameters involved in the dynamics), and $\Omega \subseteq \mathbf{R}^D$ is the PDE's spatial domain. PINNs are capable of handling two different forms of optimization tasks:

1. Data driven solution (**forward mode**): Given the fixed model parameters λ in the PDE, equation (2.5), optimize the NN's parameters to approximate $u(x, t)$;
2. Parameter estimation (**inverse mode**): Optimize the NN's parameters to approximate the parameter set λ as well as $u(x, t)$.

Recent studies show that PINNs can efficiently identify high-resolution solutions (forward problem) and operate systems identification (inverse problem) [24].

PINNs based on the system dynamics described in the PDE equation (2.5) are trained

using the observable measurements $u(x, t)$, at chosen N spatial-temporal locations (x_n, t_n) (see Figure 2.3).

In inverse mode, PINNs attempt to estimate parameters λ as well as predicting $u(x, t)$.

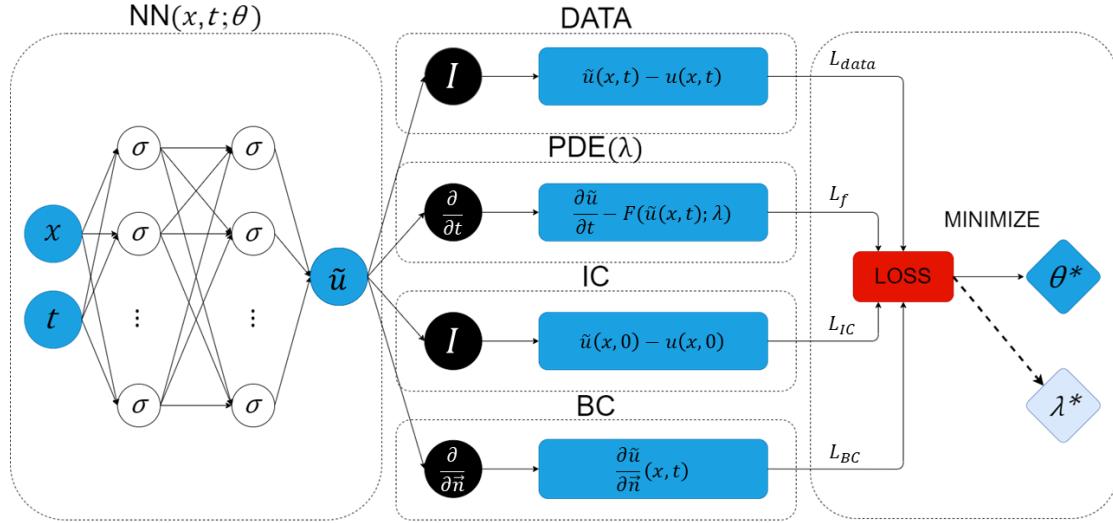


Figure 2.3: A general PINNs structure: a feed-forward NN with a loss function that corresponds to the system's physics laws, the observable measurements, initial and boundary conditions. The network optimizes its weights θ to approximate $u(x, t)$ in forward mode and is able to also estimate parameters λ in inverse mode.

Let us define f_u by rearranging the PDE equation (2.5):

$$f_u(x, t) = \frac{\partial u}{\partial t} - F(u(x, t); \lambda) \quad (2.6)$$

Notice that by the system dynamics defined above, $f_u = 0$.

The notation $\tilde{u}(x, t)$ represents the NN's approximation of $u(x, t)$. \tilde{f}_u is the estimation of f_u based on the NN's output $\tilde{u}(x, t)$ and its derivatives:

$$\tilde{f}_u(x, t) = \frac{\partial \tilde{u}}{\partial t} - F(\tilde{u}(x, t); \lambda) \quad (2.7)$$

The loss function PINNs attempt to minimize is composed of 4 different terms:

$$L = L_{data} + L_f + L_{IC} + L_{BC} \quad (2.8)$$

Each term corresponds to a different aspect within the system dynamics: empirical measurements agreement, L_{data} ; consistency of physical laws, L_f ; initial conditions, L_{IC} ; and boundary conditions, L_{BC} . These would be calculated individually and eventually summed. There are several metrics to select from when computing losses, each system requires one that fits its properties best. Furthermore, it is possible to choose different losses for each term and even combine various metrics for an

individual expression. The below terms are calculated using mean squared error (MSE).

Data agreement condition loss:

$$L_{data} = \frac{1}{N} \sum_{n=1}^N \left(\tilde{u}(x_n, t_n) - u(x_n, t_n) \right)^2 \quad (2.9)$$

As in any supervised NN, the network's outputs should fit the ground truth data.

Consistency of physical laws loss:

$$L_f = \frac{1}{N} \sum_{n=1}^N \left(\tilde{f}_u(x_n, t_n) \right)^2 \quad (2.10)$$

This term expresses the core novel concept of PINNs, a requirement to conform with the physics laws. The better the network complies with the system equation (2.5), the lower this loss term will get.

Initial condition loss:

$$L_{IC} = \frac{1}{N_i} \sum_{n=1}^{N_i} \left(\tilde{u}(x_n, 0) - u(x_n, 0) \right)^2 \quad (2.11)$$

N_i is the number of data points that comply with the condition $\{t = 0\}$.

Other possible initial conditions are available, depending on the physics dynamics at hand.

Boundary condition loss:

$$L_{BC} = \frac{1}{N_b} \sum_{n=1}^{N_b} \left(\frac{\partial \tilde{u}}{\partial \vec{n}}(x_n, t_n) \right)^2 \quad (2.12)$$

N_b is the number of data points that comply with the boundary condition. The above loss L_{BC} is based on the Neumann boundary condition mentioned before (chapter 2.1). The vector \vec{n} is the local normal to the domain Ω .

Other possible boundary conditions are available, depending on the physics dynamics at hand.

An issue that can arise when designing PINNs is the possible difference in magnitude among the different loss terms. Automatic scaling layers can amend the problem if necessary [27]. It is worth noting that there has been little study of theoretical and mathematical proofs regarding PINNs convergence and the conditions with which they can achieve a global optimum [28].

Lastly, there are currently no automatic tools to approximate the uncertainty of PINNs' output. Bayesian-PINNs may be suitable in certain dynamics to resolve that [29]. Our strategy was to estimate solutions several times and compute the mean, standard deviation, and range.

2.4 Cardiovascular Applications of PINNs

Several studies in cardiovascular medicine have recently used PINNs, notably for blood flow applications. Recent researches include the estimation of myocardial perfusion from contrast-enhanced MRI data by [30] and the prediction of hemodynamic parameters from microscopic images of aneurysms-on-a-chip by [31].

A study in the field of cardiac electrophysiology used PINNs for estimating activation times mapping in the left atrium at high spatial resolution [32]. Activation times (ATs) are the arrival times of the action potential. The study attempts to solve the forward mode described by the (curvature) eikonal equation, which mathematically represents the relation between ATs and the spatial gradient of conduction velocity. Their PINN implementation can capture the collision of wavefronts, even though it was not possible by previous algorithms. Additionally, Grandits et al [33] utilized the eikonal equation to design PINNs and estimated high resolution AT maps (and fiber directions) from synthetic (*in silico*) and patient datasets. Their PINN's performance was worse, only on synthetic data, compared with a current (numerical) inverse method.

However, the eikonal equation is a limited EP model concerning this project's core objectives: first, it cannot express the CV of re-entrant loops present in cardiac arrhythmias, and second, the only EP parameters possible for inference are AT and CV [32].

This project investigates PINNs' capability to infer various cardiac EP attributes using both synthetic (*in silico*) and experimental data. Selecting the Aliev Panfilov model [12] increases the EP model's complexity compared with previous studies in an attempt to capture the biological mechanisms of AF.

Chapter 3

Methods

3.1 Datasets

We utilized two types of datasets in this project: synthetically generated data (*in silico*) and experimental optical mapping data.

3.1.1 Generating Ground Truth EP Data

To assess the performance of PINNs on various tasks, we used ground truth (GT) synthetic EP data. These generated datasets consist of values for the membrane potential, V , and the latent variable W for every spatial-temporal step (x, t) . In order to generate the GT data, we use the central finite differences and an explicit 4-stage Runge-Kutta method [34] to solve the monodomain Aliev Panfilov model, described in equations (2.2) and (2.3). Moreover, Neumann boundary conditions are utilized and the spatial and temporal step sizes are set to $dx = 100\mu\text{m}$ and $dt = 5\text{AU}$ (arbitrary units), respectively.

Table 2.1 in chapter 2.1 lists the model parameters used for generating GT data.

We used two basic tissue geometries in this project: 1D (a cable) and 2D (a rectangle). For both domains, we programmed the excitation waves by adding an external stimulus current $V_{stim} = 0.1 \text{ AU}$ to the right-hand side of the monodomain PDE, equation (2.2), for 1 time unit.

Five different datasets are generated, for various types of waves:

1. **1D left wave** - stimulus current added to the 5-most left cells;
2. **1D right wave** - stimulus current added to the 5-most right cells;
3. **2D corner wave** - stimulus current added to the 5×5 upper-right corner cells;
4. **2D vertical planar wave** - stimulus current added to the 5-most left cells in each row in the rectangle;
5. **2D spiral wave** - created using the cross-field protocol [35]: after a first planar excitation wave propagates, a second planar wave, orthogonal to the first, is initiated. Accurate timing between both waves will produce a sustained spiral wave.

All 5 cases listed were simulated with homogeneous and isotropic conduction, whereas only two cases, 2D corner wave and 2D spiral wave, were additionally generated with heterogeneous geometry.

In order to create homogeneous geometry, the diffusion parameter D is fixed ($D = 0.1$) throughout the space and time domain. Heterogeneous geometry is defined by assigning in a small rectangle ($2 \times 2 \text{ mm}^2$) with a diffusion factor of $D = 0.02$, whereas outside the rectangle $D = 0.1$. The parameter D is time-independent in both heterogeneous and homogeneous geometries.

Canine Atrial GT Data

Another EP biophysical model for which data was generated was the canine atrial model by Varela et al [13] discussed earlier. We use different cell types, left atrial (LA) and crista-terminalis/Bachmann's bundle (BB), and different AF remodeling degrees (no remodeling, moderate, and severe) to produce several datasets. Four datasets for canine atrial model are generated:

1. LA cell type with no remodeling;
2. LA cell type with moderate remodeling;
3. LA cell type with severe remodeling;
4. BB cell type with no remodeling.

Essentially, the objective is to perform inference on these datasets, which mainly vary in action potential duration (closely related to the b parameter). As remodeling factor increases, the action potential duration decreases. Additionally, without any remodeling, the APD of LA cell type is shorter than BB's APD. This model's equations are solved with central finite differences and explicit forward Euler schemes [36], with the same 1D spatial-temporal properties and the same step sizes $dx = 100\mu\text{m}$ and $dt = 5\mu\text{s}$.

3.1.2 Experimental Optical Mapping Data

Experimental optical mapping data was supplied and acquired by Chowdhury et al [37]. It was conducted using neonatal rat myocardial preparations stained with a voltage-sensitive dye by the relevant protocols. Essentially, the optical mapping data is a time sequence of 2D spatial images, and to convert them to datasets suitable to our EP-PINNs, we adopted the following procedure. We marked two regions of interest (ROIs), squares of $0.13 \times 0.13 \text{ mm}^2$, within the optical mapping image sequence and computed their average. The signal was then sampled with a 2ms rate for roughly 300ms . We implemented two pre-processing steps to improve the signal-to-noise ratio and decrease experimental artifacts' effect:

- Identifying the APs' phase-1 to align and average the temporal sequence of three action potentials;

- Applying twice a moving average kernel.

Lastly, the detailed process results in data for one spatial point, and to obtain a second, we repeat the detailed procedure with two other ROIs (0.7mm apart) in the activation's direction. Consequently, the outcome is a 1D dataset for two cells in a small distance throughout time.

Each experiment recorded the potentials of the cells throughout the spatial-temporal space in two distinct observations: resting conditions and anti-arrhythmic drug-stimulated conditions. We utilized data from two such experiments, one with Nifedipine drug and the other with E-4031, leading to a collection of four datasets. Nifedipine and E-4031 oppositely affect the action potential duration by blocking the L-type calcium channel I_{CaL} (decreasing APD) and the rapid potassium channel I_{Kr} (increasing APD), respectively.

3.2 Design of EP-PINNs

EP-PINNs are designed and trained via the DeepXDE Python library [38]. DeepXDE enables the user to assemble PINNs in a few lines of code, as it already contains the necessary automatic differentiation tools needed for any PDE/ODE in question. The library includes several pre-defined boundary conditions, geometries, network architectures, loss functions, optimizers, and more. Users can define training objectives to achieve either model solution (forward mode) or parameter estimation (inverse mode) [38].

EP-PINNs implemented in our project are based on the system dynamics described in Aliev Panfilov model equations (2.2) and (2.3). They are trained using the observable measurements $V(x, t)$, the membrane potentials, at chosen N spatial-temporal locations (x_n, t_n) . The network's x input may be composed of one (in 1D) or two (in 2D) coordinates according to the relevant geometry within the task.

The designed EP-PINNs for the inverse mode are capable of estimating one or more of the model parameters: tissue excitability (a in Eq. 2.2), electrical conductivity (D in Eq. 2.2), and action potential duration (b in Eq. 2.3).

Let us define f_1 and f_2 by rearranging Aliev Panfilov equations (2.2) and (2.3).

$$f_1(x, t) = -\frac{\partial V}{\partial t} + \left[\vec{V}(D\vec{V}V) - kV(V - a)(V - 1) - VW \right] \quad (3.1)$$

$$f_2(x, t) = -\frac{\partial W}{\partial t} + \left[\left(\epsilon + \frac{\mu_1 W}{V + \mu_2} \right) (-W - kV(V - b - 1)) \right] \quad (3.2)$$

Notice that by Aliev Panfilov model, $f_1(x, t) = 0$ and $f_2(x, t) = 0$. The notations $\tilde{V}(x, t)$ and $\tilde{W}(x, t)$ are the NN's approximation of $V(x, t)$, $W(x, t)$ respectively. \tilde{f}_1 and \tilde{f}_2 are the estimations of f_1 and f_2 respectively, based on the NN's outputs $\tilde{V}(x, t)$, $\tilde{W}(x, t)$ - see Figure 3.1.

As described before, the training loss function this EP-PINN attempts to minimize is composed of four different terms:

$$L = L_{data} + L_f + L_{IC} + L_{BC} \quad (3.3)$$

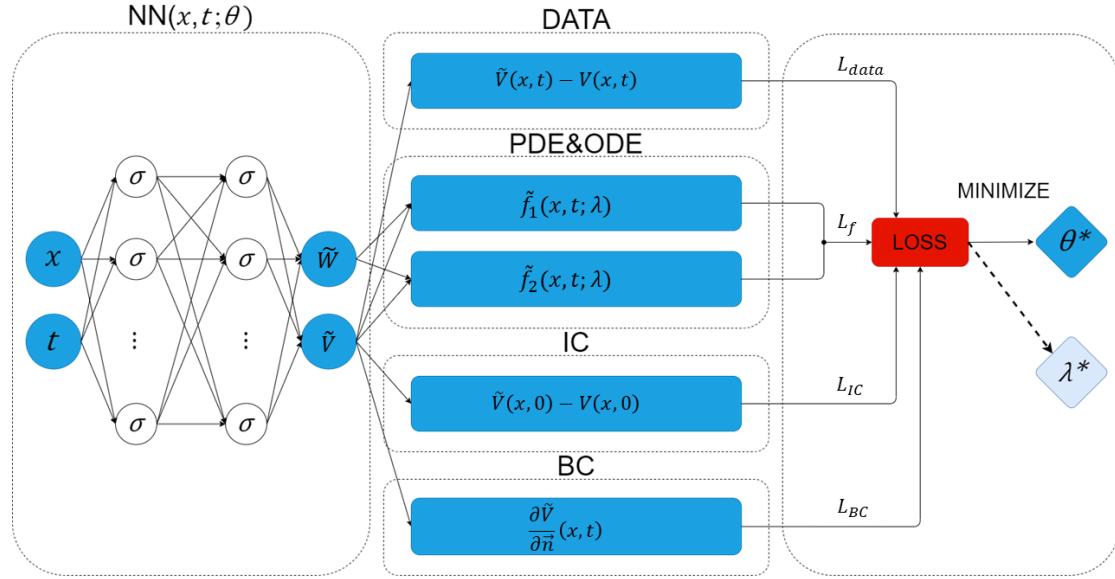


Figure 3.1: EP-PINN design based on the Aliev Panfilov model: a feed-forward NN with the spatial-temporal point as an input and outputs estimations for V and W . The loss function is composed of various terms for data agreement, physics laws, initial and boundary conditions. The EP-PINN optimizes its weights to solve the coupled PDE-ODE system, and if requested, it can also solve the inverse mode. λ is the general term for the target set of parameters to estimate, and it can be one or more of a , b , and D .

Each term corresponds to a different aspect within the system dynamics, empirical measurements agreement, L_{data} , consistency of physical laws, L_f , initial conditions, L_{IC} , and boundary conditions, L_{BC} and would be calculated individually using mean squared error (MSE).

Data agreement condition loss:

$$L_{data} = \frac{1}{N} \sum_{n=1}^N \left(\tilde{V}(x_n, t_n) - V(x_n, t_n) \right)^2 \quad (3.4)$$

Consistency of physical laws loss:

$$L_f = \frac{1}{N} \sum_{n=1}^N \left(\tilde{f}_1(x_n, t_n)^2 + \tilde{f}_2(x_n, t_n)^2 \right) \quad (3.5)$$

The two terms correspond respectively to the PDE, equation (2.2), and the ODE, equation (2.3).

Initial condition loss:

$$L_{IC} = \frac{1}{N_i} \sum_{n=1}^{N_i} \left(\tilde{V}(x_n, 0) - V(x_n, 0) \right)^2 \quad (3.6)$$

N_i is the number of data points comply with the condition $\{t = 0\}$.

Boundary condition loss:

$$L_{BC} = \frac{1}{N_b} \sum_{n=1}^{N_b} \left(\frac{\partial \tilde{V}}{\partial \vec{n}}(x_n, t_n) \right)^2 \quad (3.7)$$

N_b is the number of data points that comply with the boundary condition. The above loss L_{BC} is based on the Neumann boundary condition described in previous chapters.

Note that the EP-PINN described is trained to predict both $V(x, t)$ and $W(x, t)$ even though the training ground truth (GT) data includes only $V(x, t)$. That structure is more closely resembles the experimental setup as W is a latent variable and immeasurable in practice.

A feed-forward fully connected NN was constructed, with a \tanh activation function for each layer, except the output layer. We initialize the network's weights with the Xavier initialization scheme [39].

As discussed in previous sections, cardiac data acquisition in our experimental setup is complex and expensive, which means that in many cases, the amount of training data available is limited. Therefore, EP-PINNs for the 1D model used 10% of the generated GT data (randomly sampled) for training and 90% for testing, whereas the 2D model used 20% and 80%, respectively.

We designed the detailed EP-PINN with three architectures for the feed-forward NN, all accept (x, t) as an input (either 2 or 3 values for 1D or 2D, respectively) and output \tilde{V} and \tilde{W} . The difference lies with the network's width and depth through the hidden layers:

- Architecture A: 4 layers with 32 neurons each;
- Architecture B: 5 layers with 60 neurons each;
- Architecture C: 6 layers with 100 neurons each.

Heterogeneity NN Modification

There are two challenges that heterogeneous setup imposes:

1. D is not a scalar as in homogeneous geometry but varies over the 2D space;
2. D is time-independent. Incorporating that property in the EP-PINN is critical.

The 2D heterogeneous data includes a small rectangle with a reduced $D = 0.02$ value, whereas elsewhere within the 2D space, $D = 0.1$ (detailed in chapter 3.1.1). In that geometry, x carries two coordinates as we generated and tested heterogeneous setup only in 2D.

We designed two parallel-trained feed-forward fully connected NNs to both comply with the system dynamics (accurately predict $V(x, t)$ and $W(x, t)$) and estimate

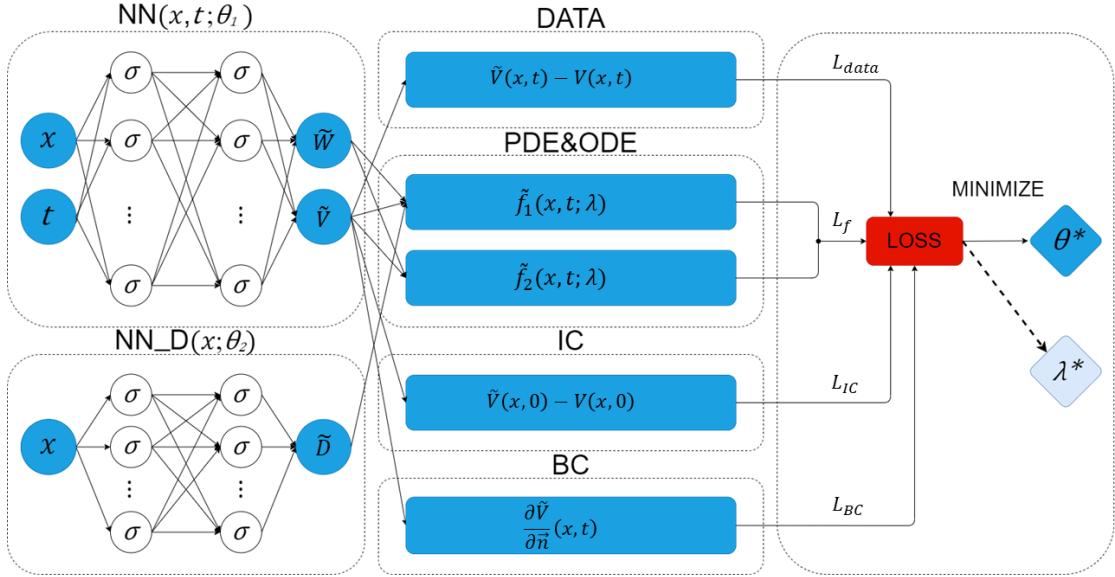


Figure 3.2: EP-PINN for heterogeneity in D . The added NN (NN_D) approximates $\tilde{D}(x)$, and uses that value to compute the term $\tilde{f}_1(x, t; \lambda)$.

$D(x)$. The modified EP-PINN's structure is similar to the one described in Figure 3.1 with Architecture A. The second network takes only the spatial inputs x to approximate $D(x)$. It consists of 5 hidden layers with 60 neurons each and a \tanh activation function. The networks' outputs ($\tilde{V}(x, t)$, $\tilde{W}(x, t)$, and $\tilde{D}(x)$) are then incorporated into the relevant expressions within the loss function, and back-propagation is carried accordingly.

Forward and inverse mode tasks utilized that EP-PINN differently:

Forward mode: $D(x)$ is not to be estimated but assumed as known for every coordinate x . Therefore, we modify the network's output $D(x)$ to match the values in the GT rectangle geometry, DeepXDE library enables that design.

Inverse mode: We incorporate into the EP-PINN that the value of $\tilde{D}(x)$ in every spatial coordinate x should either be 0.02 or 0.1. The sigmoid activation function operates as a smooth discriminator. Therefore, we perform the scaled sigmoid version on the output of $\tilde{D}(x)$: $A(\tilde{D}) = \text{sigmoid}(\tilde{D}) \cdot 0.08 + 0.02$.

Note: Each experiment may require a different training dataset (the synthetic Aliev Panfilov GT, the synthetic Canine Atrial GT, or the experimental data), geometry, and NN architecture. Regardless, all share the same core EP-PINN design based on the monodomain Aliev Panfilov EP model.

3.3 Training and Evaluation

We designed many tasks in this project with different objectives. Each experiment required adjustments to the input data, NN architecture, training schemes, learning rates, optimizers, and more. Two general training methods are described in this

chapter. However, specific details may vary for individual tasks and will be reported appropriately.

Auxiliary Datasets

DeepXDE library allows the users to train the network with another type of data besides the GT dataset, called the auxiliary dataset. Utilizing that dataset is expected to assist the PINN in optimizing its outputs to the physics equations and conditions. The typical training inputs, the observable measurements, may not cover the entire spatial-temporal domain sufficiently. Thus, the auxiliary dataset would supply the network with points across the domain in any selected resolution.

To see how this is possible, note that the GT values of V are not associated with the consistency of physical laws loss term, L_f , and the boundary condition term, L_{BC} in the network's loss function. Auxiliary data consists of (x_i, t_i) pairs and no GT values for $V(x_i, t_i)$ but instead, the estimations $\tilde{V}(x_i, t_i)$. If auxiliary data is requested, each training batch will include both ground truth and auxiliary datasets.

Furthermore, it is possible to train up to two classes of auxiliary datasets: domain and boundary. Auxiliary domain data would be inside the spatial-temporal domain and auxiliary boundary data at the set corresponds to the boundary conditions. The auxiliary training dataset does not carry ground truth values for V . Therefore, it accounts only for the physical dynamics and boundary condition loss terms. Loss back-propagation and updates for the NN's weights are carried out accordingly.

We have used auxiliary domain and boundary data points in all experiments, sizes are detailed in the below training schemes section.

EP-PINN Initialization

The PINNs' initialization phase requires special attention as the network's performance is closely affected by it. By our experiments, the trained NN fails to converge if the initialized network's loss is too large. That event happened in a non-negligible amount of trial runs, and we assume that the loss function's complex geometry is mainly responsible for that. To address this issue, we execute a recurrent initialization phase that re-initializes the network until the loss is sufficiently tolerable ($\leq 4 AU$). That method assisted the network to converge consistently.

In inverse mode, one or more of the parameters are optimized through training. We initialize their values with: $a = 0.02$, $b = 0.3$, and $D = 0.2$. These values are relatively realistic within Aliev Panfilov model and equate to twice the GT values.

Training Schemes

Scheme 1:

Designed for the less complex cases, it consists of one training phase with Adam optimizer and a learning rate of $lr = 5E-04$.

Scheme 2:

Designed for complex cases, it consists of three training phases: initial, main, and

final.

The initial phase computes the loss function only for the data agreement term, L_{data} . To achieve that, we assign the weights $[1, 0, 0, 0]$ to the 4 different loss terms in L - equation (3.3). This phase runs for 5-15k iterations using the Adam optimizer with a learning rate $lr = 5E-04$.

The main training phase utilizes all the aspects of the training loss function. It uses the Adam optimizer and requires 120k-150k iterations. Its objective is to locate the desired local optimum and approach it. Occasionally, the first 2 stages are enough for the EP-PINN to converge, but that is not consistent and far from guaranteed. For that reason, we designed a third stage to follow the main one [38].

The final phase trains the network with the L-BFGS optimizer until convergence [40]. It can run for as low as several dozen iterations and as many as several thousand (its upper limit is 15k iterations).

Essentially, the final stage proved advantageous as L-BFGS shortened the overall number of iterations needed for convergence, compared with the two-phase training process using only the Adam optimizer [38].

Evaluation

We evaluated the EP-PINNs' performance in reconstructing $V(x, t)$ with root mean square error (rMSE):

$$rMSE = \sqrt{\frac{1}{N} \sum_{n=1}^N (V(x_n, t_n) - \tilde{V}(x_n, t_n))^2} \quad (3.8)$$

Experiments were repeated five times to examine the variability in both rMSE and parameter estimations.

For parameter estimation experiments, we report the relative error (RE) between the parameter reference (one of a , b , and D) and its estimation:

$$RE(\lambda_{ref}, \lambda_{est}) = \frac{\lambda_{ref} - \lambda_{est}}{\lambda_{ref}} \quad (3.9)$$

To evaluate the EP-PINN's prediction of D with heterogeneous geometry, we computed the rMSE value between the GT values and the predicted values over the 2D rectangle:

$$rMSE(D) = \sqrt{\frac{1}{N} \sum_{n=1}^N (\tilde{D}(x_n) - D(x_n))^2} \quad (3.10)$$

Resources

The computing resources used to train our EP-PINN are the RTX6000 GPU and 4 AMD EPYC-7742 CPUs. Training time varied as the 1D experiments took approximately 15 minutes, the 2D several hours and heterogeneous 2D geometry trained for up to 16 hours. These were implemented and executed on Imperial College's High-Performance Computer (HPC).

The project's software design is detailed in Appendix A. The software archive and data files are available at the Github repository [alon01/PINNs-for-Cardiac-EP](https://github.com/alon01/PINNs-for-Cardiac-EP).

3.4 EP-PINNs for Monodomain Forward Mode

3.4.1 1D

We used EP-PINNs to solve the 1D forward mode. For the 1D monodomain Aliev Panfilov model, we assume homogeneous and isotropic conduction. That simplifies the PDE equation (2.2) as D becomes a scalar.

The 1D geometry is a 200 cells cable (equivalent to 20-mm) simulated for 70 time-steps (equivalent to 903 ms). Therefore, the overall GT 1D synthetic data size is $200 \times 70 = 14k$. Each data point carries a GT value for the membrane potential V . All the experiments in the section were trained with EP-PINN architecture A and training scheme 1 for 60k iterations. Training data consists of 10% of the GT data (1.4k points), 20k auxiliary domain points, and 1k auxiliary boundary points.

We designed experiments to assess the network's robustness and performance in solving the forward mode in several scenarios:

Extended GT Data

To establish the effect of the added GT values of W (the latent variable) on the network performance, we trained the network with GT data (left pacing) of both V and W . This case requires an additional MSE loss for W to be added to the L_{data} term.

$$L_{data} = \frac{1}{N} \sum_{n=1}^N \left(\tilde{V}(x_n, t_n) - V(x_n, t_n) \right)^2 + \frac{1}{N} \sum_{n=1}^N \left(\tilde{W}(x_n, t_n) - W(x_n, t_n) \right)^2 \quad (3.11)$$

No other experiment in the project contains GT data for W in the training dataset.

Excitation Direction

To observe the EP-PINN's capability to converge for different pacing directions, we trained the network with two distinct datasets. The first was generated with a left excitation wave and the second with a right one. The data generation scheme is described in chapter 3.1.1 - wave types 1 and 2. All 1D experiments use exclusively left excitation wave data except this one.

Electrical Signal

To test the impact of different electrical signals on the EP-PINNs' performance, we trained the EP-PINN with 3 distinct training datasets, each with one parameter change:

1. $a = 0.002$

2. $b = 0.075$

3. $D = 0.02$

Noise

To evaluate how GT data corrupted with noise affects the EP-PINN's performance, we added to the membrane potential V a zero-mean Gaussian noise with a standard deviation of 0.05, 0.1, 0.5, 1.0, or 10.0 AU (same units as V).

Training Data Size

To test how the training GT dataset size affects the EP-PINN's performance, we trained the network with $10k$, $5k$, $1k$, or 100 randomly sampled points and no auxiliary datasets (domain and boundary).

3.4.2 2D

We used EP-PINNs to solve the 2D forward mode. The 2D geometry is a 100×100 -cell (equivalent to 10×10 -mm) rectangle simulated for 70 time-steps (equivalent to 903 ms). Therefore, the overall GT synthetic data size is $100 \times 100 \times 70 = 700k$. The training dataset consists of 20% of the GT data (140k points), 40k auxiliary domain points, and 4k auxiliary boundary points.

We designed 5 experiments to analyze the EP-PINNs' performance in the 2D domain based on different 2D wave propagation: homogeneous corner wave, homogeneous vertical planar wave, homogeneous spiral wave, heterogeneous corner wave, and heterogeneous spiral wave. Each experiment utilized the corresponding GT synthetic data generated as described in chapter 3.1.1.

The five tasks presented varying degrees of complexity that challenged us to plan and execute different training strategies and EP-PINN architectures. Intuitively, the spiral wave is the most complex one and empirically required a deeper and wider EP-PINN. Therefore, corner wave and planar wave experiments are carried out with the same network architecture and learning setup while spiral wave with another:

- **Group 1:** homogeneous corner wave, heterogeneous corner wave, homogeneous vertical planar wave;
- **Group 2:** homogeneous spiral wave, heterogeneous spiral wave.

EP-PINNs for Group 1 utilized architecture A and training scheme 2 with 15k iterations in the first stage, 120k iterations in the main phase with a learning rate of $lr = 5\text{E-}04$.

EP-PINNs for Group 2 utilized architecture B and training scheme 2 with 15k iterations in the first stage, 150k iterations in the main phase with a learning rate of $lr = 1.0\text{E-}04$.

3.5 EP-PINNs for Parameter Estimation

In accordance with the project's main objective, we implemented EP-PINNs that solve the inverse monodomain Aliev Panfilov model to predict various EP parameters. Experiments detailed in this chapter utilize various training datasets (monodomain - 3.5.1, canine atrial - 3.5.2, experimental - 3.5.3) and different geometries (1D/2D, homogeneous/heterogeneous). Throughout all experiments, the loss function for the EP-PINNs is the one described in equation (3.3).

We seek to estimate the following EP parameters: a , which is related to tissue excitability; b , which controls the action potential duration (APD); and D , the diffusion coefficient associated with electrical conductivity. Every experiment aims to estimate one of the following parameter sets:

- $\{a\}$
- $\{b\}$
- $\{D\}$
- $\{a, D\}$
- $\{b, D\}$

We assume that parameters a and b are constant scalars with respect to space and time. Although D is constant through space and time in homogeneous geometry, it varies through space in heterogeneous geometry. These 3 parameters were selected for their known biophysical interpretation, their established susceptibility to disease remodeling, and the relatively simple mathematical relationship between them.

Each task is described by the parameter set to estimate and the data used for training. The EP-PINNs simultaneously optimize for the model solutions (V and W) and the required parameters. We report the computed rMSE (equation (3.8)) for the variable V and the RE for the parameter set. In heterogeneous geometry, we also report the rMSE for D (equation (3.10)).

3.5.1 Monodomain Model

All EP-PINNs in this section are trained with the monodomain Aliev Panfilov GT synthetic data, generated as described in 3.1.1. We tested each such EP-PINN in estimating, separately, all 5 parameter sets.

Homogeneous 1D Geometry

We adopted the same EP-PINNs setup as the forward 1D experiments (architecture, geometry, training process, data sizes), discussed in 3.4.1. We trained all EP-PINNs with the GT synthetic data (1D left wave).

To further assess its performance, we trained the network with datasets corrupted by noise. Two such GT training datasets were generated by adding to V a zero-mean Gaussian noise with a standard deviation of 0.05 or 0.1 AU (same units as V).

Homogeneous 2D Geometry

Three different training datasets were used for EP-PINNs with homogeneous 2D geometry: corner wave, vertical planar wave, and spiral wave. We used the same EP-PINNs' setup as the forward 2D respective scenarios, described in 3.4.2.

Heterogeneous 2D Geometry

EP-PINNs for heterogeneous 2D geometry use two different training datasets: corner wave and spiral wave. We used the modified EP-PINN architecture described in 3.2 and the same training scheme as the forward 2D corresponding scenarios.

3.5.2 Canine Atrial Model

We designed a more advanced task to test the designed EP-PINNs' ability to estimate EP parameters when training with ground truth data from the Canine Atrial model. These action potentials' shapes are markedly different from those of the Aliev Panfilov model, which was designed to model ventricular APs.

We generated the 1D canine Atrial data as described in chapter 3.1.1. The geometry setup is a 200 cells cable (equivalent to 20-mm) simulated for 70 time-steps (equivalent to 903 ms). Therefore, the overall GT 1D synthetic data size is $200 \times 70 = 14k$. All the experiments in the section were trained with EP-PINN architecture A and training scheme 1 for 60k iterations. Training data consists of 10% of the GT data (1.4k points), 20k auxiliary domain points, and 1k auxiliary boundary points.

We use two cell types: left atrial (LA) and crista-terminalis/Bachmann's bundle (BB) [13]. We also use several remodeling conditions: 0 - no remodeling, 1 - moderate remodeling, 2 - severe remodeling.

In Aliev Panfilov Model, the relationship between Action Potential Duration (APD) and the value of b is inverted [41]. Based on that, We set two objectives:

1. Identify the reduction in APD in LA cell type caused by increasing remodeling conditions by predicting larger b values to remodeled APs;
2. Identify the increase in APD in BB cell type compared with LA cell in non-remodeled conditions.

By switching between the remodeling conditions and the cell types in the Canine Atrial model, we acquire datasets representing different APDs. We do not know the equivalent b values corresponding to these APDs. Therefore, we essentially expect to achieve the correct relative b estimations (increase/decrease) between the various datasets rather than explicit values.

3.5.3 Experimental Data

We used EP-PINNs to estimate parameters with experimental data from two separate experiments. Each experiment is composed of two datasets, one for baseline

measures and the other for drug-stimulated results. The 1D geometry is a 2-cell cable recorded for different temporal lengths between the two experiments. The first, using Nifedipine drug, logged 145 time steps, and the second, using E-4031 drug, for 250 (detailed in chapter 3.1.2). Therefore, the overall experimental data size is $2 \times 145 = 290$ for the first experimental set (two datasets) and $2 \times 250 = 500$ for the second (two datasets as well). The training dataset utilized 90% of the data (261 and 450 points, respectively), 250 auxiliary domain points, and 200 auxiliary boundary points. We used the remaining 10% experimental data points for testing evaluations.

We designed two experiments:

1. Nifedipine: Identify the reduction in APD in drug-stimulated tissue as compared with the baseline (resting conditions) data by predicting larger b values;
2. E-4031: Identify the increase in APD in drug-stimulated tissue as compared with the baseline data by predicting smaller b values.

We utilized the EP-PINNs with architecture C and the 3-phase training scheme 2. The initial phase was carried with 5k iterations with a learning rate $lr = 5\text{E-}04$ and the main phase with 150k iterations with $lr = 1.0\text{E-}04$.

We tested each experiment 10 times to acquire a mean and a standard deviation for b estimations. We note the two approximated means b_{base} and b_{drug} for baseline and drug-stimulated datasets, respectively, and evaluate the results with a t-test to obtain a p-value:

$$H_0 : b_{base} = b_{drug}$$

$$H_1 : b_{base} \neq b_{drug}$$

Chapter 4

Results

Our experiments show that EP-PINNs can accurately solve the forward monodomain Aliev Panfilov model in 1D and 2D. The networks and methods designed are able to perform parameter estimation in diverse settings, including the Canine Atrial data and, more impressive, the experimental data.

4.1 EP-PINNs for Monodomain Forward Mode

4.1.1 1D

Using only 10% of the spatial-temporal GT data, the designed EP-PINNs can correctly solve the monodomain Aliev Panfilov model, precisely reproducing the associated AP features, morphology, and conduction properties. The trained model attained a minimal rMSE ($9.0\text{E-}03 \pm 4.0\text{E-}03$).

Table 4.1 details the rMSE values for the various scenarios tested. As expected, when adding the GT values of W for training, the error slightly decreases. Furthermore, the EP-PINN's performance was independent of the GT parameter values and the pacing direction with $\text{rMSE} \leq 1.0\text{E-}02$ in all cases.

Table 4.1: rMSE values for 1D Forward mode

Experiment	rMSE \pm std
Extended GT data	$6.0\text{E-}03 \pm 2.0\text{E-}03$
Right pacing	$7.0\text{E-}03 \pm 4.0\text{E-}03$
Change parameter a value	$6.0\text{E-}03 \pm 4.0\text{E-}03$
Change parameter b value	$1.0\text{E-}02 \pm 3.0\text{E-}03$
Change parameter D value	$6.0\text{E-}03 \pm 4.0\text{E-}03$

Figure 4.1 demonstrates the predicted solution next to the GT data. Figure 4.1 (a) shows the EP-PINN's predicted membrane potential, V (in arbitrary units - AU), for every spatial-temporal point via a heat map, whereas figure 4.1 (b) shows the GT values. While the two figures clearly show the model's capability of predicting the membrane potential's shapes and values throughout the spatial-temporal domain,

small boundary dissimilarities exist.

Figure 4.1 (c) displays a specific cell's action potential through time within the 1D cable ($x = 15mm$). The EP-PINN predicted the V values precisely, and they coincide with the GT values almost entirely. It's also visible to see the actual training GT data points used for this specific cell (x marks). Lastly, figure 4.1 (d) represents the cable's V values for an arbitrary time-step ($t = 35 AU$). It portrays a frozen-in-time cable's potentials across its complete spatial length.

The 4 figures demonstrate several perspectives of the EP-PINN's accuracy in predicting the AP propagation while training with a small dataset.

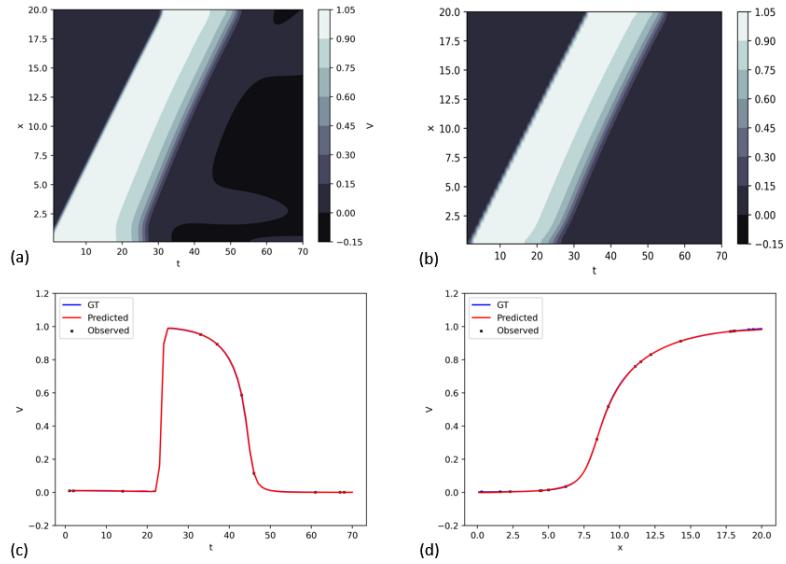


Figure 4.1: A comparison of the PINN solution for the Forward mode and the ground truth data. (a) the predicted membrane potential, V , heat map, (b) the GT V heat map, (c) a random cell's AP, (d) the cable's V values in a specific time step.

As expected, we found that increasing the noise factor leads to higher error rates - see Figure 4.2(a). The EP-PINN failed to converge when trained with data corrupted with noise with a standard deviation larger than 1 (the peak-to-peak amplitude of the AP).

Interestingly, we have additionally found that EP-PINNs can accurately solve the forward mode using even small numbers of GT V points for training (and no auxiliary points). rMSE values are as low as 7.4E-03 with only 1000 data points and rising to 2.1E-03 with merely 100 GT values - see Figure 4.2 (b).

4.1.2 2D

The designed EP-PINNs solved the forward mode with great accuracy (see Table 4.2). EP-PINNs trained with group 1 datasets (homogeneous corner, homogeneous planar, and heterogeneous corner) were able to reproduce AP propagation with excellent precision, with the highest mean rMSE values at 7.E-03 (in corner wave).

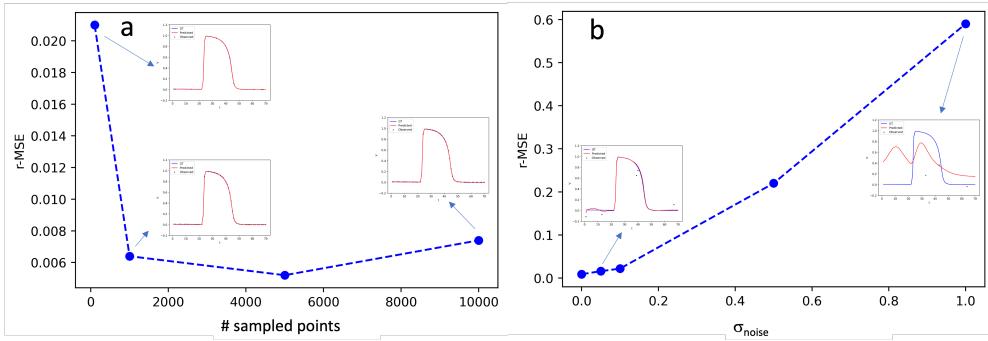


Figure 4.2: The impact of training data size (a) and noise (b) on the rMSE values of the EP-PINNs' V predictions. The corresponding cell's AP estimations are added to the diagrams.

EP-PINNs trained with the more complex group 2 datasets (homogeneous and heterogeneous spiral wave) achieved great accuracy, although rMSE values were higher than other tasks. Figure 4.3 shows PINN's wave predictions. Several animations were also generated to show the network's impressive estimations (located at the Github repository).

Table 4.2: rMSE values for 2D Forward mode

Experiment	rMSE \pm std
Homogeneous corner wave	7.7E-03 \pm 7.7E-05
Heterogeneous corner wave	6.2E-03 \pm 1.2E-03
Homogeneous planar wave	9.0E-04 \pm 6.0E-04
Homogeneous spiral wave	2.5E-02 \pm 2.3E-03
Heterogeneous spiral wave	4.4E-02 \pm 4.3E-03

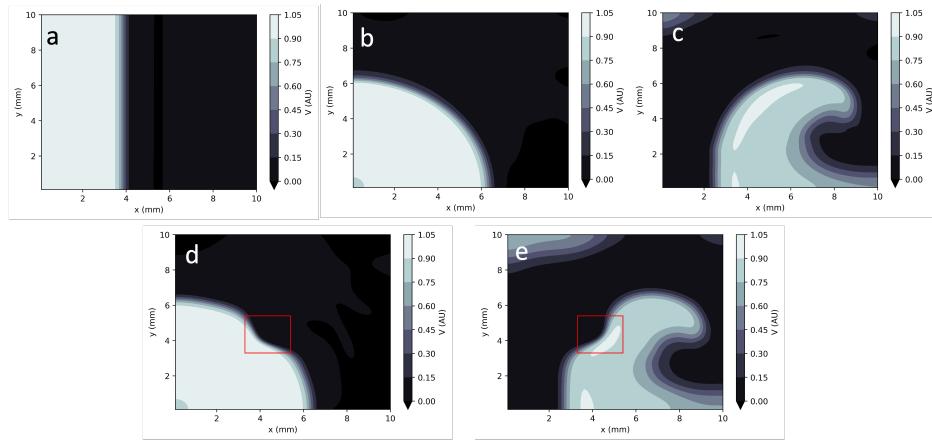


Figure 4.3: EP-PINNs solutions to the Aliev Panfilov model in 2D forward mode. The prediction of V for a cell frozen-in-time. (a) Planar wave; (b) Corner wave; (c) Spiral wave; (d) Heterogeneous corner wave; (e) Heterogeneous spiral wave.

4.2 EP-PINNs for Parameter Estimation

4.2.1 Monodomain Model

Homogeneous 1D Geometry

EP-PINNs estimated the parameter sets with mixed accuracy rates as seen in Figure 4.4. The parameter b was estimated with the lowest absolute relative error (RE) of 1.3E-02. Surprisingly, adding noise to the training data had barely impacted its scores with $|RE| = 4.0E-03$ and $|RE| = 2.4E-02$ for added noise $\sigma = 0.05$ and $\sigma = 0.1$, respectively. When jointly approximated with D , it showed the same performance pattern with similar minimal RE rates (For full results table, see Appendix B).

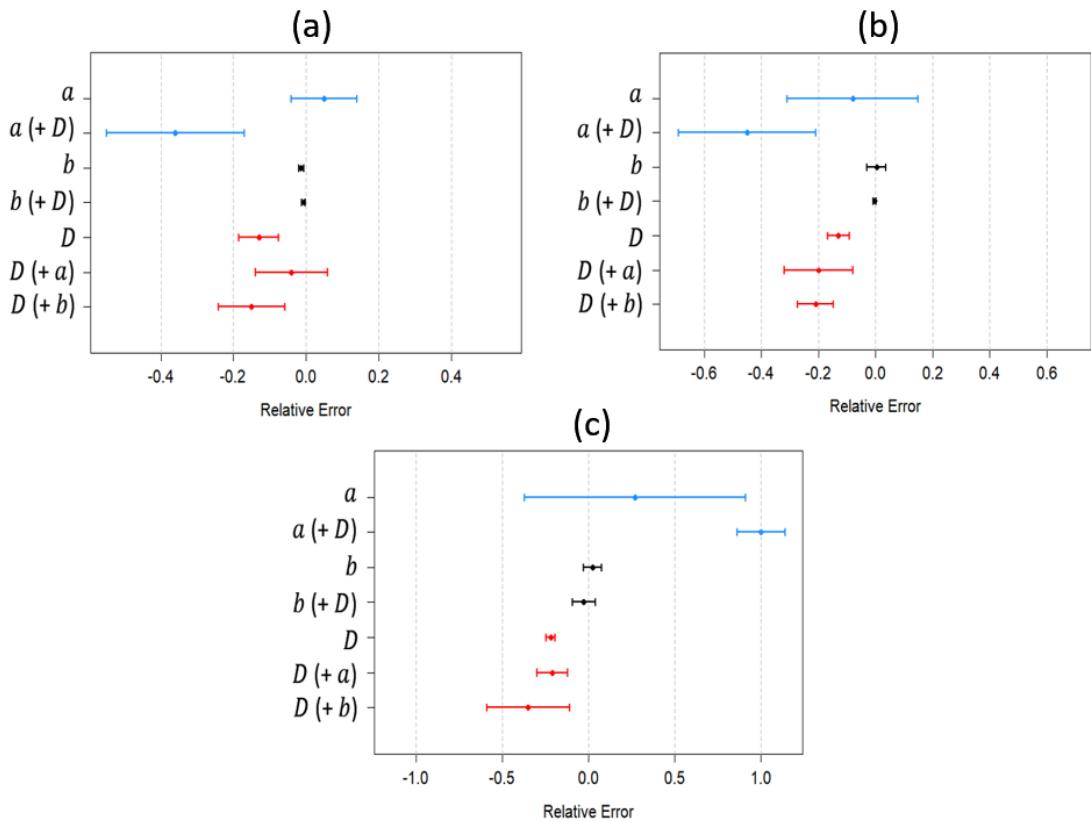


Figure 4.4: Relative error rates (mean \pm std) for the 1D parameter estimation tasks trained with: a) GT data; b) GT data + moderate noise ($\sigma = 0.05$); c) GT data + severe noise ($\sigma = 0.1$). The $(+P)$ sign notes the coupled parameter P estimated alongside the relevant parameter.

The predictions for a were not consistent and were severely affected by noise. EP-PINN accurately predicted a with $|RE| = 3.6E-02$ for GT data and $|RE| = 8.0E-02$ for the corrupted data ($\sigma = 0.05$). When either trained with a greater noise factor ($\sigma = 0.01$) or learned simultaneously with D , the network was unable to produce a meaningful prediction. In these cases, its $|RE|$ rates ranged from 2.7E-01 at best to 1.0E-00 at worst.

D was best estimated when coupled with a as its $|RE|$ score was 4.0E-02 for that case. Other scenarios (solely D , joint set $b + D$, and added noise) resulted in $|RE|$ rates in the range [1.3E-01 - 3.5E-01].

When trained with the GT 1D datasets, EP-PINNs have achieved great accuracy in solving the Aliev Panfilov equations with rMSE for membrane potential, V , values less than 9.0E-03. With added noise, the precision slightly decreased as the EP-PINNs achieved a maximum rMSE of 2.5E-02.

Homogeneous 2D Geometry

Theoretically, vertical planar waves carry the simplest propagation patterns, and spiral waves have the most complex. Therefore, it is no surprise to observe the predictions' accuracy decrease for the spiral wave data (For full result table, see Appendix C).

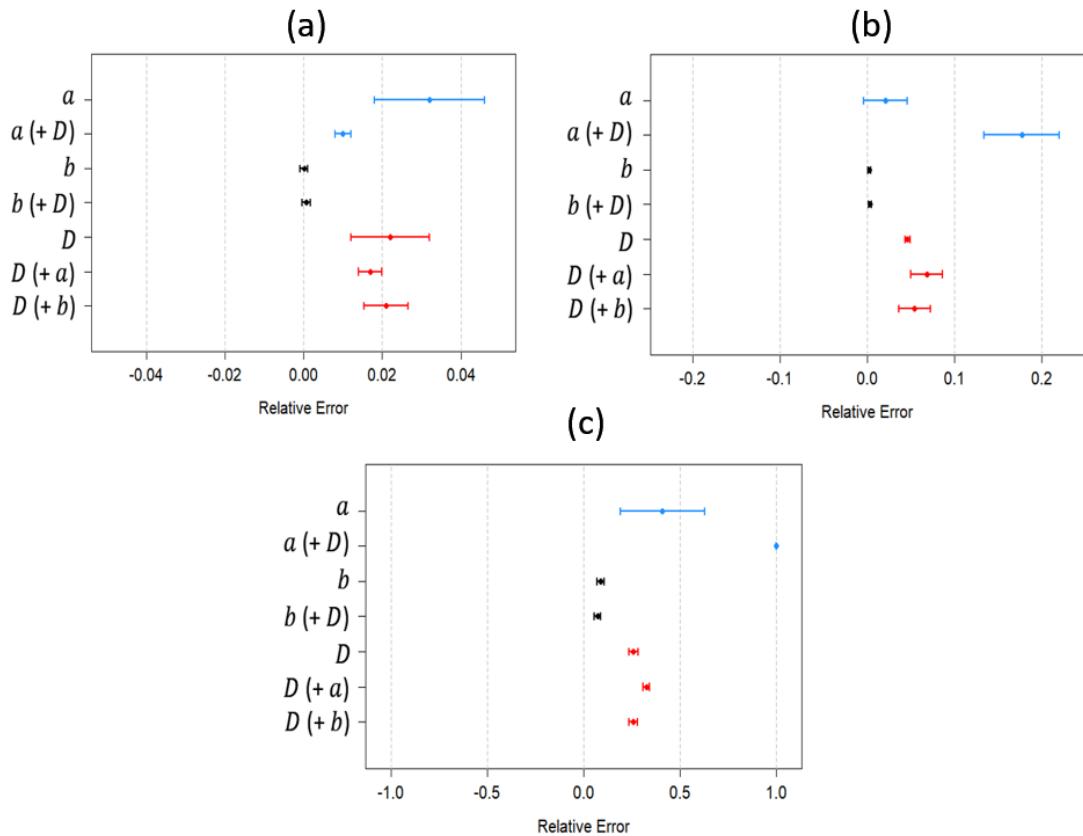


Figure 4.5: Relative error rates (mean \pm std) for the 2D parameter estimation tasks trained with the homogeneous GT datasets: a) Vertical planar wave; b) Corner wave; c) Spiral wave. The $(+P)$ sign notes the coupled parameter P estimated alongside the relevant parameter.

As Figure 4.5 shows, EP-PINNs for homogeneous 2D cases produce similar prediction trends as in the 1D ones, prominently the excellent accuracy for the estimations of b . Its mean $|RE|$ rates vary from 1.9E-04 in the corner wave to 8.6E-02 in the spiral wave dataset. There were no significant changes to the error rates when it was

jointly approximated with D .

Mean $|RE|$ values for a , when estimated individually, ranged between 1.6E-02 and 4.3E-01. As expected, the upper bound is for the spiral wave dataset. For parameter D , results were noticeably similar with a range of 1.8E-02 and 2.6E-01. Again, spiral wave propagation logged the highest error rate.

In cases where the EP-PINN estimated a couple of parameters simultaneously, spiral wave data proved the toughest task for the designed network. a and D rated relative errors of 1.0E-00 and 3.2E-01 respectively on that task, whereas when trained with the vertical planar wave data, they scored 9.2E-03 and 1.7E-02, respectively. For the coupled b and D case, results for D were almost identical to its RE scores when estimated individually.

The precision of the solution of the monodomain Aliev Panfilov equations with homogeneous 2D data expresses similar patterns as the forward mode. rMSE values for predicting V were lowest when training the EP-PINN with the planar wave dataset. Across the various parameter sets, its worst rMSE score was 1.6E-03. For corner and spiral waves, rMSE peaked at 3.9E-03 and 2.7E-02, respectively.

Heterogeneous 2D Geometry

As described in chapter 3.5.1, we use two metrics for heterogeneous 2D tasks: RE for the parameters a and b , and $rMSE(D)$ for parameter D (equation (3.10)). When assigned the task, EP-PINN predicts a value $\tilde{D}(x, y)$ for each coordinate on the rectangle. Therefore, to estimate D 's overall error for the entire rectangle, we use the rMSE metric - see Figure 4.6.

Table 4.3: Mean error rates for Parameter estimation in heterogeneous 2D geometry.

Parameter Set	Corner		Spiral	
	RE	rMSE	RE	rMSE
a	6.1E-02	-	2.9E-01	-
b	1.4E-03	-	8.3E-02	-
D	-	1.4E-02	-	2.8E-02
$a + D$	a	1.0E-01	-	1.0E-00
	D	-	1.4E-02	-
$b + D$	b	3.0E-03	-	8.2E-02
	D	-	1.4E-02	-

Maintaining the previously acknowledged patterns, estimating b was carried with an exceptional low rates: 1.4E-03 and 2.8E-02 for corner and spiral waves, respectively. Even when jointly approximated with D , accuracy levels held firm.

EP-PINN recorded mixed RE rates for predictions of a , 6.1E-02 when trained with the corner dataset, and 2.9E-01 with the spiral set. Moreover, when coupled with D , the estimation's RE jumped to 1.0E-00 (utilizing the spiral data).

rMSE scores for D estimations kept steady regardless of the targeted parameter set (solely D or coupled with either a or b). The EP-PINN achieved 1.4E-02 and 3.0E-02 when trained with the corner and spiral datasets, respectively.

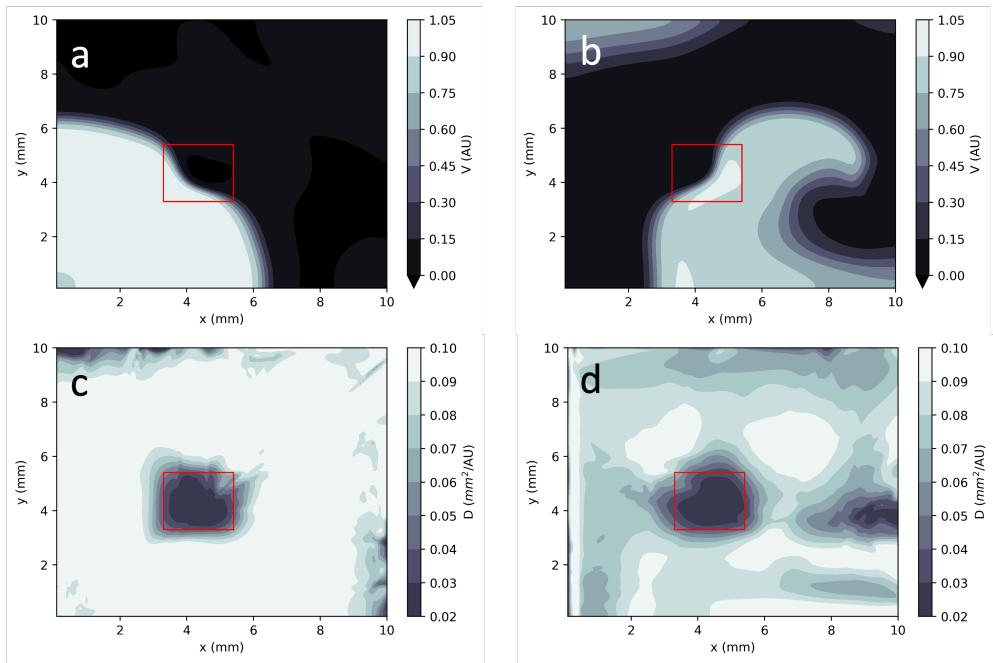


Figure 4.6: EP-PINNs' solution for heterogeneous geometry in the parameter estimation tasks. Estimation of the cell's potentials via heat maps for: a) corner wave, and b) spiral wave. Additionally, c) and d) are the corresponding maps for $\tilde{D}(\vec{x})$ approximations. The successfully estimated heterogeneity is visible in both maps.

4.2.2 Canine Atrial Model

EP-PINNs achieved impressive results in capturing the APD reduction in the left atrial (LA) cell type with increased remodeling conditions (see TABLE 4.4). The three mean predictions for b increased as the remodeling conditions intensified, depicting shorter APDs. Predicted values started at 3.31E-02 (no remodeling), raised to 4.2E-02 (moderate remodeling), and peaked at 8.0E-02 (severe remodeling). In addition, there is a significant gap between the three estimated b ranges (and no overlap), which supports the assertion of the EP-PINN's ability to distinguish between the 3 signals' APD.

Table 4.4: Parameter b estimation under AF remodeling in LA cell type

Remodeling	Mean	Range	rMSE
None	3.31E-02	[3.26 – 3.36]E-02	2.9E-02
Moderate	4.2E-02	[4.15 – 4.33]E-02	4.0E-02
Severe	8.0E-02	[7.83 – 8.38]E-02	4.2E-02

Additionally, even though the EP-PINN was trained with the Canine Atrial model GT data, it reproduced the AP propagation mostly accurately with rMSE rates capped at 4.2E-02. Figure 4.7 displays the estimated cell's potential with time, at the spatial coordinate $x = 15\text{mm}$, for each remodeling factor. Although some small discrepancies exist between the model prediction and the GT data, the EP-PINN correctly captures the shapes and properties of the signals.

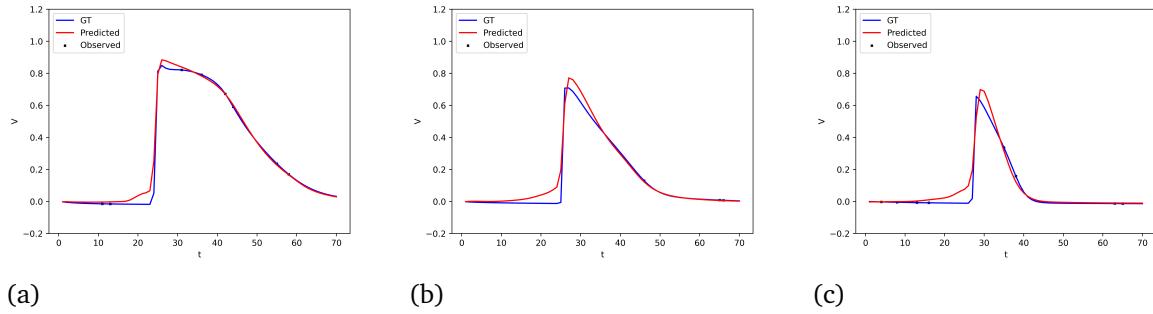


Figure 4.7: Canine Atrial model cell plots (LA type) for different remodeling conditions: (a) none; (b) moderate; (c) severe.

In the second task, we expected the PINN to identify that under no remodeling conditions, APD is longer in BB cell type than in LA cell type. The results do not show that relationship, the mean estimation of b for the BB cell type was $3.51\text{E-}02$ while the mean approximation for the LA cell was $3.31\text{E-}02$. Moreover, the prediction range for BB was $[3.45 - 3.57]\text{E-}02$ and $[3.26 - 3.36]\text{E-}02$ for LA. Opposed to the objective, the EP-PINN has predicted a decrease in APD between LA cells and BB cells, although the differences in the estimated b are fairly minimal, $2.0\text{E-}03$.

4.2.3 Experimental Data

The designed EP-PINNs, trained with the experimental optical mapping datasets, were able to identify the effect of both drugs on the action potentials duration relative to the baseline (see Figure 4.8). In the first task, EP-PINNs detected the decrease in APD caused by Nifedipine (as an increase of b estimation). The relative increase was $\Delta b = (9.7 \pm 5.7)\text{E-}02$, and the result is statistically significant with p -value of $p < 1.0\text{E-}16$. EP-PINNs were also able to detect the rise in APD caused by E-4031. The estimated reduction in b was not statistically significant with a p -value of $p = 0.38$. The relative change in b estimations was $\Delta b = (-2.1 \pm 3.0)\text{E-}02$.

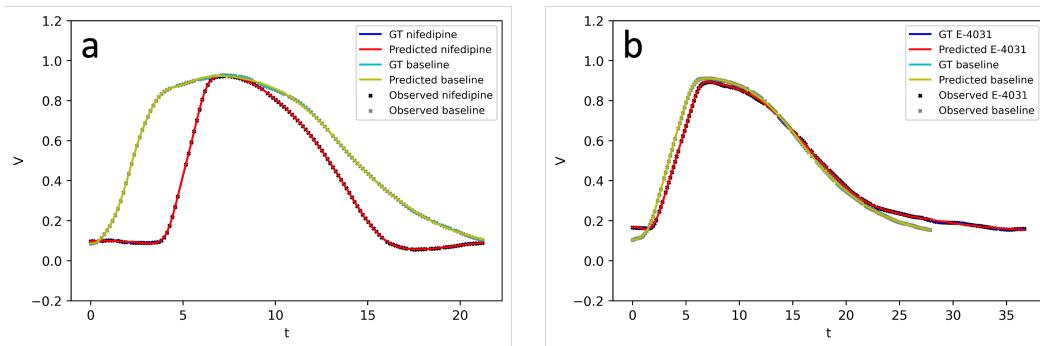


Figure 4.8: EP-PINNs solution for the experimental optical mapping data. Figures show the APs as estimated by EP-PINNs and the training data for the two anti-arrhythmic drug experiments: (a) Nifedipine; (b) E-4031.

Chapter 5

Discussion

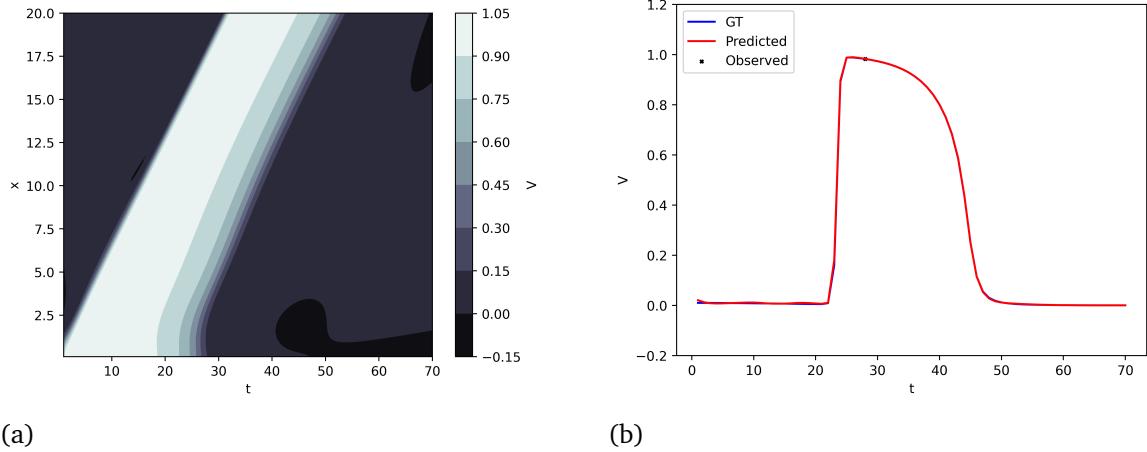
We demonstrated EP-PINNs' ability to reproduce action potential propagation and perform parameter inference in both 1D and 2D. By utilizing sparse measurements of membrane potential (V), EP-PINNs successfully estimated cardiac EP parameters such as action potential duration (b), excitation threshold (a), and conductivity (D). Additionally, we demonstrated EP-PINNs' potential in clinical applications by accurately identifying heterogeneities, depicted with spatially reduced D , in arrhythmic conditions.

5.1 Forward Solution

Our project builds on the works of Sahli-Costanal et al [32] and Grandits et al [33], who established EP-PINNs' promising capacity to approximate high-resolution left atrial activation time and conduction velocity maps using cardiac EP models (both forward and inverse modes). We developed EP-PINNs on a more complex EP model, the monodomain Aliev Panfilov [12]. This model is able to additionally capture restitution properties of the cardiac tissue and represent the AP's shape. Emulating arrhythmic conditions, such as spiral waves, is enabled by including the latent variable, W . In clinical cardiac experiments, data for W is typically unavailable. Thus, the EP-PINNs' excellent performance in reproducing APs and parameter inference without using any data for W encourages the development of EP-PINNs for even more complicated EP models with more latent variables. The study of Yazdani et al [27], which deployed PINNs for several biological systems based on sets of coupled ODEs, may support that assertion.

We designed several stress tests in 1D geometry to analyze the EP-PINNs' performance with a small-sized dataset and another dataset corrupted by noise. In both scenarios, the EP-PINNs have successfully reproduced the action potential propagation. Figure 5.1 displays the EP-PINN's prediction heat map and the cell plot when trained with only 100 ground truth data points for V . Similarly, Figure 5.2 shows the two exact figures but in the case of dataset corrupted with a large noise factor $\sigma = 0.1$. These results show the robustness of our EP-PINNs to noise and sparse data that was achieved by explicitly enforcing the system's biophysical equations in the

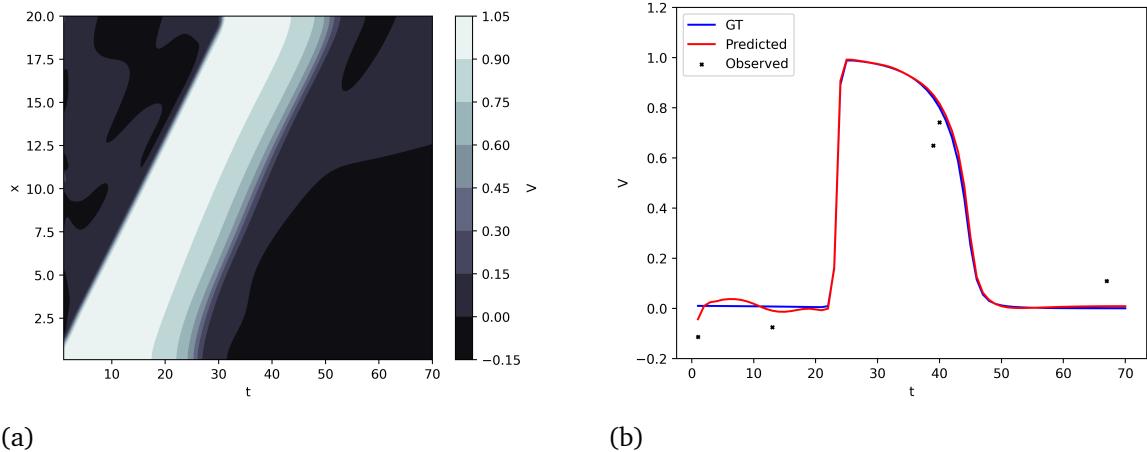
loss function. Essentially, these added expressions function as regularizers.



(a)

(b)

Figure 5.1: The EP-PINN’s predictions when trained with only a 100 ground truth data points: (a) a heat map, (b) a cell plot.



(a)

(b)

Figure 5.2: The EP-PINN’s predictions when trained with data corrupted with $\sigma = 0.1$ noise factor: (a) a heat map, (b) a cell plot.

Solving the forward mode in 2D geometry proved more difficult and required a more complex training scheme than the 1D. Furthermore, to reproduce the most convoluted 2D dataset, the spiral wave (with either homogeneous or heterogeneous diffusion), we utilized a larger feed-forward fully connected NN. These additions to the network’s capacity and the training phases were necessary for robustness and convergence of the 2D EP-PINNs. Clearly, we could have set one unified setup for the project with the more complex architecture and training method. It would enable solving the 1D tasks as well, but at the cost of redundant computational resources.

5.2 Parameter Inference

For parameter inference tasks, the main advantage PINNs present is dealing with sparse and small-sized datasets. The EP-PINNs designed for the project correctly estimated the parameters with only 10% and 20% of the geometry domain in 1D and 2D, respectively. The parameters we established to approximate throughout the project are closely related to observable cardiac electrophysiology properties: action potential duration (parameter b), excitation threshold (parameter a), and conduction velocity (parameter D). Training with the generated synthetic GT V values, the EP-PINNs were able to precisely estimate b . Impressively, regardless of the domain geometry (1D or 2D), noisy data, and heterogeneities in the diffusion tensor, prediction errors ($|RE|$ values) for b have not exceeded 8.6E–02.

However, predicting a proved to be more challenging in several scenarios. In the presence of significant noise and when approximated jointly with D , estimations were essentially unreliable. One explanation for these results may relate to the magnitude of a . In such cases (e.g., two parameters to estimate), the loss function's rMSE magnitude was similar to the value of the simulated $a = 1.0\text{E}-02$. The additive form with which a appears in the model equation (2.2) may obstruct its optimization with such a small value. Training the EP-PINNs with datasets generated with a larger a would resolve that supposition.

The results for a may also be associated with the weak dependency between the GT values of $V(x, t)$ and the supra-threshold in certain electrical wave forms. The spiral waves epitomize a unique case as their morphologies strongly depend on the excitation threshold and other EP parameters. Another significant attribute of spiral waves is the inherent dependency within the model parameters. That emphasizes two challenges: designing a single EP-PINN to handle many forms of electrical wave propagation and estimating two model parameters simultaneously. We suggest a different EP-PINN architecture and training plan to confront these issues by simultaneously utilizing several datasets. The datasets are acquired from the same system but with various wave structures. The idea is to Connect several NNs in parallel, each trained with a different type of dataset and contributes to the PINN's loss function.

Relative error scores for D estimations were not as precise as the ones for b , but in all of the tasks, rates did not exceed 3.5E–01. Moreover, EP-PINNs accurately recognized heterogeneities in D within the 2D domain. As presented before, identifying heterogeneities in EP parameters (e.g., fibrosis) may be essential in atrial fibrillation treatment as these may become prime locations for ablation [8]. In the monodomain Aliev Panfilov model, regions of fibrosis were modeled with a decreased diffusion coefficient, D . We have demonstrated that EP-PINNs can aid in personalizing AF ablation strategies as they are capable of reproducing these maps and identify heterogeneities.

Canine Atrial Model

We demonstrated the ability of our EP-PINNs to infer parameters for data generated from a different, more complex cardiac EP model, the Canine Atrial model

[13]. Even though we constructed the EP-PINN’s loss function to fit the data to the Aliev Panfilov model, it was able to identify the reduction in action potential duration (APD) by predicting an increase in b . Originally, the decrease in APD was achieved by increasing the AF-induced remodeling rates in the left atria (LA) cells. EP-PINN wrongly estimated a small (6%) increase in APD between LA cells and crista-terminalis/Bachmann’s bundle (BB) cells, whereas experiments show a subtle decrease in APD [13]. This result may stem from uncalibrated system EP parameters. In this task, the EP-PINNs estimate only b and set the other parameters as constants. We might need to approximate more model parameters in order to observe the correct APD relationship between the two cell types (LA and BB).

Experimental Optical Mapping Data

Despite the inevitable existence of noise and artifacts in the experimental optical mapping datasets, EP-PINNs were able to identify the impact of the anti-arrhythmic drugs (Nifedipine and E-4031) on the action potential duration (relative to its baseline). Utilizing experimental data to perform parameter estimation provides a valuable perspective of EP-PINNs’ performance. Most of the previous studies assess PINNs solely with synthetically generated datasets [24, 27, 32, 38], whereas we have complemented the *in silico* tasks with the experimental data tests.

The two chosen experiments mainly affect the cells’ APD, recognized as a change in b in the Aliev Panfilov model. Future works can utilize other experimental drugs that have an extensive impact on various cardiac electrophysiology properties. EP-PINNs are capable to simultaneously infer several model parameters, as demonstrated on the synthetic datasets, and therefore may support the characterization of these anti-arrhythmic drugs. In addition, more complex and expressive EP models may suit such tasks as they possess a wider characterization potential.

5.3 Limitations and Future Modifications

Clinically acquired data would naturally be noisy, and our analysis shows that EP-PINNs are relatively robust to moderate noise, both in reconstructing AP propagation and parameter estimation. In order to improve the designed network’s predictions under large noise factors, we suggest a weighted loss function. Currently, the various loss expressions are evenly weighted. Reducing the weight of the data agreement loss term (L_{data}) and increasing the physical laws term’s weight (L_f) may enhance the network’s robustness to noise. Further investigation is required to determine the optimized relative loss weights.

The current EP-PINNs’ training data is the membrane potentials, $V(x, t)$, that are not usually measured in the clinical environment. During clinical procedures, extracellular electrical potentials, $\phi_e(t)$, are typically measured using contact electrodes (EGM) and can be utilized to adjust the proposed EP-PINNs. Mathematically, in theory [42], the values of $\phi_e(t)$ are an integration of the weighted values of the term $\vec{\nabla} D \vec{\nabla} V$ from the monodomain PDE (equation (2.2)). $\phi_e(t)$ is the measured electrical

potential in a specific electrode e (in arbitrary units) and a specific time. The complete dataset is composed of the electrical potentials evaluated in several spatially spread electrodes for a series of temporal steps. To utilize that new data format, we suggest designing a modified EP-PINNs, with the addition of ϕ_e as inputs and a convolutional NN instead of the feed-forward one. These adjustments may suit the new approach, as the equations characterize an integrated weighted relationship, similar to convolution/cross-correlation.

Several other aspects of the current project should be modified in order to determine the full potential of EP-PINNs to serve future clinical applications. The Aliev Panfilov model is relatively simple, with only 6 parameters, and a more complex EP model will be required. Furthermore, we approximated sets of 1 or 2 parameters, while to sufficiently characterize the cardiac activity, we should simultaneously estimate a broader set of model properties. Since the designed EP-PINNs used the simple 1D and 2D domains, future works would need to construct 3D (atrial-like) geometries to generalize our limited representation of the heart. Increasing the problem's dimension would require NNs with greater capacity and consequently a significant added burden on computational resources. A possible plan to alleviate these would be to acquire datasets of electrogram (EGM) signals instead of the membrane potentials.

We suggest a thorough development and investigation of the uncertainty measures for the estimated model parameters. The simple tools used in this project, such as mean, standard deviation, and range, provide little insight on sensitivity to noise, data size, and other conditions. Lastly, the current work did not evaluate EP-PINNs with spiral wave break-up data in the chaotic or pseudo-chaotic scenarios [43], which occur in some arrhythmias. These cases are excessively complex and clinical data able to characterize them are not typically acquired.

References

- [1] Rui Shi, Mark Norman, Zhong Chen, and Tom Wong. Individualized ablation strategy guided by live simultaneous global mapping to treat persistent atrial fibrillation. *Future cardiology*, 14(03):237–249, 2018. pages 1
- [2] Faisal Rahman, Gene F Kwan, and Emelia J Benjamin. Global epidemiology of atrial fibrillation. *Nature Reviews Cardiology*, 11(11):639–654, 2014. pages 1
- [3] Sumeet S Chugh, Rasmus Havmoeller, Kumar Narayanan, David Singh, Michiel Rienstra, Emelia J Benjamin, Richard F Gillum, Young-Hoon Kim, John H McAnulty Jr, Zhi-Jie Zheng, et al. Worldwide epidemiology of atrial fibrillation: a global burden of disease 2010 study. *Circulation*, 129(8):837–847, 2014. pages 1
- [4] Mayo Clinic. atrial fibrillation. URL <https://www.mayoclinic.org/diseases-conditions/atrial-fibrillation/symptoms-causes/syc-20350624>. pages 1
- [5] Stephan Willems, Atul Verma, Timothy R Betts, Steven Murray, Petr Neuzil, Huseyin Ince, Daniel Steven, Arian Sultan, Patrick M Heck, Mark C Hall, et al. Targeting nonpulmonary vein sources in persistent atrial fibrillation identified by noncontact charge density mapping: Uncover af trial. *Circulation: Arrhythmia and Electrophysiology*, 12(7):e007233, 2019. pages 2
- [6] Atul Verma, Chen-yang Jiang, Timothy R Betts, Jian Chen, Isabel Deisenhofer, Roberto Mantovan, Laurent Macle, Carlos A Morillo, Wilhelm Haverkamp, Rukshen Weerasooriya, et al. Approaches to catheter ablation for persistent atrial fibrillation. *New England Journal of Medicine*, 372(19):1812–1822, 2015. pages 2
- [7] Gerhard Hindricks, Tatjana Potpara, Nikolaos Dagres, Elena Arbelo, Jeroen J Bax, Carina Blomström-Lundqvist, Giuseppe Boriani, Manuel Castella, Gheorghe-Andrei Dan, Polychronis E Dilaveris, et al. 2020 esc guidelines for the diagnosis and management of atrial fibrillation developed in collaboration with the european association for cardio-thoracic surgery (eacts) the task force for the diagnosis and management of atrial fibrillation of the european society of cardiology (esc) developed with the special contribution of the european heart rhythm association (ehra) of the esc. *European heart journal*, 42(5):373–498, 2021. pages 2

- [8] A Roy, M Varela, and Oleg Aslanidi. Image-Based Computational Evaluation of the Effects of Atrial Wall Thickness and Fibrosis on Re-entrant Drivers for Atrial Fibrillation. *Frontiers in Physiology*, 9:1352, 10 2018. ISSN 1664-042X. doi: 10.3389/fphys.2018.01352. URL <https://www.frontiersin.org/article/10.3389/fphys.2018.01352/full>. pages 2, 6, 7, 35
- [9] RH Clayton, Olivier Bernus, EM Cherry, Hans Dierckx, Flavio H Fenton, Lucia Mirabella, Alexander V Panfilov, Frank B Sachse, Gunnar Seemann, and H Zhang. Models of cardiac tissue electrophysiology: progress, challenges and open questions. *Progress in biophysics and molecular biology*, 104(1-3):22–48, 2011. pages 4, 5, 6
- [10] Greg Ikonnikov and Dominique Yelle. Physiology of cardiac conduction and contractility. URL <http://www.pathophys.org/physiology-of-cardiac-conduction-and-contractility/>. pages 4
- [11] cellML. Electrophysiology. URL <http://models.cellml.org/electrophysiology>. pages 5
- [12] Rubin R Aliev and Alexander V Panfilov. A simple two-variable model of cardiac excitation. *Chaos, Solitons & Fractals*, 7(3):293–301, 1996. pages 5, 11, 33
- [13] Marta Varela, Michael A Colman, Jules C Hancock, and Oleg V Aslanidi. Atrial heterogeneity generates re-entrant substrate during atrial fibrillation and anti-arrhythmic drug action: mechanistic insights from canine atrial models. *PLoS computational biology*, 12(12):e1005245, 2016. pages 7, 13, 23, 36
- [14] Steven A Niederer, Joost Lumens, and Natalia A Trayanova. Computational models in cardiology. *Nature Reviews Cardiology*, 16(2):100–111, 2019. pages 7
- [15] Paris Perdikaris and George Em Karniadakis. Model inversion via multi-fidelity bayesian optimization: a new paradigm for parameter estimation in haemodynamics, and beyond. *Journal of The Royal Society Interface*, 13(118):20151107, 2016. pages 7
- [16] Richard C Aster, Brian Borchers, and Clifford H Thurber. *Parameter estimation and inverse problems*. Elsevier, 2018. pages 7
- [17] Jim O Ramsay, Giles Hooker, David Campbell, and Jiguo Cao. Parameter estimation for differential equations: a generalized smoothing approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 69(5):741–796, 2007. pages 7
- [18] Socrates Dokos and Nigel H Lovell. Parameter estimation in cardiac ionic models. *Progress in biophysics and molecular biology*, 85(2-3):407–431, 2004. pages 7
- [19] Go W Beeler and H Reuter. Reconstruction of the action potential of ventricular myocardial fibres. *The Journal of physiology*, 268(1):177–210, 1977. pages 7

- [20] Sam Coveney, Cesare Corrado, Jeremy E Oakley, Richard D Wilkinson, Steven A Niederer, and Richard H Clayton. Bayesian calibration of electrophysiology models using restitution curve emulators. *Frontiers in Physiology*, page 1120, 2021. pages 7
- [21] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6):84–90, 2017. pages 8
- [22] Tom Young, Devamanyu Hazarika, Soujanya Poria, and Erik Cambria. Recent trends in deep learning based natural language processing. *ieee Computational intelligenCe magazine*, 13(3):55–75, 2018. pages
- [23] Babak Alipanahi, Andrew Delong, Matthew T Weirauch, and Brendan J Frey. Predicting the sequence specificities of dna-and rna-binding proteins by deep learning. *Nature biotechnology*, 33(8):831–838, 2015. pages 8
- [24] Maziar Raissi, Paris Perdikaris, and George E Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, 2019. pages 8, 36
- [25] Eitan Tadmor. A review of numerical methods for nonlinear partial differential equations. *Bulletin of the American Mathematical Society*, 49(4):507–554, 2012. pages 8
- [26] Atilim Gunes Baydin, Barak A. Pearlmutter, Alexey Andreyevich Radul, and Jeffrey Mark Siskind. Automatic Differentiation in Machine Learning: a Survey. *Journal of Machine Learning Research*, 18(153):1–43, 2018. ISSN 1533-7928. URL <http://jmlr.org/papers/v18/17-468.html>. pages 8
- [27] Alireza Yazdani, Lu Lu, Maziar Raissi, and George Em Karniadakis. Systems biology informed deep learning for inferring parameters and hidden dynamics. *PLoS computational biology*, 16(11):e1007575, 2020. pages 10, 33, 36
- [28] Yeonjong Shin, Jerome Darbon, and George Em Karniadakis. On the convergence of physics informed neural networks for linear second-order elliptic and parabolic type pdes. *arXiv preprint arXiv:2004.01806*, 2020. pages 10
- [29] Daniele Ceccarelli. Bayesian physics-informed neural networks for inverse uncertainty quantification problems in cardiac electrophysiology. 2021. pages 10
- [30] Rudolf LM van Herten, Amedeo Chiribiri, Marcel Breeuwer, Mitko Veta, and Cian M Scannell. Physics-informed neural networks for myocardial perfusion mri quantification. *arXiv preprint arXiv:2011.12844*, 2020. pages 11
- [31] Shengze Cai, He Li, Fuyin Zheng, Fang Kong, Ming Dao, George Em Karniadakis, and Subra Suresh. Artificial intelligence velocimetry and

- microaneurysm-on-a-chip for three-dimensional analysis of blood flow in physiology and disease. *Proceedings of the National Academy of Sciences*, 118(13), 2021. pages 11
- [32] Francisco Sahli Costabal, Yibo Yang, Paris Perdikaris, Daniel E Hurtado, and Ellen Kuhl. Physics-informed neural networks for cardiac activation mapping. *Frontiers in Physics*, 8:42, 2020. pages 11, 33, 36
- [33] Thomas Grandits, Simone Pezzuto, Francisco Sahli Costabal, Paris Perdikaris, Thomas Pock, Gernot Plank, and Rolf Krause. Learning atrial fiber orientations and conductivity tensors from intracardiac maps using physics-informed neural networks. In *International Conference on Functional Imaging and Modeling of the Heart*, pages 650–658. Springer, 2021. pages 11, 33
- [34] John Charles Butcher. A history of runge-kutta methods. *Applied numerical mathematics*, 20(3):247–260, 1996. pages 12
- [35] Catalina Tobón, Laura C Palacio, Juan E Duque, Esteban A Cardona, Juan P Ugarte, Andrés Orozco-Duque, Miguel A Becerra, Javier Saiz, and John Bustamante. Simple ablation guided by a pen mapping in a 2d model during permanent atrial fibrillation. In *Computing in Cardiology 2014*, pages 1029–1032. IEEE, 2014. pages 12
- [36] Timothy Bui. Explicit and implicit methods in solving differential equations. 2010. pages 13
- [37] Rasheda A Chowdhury, Konstantinos N Tzortzis, Emmanuel Dupont, Shaun Selvadurai, Filippo Perbellini, Chris D Cantwell, Fu Siong Ng, Andre R Simon, Cesare M Terracciano, and Nicholas S Peters. Concurrent micro-to macro-cardiac electrophysiology in myocyte cultures and human heart slices. *Scientific reports*, 8(1):1–13, 2018. pages 13
- [38] Lu Lu, Xuhui Meng, Zhiping Mao, and George Em Karniadakis. Deepxde: A deep learning library for solving differential equations. *SIAM Review*, 63(1): 208–228, 2021. pages 14, 19, 36
- [39] Xavier Glorot and Yoshua Bengio. Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, pages 249–256. JMLR Workshop and Conference Proceedings, 2010. pages 16
- [40] Dong C Liu and Jorge Nocedal. On the limited memory bfgs method for large scale optimization. *Mathematical programming*, 45(1):503–528, 1989. pages 19
- [41] M Osman Gani and Toshiyuki Ogawa. Stability of periodic traveling waves in the aliev–panfilov reaction–diffusion system. *Communications in Nonlinear Science and Numerical Simulation*, 33:30–42, 2016. pages 23

- [42] Marta Varela and Oleg V Aslanidi. Role of atrial tissue substrate and electrical activation pattern in fractionation of atrial electrograms: a computational study. In *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 1587–1590. IEEE, 2014. pages 36
- [43] Flavio Fenton and Alain Karma. Vortex dynamics in three-dimensional continuous myocardium with fiber rotation: Filament instability and fibrillation. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 8(1):20–47, 1998. pages 37

Appendix A

Software Design

This project attempt to analyze EP-PINNs on a variety of tasks, with different geometries, data inputs, modes (forward/inverse), and more. All synthetic data generation programs were written and executed in Matlab. These files were developed by Dr. Marta Varela and used in this project for the sole purpose of generating GT datasets. We implemented the various experiments and EP-PINNs entirely using Python software. To design and train the EP-PINNs, we utilized the DeepXDE library. The software archive, data files, and several animations are available at the Github repository [alono1/PINNs-for-Cardiac-EP](https://github.com/alono1/PINNs-for-Cardiac-EP).

The project's program files are modular and divided into 5 modules:

1. Main module - main.py;
2. Utility module - utils.py;
3. PINN module - pinn.py;
4. 1D figures module - generate_plots_1d.py;
5. 2D figures module - generate_plots_2d.py.

Main Module

The main module is responsible for collecting the arguments from the user to initiate a specific scenario with its designed properties. After establishing the requested task, it defines the relevant system dynamics, EP-PINNs, and the training scheme. Lastly, it executes the training process and computes the metric losses for the testing dataset. A user can control the following arguments:

- input data file - the training/testing dataset;
- folder name - the folder to save the trained EP-PINN;
- dimension - the required geometry, either 1D or 2D;
- noise - add noise to the training dataset;

- w input - add GT W values to the training process;
- heterogeneity - create a 2D geometry with heterogeneity in D ;
- plot - create the figures for the relevant geometry(1D/2D);
- animation - generate an animation for the 2D geometry.

A user must state the input file data, folder name, and dimension, whereas the other arguments are not compulsory. This design enables the user to perform many different tasks without changing a single line of code. The main module is the only module needed to execute in order to simulate any selected task.

Utility Module

The Utility module is essentially gathering the system dynamics required for the task. It sets the geometry dimension and the spatial-temporal grid space on which the problem is defined. It formulates the equations at the heart of the EP-PINN, the coupled Aliev Panfilov PDE-ODE model. These equations are coded differently in different geometries (e.g., the 1D gradient is composed of one value for x coordinate, but the 2D consists of two values, for x and y).

Another critical property that affects the equations' structure is the existence of heterogeneity. Initial and boundary conditions are also set through the Utility module. Finally, it differentiates between forward and inverse modes. In the forward mode, it assigns all parameters to be constants. Conversely, in inverse mode, a set of parameters are required to be learned through training. Therefore, it assigns them with learnable tensors.

PINN Module

PINN module is responsible for the design and training of the request EP-PINN, it consists of the class called PINN. Based on the system dynamics defined before, it selects the NN architecture, with its relevant parameters. It executes the stabilizing scheme detailed before until the network is ready for training. Lastly, based on the current task, it chooses between the 1-phase or the 3-phase scheme and executes it.

1D Figures Module

1D figures module is composed of functions to create three figures:

1. Cell plot: a plot for a single location ($x = 15 \text{ mm}$) within the 1D cable. It details the cell's potential throughout time;
2. Cable plot: a plot of the full 1D cable's potential in a single time step ($t = 36 \text{ AU}$);
3. Grid plot: a heat-map figure - the spatial domain is expressed in the y-axis, the time domain in the x-axis, the colors represents the V values at every spatial-temporal point (higher values mean brighter colors).

The user may request to generate plots. In such a case, it produces all three figures.

2D Figures Module

2D figures module is composed of functions to create three figures:

1. Cell plot: a plot of a specific ($x = 7.5, y = 7.5$) coordinate within the 2D rectangle space. It shows the cell's potential throughout time;
2. Grid plot: a heat-map figure - the (x, y) spatial space is expressed with the figure's x-y axis system. It displays a random time step ($t = 36 AU$), and the colors portray the V values in the entire 2D space (higher values with brighter colors);
3. Animation: An animation of the grid plot throughout the full time domain to create a video file showing the wave propagation.

The user may request to generate plots. In such a case, it produces the two figures. In addition, a user can separately request the animated video file to be created.

Appendix B

Inverse Mode - 1D

1D problem	No noise	Noise $\sigma = 0.05$	Noise $\sigma = 0.1$
a	GT $a = 0.01$		
mean	1.05E-02	9.20E-03	1.27E-02
std	9.10E-04	2.30E-03	6.40E-03
range	[9E-03 - 1.14E-02]	[6.2E-03 - 1.1E-02]	[5E-03 - 2E-02]
mean RE (%)	5	-8	27
std RE (%)	9.1	23	64
rMSE V	7.70E-03	1.40E-02	2.30E-02
b	GT $b = 0.15$		
mean	1.48E-01	1.51E-01	1.54E-01
std	9.70E-04	4.90E-03	8.00E-03
range	[1.46E-01 - 1.49E-01]	[1.44E-01 - 1.56E-01]	[1.43E-01 - 1.64E-01]
mean RE (%)	-1.3	0.4	2.4
std RE (%)	0.6	3.3	5.3
rMSE V	9.00E-03	1.70E-02	2.40E-02
D	GT $D = 0.1$		
mean	8.70E-02	8.70E-02	7.80E-02
std	5.50E-03	3.80E-03	2.50E-03
range	[7.81E-02 - 9.36E-02]	[8.3E-02 - 9.2E-02]	[7.6E-02 - 8.1E-02]
mean RE (%)	-13	-13	-22
std RE (%)	5.5	3.8	2.5
rMSE V	5.40E-03	1.40E-02	2.50E-02
a & D	GT $a = 0.01$ GT $D = 0.1$		
mean $a \mid D$	0.0064 0.096	0.0055 0.08	0.02 0.079
std $a \mid D$	0.0019 0.01	0.0024 0.012	0.0014 0.0087
range a	[4.0E-03 - 8.8E-03]	[2.1E-03 - 7.5E-03]	[1.98E-02 - 2.2E-02]
range D	[7.5E-02 - 1.3E-01]	[6.3E-02 - 9.3E-02]	[6.8E-02 - 9E-02]
mean RE (%) a	-36	-45	100
mean RE (%) D	-4	-20	-21
std RE (%) $a \mid D$	19 10	24 12	14 8.7
rMSE V	9.00E-03	1.20E-02	1.90E-02
b & D	GT $b = 0.15$ GT $D = 0.1$		
mean $b \mid D$	0.149 0.085	0.1496 0.079	0.146 0.065
std $b \mid D$	0.0007 0.0091	0.0009 0.0063	0.01 0.024
range b	[1.48E-01 - 1.5E-01]	[1.49E-01 - 1.51E-01]	[1.34E-01 - 1.6E-01]
range D	[6.72E-02 - 9.08E-02]	[7.36E-02 - 8.8E-02]	[3.76E-02 - 9.65E-02]
mean RE (%) b	-0.7	-0.3	-2.7
mean RE (%) D	-15	-21	-35
std RE (%) $b \mid D$	0.46 9.1	0.6 6.3	6.6 24
rMSE V	8.80E-03	1.30E-02	2.40E-02

Table B.1: Inverse 1D results

Appendix C

Inverse Mode - 2D

2D (homogeneous)	Planar	Corner	Spiral
a		GT $a = 0.01$	
mean	1.00E-02	1.02E-02	5.90E-03
std	1.40E-04	2.50E-04	2.19E-03
range	[1.0E-02 - 1.01E-02]	[9.9E-03 - 1.04E-02]	[2.4E-03 - 8.3E-03]
mean RE (%)	3.2	2.1	40.7
std RE (%)	1.4	2.5	21.9
rMSE V	1.00E-03	3.70E-03	2.25E-02
b		GT $b = 0.15$	
mean	1.50E-01	1.50E-01	1.37E-01
std	1.60E-04	2.30E-04	2.85E-03
range	[1.4E-01 - 1.5E-01]	[1.4E-01 - 1.5E-01]	[1.3E-01 - 1.4E-01]
mean RE (%)	0.01	0.2	8.7
std RE (%)	0.1	0.15	1.9
rMSE V	9.30E-04	3.44E-03	2.69E-02
D		GT $D = 0.1$	
mean	9.77E-02	9.54E-02	7.41E-02
std	1.00E-03	3.00E-04	2.40E-03
range	[9.6E-02 - 9.8E-02]	[9.4E-02 - 9.5E-02]	[7.0E-02 - 7.7E-02]
mean RE (%)	2.2	4.6	25.8
std RE (%)	1	0.3	2.4
rMSE V	1.10E-03	2.60E-03	2.18E-02
a & D		GT $a = 0.01$ GT $D = 0.1$	
mean	0.0099 0.0983	0.0082 0.0931	0.0001 0.0675
std	0.0002 0.0003	0.0004 0.0018	0.0001 0.0017
range a	[9.6E-03 - 1.0E-02]	[7.7E-03 - 8.8E-03]	[1.0E-06 - 1.6E-06]
range D	[9.7E-02 - 9.8E-02]	[9.1E-02 - 9.5E-02]	[6.6E-02 - 7.0E-02]
mean RE (%)	0.9780 1.7189	17.7376 6.8573	99.9875 32.5105
std RE (%)	0.2 0.3	4.3 1.8	0.002 1.7
rMSE V	1.00E-03	3.00E-03	2.00E-02
b & D		GT $b = 0.15$ GT $D = 0.1$	
mean	0.1498 0.0978	0.1496 0.0946	0.1391 0.0743
std	0.0001 0.0005	0.0002 0.0018	0.0023 0.0021
range b	[1.4E-01 - 1.5E-01]	[1.5E-01 - 1.5E-01]	[1.3E-01 - 1.4E-01]
range D	[9.7E-02 - 9.8E-02]	[9.1E-02 - 9.6E-02]	[7.1E-02 - 7.6E-02]
mean RE (%)	0.07 2.2	0.28 5.4	7.2 25.6
std RE (%)	0.1 0.56	0.2 1.8	1.6 2.1
rMSE V	1.39E-03	3.54E-03	3.30E-02

Table C.1: Inverse 2D homogeneous results

2D (homogeneous)	Corner	Spiral
a	GT $a = 0.01$	
mean	1.00E-02	7.10E-03
std	5.00E-04	8.00E-04
range	[9.9E-03 - 1.1E-02]	[5.9E-03-7.9E-03]
mean RE (%)	7.3	28.6
std RE (%)	5.1	8.3
rMSE V	9.00E-03	2.75E-02
b	GT $b = 0.15$	
mean	1.49E-01	1.37E-01
std	5.00E-04	2.50E-03
range	[1.4E-01 - 1.5E-01]	[1.3E-01 - 1.4E-01]
mean RE (%)	0.2	8.2
std RE (%)	0.3	1.7
rMSE V	7.00E-03	2.69E-02
D	GT $D = 0.1$	
mean rMSE D	1.40E-02	2.70E-02
std rMSE D	1.70E-03	2.10E-03
rMSE V	7.10E-03	2.51E-02
a & D	GT $a = 0.01$, GT $D = 0.1$	
mean a	8.90E-03	1.00E-06
std a	2.00E-04	2.00E-06
range a	[8.7E-03 - 9.2E-03]	[9.7E-07 - 1.6E-06]
mean RE (%) a	10.6	99.9
std RE (%) a	2.7	0.002
mean rMSE D	1.30E-02	3.30E-02
std rMSE D	1.00E-03	1.10E-03
rMSE V	6.80E-03	2.80E-02
b & D	GT $b = 0.15$, GT $D = 0.1$	
mean b	1.49E-01	1.37E-01
std b	7.60E-04	2.70E-03
range b	[1.48E-01 - 1.49E-01]	[1.33E-01 - 1.4E-01]
mean RE (%) b	0.5	8.5
std RE (%) b	0.5	1.8
mean rMSE D	1.50E-02	2.80E-02
std rMSE D	2.80E-03	2.21E-03
rMSE V	1.00E-02	2.90E-02

Table C.2: Inverse 2D heterogeneous results