Modelling and simulation of the electric activity of the heart using Graphic Processing Units

Instituto de Investigación en Ingeniería de Aragón (I3A)  
Universidad de Zaragoza

Andrés Mena Tobar

Supervisor  
José Félix Rodríguez Matas
Mr. José Félix Rodríguez Matas,

CERTIFY

That the thesis entitled “Modelling and simulation of the electric activity of the heart using Graphic Processing”, by Mr. Andrés Mena Tobar to be entitled to PhD has been done on his supervision.

Zaragoza, March, the 21st 2017.

Signature: José Félix Rodríguez Matas.
To my family,

for staying by my side all the time.

Thank you very much:

Lucia, Luisa,

Paula, María,

Pili, Andrés,

Maribel, Luis

and José Felix
Acknowledgments

Me gustaría expresar mi agradecimiento, en primer lugar, a José Félix Rodríguez, sin él esta tesis no se hubiese realizado. No puedo olvidar en este punto a Elvio H., que junto a José Félix, me animaron a comenzar este proyecto. JF más que director ha sido y es un gran amigo.

Este trabajo me gustaría dedicárselo a mis padres (Pili y Andrés) que estuvieron y están apoyándome y ayudándome en todos los momentos, y animándome siempre a seguir adelante.

Pero muy especialmente se lo quiero dedicar a mi mujer y compañera, Luisa. Durante estos años de convivencia siempre me ha dado su apoyo y comprensión. A que me lanzara a este empeño que sin haber estado ella allí, no se hubiera realizado.

Me gustaría agradecérselo también, a mis hermanos Luis y Maribel, que son mi familia. Y a mis hermanos políticos Silvia, Nicolás, David y Jessica. A mi amigos Inma y Santi. A Nicolás y Pili. Y a mi entrenador Miguel Ángel de Arce.

Quiero nombrar a aquellas personas que me han animado de una forma u otra durante este periodo, Javier Bayod, Andrés Galavis, Luis Peiro, José Antonio Bea, Olfa, Mauro, Víctor Acosta, Miguel Ángel Martínez, David González, Elías, Esperanza Tabuenca, Manolo, Conchi, Begoña, Manu, María Ángeles, Fany, Berto ... y a todos que he tenido de compañeros estos años.

Y muy especialmente a una personita que me da la vida, a mi hija Lucía.

Andrés Mena Tobar.
Zaragoza, March 2017.
Abstract

Mathematical modelling and simulation of the electric activity of the heart (cardiac electrophysiology) offers an ideal framework to combine clinical and experimental data in order to help understanding the underlying mechanisms behind the observed respond under physiological and pathological conditions. In this regard, solving the electric activity of the heart possess a big challenge, not only because of the structural complexities inherent to the heart tissue, but also because of the complex electric behaviour of the cardiac cells. The multi-scale nature of the electrophysiology problem makes difficult its numerical solution, requiring temporal and spatial resolutions of 0.1 ms and 0.2 mm respectively for accurate simulations, leading to models with millions degrees of freedom that need to be solved for thousand time steps. Solution of this problem requires the use of algorithms with higher level of parallelism in multi-core platforms. In this regard the newer programmable graphic processing units (GPU) has become a valid alternative due to their tremendous computational horsepower. This thesis develops around the implementation of an electrophysiology simulation software entirely developed in Compute Unified Device Architecture (CUDA) for GPU computing. The software implements fully explicit and semi-implicit solvers for the monodomain model, using operator splitting and the finite element method for space discretization. Performance is compared with classical multi-core MPI based solvers operating on dedicated high-performance computer clusters. Results obtained with the GPU based solver show enormous potential for this technology with accelerations over $50\times$ for three-dimensional problems when using an implicit scheme for the parabolic equation, whereas accelerations reach values up to $100\times$ for the explicit implementation. The implemented solver has been applied to study pro-arrhythmic mechanisms during acute ischemia. In particular, we investigate on how hyperkalemia affects the vulnerability window to reentry and the reentry patterns in the heterogeneous substrate caused by acute regional ischemia using an anatomically and biophysically detailed human biventricular model. A three-dimensional geometrically and anatomically accurate regionally ischemic human heart model was created. The ischemic region was located in the inferolateral and posterior side of the left ventricle mimicking the occlusion of the circumflex artery, and the presence of a washed-out zone not affected by ischemia at the endocardium has been incorporated. Realistic heterogeneity and fiber anisotropy has also been considered in the model. A highly electrophysiological detailed action potential model for human has been adapted to make it suitable for modeling
ischemic conditions (hyperkalemia, hypoxia, and acidic conditions) by introducing a formulation of the ATP-sensitive $K^+$ current. The model predicts the generation of sustained re-entrant activity in the form single and double circus around a blocked area within the ischemic zone for $K^+$ concentrations below 9mM, with the reentrant activity associated with ventricular tachycardia in all cases. Results suggest the washed-out zone as a potential pro-arrhythmic substrate factor helping on establishing sustained ventricular tachycardia.
Resumen

El modelado matemático y simulación de la actividad eléctrica del corazón (electrofisiología cardíaca) constituyen un marco ideal para combinar datos clínicos y experimentales con el fin de ayudar a comprender los mecanismos subyacentes detrás de la respuesta observada en condiciones fisiológicas y patológicas. En este sentido, la resolución de la actividad eléctrica del corazón presenta un gran reto, no sólo por las complejidades estructurales inherentes al tejido cardíaco, sino también por el complejo comportamiento eléctrico de las células cardíacas. La naturaleza multiscale del problema electrofisiológico dificulta su solución numérica, requiriendo resoluciones temporales y espaciales de 0.1 ms y 0.2 mm respectivamente para simulaciones precisas, llevando a modelos con millones de grados de libertad que necesitan ser resueltos por miles de pasos de tiempo. La solución de este problema requiere el uso de algoritmos con mayor nivel de paralelismo en plataformas multi-core. En este sentido, las nuevas unidades de procesamiento gráfico (GPU) programables se han convertido en una alternativa válida debido a su enorme potencia computacional. Esta tesis se desarrolla en torno a la implementación de un software de simulación electrofisiológica totalmente desarrollado en “Compute Unified Device Architecture” (CUDA) para la computación GPU. El software implementa soluciones totalmente explícitas y semi-implícitas para el modelo de monodominio, utilizando la división del operador y el método de elemento finito para la discretización del espacio. El rendimiento se compara con los métodos clásicos basados en MPI de varios núcleos que operan en clusters dedicados de computación de alto rendimiento. Los resultados obtenidos con el software basado en GPU muestran un enorme potencial para esta tecnología, con aceleraciones de más de $50\times$ para problemas tridimensionales cuando se utiliza un esquema implícito para la ecuación parabólica, mientras que las aceleraciones alcanzan valores de hasta $100\times$ para la implementación explícita. El solver implementado se ha utilizado para estudiar los mecanismos pro-arrítmicos durante la isquemia aguda. En particular, investigamos cómo la hiperkalemia afecta a la ventana de vulnerabilidad, a la reentrada y a los patrones de reentrada en el sustrato heterogéneo causados por isquemia regional aguda usando un modelo biventricular humano, anatómicamente y biofísicamente detallado. Se creó un modelo tridimensional geométricamente y anatómicamente exacto del corazón humano isquémico. La región isquémica se localizó en el lado inferolateral y posterior del ventrículo izquierdo, imitando la oclusión de la arteria circunfleja y se ha incorporado la presencia de una zona de lavado, no afectada por la isquemia endocárdica.
También se han considerado en el modelo la heterogeneidad realista y la anisotropía de la fibra. Se ha adaptado un modelo de potencial de acción detallado electrofisiológico para el ser humano, adecuado para modelar las condiciones isquémicas (hipercaliemia, hipoxia y acidosis) introduciendo una formulación de la corriente $K^+$ sensible al ATP. El modelo predice la generación de actividad reentrante sostenida en forma de rotor simple y doble alrededor de un área bloqueada dentro de la zona isquémica para concentraciones de $K^+$ inferiores a 9 mM, con la actividad reentrante asociada con taquicardia ventricular en todos los casos. Los resultados sugieren que la zona de lavado como un posible factor de sustrato pro-arrítmico, que ayuda en el establecimiento de taquicardia ventricular sostenida.
1 Introduction
1.1 Anatomy ................................................. 1
1.2 Electrical activity of the heart, stimulation and conduction. .......... 4
1.3 The action potential ...................................... 7
  1.3.1 Threshold Potential .................................. 8
  1.3.2 Phases of Action Potential ........................... 9
  1.3.3 Action Potential in different tissues .................. 10
1.4 Excitability of the cardiac cells. ................................ 11
1.5 The electrocardiogram .................................... 13
1.6 Myocardical Ischemia ..................................... 15
  1.6.1 Reentries ............................................. 17
  1.6.2 Refractory period in ischemic cells ................... 19
1.7 Numerical approaches .................................... 21
  1.7.1 Finite Element Method ............................... 22
1.8 Motivation ................................................ 28
1.9 Objectives ............................................... 31
1.10 Hypothesis of work and Methodology .......................... 32

2 Equations that govern the electrical activity of the heart .......... 33
2.1 Governing equations .................................... 33
## CONTENT

2.1.1 Bidomain Model ........................................... 34  
2.1.2 Monodomain Model ....................................... 36  
2.2 Myocardium conductance .................................... 37  
2.3 Action Potential Models .................................... 38  
  2.3.1 Structure of an Action Potential model ................ 39  
  2.3.2 The Cell Membrane ..................................... 39  
  2.3.3 Gates .................................................. 41  
  2.3.4 Ionic Channels ......................................... 43  
  2.3.5 The tenTusscher Action Potential Model .............. 44  
2.4 Numerical solution of the electric activity of the heart .......... 46  
  2.4.1 Spatial-temporal discretisation ....................... 47  
  2.4.2 Integration of the mass matrix ....................... 49  

3 GPU implementation ........................................... 51  
  3.1 The CUDA programming model ............................ 51  
  3.2 Parallel implementation in GPU .......................... 52  
    3.2.1 ODE solver ........................................ 53  
    3.2.2 Parabolic solver .................................... 55  
  3.3 Benchmarking ............................................. 56  
  3.4 Results .................................................. 58  
  3.5 Discussion and conclusions .............................. 64  

4 Vulnerability in Regionally Ischemic Human Heart. Effect of the Extracellular Potassium Concentration .......................... 71  
  4.1 Methods .................................................. 72  
    4.1.1 Mathematical model ................................. 72  
  4.2 Model of acute ischemia .................................. 73  
    4.2.1 Action Potential model under ischemic conditions .... 73  
    4.2.2 Heart model ........................................ 76  
    4.2.3 Electrophysiological heterogeneities under acute ischemia . 77  
    4.2.4 Stimulation protocol ............................... 78  
    4.2.5 Numerical simulations .............................. 79  
  4.3 Results .................................................. 80  
  4.4 Discussion and conclusions .............................. 83
5 Conclusions and future work

5.1 Conclusions

5.1.1 GPU implementation

5.1.2 Acute ischemic simulations

5.2 Conclusiones

5.2.1 Implementación en GPUs

5.2.2 Simulaciones de la isquemia aguda

5.3 Future work

5.4 Contributions

5.4.1 Projects

5.4.2 Journal papers
List of Figures

1.1 Anatomic location of the heart. . . . . . . . . . . . . . . . . . . . . . . 2
1.2 Internal Morphology. . . . . . . . . . . . . . . . . . . . . . . . . . . . . 3
1.3 Electrical system of the heart, the whole is shown from the SA node
to the arborization of Purkinje. . . . . . . . . . . . . . . . . . . . . . . . . 5
1.4 Main currents that contribute to the action potential and phases of
the same. Image taken from [70]. . . . . . . . . . . . . . . . . . . . . . . 8
1.5 Refractory periods. Where, A.R.P: Absolute Refractory Period; E.R.P:
Effective Refractory Period; R.R.P : Relative Refractory Period;
S.N.P : Supernormal phase. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 12
1.6 Conducting systems of the heart and (interrelationship) correlation
with the ECG/ERG. Figure taken from [39] . . . . . . . . . . . . . . . . . . 13
1.7 Shape, intervals and segments of the ECG . . . . . . . . . . . . . . . . 15
1.8 Patterns of excitation on a pig animal model. A single spiral is a
responsible for the continuation of the tachycardia with a pattern
from a single rotor to a figure of eight changing from beat to beat.
Figure taken from [50] . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 19
1.9 AP of a healthy cell and an ischemia one cell and its respective re-
fractory periods. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 20
2.1 Detail of the cell membrane. . . . . . . . . . . . . . . . . . . . . . . . . . 40
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.2</td>
<td>The equivalent circuit of the cell membrane</td>
<td>44</td>
</tr>
<tr>
<td>2.3</td>
<td>Ten Tusscher model 2006[102]</td>
<td>45</td>
</tr>
<tr>
<td>3.1</td>
<td>CUDA program model</td>
<td>52</td>
</tr>
<tr>
<td>3.2</td>
<td>CUDA program model</td>
<td>53</td>
</tr>
<tr>
<td>3.3</td>
<td>Structure of the vector of state variables for the ionic model, ( u ), and the transmembrane potential, ( V )</td>
<td>54</td>
</tr>
<tr>
<td>3.4</td>
<td>Model problems used to test the performance of the GPU algorithm.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>a) ODE solver, b) Parabolic solver</td>
<td>57</td>
</tr>
<tr>
<td>3.5</td>
<td>Performance of the ODE solver: a) Relative computing time for the GPU-ODE solver as the number of nodes increases for models with different number of state variables (times have been normalised with respect to the 4 SV model; b) Speed-up against a single CPU core for ionic models with different number of state variables.</td>
<td>59</td>
</tr>
<tr>
<td>3.6</td>
<td>Performance of the parabolic solver: a) Explicit algorithm (( \theta = 0 )); b) Semi-implicit algorithm (( \theta = 1 )).</td>
<td>61</td>
</tr>
<tr>
<td>3.7</td>
<td>Average relative computing time of the implicit parabolic solver for different sparse matrix formats on structured meshes (a), and unstructured meshes (b). Note that the DIA sparse format is not suitable for unstructured meshes.</td>
<td>62</td>
</tr>
<tr>
<td>3.8</td>
<td>Average relative computing time of the implicit parabolic solver for different iterative schemes. HYB sparse solver and ILU preconditioning have been used for the computations.</td>
<td>63</td>
</tr>
<tr>
<td>3.9</td>
<td>Depolarization front after 20ms of activation. a) Human bi-ventricular model discretised with hexahedral elements; b) Human atria discretised with tetrahedral elements.</td>
<td>64</td>
</tr>
<tr>
<td>3.10</td>
<td>A coupled problem with large bandwidth. a) Detailed of the one-dimensional network representing the fast conduction system of the heart (left panel) coupled to the human bi-ventricular model (right panel). Coupling between the two meshes takes place at discrete points only (black markers in the right panel); b) Sparsity pattern of the resulting linear system.</td>
<td>65</td>
</tr>
<tr>
<td>4.1</td>
<td>Fraction of open channels for different cell types. Symbols represent experimental values for different cell types in different species. Solid lines represent the model fit Eq. 4.5.</td>
<td>74</td>
</tr>
</tbody>
</table>
4.2 Electrophysiologic heterogeneity in the acute regional ischemic heart. A) Resting potential in the heterogeneous heart showing higher transmembrane potential in the ischemic tissue (CZ) and the transition through the border zone (BZ). The right panel shows a detail of the washed-zone at the endocardium define in the model as reported in [112]. B) Stimulation sites for the normal SA stimulation according to [24]. C) Location of the pseudo-ECG probes corresponding to the six derivations of the standard ECG. A pseudo-ecg is depicted in the bottom panel, exhibiting the ST elevation in V5-V6, with an acute T-wave in V6, and ST depression in V1-V4 consistent with an infarction involving the inferior, lateral and posterior walls caused by the occlusion of the proximal circumflex artery. In addition, the positive T wave following the changes in the ST segment is consistent with the inverse relationship between APD and activation times.

4.3 Depolarization pattern through the ischemic tissue (delimited by the white line) during basic activation (A), and right after the ectopic beat (B). The ischemic tissue is depolarized by three different fronts coming from the normal activation sites shown in Fig 4.2. The ectopic beat was delivered at the first point being depolarized within the ischemic tissue (point F in the right panel).

4.4 Vulnerability window for different levels of $[K^+]_o$. Early stimulation starts at point F in Figure 4.3.

4.5 Reentrant patterns at endocardium (A) and epicardium (B) for the last recorded beat for the case $[K^+]_o = 7.0$ mM and CI=260 ms. Panel C shows the pseudo-ECG corresponding to the derivation V3, the arrows indicate the the window corresponding to the patterns shown in panels (A) and (B).

4.6 Transmembrane potential at t=2500 ms after initiation of earlier stimulation ($[K^+]_o = 7.0$ mM and CI=260 ms). A) Depolarization map at the epicardium and endocardium. B) Transmural depolarization map.

4.7 Evolution of the reentry pattern from beat to beat. The sequence indicates an evolution from a single circus to a double circus (figure of eight) to then come back to a single circus patter ($[K^+]_o = 7.0$ mM and CI=260 ms).
4.8 Vulnerability window for different levels of $[K^+]_o$ in a model without
the washed-out zone. Early stimulation starts at point F in Figure 4.3. 86
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Normal range interval values of the ECG</td>
<td>15</td>
</tr>
<tr>
<td>3.1</td>
<td>Coefficient of determination between the state variables corresponding to the CUDA implementation and the C++ implementation. The coefficient of determination has been calculated with the entire time course of each state variable during the last beat of the stimulation protocol.</td>
<td>58</td>
</tr>
<tr>
<td>3.2</td>
<td>Computation overhead of the single kernel versus the multiple kernel implementation for the TP06 ionic model.</td>
<td>60</td>
</tr>
<tr>
<td>4.1</td>
<td>Extracellular Potassium and Intracellular ADP and ATP concentrations defining physiologic and ischemic conditions for the simulations</td>
<td>76</td>
</tr>
<tr>
<td>4.2</td>
<td>Value of the parameters for the $I_{KATP}$ current adapted to the ten Tusscher model for different cell types. Also shown, the APD under physiologic and ischemic conditions.</td>
<td>77</td>
</tr>
</tbody>
</table>
CHAPTER 1

Introduction

1.1 Anatomy

The heart is located in the chest cavity, in the middle mediastinum, between the two lungs and immediately retro-sternal, i.e., ahead of the sternum and costal cartilages of the third, fourth and fifth ribs, right and left. The right third of the Heart, barely protruding from the right sternal border and the remaining two thirds are located to the left, ending in a tip (see Figure 1.1). Its lower face rests on the diaphragm muscle which separates the chest cavity from the abdominal cavity. It is oriented in space, from top to bottom, from right to left and from the back to get ahead. The size and weight vary considerably by age and sex, but in a young adult of average height, the heart weighs between 270 and 300 grams.

The heart is covered externally by two sheets of serous tissue called pericardium, one intimately attached to the body (epicardium) and another, continuing with the first, is reflected in the base around the heart to completely surround it (pericardium itself); between the two sheets, which are not adhered together, there is a virtual cavity allowing free movements of cardiac contraction. It is supported from the top by the large arterial trunks. These are the aorta, pulmonary artery, Superior Vena Cava, Inferior Vena Cava and four Lung veins [38].
Internal Morphology. Four chambers, two upper-called right atrium and left atrium, and two lower can be seen inside. They are called the right ventricle, and left ventricle respectively (see Figure 1.2). The atria are separated by a wall or inter-atrial septum and ventricles by the inter-ventricular septum. Both partitions are continued to each other, forming a true membranous-muscular wall separating the heart into two right chambers and two left chambers. This separation is functional, since the right cavities are connected to the pulmonary circulation and left cavities, with the systemic general circulation.

Veins Cavas Upper and Lower bringing carbo-oxygenated blood throughout the body reach the right atrium. The blood passes to the right ventricle, which by contracting (systole Ventricular), sends it to the pulmonary artery leading to both lungs to perform gas exchange. The oxygenated blood returns to the left atrium through the four pulmonary veins and the left ventricle and is pumped into the aorta for distribution throughout the body.

The atria to the ventricles communicate through respective valves whose function is to open to allow entry of blood into the cavity, then during systole the valves get sealed to prevent blood from flowing back. These valves are Mitral between left atrium and ventricle and the tricuspid, between right atrium and right ventricle.

The left ventricle dumps its contents into the aorta through the aortic valve.
and the right ventricle discharge to the pulmonary artery through the pulmonary valve. Both valves have three leaflets called semilunar or sigmoidal forming a kind of three-pointed star. Unlike the atrioventricular valves, they lack of chordae the support them and are sealed because they resemble tiny parachutes, which bulge and contact each other.

Anatomically the right ventricle is thin, since it must contract against very low pressure, its wall is between 4 and 5 mm. of thickness. The left ventricle must overcome resistance or systemic blood pressure, therefore its force of contraction should be higher. For this reason, the walls are thicker, having a thickness of between 8 and 15 mm.

The cardiac cycle consists of three phases:

1. Atrial systole: phase in which the atria contract, electrically corresponds to the atrial depolarization.

2. Ventricular systole: phase in which the ventricles contract, electrically corresponds to ventricular depolarization.

3. Diastole: The phase in which the heart muscle relaxes, electrically corresponds
to ventricular repolarization.

1.2 Electrical activity of the heart, stimulation and conduction.

Cardiac cells, as a whole, generate their own stimuli lead quickly and then activate the contractile mechanisms within them. These are the excitation and conduction system of the heart, whose functionality is explained below. Figure 1.3 shows the stimulation and conduction system of the heart.

Cardiac stimuli normally originated in the sinus node (SA), or Keith and Flack (normal heart pacemaker), located at the junction of the superior vena cava with the right atrium. The electrical stimulus generated travels by the inter-nodal anterior, middle and posterior beams to the atrioventricular node (AV), or Aschoff- Tawara, located at the endocardium of the same right atrium but in the septal region, immediately above the insertion of the septal leaflet of the tricuspid valve. The AV node extends into the ventricle forming the bundle of His. The area, where the three intermodal beams and the His bundle meet, is known as atrio-ventricular union. The bundle of His descends through the interventricular septum, where it divides into two main branches: trunk of the left branch and trunk of the right branch.

The right branch descends on the right side of the inter-ventricular septum to the base of the anterior papillary muscle, where it ends in a profuse branching or Purkinje network. The trunk of the left branch descends through the septum and on its left side, divided in a previous branch which immediately activates the anterior superior aspect of the left ventricle.

This branch is in contact with the aortic valve and is longer and thinner than its later companion. The much shorter and thicker, posterior branch ends in the back and underside of the Heart. All branches end in the Purkinje network, located in the subendocardial miocardium region inter-connecting each other, so that interruption of stimuli conduction by one does not prevent muscle activation.

The stimulus conduction velocity within this system is not uniform, accepting the following average values:

1. Purkinje Fibers: 2 m/s.
2. Atrial Musculature: 0.8 to 1 m/s.
3. Ventricular Musculature: 0.5 to 0.7 m/s.
Figure 1.3: Electrical system of the heart, the whole is shown from the SA node to the arborization of Purkinje.
4. Nodule AV: 0.2 m/s.

The low rate of spread in the AV node seems established to protect the ventricles from too fast pulses that may be originated in the atria.

Sinus Node (SA). It is an accumulation of myofibrils of about 15 mm in length; 5 mm to 7 mm wide and 2 mm thick, rich in collagen, and located at the top of the right atrium [70]. It originates the normal excitation of the myocardium, from which the name of sinus rhythm is derived for said pacemaker (Figura 1.3). SA is the heart’s primary automatic center. It has a generous blood supply that is achieved mainly by the right coronary artery but also receives blood from the left coronary artery. It also has a large number of nerve endings that also surround it, hence the possibility of influencing its functions by means of drugs.

Internodals Bundles. In the atrial myocardium there are 3 specific tissue pathways that establish an anatomic-functional bridge between the SA and AV nodules, the anterior, middle and posterior internodal bundles. Of these three beams, those of Bachmann and Wenckebach have been very well studied. These beams establish the anatomical pathways that guarantee electrical communication between the two atria and the two nodules, ensuring the progression of the stimulus from its origin in the atria to its entry into the ventricles.

Ventricular Atrium Node (AV). The AV node measures 5 to 6 mm in length, 2 to 3 mm in width and 1.5 mm in thickness [70]. It is located more in the atrial septal defect, in its most posterior and inferior portion. It is surrounded by the coronary sinus and the mouth of the inferior vena cava. Its irrigation depends on the right coronary artery. The AV node is considered as an integral part of a much more extensive anatomical system called the union composed of the following anatomical structures:

- Atrial regions adjacent to the AV node.
- The AV node itself.
- The His bundle.
- The proximal portions of the bundle.
Chapter 1. Introduction

The bundle of His.  It is an extension of the AV node; It is located in the muscular portion of the interventricular septum; It is approximately 10 to 20 mm in length and has a diameter between 1 and 4 mm. It receives its vascularization of the left coronary.

Left branch of the bundle of His.  It is the result of the bifurcation of the His bundle directed to the left ventricle. Its fibers end in the apex and in the anterior and posterior papillary muscles. Its anatomical position is subendocardial.

Anterior and posterior fascicles of the left branch.  The left branch bifurcates, almost immediately, in two fascicles: one anterior and one posterior, which, apparently, already from the bundle of His, have delimited the fibers for each fascicle. The anterior fascicle is longer and thinner than the posterior fascicle. Both end up inserting into the anterior and posterior papillary muscles of the left ventricle. The 2 fascicles receive blood from both coronary arteries.

Right branch of the bundle of His.  It goes to the right ventricle and also terminates in the papillary muscles. It appears as a continuation of the His bundle after the left branch has been detached. It is irrigated by the left coronary.

Purkinje Network.  It forms a dense mesh that invades the ventricular walls and ends up being confused with the myocardial fibers. It receives its irrigation from the arteries that nourish the myocardial regions where it lies. The anatomical arrangement of the Purkinje network is fundamental for understanding the morphogenesis of electrocardiogram waves.

1.3 The action potential.

The action potential (AP) is a bioelectrical signal corresponding to the potential difference between the intracellular and extracellular medium, separated by the cell membrane [38].

In addition to separate the intracellular space from the extracellular space, the cell membrane contains ion channels exchangers and electrogenic pumps that allow both passive and active ion mobility between the two media. This exchange of ions
mainly occurs due to the difference in ion concentrations between the two media and therefore the forces due to diffusion and electric field. Thus, this dynamic flow of ions through the membrane is responsible for the generation mechanism of the action potential. Figure 1.4 represents the action potential of different heart cells and the different phase of the AP. You can also appreciate the different currents that influence the AP.

![Depolarization and Repolarization Currents](image)

Figure 1.4: Main currents that contribute to the action potential and phases of the same. Image taken from [70]

### 1.3.1 Threshold Potential.

It corresponds to the value of the membrane potential from which an Action Potential is generated. In non-automatic cells, this potential is reached through ionic fluxes from neighboring depolarized cells. In automatic cells the threshold potential can be reached by spontaneous diastolic depolarization of itself. Automatic cells do not have a stable membrane potential at rest, because they continuously discharge
AP, and when they do not, they are increasing their depolarization to reach the threshold of $Na^+$ or $Ca^{2+}$ channels and to discharge AP.

1.3.2 Phases of Action Potential

The AP consists of several phases (Phase 0, 1, 2, 3 and 4), which are characterized by the activation of different ionic currents. AP phases can be briefly described as:

- **Phase 0** consists of an increasing initial phase of the action potential (about 1 ms) and is accomplished by the sodium current ($I_{Na}$) and the L-type calcium current ($I_{Ca(L)}$). The first current to be activated is $I_{Na}$, whose high peak value produces a rapid rise in the membrane potential that allows the activation of $I_{Ca(L)}$, which will remain active for the rest of the depolarization and throughout phases 1 and 2.

- **Phase 1.** After the action potential has reached its peak value, rapid repolarization occurs and is attributed to the activation of the transient outward potassium current ($I_{to}$). The magnitude of $I_{to}$ varies remarkably in different locations of the heart. $I_{to}$ density is higher in epicardium. $I_{to}$ is responsible for the spike-dome shape on the AP.

- **Phase 2.** The slow repolarization phase occupies most of the action potential duration, and its dynamics is determined by the inactivation of the current associated with the L-type calcium channels ($I_{Ca(L)}$) and activation of the delayed rectifier potassium current ($I_K$). The relative contribution of $I_{Kr}$ and $I_{Ks}$ in the repolarization changes depending on the species and between different cell regions of the ventricular wall.

- **Phase 3.** Towards the end of the action potential, repolarization speed increases considerably due to the rapid increase in $I_{Kr}$ and $I_{K1}$ currents. The ion output through these channels is increased to create a positive feedback causing rapid repolarization. The current $I_{Kr}$ controls the initial portion of the repolarization while $I_{K1}$ controls the latter portion.

- **Phase 4.** The membrane potential of the cell is at rest, this being similar to the equilibrium potential potasios ion, and potassium inverse rectifier current ($I_{K1}$) is active.
1.3.3 Action Potential in different tissues

Sinus Node. It is characterized by its 4 phases of depolarization, and because it gives rise to the activity of pacemakers. The depolarization phase 4 is due to the high density of the pacemaker currents and the lack of $I_{K1}$, which also explains the relative depolarization state of such tissue. Sodium channels are scarce and AP elevation is slow, as it is mainly mediated by $I_{Ca,L}$. There is no phase 1, due to the lack of $I_{to}$, the duration of the AP is short and the frequency of depolarization is determined by the sympathetic and parasympathetic modulation of $I_f$ (responsible for the pacemaker activity).

Atrial tissue. The auricular AP has a rapid upstroke, which gives a rapid conduction of the right atrium to the left atrium and from the sinus node to the AV node. It has phase 1 followed by a short phase of plateau and a rapid repolarization.

Atrio-ventricular node. The AV node is similar to the sinus node because of the missing $I_{Na}$. The conduction through the AV node is due to the presence of the current $I_{Ca,L}$ and the propagation speed is slow. $I_{Ca,L}$ is activated by sympathetic stimulation and inhibited by parasympathetic, which are important determinants of conduction through the AV node pulse. The phase 4 of depolarization is not as prominent as in the sinus node.

His bundle and Purkinje fibers. The fibers of the His bundle and the Purkinje system have a high $I_{Na}$ density that facilitates the rapid conduction of impulses, so that ventricular myocytes can be activated synchronously. In addition, it has a high current $I_{K1}$ and a weak current of pacemaker. Therefore, this tissue is characterized by a resting potential of -90 mV, which is close to potassium resting potential, and a slow diastolic depolarization.

Ventricular Tissue. Ventricular myocytes have a density of $I_{Na}$ and $I_{K1}$ that is higher than that of atrial myocytes, and the resting potential of these cells is close to -90 mV. The configuration of the Action Potential varies according to the location within the ventricle. The myocytes of the epicardial layer have a high density of $I_{to}$; This leads to marked repolarization in phase 1 followed by a depolarization of the calcium currents, which generates the characteristic spike-dome configuration.
In contrast, the endocardial layer has much lower $I_{to}$ density with an appreciable reduction of the amplitude of phase 1. The middle layer, myocardium, is formed by the so-called M cells, which have strong currents $I_{to}$ but weak and delayed $I_{K}$ currents, as well as sodium currents activated by slower gates. For these reasons, the M-cell Action Potential is characterized by a spike-dome configuration and a longer AP duration than the Action Potential of epicardial and endocardial cells.

1.4 Excitability of the cardiac cells.

The excitability of cardiac cells corresponds to the property they possess to generate an Action Potential as a consequence of a stimulus applied to them. During much of the action potential, the cell is non-excitable regardless of the intensity of the stimulation applied. This phase of the AP is called the refractoriness phase.

Refractory periods of cardiac cells can be classified as follows: (see Figure 1.5):

- **Absolute refractory period.** It corresponds to a state of total inexcitability. This state exists during phases 0, 1, 2 and part of phase 3 in fast cells (atrial cells, ventricular cells, bundle of His cells and Purkinje fibers). Under normal conditions from a membrane potential value of $-55 \text{ mV}$, the cell partially regains its excitability, which is demonstrated by stimulating the cell with supra-normal currents. The first generated responses are not propagated.

- **Effective refractory period.** It includes the absolute refractory period and the one with generated non-propagated responses. This ends when the first propagated AP appear. The latter occur from low (negative) potentials and on the other hand, at times, when the conductivities of the depolarization channels are not yet recovered. Because of this, the action potentials that may occur at this stage have a slow or low amplitude phase.

- **Relative refractory period.** It is one in which cells are only excitable with supra-normal currents, but they generate propagated responses.

- **Supernormal phase of excitability.** A short period in which sub threshold stimuli are capable of generating an action potential. They appear at the end of phase 3 only in fast cells. However, This higher degree of excitability of the fibers occurs during non-negative membrane potentials, and therefore,
the generated responses are slow. On the other hand, the degree of recovery of the excitability in this period is heterogeneous, which carries the risk of desynchronized responses from a single stimulus (vulnerable phase).

- **Total refractory period.** It ends with the return of complete excitability after the supernormal phase.

- **Functional refractory period.** It is the shortest interval that separates 2 normally propagated responses.

In slow fibers, the effective refractory period goes beyond the duration of the
action potential. This is due to the longer time constant of the slow channel (compared to the fast channel), which makes it unable to reactivate even though the action potential has been completed.

1.5 The electrocardiogram.

In 1893, the Dutch physiologist Willem Einthoven presented, at a meeting of the Dutch Medical Association, the term *electrocardiograma* (ECG or EKG) to name a new non-invasive cardiac method of investigations. The data delivered by the electrocardiograph indicated the electrical activity in connection with the heart beat. A path of a normal ECG heart beat consists of a P wave, a QRS complex and the T wave (see Figure 1.7).

Figure 1.6: Conducting systems of the heart and (interrelationship) correlation with the ECG/ERG. Figure taken from [39]
The characteristic waves of the ECG can be related to events linked to the propagation of the action potential along the heart from the SA node to the myocardial wall. As shown in Figure 1.6, each action potential of the heart is produced, in the SA node. This generates action potentials at a steady pace spontaneously, although the heart rate is controlled by nerves belonging to the sympathetic nerve system and the vagus nerve from the parasympathetic nerve system that produce, respectively, its acceleration and deceleration. The AP then propagates though the atria and it stops in the AV node. A delay is produced, in this point, to obtain a proper synchronization between the action of the atria and the ventricles’ action. After that, the electric stimulation is distributed into both ventricles through the bundle of His and its right and left branches, and the system of Purkinje in order to depolarize the ventricles endocardium to epicardium.

On the contrary ventricular recovery takes place from the epicardium to endocardium, taking an opposite sequence with respect to the activation. The repolarization wave is produced when each cell comes back to its own resting potential.

The characteristic waves of the ECG can then be related to events linked to the method of the action potential propagation, which reflect the cycle of depolarization and re-polarization of the heart.

- The P wave represents the depolarization of the muscle system of the atrium.
- The QRS complex is the result of the combination of the atria repolarization and the ventricular depolarization, which is produced almost simultaneously.
- The T wave represents the ventricular repolarization and the U wave seems to be the result of the repolarization of the system of the intraventricular conduction (fibers of Purkinje), but its production mechanism is unknown.

The time intervals among the different ECG waves are relevant for diagnosis, because they show the electrophysiological processes and are clinic implications when they are out of the normal variation range.

The table 1.1 shows the normal values for the intervals of the ECG.

A number of heart features can be evaluated/measured by the ECG as for instance: atrial tachycardia, ventricular hypertrophy, presence of ischemia, damaged
Chapter 1. Introduction

Figure 1.7: Shape, intervals and segments of the ECG

<table>
<thead>
<tr>
<th>ECG Parameter</th>
<th>Normal Range [seconds]</th>
</tr>
</thead>
<tbody>
<tr>
<td>PR Interval</td>
<td>0.12-0.20</td>
</tr>
<tr>
<td>QRS Interval</td>
<td>0.06-0.10</td>
</tr>
<tr>
<td>ST Segment</td>
<td>0.05-0.15</td>
</tr>
<tr>
<td>QT corrected Interval</td>
<td>0.35-0.44</td>
</tr>
<tr>
<td>RR Interval</td>
<td>0.60-1.00</td>
</tr>
</tbody>
</table>

Table 1.1: Normal range interval values of the ECG

parts of the heart, among others. In addition, the ECG has a number of applications in diagnosis and monitoring of cardiac pathologies: bradycardia, myocardial infarction, congenital defects, abnormal activation, among others.

1.6 Myocardial Ischemia

The Myocardial Ischemia is a pathological condition which occurs before a myocardial infarction. This situation is caused by a reduction the supply of oxygen and nutrients to the affected cells. This introduces a number of electrophysiological changes that influence the action potential and that may lead to fatal arrhythmias, such as those produced by the reentries [113] may trigger the ventricular fibrillation
The myocardial ischemia can be divided into several steps, whose features depend on the animal species it is. Thus, in big size hearts (dogs), the ventricular arrhythmias take place in two different periods of time, the first, the early ischemia, which starts with the occlusion of the artery and whose duration is over 10-15 minutes and the second which start 15-20 minutes later and it lasts from 30 to 50 minute. Whereas, in small size hearts (rabbits or guinea pigs), the period of arrhythmias starts after 6 minutes lasting for about half an hour, with maximum arrhythmias after 10 or 12 minutes.

The electrophysiological effects of the acute ischemia can be analyzed through its three components: the hypoxia, the acidosis and the hyperkalemia [56].

**Hypoxia.** The reduction of oxygen in the cells, which leads to a decrease of the oxidative phosphorylation which, on the one hand, causes the decrease in the concentration of Adenosine Tri-Phosphate (ATP) and an increase of the Adenosine Di-Phosphate (ADP) and other metabolites. In normoxia, the intracellular ATP concentration is from 5 to 10 mmol/L and ADP concentration is from 5 to 70 µmol/L and after 10 minutes of the ischemia, the ATP concentration is of 45% [108] and the ADP(concentration) is between 630 and 3400% [2]. Specifically, for the myocyte of a rabbit, in normaxia, the ATP concentration is of 6.8 mmol/L and after 10 minutes of ischemia such concentration decreases to 4.6 mmol/L [108].

**Acidosis** It is the reduction of intra cellular and extra cellular PH. In normoxia, the intracellular PH is of 7.3, a figure that decreases by one 10 minutes after the beginning the ischemic [54, 114].

**Hyperkalemia** It consists of the increase of the extracellular potassium concentration which takes place into two stages. The first stage takes places in the first 4-8 minutes, when the concentration is doubled, stabilizing at 10-12 mmol/L, then such concentration is between 11 to 14 mmol/L [44], and finally, after 20 minutes, it increases again [108]. This increase in the extracellular potassium concentration is a highly arrhythogenic factor, because, on the one hand, it causes an increase in the resting potential reducing the excitability of the tissue, and on the other hand, it decreases the conduction velocity(CV).
It also causes an increase in the intra-cellular and extra cellular concentration of other ions, as in the sodium, calcium and magnesium ions. \[113\].

The ischemia causes tissue inhomogeneities because the electrophysiological changes, previously discussed, are only taking place in the ischemic tissue and they don’t affect the healthy cells. In turn, not all cells affected by ischemia have the same features. In fact, there is a zone, which is called the border zone (BZ), between the ischemia central zone (CZ) and the normal zone (NZ).

The composition of this area has been analyzed by different groups. Janse et al. \[49\] studied the ischemic heart of a pig experimentally and described the border zone (BZ) as an area consisting of well-defined sets of purely ischemic cells and sets of completely healthy cells \[49\].

On the other hand, other authors describe this zone as a set of cells affected by the ischemia in less extent than cells belonging to the central zone (CZ). Thus, the electrophysiological alterations gradually vary along the central zone from the normal conditions in the normal zone (NZ) to the ischemia conditions in the central zone (CZ) \[20\].

The extra cellular potassium variation has been experimentally studied in pig’s hearts and this variation has been observed to be progressive over 1 centimeter of the tissue.

All these electrophysiological changes modify the action potential of the ischemic cells as described below. The resting potential increases due to the greater extracellular potassium concentration \[32\]. When the resting potential increases, the excitability of the cell decreases and the conduction velocity also decreases. This effect is observed after some minutes the ischemia has started \[51\]. The AP duration (APD) \[44, 114\] decreases due to the higher concentration of potassium which causes an increase in the conductance of the inverse rectifier potassium \[91\] channel and to the $I_{K(\text{ATP})}$ channels \[28\]. Other important consequences are the reduction of the maximum membrane potential and the reduction of the cell depolarization speed \[17\] and the increase of the refractory period (RP).

### 1.6.1 Reentries

During the acute stage of the myocardial ischemia reentries occur being the main cause of ventricular tachycardia (VT) and ventricular fibrillation (VF).

In the atrium of the rabbit, different types of reentries occur \[3\]. One reentry
may be anatomical or functional and orderly or disorderly:

- A reentry is anatomical when the excitatory movement is fixed and surrounds a non-exitable anatomical obstacle.

- A reentry is functional when it surrounds a zone of functional occlusion which is a non-exitable zone due to the refractoriness at the moment when the stimulus arrives [3].

- The reentry is orderly if the place and the size of the reentry circuit is well located and defined, otherwise it is considered a disorderly reentry.

The following conditions are necessary to be met for a reentry to occur:

1. The length of the circuit \( L \) must be higher than the result of multiplying the conduction velocity \( \text{CV} \) and the refractory period \( \text{RP} \), which is called excitation wavelength \( \lambda \) ( \( L > \text{CV} \cdot \text{RP} = \lambda \) ) [84].

2. A unidirectional block must occur.

Thus, the mechanism responsible for the emergence and repetition of the reentries are, on the one hand, the decrease in the conduction velocity, which causes a decrease in \( \lambda \), and, on the other hand, the emergence of inhomogeneities in the refractory period and the duration of the AP, which are key in the unidirectional conduction block.

Different types of reentries have been experimentally registered. Reentries with the shape either of an eight [51] or as a spiral [6] during the early stage of the ischemia (see Figure 1.8).

Both types of reentries have been reproduced in simulation by previous stimuli which are applied to two dimensions (2D) tissues, which are either healthy or ischemic: The number eight shaped circular have been obtained by stimulating in the center of the simulated tissue, for healthy tissues models [88], and by stimulating from the ends for the ischemic ones [29].

The reentries in the shape of a spiral can be obtained by applying "cross-shock", which is about applying a previous stimulus perpendicularly to the former stimulus [23, 6].
Figure 1.8: Patterns of excitation on a pig animal model. A single spiral is a responsible for the continuation of the tachycardia with a pattern from a single rotor to a figure of eight changing from beat to beat. Figure take from [50]

1.6.2 Refractory period in ischemic cells

As seen in previous section, the conjunction of a early and special tissue conditions, which are the base for the reentry is necessary to generate reentries. In this sense, the factors that most influence in the generation of reentries are heterogeneity in: the conduction velocity, the excitability of the cell membrane and the cell refractoriness [21, 55]. The alteration induced by these factors may cause unidirectional block, which is an essential requirement to generate reentries.

In normal conditions, the RP generally matches the APD, so it may be thought that the repolarization of the cell set an end to the RP, as it can be observed in Figure 1.9. However, under some circumstances such as the acidosis or the hyperkalemia, the RP is significantly higher than the APD [51]. This phenomena is called Post
Repolarization Refractoriness (PRR).

![Diagram showing AP and E.R.P](image)

Figure 1.9: AP of a healthy cell and an ischemia one cell and its respective refractory periods.

In acute ischemia the conduction velocity is modified, the excitability of the cell membrane and the refractoriness of the cells, which are the factors whose heterogeneity influences the most on the generation of the reentries.

On the one hand, as regards as the RP, it is experimentally known that, in acute ischemia such parameter increases in the 5-10 minutes after the occlusion of the coronary artery, developing the PRR, and that in those partially depolarized cells, the parameter depends on the resting potential [51]. This refractoriness is related to the increase of the extracellular potassium concentration, the hypoxia and the acidosis [55].

On the other hand, the reduction of the excitability in the acute ischemia due to the decrease of the fast sodium current and the slowing of its recovery increases the vulnerability window (VV) [21, 83, 55]. The vulnerability window is the period in which if a extra-stimulus is applied, re-entrant activity is generated.

Therefore, the ischemia progresses, the vulnerability window may increase or decrease due to the large increase of the scattering of refractoriness linked to a decrease in excitability [83]. As far as regarding the ionic currents, it is known that the vulnerability window is affected by the changes in availability of the sodium channels and the potassium currents [55].

Another aspect to take into account during ischemia is the spatial distribution
of refractory periods. During the ischemia a gradient appears from the border to the center of the affected area by that pathology (ischemia), mostly due to the accumulation of extracellular potassium, in either acute ischemia [20, 21, 22] or ischemia 1b [37].

All aspects discussed in previous sections may be represented in mathematical equations. These will be discussed in detail in chapter 2. The nature of these systems of equations makes the use of numerical methods necessary to be solved.

1.7 Numerical approaches

Solving the electrophysiology problem represents a real challenge due to the different time scales involved in the problem, aspects that will be explained in detail in the next chapter. In general lines, the numerical solution of the mathematical equations implicated in this problem asks for time and space discretizations. The main techniques used to solve this problem can be summarized in three groups:

- **Finite difference method.** The spatial derivatives are approximated by finite difference expressions [26, 82, 92, 106]. It stands out for its simplicity of implementation, as well as by low memory requirements, but its counterpart is that its application is very difficult in the case of real geometries.

- **Finite elements method.** It is an integral formulation of the problem where the result is approximated by using a linear combination of test functions which are generally polynomials and temporary approximation is made by finite differences [18, 19, 104]. This method is known for its strength and to deal with any type of geometry.

- **Finite volume method.** It is similar method to the finite element method as it also begins from an integral formulation and like this, this method can be used in arbitrary geometries. An additional characteristic is low continuity (ionic current in our case) from one cell to its neighbor cell, a fact that is not necessarily guaranteed in a finite element formulation [40, 48].
1.7.1 Finite Element Method.

The Finite Element Method is a method to numerically solve differential equations and is based on an approximation of the exact solution for a linear combination of test functions. These functions are typically linearly independent polynomials that satisfy the boundary conditions. Mathematically the finite element method can be treated as a weighted residual procedure of Galerkin type, or as a means of obtaining an approximate solution to a weakly formulated problem. Finally, it can also be interpreted as a case of applying the method of Ritz, i.e., as a rough way to get the solution of problems relating to a variational principle [46, 116].

The method of the weighted residuals to the problem of one-dimensional diffusion reaction is to find a function \( V(x, t) : \Omega \times [0, T] \longrightarrow \mathbb{R} \), with \( \Omega = [0, L] \) and \( t \in [0, T] \), such that:

\[
C_m V_t = D V_{xx} - I_{ion}(V, t),
\]

where \( I_{ion} : \Omega \times ]0, T[ \longrightarrow \mathbb{R} \) is a continuous function. The solution should, in turn, meet the initial and boundary conditions.

\[
V(x, 0) = f(x), \quad 0 < x < L, \ t = 0
\]

\[
V_x = 0, \quad x = 0, \ x = L, \ t \geq 0.
\]

As seen above, the finite difference solution is based on an approach by points to the unknown function and the approach extends to the entire domain of calculation by the interpolation between the values obtained at these points.

In the weighted residue method, it is thought that a function is extended to the entire domain, which is defined in terms of unknown coefficients. The procedure could be:

1. Approximate the solution \( V \) in the entire domain of calculation by:

\[
V^h = N_o + \sum_{j=1}^{n} c_j N_j,
\]

where \( c_j \in \mathbb{R} \times ]0, T[ \) are unknown variable coefficients, \( N_o : \Omega \times ]0, T[ \longrightarrow \mathbb{R} \) is a known function satisfying the boundary conditions of the problem (equation (1.2)) and \( N_j : \Omega \times ]0, T[ \longrightarrow \mathbb{R} \) are also known functions but they meet the homogeneous boundary conditions.
2. Determine the variable coefficients $c_j$, so that the differential equation (1.1) is met in weighted basis in the entire domain $\Omega$.

3. One way to do this is to take $n$ functions $w_k$, $k = 1, \cdots, n$ (weighting functions) and raise the $n$ equations:

$$
\int_0^L w_k \left( C_m V^h_t - D V^h_{xx} + I_{ion} \right) \, dx
$$

(1.4)

This system of equations provides a means to determine the variable coefficients $c_j$, $j = 1, \cdots, n$.

This procedure is known with the name of the weighted residue method, as the second factor is the subintegral amount is called the residue function. The method of the weighted residues is very general and very powerful. This technique includes procedures such different, at first sight, as the FEM (finite element Method), and the boundary element method or the least squares adjustment. The generality is that there are many possibilities to choose the $N_j$ approximation functions and the weighting functions $w_k$.

The Galerkin method is the weighted residue method corresponding to the more classic formulation of the Finite Element Method. According to this method, the $V^h$ approximation is constructed as:

$$
V^h = g^h + \sum_{j=1}^{n} c_j N_j
$$

(1.5)

where: $c_j \in \mathbb{R}$, are variable coefficients, $g^h : \Omega \times ]0,T[ \rightarrow \mathbb{R}$ is a function which meets the boundary condition, $N_j : \Omega \rightarrow \mathbb{R}$ are functions that are Zero in the boundary. The weighting functions $w_k : \Omega \rightarrow \mathbb{R}$, are chosen equal to the $N_j : w_j = N_j$, $j = 1, \cdots, n$. In the Galerkin method, the same space as in the approximation functions is used as a weighting function space.

To simplify, grouping in $F(V,t)$ the reactive term and the capacitive term, the weighted residues equation (1.4) results into:

$$
\int_0^L N_k \left[ F(V,t) - D V^h_{xx} \right] \, dx = 0.
$$

(1.6)

ponderación en el dominio
Integrating by parts and considering we have homogeneous Neumann boundary conditions, we get:

\[ \int_0^L \left[ N_k F(V,t) + (N_k)_x D V_x^h \right] \, dx = 0 \quad (1.7) \]

where the integration formulae considered by parts is

\[ \int_0^L w \, dv = - \int_0^L v \, dw + [w v]_0^L. \]

From the equation (1.7) a system of de n linear equations with n variables which allows to determine by the \( c_j \) coefficients of the equation is obtained (1.3).

The Galerkin method dates from 1915 and it is a first way of justifying the Finite Element Method from mathematical point of view. The contribution of the modern Finite element method to the Galerkin method consists of easily automatable systematic way, of constructing the \( N_j \) and \( g_h \) functions from locally defined functions.

**Steps to solve a problem using the Finite Element Method.**

In the FEM the domain is divided in a finite number of sub-domains, called elements and the Galerkin method is applied to each element. The global solution is obtained by assembling all the elements. The Finite Element Method can be subdivided into the following steps:

- Discretization of the spatial Domain.
- Functions of interpolation or approximation to solution.
- Determination of the elemental equation.
- Assembly of the algebraic system of equation.
- Solving the system equations.
- Visualization of the results.

**Spatial Discretization.** The spatial discretization os mesh in the Finite Element Method consists of subdividing the domain of the \( \Omega \) calculation with a finite number of elements of the \( \Omega^e \) size which take the entire domain:
\[ \Omega = \bigcup_{e \in \{1, \ldots, N_{elm}\}} \Omega^e \]

where \( N_{elm} \) is the number of elements. Moreover, we must add that the elements don’t overlap or intersect. The shape of the element is very simple and they only vary in shape and size. The 2D most commonly used elements are the triangle and the quadrangle, whereas 3D tetrahedrons and hexahedrons are used.

Spatial discretization is very important in the accuracy of the analysis by the Finite Element Method, the smaller the element is, the more accurate the solutions is. When items are added to increase the accuracy, is called refinement and this type of refinement is called \( h \).

**Interpolation Functions.** The interpolation functions allow to determine the solution in an arbitrary point of the mesh. In the FEM, these functions are local and they are only defined within the domain of the element \( \Omega^e \) and weighted by the value of the nodal variables. The election of the Shape functions is subject to the desired order approach and to the spatial dimension of the problem. The interpolation or shape functions must meet the following conditions:

- The value of the shape function in its node must be one, that is:
  \[
  N^e_i(x) = \delta_{i,j} = \begin{cases} 
  1 \Rightarrow x = x_i \\
  0 \Rightarrow x = x_j, \ j \neq i.
  \end{cases}
  \]

- The sum of the shape functions in an arbitrary node is one, that is:
  \[
  \sum_{i}^{N} N^e_i(x) = 1
  \]

Thus, for a space of dimension 1, i.e., \( \Omega = [-1, 1] \), the lineal shape functions are:

\[
\{ \Psi^e \} = \frac{1}{2} \begin{cases} 
  1 - x \\
  1 + x
  \end{cases}
\]
Elemental equation. Considering again the one dimensional problem, the weak formulation of the equation (1.2) in an $\Omega^e$ element is obtained multiplying such equation by the weight function $w^e$ and integrating over the element

$$
\int_{\Omega^e} w^e [(C_m V_t + I_{ion}) - D V_{xx}] d\mathbf{x} = 0. \tag{1.8}
$$

Integrated by parts, we obtain the weak formulation of the equation (1.8)

$$
\int_{\Omega^e} [w^e (C_m V_t + I_{ion}) + D w^e_x V_x] d\mathbf{x} - \oint_{\Gamma^e} D w^e V_x d\mathbf{s} = 0, \tag{1.9}
$$

The discretization of the equation (1.9) is obtained by replacing the approximation of $V$ as

$$
V(x, t) \simeq \sum_{j=1}^{n^e} N_j^e(x) V_j^e(t) \tag{1.10}
$$

where $n^e$ is the number of nodes per element. In the matrix form, this approximation can be written as:

$$
V(x, t) = (N^e)^T \mathbf{V}^e.
$$

Replacing the $V$ approximation in (1.9) and solving $w_j^e = N_j^e$ the discretizing form in a time $t$ for the node $i$ of the element is obtained.

$$
\sum_{j=1}^{n^e} \left( M_{i,j}^e \dot{V}_j^e + K_{i,j}^e V_j^e \right) = Q_i^e. \tag{1.11}
$$

In the matrix form we observe

$$
[M^e] \left\{ \dot{\mathbf{V}}^e \right\} + [K^e] \{\mathbf{V}^e\} = \mathbf{Q}^e \tag{1.12}
$$

where:

$$
M_{ij}^e = \int_{\Omega^e} C_m N_i N_j d\mathbf{x}, \tag{1.13}
$$

$$
K_{ij}^e = \int_{\Omega^e} D (N_i)_x (N_j)_x d\mathbf{x}. \tag{1.14}
$$
\[ Q_i^e = \int_{\Omega_e} N_i I_{iom} \, dx \] 

The term obtained because of the boundary integral in (1.9) is not included in the discretized equations as it does not contribute in the assembled system. Note that this term is void among the elements which share faces and it is exactly satisfied in the boundaries with homogeneous Neumann conditions as it is the case under consideration.

**Assembly of the system of equations.** The elementary matrix and vectors are assembled to generate a system of equations which describes the entire domain. The assembled matrix are represented as:

\[
M = \mathbb{A}^{N_{elm}}_{e=1} M^e \\
K = \mathbb{A}^{N_{elm}}_{e=1} K^e \\
F = \mathbb{A}^{N_{elm}}_{e=1} Q^e
\]  

where \( N_{elm} \) is the number of elements. The assembled equation has the following form:

\[
M \dot{V} + KV = F
\]

**System solution** To solve the system of equations in the equation 1.17, you can use direct or iterative methods, but we will refer to them later in this chapter.

**Advantages and disadvantages**

One of the biggest advantages of the FEM is that unstructured meshes can be used, with which the time of generating a mesh is considerably reduced. On the other hand, the size of the element is not necessary need not be uniform. The latter is important when there are regions with large gradients when it is necessary to use small elements, whereas in regions with smooth gradients, bigger size elements can be used. A further and important advantage is that no changes need not be made when the geometry is irregular and complex. In addition, the method is more accurate when there are problems with the irregular gradients, since you can take advantage of the above mentioned.

As disadvantage, the unstructured mesh requires a greater use of memory for the storage of the data, which affects CPU times as regards as the order of data.
1.8 Motivation

In the last decades, mathematical modelling and computer simulations have become a useful tool for tackling problems in science and engineering, among which heart electrophysiology. In this regard, modeling the electric activity of the heart, under physiological and pathological conditions, has attracted the attention of a number of researcher [47] since ventricular tachycardia and fibrillation are among the major causes of sudden death [69]. Since direct measurements are many times limited to only surface signals, multi-scale numerical simulations where, the electrical activity at the surface, as well as in the myocardium, can be related to the underlying electro-chemical behavior of the cell, helps to gain further insights into the problem.

The electric activity of the heart is usually studied using the well known bidomain model [33, 43]. It consists of an elliptic partial differential equation and a parabolic partial differential equation coupled to a system of stiff non-linear ordinary differential equations describing the ionic current through the cellular membrane. This model can be simplified to the so called anisotropic monodomain equation [33], a parabolic reaction-diffusion equation describing the propagation of the transmembrane potential coupled to a system of ordinary differential equations describing the cellular ionic model. The monodomain model represents a much less computationally expensive model for the electric activity of the heart, and has been extensively used [18, 27, 41, 102].

The high computational cost of the bidomain and monodomain models is due to the stiffness of the system of ordinary differential equation describing the transmembrane ionic current which introduces different space and time scales. The depolarization front is localised in a thin layer of less than a millimetre, requiring therefore discretizations of the order of tenth of a millimetre in order to accurately resolve the depolarization front, implying models with millions of degrees of freedom to simulate the heart. The time scale is another fundamental issue in cardiac simulations. The time constants involved in the kinetics of cellular models range from 0.1 to 600 ms, requiring in some phases of the process the use of time steps of the order of a hundredth of a millisecond. Hence, solving a single heart beat requires thousands time steps.

A number of alternatives have been proposed to solve this problem. In this particular, the multi-length scale nature of the problem has inspired the development of adaptive techniques where the mesh is allowed to change with time coupled with
adaptive time integration schemes to improve the computational performance [15, 19, 16]. However, dynamic loading for these adaptive schemes is still cumbersome, limiting their application in massive parallel architectures. Recent efforts [7, 8, 110, 111] suggest the use of multilevel meshes, fixed in time, along with adaptive time schemes which take advantage of the different kinetics of the ionic currents. This allows for reductions of up to two orders of magnitude in CPU time with respect to traditional explicit algorithms. However, these techniques use the fine mesh (lower level mesh) for solving the partial differential equations (responsible for the propagation of the depolarization front).

Despite of the efforts of designing more efficient schemes, the solution of the electrophysiology problem requires the use of algorithms with higher level of parallelism in multi-core platforms. In this regard, the next generation of high-performance computing (HPC) platforms promise to deliver better performance in the PetaFLOPS range. However, achieving high performance on these platforms relies on the fact that strong scalability can be achieved, something challenging due to the performance deterioration caused by the increasing communication cost between processors as the number of cores increases. That is, with increasing number of cores, the load assigned to each processor decreases, but the communication between different processors associated with the boundaries of a given partitioned domain increases. Therefore, when communication costs domain, no further benefits are obtained from adding additional processors. An alternative to the multi-core platforms is emerging in the newer programmable graphic processing units (GPU) which in recent years has become a highly parallel, multithreaded, many-core processor with tremendous computational horsepower [10, 73]. GPUs outperform multi-core CPUs architectures in terms of memory band width, but underperforms in terms of double precision floating point arithmetic. However, GPUs are built to schedule a large number of threads, thus, reducing latencies in their multi-core architecture.

Sanderson et al. [93] proposed a general purpose, graphics processing unit (GP-GPU)-based approach for the solution of advection-reaction-diffusion models. They report an increase of performance of up to 27 times for an explicit solver when used on three dimensional problems. Regarding cardiac electrophysiology, previous studies have reported speedups by a factor of 32 for the monodomain model [94] using an explicit finite difference scheme with a rather simple transmembrane ionic model. In their study Sato et al. [94] have established the solution of the PDE equation as the bottle neck of the computation with GPU. However, in their studies older NVidia
Modelling and simulation of the electric activity of the heart using GPU

Andrés Mena Tobar

GT8800 and GT9800 GX2 cards only supporting single precision floating point operations were used, that greatly limited the computations of the parabolic system. Chai et al. [14] successfully solved a 25 million nodes problem on a multi-GPU architecture using the mono domain model and a four state variables model. Bartocci et al. [4] have performed an implementation of a finite difference explicit solver for cardiac electrophysiology. They evaluate the effect of the ionic model size (number of state variables) on the performance in simulating 2D tissues, and compare single precision and double precision implementation. They provided acceleration with respect to real time. For small ionic models and the single precision implementation they report simulations faster that real time for small problems, whereas for highly detail models, with larger number of state variables, they report simulation times between 35 and 70 times larger than real time. Rocha et al. [85] implemented and an implicit method on the GPU. Spatial discretization of the parabolic equation was performed by mean of the Finite Element Methods (FEM) keeping full stiffness matrices. Promising acceleration ratios were achieved with two dimensional (2D) bidomain tissue models using an unpreconditioned CG method. However, with unstructured three dimensional (3D) bidomain simulations, the number of iterations required for convergence became prohibitive. In a more recent work Neic et al. [71] showed that 25 processors were equivalent to a single GPU when computing the bidomain equations. This new capability to solve the governing equations on a relatively small GPU cluster makes it possible to one day introduce simulation using patient specific computer models into a clinical workflow. On a more recent work, Vigueras et al. [107] have port to the GPU a number of components of a parallel c-implemented cardiac solver. They report accelerations of 164 times of the ODE solver and up to 72 times for the PDE solver. They have also achieved accelerations of up to 44 times for the mechanics residual/Jacobian computation in electromechanical simulations.

When dealing with the pathological heart, ventricular tachycardia and fibrillation are known to be two types of cardiac arrhythmias that usually take place during acute ischemia and frequently lead to sudden death [89]. Even though these arrhythmias arise from different conditions, ischemia is the most important perpetrator among them. During ischemia, the delivery of nutrients to the myocardium diminishes, causing metabolic changes, which result in a progressive deterioration of the electric activity in the injured region [51]. These metabolic changes are mainly hypoxia, increased concentrations of the extracellular potassium $[K^+]_o$ (hyperkalemia),
decrease of intracellular ATP (hypoxia), and acidosis [12]. From an electrophysiological point of view, these metabolic changes simply alter the action potential (AP), excitability, conduction velocity (CV), and effective refractive period (ERP) among others, creating a substrate for arrhythmias and fibrillation [12, 51]. In addition, the impact of ischemia in the myocardium is characterized by a high degree of heterogeneity both intramurally and transmurally. In the tissue affected by acute ischemia, two zones can be distinguished: i) the central ischemic zone (CIZ) corresponding to the core of the tissue suffering from the lack of blood, and ii) a border zone (BZ) which comprises changes in electrophysiological properties between the healthy and ischemic regions [20, 49]. Pro-arrhythmic mechanisms of acute ischemia have been extensively investigated, although often in animal models rather than in human ventricles. Seminal studies by Janse et al. [50, 51] in pig and dog hearts highlight the complexity of the proarrhythmic and spatio-temporally dynamic substrate in acute ischemia. Heterogeneity in excitability and repolarization properties across the border zone leads to the establishment of reentry around the ischemic region following ectopic excitation [51, 112]. The same studies also showed intramural reentry in certain cases (highlighting the potential variability in the mechanisms). However, the mechanisms that determine reentry formation and intramural patterns in acute ischemia in the three-dimensional human heart remain unclear, due to low resolution of intramural recordings.

1.9 Objectives

The main objective of the thesis is to develop an efficient solver for the solution of electrophysiology applications using the monodomain model in parallel architectures based on Graphic Processing Units (GPU).

The developed software will be used to study the acute ischemic heart. Specific objectives related to this application are:

- To study the effect of different extracellular potassium, $[K^+]_o$, on the vulnerable window of acute ischemic heart.

- To study the influence of the location of the ectopic beat on the vulnerable window in the acute ischemic heart.
• To study the different depolarization patterns and reentry mechanisms on a human acute ischemic heart.

1.10 Hypothesis of work and Methodology.

The new contributions on this work are:

• 1. GPU solver. The proposed monodomain solver will be entirely developed in CUDA at both levels: i) integration of the ionic model, and ii) solution of the algebraic system of equations using either explicit and implicit schemes. The ionic model will implement the ionic models using the structure adopted by the CellML standard (www.cellml.org) in order to facilitate the incorporation of new models in the current solver database. The full CUDA implementation will increase compatibility while reducing developing time by minimizing the interaction with third-party libraries used for most common parallel applications from the partition of the computational domain (METIS and ParMETIS) and the solution of the system of equations (PETC, PSBLAS).

• 2. Depolarization patterns in the acute ischemic heart. The reentry patterns and vulnerability window on a acute ischemic heart will be computed, for the first time, as a function of the extracellular potassium concentration, $[K^+]_o$ and the location of the ectopic beat. In addition, the effect of the presence of a washed zone in the endocardium (as reported by Wilensky et al. [112]) on the vulnerable window and reentry patterns will also be studied and reported. This study will contribute to a better understanding of the underlying mechanisms of the arrhythmias caused by acute ischemia.
Equations that govern the electrical activity of the heart

The heart shows two types of behavior, one is electrical and the other is mechanical. All myocardic cells are similar as regards as mechanic function. However, from an electrical point of view, these (the cells) may be of several types. The electric impulses transmitted through the heart are responsible for the rhythmic contraction of the heart muscle/cardiac muscle. When the system working normally, the atria are contracted approximately a sixth of a second before the ventricles, which enables the filling of the ventricles before pumping the blood into the lungs and the peripheral circulation [38]. Another important point of the system is that the ventricles contract synchronously to generate a proper blood pumping. Therefore, all cells need to develop an action potential in an orderly manner, for which the cells must be excited conveniently along the cardiac cycle. To fully understand these phenomena, this chapter describes how the electrical activity in the heart takes place, how synchronizes and the mathematical equations that rule them.

2.1 Governing equations

This section describes the equations governing the propagation of the electric activity of the heart.
2.1.1 Bidomain Model

The electrical coupling of the cardiomyocytes and the conduction through the ventricles is mathematically described by the bidomain model [105]. In this model, the cardiac heart tissue is represented by two continuous domains which share the space, that is, intracellular and extracellular domain coexist spatially. This oppose to reality because each of them physically takes a fraction of total volume. In this model, each domain acts as a volume conductor with a different conductivity tensor and different potential, and the ionic currents flow from one domain to another through the cell membrane which acts as a condenser.

The currents in the two domains are given by Ohm’s Law:

\[ J_i = -M_i \nabla V_i, \]  
\[ J_e = -M_e \nabla V_e, \]  

(2.1)
(2.2)

where \( J_i \) is the intracellular current, \( J_e \) is the extracellular current, \( M_i \) and \( M_e \) are the conductivity tensors and \( V_i, V_e \) are the intracellular and extracellular potentials respectively.

The cell membrane acts as a condenser and owing to its small thickness the charge stored on one side is compensated immediately on the other side, by which the accumulation of charge at any point is zero, that is:

\[ \frac{\partial}{\partial t} (q_i + q_e) = 0, \]  

(2.3)

where \( q_i \) and \( q_e \) are the charges in the intracellular and extra-cellular space respectively.

In each domain, the flow of current in a point must equal the rate of accumulation plus the ionic current coming out the point, that is:

\[ -\nabla \cdot J_i = \frac{\partial q_i}{\partial t} + \chi J_{ion}, \]  
\[ -\nabla \cdot J_e = \frac{\partial q_e}{\partial t} - \chi J_{ion}, \]  

(2.4)
(2.5)

where \( J_{ion} \) is the current through the membrane. As the ionic current is measured by unit of area of the cellular membrane, whereas the density of charge and the flow of current are measured by unit of volume. The constant \( \chi \) represents the cell
membrane area to volume ratio. On the other hand, the sign of the ionic current is defined as positive when the current leaves the intracellular space and gets into the extracellular.

Introducing (2.4) and (2.5) in (2.3) we get the current conservation equation:

\[ \nabla \cdot J_i + \nabla \cdot J_e = 0. \] (2.6)

Replacing the equations (2.1) and (2.2) in (2.6) we obtained:

\[ \nabla \cdot (M_i \nabla V_i) + \nabla \cdot (M_e \nabla V_e) = 0. \] (2.7)

The charge of the cell membrane directly depends on the difference of the membrane potential, \( V = V_i - V_e \) and the capacitance of the membrane

\[ V = \frac{q}{\chi C_m}, \] (2.8)

where \( C_m \) is the membrane capacitance and

\[ q = \frac{q_i - q_e}{2}. \] (2.9)

Combining the equation (2.8) and (2.9) and using the derivative with respect to time we have:

\[ \chi C_m \frac{\partial V}{\partial t} = \frac{1}{2} \frac{\partial (q_i - q_e)}{\partial t}. \]

Using the equation (2.3) we get the relation:

\[ \frac{\partial q_i}{\partial t} = -\frac{\partial q_e}{\partial t} = \chi C_m \frac{\partial V}{\partial t}. \]

Replacing this latter expression in equation (2.4) and using equation (2.1) we obtain

\[ \nabla \cdot (D_i \nabla V_i) = C_m \frac{\partial V}{\partial t} + J_{\text{ion}}, \] (2.10)

where \( D_i = M_i / \chi ). \)

The equations (2.7) and (2.10) depend on three potentials \( V_i, V_e \) and \( V \). Eliminating \( V_i \) from the equations (2.7) and (2.10) the equations, for the bidomain model are obtained:

\[ \nabla \cdot (D_i \nabla V) + \nabla \cdot (D_i \nabla V_e) = C_m \frac{\partial V}{\partial t} + J_{\text{ion}}, \] (2.11)
\[ \nabla \cdot (D_i \nabla V) + \nabla \cdot \left( (D_i + D_e) \nabla V_e \right) = 0. \] (2.12)

Assuming that the heart is surrounded by a non-conductive medium, the normal component of both currents (intracellular and extracellular) are zero at the boundary, by which we have:

\[ \mathbf{n} \cdot J_i = 0, \]
\[ \mathbf{n} \cdot J_e = 0 \] (2.13)

where \( \mathbf{n} \) is the outer normal. Using the expression for both currents and eliminating \( V_i \) the boundary conditions of the model are obtained

\[ \mathbf{n} \cdot (D_i \nabla V + D_i \nabla V_e) = 0, \] (2.14)
\[ \mathbf{n} \cdot \nabla (D_e \nabla V_e) = 0. \] (2.15)

## 2.1.2 Monodomain Model

As it can be observed, the bidomain model represents the electric currents in both the intracellular and extracellular medium, and it consists of a non-linear parabolic equation coupled with an elliptical equation. Under particular conditions, the bidomain model can be decoupled allowing to calculate the transmembrane potential independently of the extracellular potential. Assuming that the conductivity tensors have the same variation in the anisotropy, that is \( D_e = \lambda D_i \), where \( \lambda \) is a scalar, so \( D_e \) can be eliminated from the equations (2.11) and (2.12), obtaining:

\[ \nabla \cdot (D_i \nabla V) + \nabla \cdot (D_i \nabla V_e) = C_m \frac{\partial V}{\partial t} + J_{ion}, \] (2.16)
\[ \nabla \cdot (D_i \nabla V) + (1 + \lambda) \nabla \cdot (D_i \nabla V_e) = 0. \] (2.17)

from the equation (2.17) we have

\[ \nabla \cdot (D_i \nabla V_e) = -\frac{1}{1 + \lambda} \nabla \cdot (D_i \nabla V), \]

replacing in the equation (2.16) and operating we get to the standard formulation for the monodomain model:

\[ \nabla \cdot (D \nabla V) = C_m \frac{\partial V}{\partial t} + I_{ion} \] (2.18)
where $D = \frac{\lambda}{1+\lambda} D_1$. With the following boundary condition:

$$\mathbf{n} \cdot \nabla (D \nabla V) = 0,$$  \hspace{1cm} (2.19)

The monodomain model takes a numerical and computational complexity which is less than the bidomain. For this reason, the monodomain model is often used to study the propagation of the action potential in the Heart.

The equation (2.18) is a parabolic equation describing a reaction-diffusion phenomenon. The part associated with the reaction is determined by the term $J_{\text{ion}}$ which is governed by the cellular model. The diffusive (or conductive in this case) part models the propagation of the AP in the tissue.

### 2.2 Myocardium conductance.

Within the heart, the fibers are organized transmurally with orientations varying from $-60^\circ$ (regarding the circumferential axis) to $+60^\circ$ from epicardium to the endocardium [99]. The orientation of the muscle fibers in each point of the myocardium can be obtained either by histology, or more recently using magnetic resonance imaging (MRI), in particular a technique known as diffusion tensor magnetic resonance imaging (DTMRI) [42]. Born for neuroimaging applications, DTI, a special kind of the more general diffusion weighted MRI (DW-MRI), is an imaging method that uses the diffusion of water molecules to generate contrast in MR images. Since diffusion of water in tissues is not free, but is affected by the interaction with obstacles such as fibers and heterogeneities in general, water molecule diffusion patterns can be used to identify details about tissue microstructure. DTI in particular, enables the measurement of restricted diffusion of water in the myocardium. In DTI, each voxel contains the rate of diffusion and the preferred directions of diffusion. Therefore, assuming that the diffusion is faster along the fiber axis, the eigenvector corresponding to the largest diffusion tensor eigenvector defines the direction of the fiber axis [42]. These are the data that we have to perform our simulation in the human heart tissue \footnote{http://gforge.icm.jhu.edu/gf/project/dtMRI_data_sets/}. Even though the heart tissue is truly orthotropic [57], for this work we consider it as transversely isotropic, with the direction of maximum conduction corresponding to the cardiac fiber direction. In the material fiber system,
the conductivity tensor is
\[
D = d_o \begin{pmatrix} 1 & 0 & 0 \\ 0 & r & 0 \\ 0 & 0 & r \end{pmatrix},
\]  
(2.20)

where \(d_o\) represents the conductance in the fiber direction and \(r \leq 1\) the conductivity ratio between the transversal and longitudinal fiber. In Cartesian coordinates, under conditions of transverse anisotropy, the diffusion tensor can be written as:
\[
D = d_o [(1 - r) f \otimes f + r I],
\]  
(2.21)

where \(f\) is the fiber orientation, \(I\) is second order identity tensor and \(\otimes\) indicates the tensorial product \(((a \otimes b)_{ij} = a_i b_j\)). Expressing the equation (2.21) in components, we obtained:
\[
D = d_o (1 - r) \begin{pmatrix} f_1 & f_1 & f_1 f_3 \\ f_2 & f_2 f_1 & f_2 f_3 \\ f_3 & f_3 f_1 & f_3 f_3 \end{pmatrix} + d_o r \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.
\]  
(2.22)

\section{2.3 Action Potential Models}

Hodking and Huxley \[45\] in 1952 introduced the first mathematical model to reproduce the action potentials in the cell membrane. Since then, many cardiac cell models following the formalism established by these researchers have been developed. AP models can be divided in two main families: i) phenomenological models, and ii) electrophysiological detailed models.

Phenomenological models macroscopically reproduced the behavior of the cell in terms of the shape and duration of the action potential, restitution properties, and conduction velocity. Electrophysiological detailed models offer a detailed description of the cellular physiology. They not only include more currents, but they also include pumps and exchangers as well as intracellular ion concentration dynamics. Models have been developed for a number of species as well as cell-types within the conduction system of the heart, as for example: Stewart et al. \[97\] for human Purkinje cells, Maleckar et al \[63\] and Nygren et al. \[76\] for human atria, tenTusscher et al. \[101, 102\], O’Hara et al. \[78\], Carro et al. \[13\] for human ventricle, Luo and Rudy \[61, 62\] for guinea pig ventricle, and Shannon et al. \[95\] for rabbit ventricle among others.
In this thesis, we have implemented a number of AP models, namely: Maleckar et al. (MA09) and Nygren et al. (NY98) AP models for atria, tenTusscher et al. (TP06) and O’Hara et al. (OH11) for human ventricle, Stewart et al. (ST99) for Purkinje cells, and the Bueno Orovio et al. (BV08) as a type of phenomenological model. The TP06 model will be extensively used on Chapter 4 to study the acute ischemic heart. In the following, the basic structure of a modern AP model is described in detail.

2.3.1 Structure of an Action Potential model

As known, it is possible to reproduce the characteristics of AP (action potential) with simple models. However, an important objective in the modalization of physiological phenomena is to research how the changes on the cell physiology affect the tissue and finally, the organ studied. For that, it is necessary that the models study the cell physiology from membrane to the ionic channels which set the gates for the exchange between the intra-cellular and extra-cellular media, including dynamic mechanism in the mycoplasma.

2.3.2 The Cell Membrane

The cell membrane separates the extracellular medium from cytoplasm. It is formed by a very thin larger of lipid and protein molecules, which are mainly held linked by non-covalent interaction [1]. The lipid molecules are arranged in a continuous double layer of about 5 nm thick, see Figure 2.1. This bi-layer is the basic structure and acts as a relatively impermeable barrier to the passing of most water-soluble molecules.

The protein molecules are dissolved in lipid bi-layer, arranging a connection between the inside and the outside of the cell. The molecules form channels through the membrane through which the ions can pass. Figure 2.1 schematically shows the structure of the cell membrane with the transport proteins imbibed. Some proteins form pumps and ionic exchangers, needed to maintain the correct concentration of ions in the cell. Both pumps and exchangers have the ability to transport ions in an opposite direction to the gradient generated by the ionic concentration (electric gradient). This process is achieved by using either the gradient of a different ion (exchangers) or consume chemical energy stored in the form of ATP (pumps). This
type of transportation is called active.

Along with the pumps and exchangers, certain proteins form channels in the membrane through which the ions can flow. The flow of ions through these channels is passive and is governed by the concentration gradients and the electric fields. Most of the channels are alternatively selective as to the type of ion that can pass through them. Besides, the channels are capable of opening and closing themselves in response to the changes in the electric field and the ionic concentration. This characteristic is essential in order to propagate the signal in an excitable tissue.

**The Nernst Equation**

The ionic concentrations inside and outside the cell differ greatly in cardiomyocytes, i.e., a concentration gradient exists for all permeable ions contributing to net ion flux. This concentration gradient causes ions to flow, or diffuse, from regions of high to low concentration. This ion flow has as a consequence the accumulation of ions at the inner and outer membrane surfaces, establishing, therefore, an electric field (a potential difference) within the membrane. Since the ions are charged particles, this electric field exerts forces on the ions crossing the membrane that oppose the diffusional forces established by the difference in ionic concentration. Therefore, to describe membrane ion movements, electric-field forces and diffusional forces should be considered. In this regard, equilibrium is attained when the diffusional force balances the electric field force for all permeable ions.

For a membrane that is permeable to only one type of ion, the electrochemical balance between forces due to the concentration gradient and the potential gradient for a particular ion can be described by the Nernst-Planck-Einstein equation:
Chapter 2. Equations that govern the electrical activity of the heart

\[ E_k = \frac{RT}{z_k F} \ln \left( \frac{c_{k,e}}{c_{k,i}} \right) \]  

(2.23)

where \( R \), \( F \), \( T \) are: gas constant, the Faraday constant, and the absolute temperature constant respectively; \( E_k \) is the equilibrium voltage across the membrane (Nernst Potential) for the \( k^{th} \) ion, \( z_k \) is the valence of the \( k^{th} \) ion and \( c_{k,e} \), \( c_{k,i} \) are the extracellular and intracellular concentrations of the \( k^{th} \) ion.

**Goldman-Hodgkin-Katz Equation**

Assuming that the cell membrane is permeable to a single ion only is not valid. However, it is assumed that when several permeable ions are present, the flux of each is independent of the others (known as the independence principle). According to this principle, and assuming: i) the membrane is homogeneous and neutral, and ii) the intracellular and extracellular ion concentrations are uniform and unchanging, the membrane potential is governed by the well-known equation of Goldman-Hodgkin-Katz. For (the case of) \( N \) monovalent positive ion species and for (the case of) \( M \) monovalent negative ion species, the potential difference across the membrane is as follows

\[ E = -\frac{RT}{F} \log \left( \frac{\sum_{j=1}^{N} P_j [c_j^+]_i + \sum_{j=1}^{M} P_j [c_j^-]_e}{\sum_{j=1}^{N} P_j [c_j^+]_e + \sum_{j=1}^{M} P_j [c_k^-]_i} \right) \]  

(2.24)

where, \([c_j]_i\), \([c_j]_e\) are the intracellular and extracellular concentrations for the \( j^{th} \) ion, \( P_j \) is the permeability are the intracellular and extracellular concentrations for the \( j^{th} \) ion and \( E \) is the membrane potential. The permeability for the \( j^{th} \) ion defined as:

\[ P_j = \frac{D_j \beta_j}{h} \]

where \( h \) is the thickness of the membrane, \( D_j \) is the diffusion coefficient and \( \beta_j \) is the water partition coefficient of the membrane. Both \( D_j \) and \( \beta_j \) depend on the type of ion and the type of membrane.

2.3.3 Gates

Ion channels are specific units located in the cell membrane through which the ions can flow. These channels open and close in response to a potential differences or a change in ion concentration. The mechanisms by which these channels open and close are stochastic in nature and may involve complex processes. In the simplest
case, the channel, or the channel gate, is considered to be in only two possible states, either open or close. Assigned to these states, there is a possibility of opening $O$ and a possibility of closure $C$, being the transition between states stochastic. The density of opening channels is $[O]$ and the density of closed channels $[C]$. Further assuming that the density of channels, $[O] + [C]$, is constant. The change between open state and closes state can be described by a first-order reaction as:

$$C \xrightarrow{\alpha} O, \quad O \xrightarrow{\beta} C,$$

where $\alpha$ is the opening rate and $\beta$ is the closing rate of the channels. These rates depend on the membrane potential $V$ in general, even though could also be modulated by the ion concentration. By the law of mass action, the rate of change from the open state to the closed state is proportional to the concentration of channels in the open state, equally, the rate of change from the closed state to the open state is proportional to the concentration in the closed channels. Therefore, we obtain:

$$\frac{d[O]}{dt} = \alpha(V) [C] - \beta(V) [O].$$

Dividing this equation by the total density $[O] + [C]$ we obtain:

$$\frac{dg}{dt} = \alpha(V)(1 - g) - \beta(V)g,$$

where $g = [O]/([O] + [C])$ is the rate of open channels.

As $\alpha$ and $\beta$ depend on $V$, it is not possible to use a general solution of the equation (2.26). The equation (2.26) can be written as:

$$\frac{dg}{dt} = (g_\infty(V) - g)/\tau_g(V),$$

with $g_\infty = \alpha/(\alpha + \beta)$ and $\tau_g = 1/(\alpha + \beta)$. With $g_\infty$ and $\tau_g$ constants, the solution of the equation (2.28) is:

$$g(t) = g_\infty + (g_o - g_\infty) e^{-t/\tau_g}$$

where $g_o$ is the initial value of g, and $g_\infty$ is the stable state.
2.3.4 Ionic Channels

The current through an ion channel can be computed using Ohm’s law as the product of the channel conductance times the potential difference between the membrane potential and the equilibrium potential for the specific ion defined by the Nernst potential (Eq. 2.23)

\[ I_i = g_i (V - E_i) \]

where \( g \) is the permeability (or channel conductance) of the membrane to the ion \( i \). Depending on the type of ion, \( g \) can be either a constant or a time function and depending on the membrane potential, as well as on the ionic concentrations. In general, the conductance of a given channel is given as

\[ g = G_{\text{max}} O, \]

where \( G_{\text{max}} \) is the maximum channel conductance, and \( O \) is the probability that the channel is open.

As explained in the previous section, the cell membrane behaves as a condenser from an electric point of view due to its dielectric characteristics. In addition, the proteins dissolved in the cell membrane form specific units that allowed the ionic exchange between the intracellular and extracellular space. Further, some of these units, the ion channels, possess a large specificity to certain ions. Hence, from an electric standpoint, the electric current flowing across the cell membrane during activation can be described using a parallel conductance model as shown in Figure 2.2. It consists of a capacitive current plus different currents associated with the different ions under consideration.

In this model, each of these current components is assumed to independent i.e., each current utilizes its own channel. The modern notation considers a positive current, and potential, from the inside to the outside.

From the circuit in Figure 2.2, we obtain the following expression for the current through the membrane

\[ J_m = C_m \frac{dV}{dt} + J_{\text{ex}} + J_{\text{pump}} + \sum_{j=1}^{n} g_j (V - E_j), \quad (2.29) \]

where \( C_m \) is the capacitance of the membrane, \( V \) is the membrane potential, and \( g_j \) and \( E_j \) are the conductance and the Nernst potential for the ion \( j \) respectively, and \( J_{\text{ex}} \) and \( J_{\text{pump}} \) are currents associated with ion exchangers and ionic pumps (active transport elements) present in the membrane.
2.3.5 The tenTusscher Action Potential Model

Modern Action Potential Models incorporate many of the features formerly described and its formulation is reduced to the equation (2.29) in addition to differential equations (2.27) governing the dynamics of the gates and the ionic concentrations in the cytoplasm. The formulation of these models is based on experimental data collected from different animal species and are in general relatively costly from a computational view point due to the nonlinearity of their equations. Among the most widely used models of AP for human ventricle is the one proposed by tenTusscher et al. [101, 102] show schematically in Figure 2.3.

The ten Tusscher model was first introduced in 2004 (TP04) [101], this first version already specialized the model for the three main type of cardiomyocytes present in the human myocardium (endocardium, mid-myocardium and epicardium) as well as the ionic currents associated to the three main ions (sodium, potassium, and calcium), in addition to intracellular ionic homeostasis and calcium handling. The second version of the model was introduced in 2006 (TP06) [102] where the formulation of the intracellular calcium handling was modified by introducing a subspace to describe the calcium induced calcium release (CICR). Figure 2.3 shows a detailed schematic mode of the TP06 model.

The major innovation of the TP06 model with respect to its predecessor, is the modification of the calcium handling and the calcium current through the L-Type channels. The improved formulation of calcium handling includes a subspace, SS, between the SR and the cell membrane from which the release of calcium into
Chapter 2. Equations that govern the electrical activity of the heart

Figure 2.3: Ten Tusscher model 2006[102].

the cytoplasm is performed. Meanwhile, the release of calcium from the SR to SS is carried out through a mechanism controlled by the Ryanodine receptors whose dynamic is governed by a Markov chain of four states (see Figure 2.3). However, in the implementation of the TP06, the dimensionality of the Markov chain is reduced to two states in order to reduce the computational cost of the model. With the introduction of the SS, the flow of calcium through the L-type channels injects calcium in the SS, this being inactivated now by the concentration of calcium in the SS, $[Ca^{2+}]_{SS}$, instead of, $[Ca^{2+}]_{i}$. In addition, the new formulation includes two gates of inactivation depending on the potential, a slow and a fast one to accommodate additional experimental observations.

In this thesis the TP06 model has been used to study the electric activity of the human heart in normal and pathological conditions considering the trasmural heterogeneity (different proportions of cells of the epicardium, the endocardium and the mid-miocardium). Its response has also been featured under conditions of ischemia(hyperkalemia,acidosis and hypoxia). To study the effect of hypoxia a modification of the $I_{KATP}$ model of Ferrero [28]. Finally, this modified model has been used to study the electrical response of the heart under conditions of acute regional ischemia(details will be discussed in the following chapters). From
a computational point of view, the TTP06 model has 19 state variables, 14 ionic currents and requires of a minimum time integration step of 0.02 ms.

2.4 Numerical solution of the electric activity of the heart

As discussed at the beginning of the chapter, the monodomain model represents a major simplification of the bidomain model with important advantages for mathematical and computational analysis, which is suitable for studying the electrical behavior of the heart. This section discusses the numerical solution of the monodomain model described by Eq. 2.18 and Eq. 2.19.

From a mathematical and computational view, the problem defined by (2.16 - 2.17) corresponds to the solution of a linear partial differential equation, which describes the electrical conduction, coupled with a rigid nonlinear system of ordinary differential equations describing the transmembrane ion currents resulting in a problem of nonlinear reaction-diffusion. An efficient way to solve the equations (2.16 - 2.17) is by application of the splitting technique of operators [98]. The decomposition technique of operators has been applied to the mono-domain equations [100, 41]. The basic steps are summarized below.

- **Step 1:** Using $V(t)$ as the initial condition for integrating the equation

\[
\nabla \cdot (D \nabla V) = C_m \frac{\partial V}{\partial t} + J_{ion}(V,u), \text{for } t \in [t, t + \Delta t/2] \tag{2.30}
\]

- **Step 2:** Use the result obtained in Step 1 as the initial conditions to integrate

\[
C_m \frac{\partial V}{\partial t} = -J_{ion}(V,u) \tag{2.31}
\]

\[
\frac{\partial u}{\partial t} = f(u,V,t), \text{for } t \in [t, t + \Delta t] \tag{2.32}
\]

- **Step 3:** Use the result obtained in Step 2 as the initial condition for integrating

\[
\nabla \cdot (D \nabla V) = C_m \frac{\partial V}{\partial t} + J_{ion}(V,u), \text{for } t \in [t + \Delta t/2, t + \Delta t] \tag{2.33}
\]
In practice, steps 1 and 3 can be combined into one, except for the first increment. Therefore, after the initial increase, the algorithm has only two steps, Step corresponding to the integration of ordinary differential equations (Step 2) I and Step II corresponding to the integration of the homogeneous parabolic equation (Step 1 and 3).

- **Step I**: Using $V(t^k)$ as the initial condition to integrate equation

$$C_m \frac{\partial V}{\partial t} = -J_{ion}(V, u) - J_{stm}(t), \quad \text{for } t \in [t^k, t^k + \Delta t]$$

$$\frac{\partial u}{\partial t} = f(u, V, t),$$

- **Step II**: Use the result obtained in Step I as the initial conditions to integrate

$$C_m \frac{\partial V}{\partial t} = \nabla \cdot (D \nabla V), \quad \text{for } t \in [t^k, t^k + \Delta t]$$

### 2.4.1 Spatial-temporal discretisation

When performing Step II, the computational domain must be discretised in space by a mesh of either finite elements or finite differences to approximate the dependent variables of the problem, $V$ and $u$, which allows writing Eq. (2.35) as

$$M \ddot{V} + KV = 0, \quad (2.36)$$

minimise where $M$ and $K$ are the mass and stiffness matrices respectively, obtained by assembling individual element matrices over the entire computational domain.

The most well known algorithms for integrating in time the semi-discrete system (2.36) are members of the generalised trapezoidal family of methods [46]. Let $V^k$ and $\dot{V}^k$ denote vectors of the transmembrane potential and its time derivative at each nodal point of the mesh at time $t^k$, where $k$ is index of the time step, then at time $t^{k+1}$ we can write

$$M \ddot{V}^{k+1} + KV^{k+1} = 0, \quad (2.37)$$

$$V^{k+1} = V^k + \Delta t \dot{V}^{k+\theta}, \quad (2.38)$$

$$\dot{V}^{k+\theta} = (1 - \theta) \dot{V}^k + \theta \dot{V}^{k+1}, \quad (2.39)$$
where $\theta \in [0, 1]$ is a scalar parameter. Equations (2.37) through (2.39) can be combined to obtain an algebraic system of equations to determine $V^{k+1}$.

When using the operator splitting algorithm for solving the monodomain model, equations are solved in two steps. First, the electrophysiological cellular model

$$V^* = V^k - \Delta t \left( J_{ion}(V^k, u) + J_{stm}(t) \right), \quad (2.40)$$

is solved at each mesh point to obtain an intermediate transmembrane potential vector $V^*$ (Step I). Even though a forward Euler scheme has been used in (2.40), any other ODE solver can be used to calculate $V^*$. With this intermediate solution at hand, along with equations (2.38) and (2.39) and $M\dot{V}^k = -KV^k$ from the previous converged time increment, equation (2.37) becomes

$$M \frac{V^{k+1} - V^*}{\Delta t} = -K \left( \theta V^{k+1} + (1 - \theta) V^k \right), \quad (2.41)$$

or alternatively

$$\hat{K}V^{k+1} = \hat{b}, \quad (2.42)$$

where $\hat{K}$ is everything that multiplies onto $V^{k+1}$, and $\hat{b}$ contains the other terms in equation (2.41). Equation (2.42) is solved for the entire domain to obtained $V^{k+1}$ (Step II).

Hence, the basic algorithm at time $t^{k+1}$ can be summarised, as:

- **Step I**: Use $V^k$ as the initial condition to integrate equation (2.40) to obtain $V^*$

- **Step II**: Use the result obtained in Step I to solve (2.42) for $V^{k+1}$

For different values of the parameter $\theta$, different time integration schemes are obtained for integrating the discretised homogeneous parabolic equation system (2.41):

- $\theta = 0$ Forward Euler (conditionally stable).
- $\theta = 0.5$ Crank-Nicolson scheme (unconditionally stable).
- $\theta = \frac{2}{3}$ Galerkin Scheme (unconditionally stable).
- $\theta = 1$ Backward Euler (unconditionally stable).
Crank-Nicolson scheme is second order accurate in time, whereas the other are first order accurate in time. However, for \( \theta \geq 0.5 \) integration schemes are unconditionally stable.

As mentioned before, Step I can be performed using a backward difference approximation in time (implicit integration), or a forward difference approximation in time (explicit integration). Implicit integration requires the solution of a nonlinear system of equations at each point of the mesh making it computationally costly. However, it ensures the stability of the numerical solution. On the contrary, explicit integration is computationally cheaper but imposes more stringent conditions on the size of the time step in order to avoid numerical instabilities.

### 2.4.2 Integration of the mass matrix

For the standard finite element formulation, the elemental mass matrix, \( M_e \) for equation (2.41), is given by [46]

\[
M_{ij}^e = \int_{\Omega_e} N_i N_j d\mathbf{x},
\]

where \( N_j \) is the shape function of node \( j \) of the element \( e \), and \( C_m = 1 \) has been assumed, without loss of generality.

When the shape functions \( N \) used to compute \( M_e \) are the same used to approximate the potential \( V \), the resulting non-diagonal matrix is known as the consistent mass matrix. Using a consistent mass matrix implies that a linear system of equations has to be solved for \( V^{k+1} \) when an explicit scheme is used in (2.41). In order to improve numerical efficiency, the proposed algorithm evaluates \( M_e \) using a mass preserving nodal quadrature [116]. Nodal quadrature is based on the use of different base functions to those used to approximate the transmembrane potential, \( V \). This approximation with different shape functions is admissible since it satisfies the finite element criteria of integrability and completeness [116]. In the implementation, we have considered a nodal quadrature to evaluate \( M_e \) with \( N_i = J_i J \), being \( J_i \) the element Jacobian evaluated at node \( i \).
Modelling and simulation of the electric activity of the heart using GPU

Andrés Mena Tobar
This chapter presents results obtained using a novel electrophysiology simulation software entirely developed in Compute Unified Device Architecture (CUDA). The software implements semi-implicit and explicit solvers in double precision for the monodomain model, using operator splitting and FEM for spatial discretization. The potentiality of the GPU code simulations is demonstrated by running strong scalability benchmarks in 1D, 2D, and 3D problems, including realistic geometries of human heart and atria. Performance results are compared with a multi-CPU based software [41].

3.1 The CUDA programming model

The CUDA programming model (Compute Unified Device Architecture) is a sequential master program that can run parallel programs known as “kernels” in a parallel device [53, 72, 77]. A “kernel” is itself a program that is executed using a large number of parallel threads. Each thread executes the same sequential code. In this regard, the programmer arranges the threads of “kernel” in blocks of threads and these blocks in execution grids. The threads can collaborate with each other through the use of barriers synchronization and sharing memory resources in each
block, which is private to each. Figure 3.1 shows a hierarchical diagram of the thread, block and grid execution with their corresponding memory spaces.

![CUDA program model diagram](image)

Figure 3.1: CUDA program model.

### 3.2 Parallel implementation in GPU

The CUDA implementation consists of two subprograms: i) a CPU part or *host subprogram*, and ii) a GPU part or *device subprogram* as shown in Figure 3.2. The *host subprogram* prepares all structures required for the GPU execution, and then moving data from CPU main memory to the GPU memory. In addition, the *host subprogram* controls the execution and launches the device subprogram. The *device subprogram* is organised in kernels, with each kernel executed in parallel by each GPU thread. In the discretised scheme, there are two main contributors to the computational cost: solving the system of ODEs (reaction term) at each mesh point...
(Step I), and solving the linear system of equations associated with the parabolic PDE (Step II). In order to maximise performance, all vectors and matrices associated with the system of equations and the ionic model reside in GPU memory. Only the transmembrane voltage vector, $V$, is transferred back to the CPU memory when the data has to be saved on disk in order to minimise the communication time. In addition, in order to minimise memory storage, all data is stored using sparse matrix structures. The entire implementation of the solver i.e., ODE and PDE solvers, is performed in double precision, since previous studies indicate an important loss of accuracy [4] when using single precision to integrate the ionic model in Step I.

Figure 3.2: CUDA program model.

3.2.1 ODE solver.

The C/C++ code describing the ionic model for GPU implements the Rush-Larsen integration scheme for the gating variables [90] in order to improve accuracy. The exact same code was executed in CPU and GPU in order to validate the implementation.
When executed in GPU, the transmembrane potential and the state variables (SV) for all nodes in the model are stored in a unique vector in the global GPU memory. Figure 3.3 shows an example for the case of a discretised model with \( n_n \) nodes and an ionic model with \( n_{st} \) state variables. This data structure allows all GPU threads within a thread block to have access to continuous memory blocks during execution, guaranteeing data coalescence in memory and maximizing performance [73]. On the contrary, constants associated to the ionic model required to evaluate (2.34) are kept in the shared GPU memory.

In more complex cardiac models having a large number of SV, use of a single kernel to integrate the ionic model may not be possible because the number of registers available per thread is insufficient (this is the case in small cards as the Nvidia GTX295). In this case, the solution of the ionic model will require a sequence of multiple kernel invocations, with each kernel devoted to the solution of a group of SV. Because a kernel invocation may modify the input of the following kernel, it is necessary to resolve these dependencies by duplicating the variables that are common input among the kernels. Since every kernel invocation introduces an overhead in computing time, but also in memory requirements, an optimise implementation is required to find the best trade-off between kernel splitting, memory use, and kernel invocation overhead.

The influence of thread divergence was also considered in our implementation. On the GPU threads within a block run concurrently only if they execute the same instruction. In this regard, synchronisation between threads may be reduced in a conditional statement since the threads may follow different branches causing a reduction in parallelism. In the implemented models, conditional statements were substituted by Heaviside functions allowing mimicking the conditional as a product between a literal and the condition of the Heaviside function (implemented as a
3.2.2 Parabolic solver

The linear system given in (12) was solved on GPUs using the CUSP and Thrust libraries developed by Nvidia [5]. CUSP is implemented for a single GPU and natively supports a number of sparse matrix formats providing specific subroutines for an easy passage between different sparse matrix formats. The library includes highly optimised matrix-vector multiplication algorithms and iterative solvers. In addition, a variety of preconditioners based on algebraic multigrid (AMG) and approximate inverse operators are readily available in the library. In our implementation, mass and stiffness matrices, $M$ and $K$ respectively, are assembled in parallel in the GPU and stored in compressed sparse row (CSR) format at host memory. These matrices are then transformed to an efficient sparse matrix format when transferred in the GPU memory for computations.

The performance of the linear solver in CUDA is highly conditioned by the efficiency of the sparse matrix-vector multiplication, which depends on the sparse matrix format used to store the matrix during computations in the GPU memory. The CUSP library offers five different sparse matrix formats that can be used when solving a linear system of equations. Namely: i) Coordinate format (COO), a simple storage scheme for which the required storage is proportional to the number of non-zeros in the matrix; ii) CSR format, a natural extension of the COO format with a simple compression scheme applied to the repeated row indices found in the COO format; iii) Diagonal format (DIA), is a not general purpose format, but represents an efficient storage scheme for matrices arising from structured meshes; iv) ELL format represents a storage scheme well suited for vector architectures. More general than DIA, ELL efficiently encodes semi-structured and unstructured meshes when the maximum number of nonzeros per row does not greatly differ from the average; v) Hybrid (HYB) format, is a hybrid ELL/COO format which stores the majority of matrix entries in ELL and the remaining entries in COO. This allows to explore the well-suited structure of ELL for vector architecture with the storage efficiency of COO. Our implementation has considered different sparse matrix format at GPU level in order to evaluate their performance on structured and non-structured meshes.
3.3 Benchmarking

Separate benchmarking tests were performed to evaluate the ODE solver and the parabolic solver. The speed-up and accuracy of the ODE solver was evaluated on three models of action potential: i) A 4 SV model proposed by Bueno-Orovio et al. [11] (BCF); ii) the model for ventricular cardiomyocytes proposed by ten Tusscher and Panfilov (TP06) [102] comprising 19 SV; and iii) the model for atrial cardiomyocytes proposed by Maleckar et al. [63] (MLK) with 29 SV. The accuracy of the ODE solver was addressed by comparing the solution obtained using the GPU with the solution obtained with a single CPU. In this regard, the ionic model was stimulated every second for 2 ms during 100 s with as stimulation current, $J_{stm}$, of $24 \mu A/\mu F$, recording the solution for the last second of simulation with a resolution of 0.2 ms. The parallel performance was evaluated as the speedup, $S$, i.e., the execution time of the GPU implementation with respect to a single CPU core as the number of nodes, or degrees of freedom, increases (see Figure 3.4a).

The performance of the parabolic solver was tested on 1-, 2- and 3-dimensional problems. The model geometry was defined as a cable of length $L$, a rectangle with dimensions $L \times 20$ mm$^2$, and a cuboid with dimensions of $L \times 20 \times 7$ mm$^3$ as shown in Figure 3.4b. The dimension $L$ was varied in order to increase the number of nodes in the problem. The different geometries where meshed with structured (quadrilateral and hexahedral elements) and non-structured (triangular and tetrahedral elements) meshes with a characteristic element size of 0.1 mm. In addition, since ventricular cardiac muscle is anisotropic, the tissue was modelled as transversally isotropic with the fibre direction oriented along the x-axis. A propagating wave front was initiated at one of the corners of the model using a stimulation current pulse of $50 \mu A/(\mu Fcm^2)$ strength and 2 ms duration. The electrical activity was simulated over 1000 ms, with a fixed time-step of 20 $\mu s$. The parallel performance was evaluated calculating the speedup, $S$.

The efficiency of the GPU is further demonstrated on the simulation of a heart beat in realistic models of a biventricular human heart and a human atria. The biventricular human heart was discretised with 1289000 hexahedral elements and 1434129 nodes (grid points). The TP06 model has been used to simulated the action potential model of the ventricular tissue. The human atria was used to demonstrate the efficiency of the code with non-structured meshes. The human atria was discretised with 1378054 elements and 266450 nodes. The MLK model of
atrial action potential was used for simulating the electric activity. In both models, cardiac muscle anisotropy has been accounted for.

The speed-up of the GPU implementation is compared against a fortran implementation of the algorithm described in chapter 2 using the Message Passing Interface (MPI) for parallel computations \cite{41, 75}. Domain decomposition for the parallel solution has been carried out using the METIS library \cite{52}, where as the linear system of equations (2.42) has been solved using the Conjugate Gradient (CG) method with an ILU preconditioner from the PSBLAS library \cite{30}.

GPU simulations were run on a computer node with two Intel-Xeon Quad-Core CPUs E5620 clocked at 2.4GHz and 48GB DDR3 RAM. The node is equipped with four Nvidia Tesla E2090 GPUs, each with 6GB DDR5 RAM for a total of 24GB DDR5 RAM. All simulations were run in a single GPU. The CPU benchmark was run on a cluster with 8 nodes with two Intel-Xeon Quad-Core E5520 clocked at 2.26GHz and 24GB DDR3 RAM connected by a high speed infiniband network.

Figure 3.4: Model problems used to test the performance of the GPU algorithm. a) ODE solver, b) Parabolic solver.
3.4 Results

The coefficient of determination, $R^2$ between the solution obtained with the GPU and CPU was calculated for each of the state variables of all three ionic models considered. The results show a $R^2 > 0.9$ for all state variables, indicating an excellent agreement between both implementations, in addition to demonstrating the capabilities of double precision arithmetic of the E2090 GPU. Table 3.1 shows the results for the TP06 model. Similar results were found for the BCF and MLK models.

<table>
<thead>
<tr>
<th>State Variable</th>
<th>$R^2$</th>
<th>State Variable</th>
<th>$R^2$</th>
<th>State Variable</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V$</td>
<td>0.9974</td>
<td>$[Ca^{2+}]_i$</td>
<td>0.9993</td>
<td>$[Ca^{2+}]_{SR}$</td>
<td>0.9982</td>
</tr>
<tr>
<td>$[Ca^{2+}]_{SS}$</td>
<td>0.9388</td>
<td>$[Na^+]_i$</td>
<td>0.9691</td>
<td>$[K^+]_i$</td>
<td>0.9899</td>
</tr>
<tr>
<td>$m$</td>
<td>0.9981</td>
<td>$h$</td>
<td>0.9967</td>
<td>$j$</td>
<td>0.9961</td>
</tr>
<tr>
<td>$x_s$</td>
<td>1.0000</td>
<td>$r$</td>
<td>0.9971</td>
<td>$s$</td>
<td>0.9997</td>
</tr>
<tr>
<td>$d$</td>
<td>0.9976</td>
<td>$f$</td>
<td>1.0000</td>
<td>$f_2$</td>
<td>0.9998</td>
</tr>
<tr>
<td>$f_{cass}$</td>
<td>0.9973</td>
<td>$R$</td>
<td>0.9993</td>
<td>$x_{r1}$</td>
<td>0.9997</td>
</tr>
<tr>
<td>$x_{r2}$</td>
<td>0.9995</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3.1: Coefficient of determination between the state variables corresponding to the CUDA implementation and the C++ implementation. The coefficient of determination has been calculated with the entire time course of each state variable during the last beat of the stimulation protocol.

The performance of the ODE solver for the three ionic models is shown in Figure 3.5. The enormous parallel capabilities of the GPU is demonstrated in Figure 3.5a where the relative computing time required by the GPU as the number of nodes in the model increases is depicted. The figure shows that the computing overhead increases with the number of SV in the model. Furthermore, the figure shows that the overhead increases monotonically with the number of nodes in the model until reaching a steady state value depending on the complexity of the ionic model. This figure also indicates the existing a threshold (depending on the number of SV) for which the GPU is able to integrate the ionic model simultaneously without experiencing a significant degradation in computing time. After this threshold is surpassed, the computing time increases almost linearly as the number of nodes in the problem increases. Figure 3.5b shows the speed-up, $S$, obtained with the GPU as the number of nodes in the problem increases. The figure shows that a single
CPU is faster than a single GPU thread. However, the enormous computational horsepower of the GPU leads the GPU to overtake a single CPU as the number of nodes increases until attaining an asymptotic speedup, which may be up to $200\times$ for the MLK model on problems with more than a one million nodes. Note, that the maximum speedup reduces as the number of SV in the ionic model decreases.

Another important aspect to account for in the implementation of the ODE solver is the effect of splitting the kernel in multiple kernel invocations due to the limitation in the number of registers per thread. Even though, the Tesla M2090 allows the implementation of all three studied models on a single thread, a multiple kernel implementation of the TP06 model was performed in order to allow execution (in double precision) on a Nvidia GTX295 with 895Mb of memory per device. A careful optimisation of the code using the CUDA SDK profiler yield an splitting of the ionic model into 22 kernels and duplicating 15 SV. Table 3.2 shows the computation overhead introduced by the multi kernel implementation as the number of nodes in the problem increases. The computational overhead on a single CPU is also depicted for completeness. As expected, the computing overhead remains constant for the CPU implementation. However, for the GPU, the computation overhead reduces
asymptotically to zero as the number of nodes in the model increases.

<table>
<thead>
<tr>
<th>N</th>
<th>Computation overhead (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GPU</td>
</tr>
<tr>
<td>$10^9$</td>
<td>464</td>
</tr>
<tr>
<td>$10^3$</td>
<td>410</td>
</tr>
<tr>
<td>$10^4$</td>
<td>249</td>
</tr>
<tr>
<td>$10^5$</td>
<td>57</td>
</tr>
<tr>
<td>$10^6$</td>
<td>5</td>
</tr>
<tr>
<td>$2 \times 10^6$</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 3.2: Computation overhead of the single kernel versus the multiple kernel implementation for the TP06 ionic model.

Figure 3.6 shows the performance of the parabolic solver for a fully-explicit algorithm (Figure 3.6a) and a semi-implicit algorithm (Figure 3.6b) with structured meshes with the TP06 model. A CG iterative algorithm with ILU preconditioning and a hybrid sparse matrix format have been used for the computations. No significant differences were observed on the benchmarks conducted with non-structured meshes (results not shown).

The speed-up obtained with the explicit scheme (see Figure 3.6a) follows the same trend obtained for the ODE solver shown in Figure 3.5a. In fact, the results indicate that integrating the ionic model occupies more than 90% of the computing time. As observed with the ODE solver, for small problems the GPU underperforms the single CPU execution due to the lower clock-speed of the single GPU with respect to the single CPU, and the faster double precision arithmetic of the CPU. However, as the number of degrees of freedom increases the GPU outperforms the CPU (a threshold was found in 800 nodes for the benchmark problems with the TP06 model) until reaching an asymptotic behaviour with a speedup of: $\sim 100 \times$ for 1D-problems, $\sim 90 \times$ for the 2D-problems, and $\sim 70 \times$ for the 3D-problems. The reduction in acceleration is mainly due to the additional cost involved in the matrix multiplication that takes place during Step II. The cost of solving the linear system is more evident with the semi-implicit scheme (Figure 3.6b). In this case, the speed-up curve does not reach an asymptotic value as for the explicit algorithm, and the minimum problem size required for the GPU to outperform the CPU was found in 8000 nodes. The maximum speed-up for the semi-implicit scheme was found to be
Chapter 3. GPU implementation

\[ \begin{array}{cccccc}
1d & 2d & 3d & 10 & 1 & 10 \\
10 & 10 & 10 & 10 & 10 & 10 \\
\end{array} \]

\text{Number of nodes}

\text{Speedup}

\( (a) \)

\begin{array}{cccccc}
1d & 2d & 3d & 10 & 1 & 10 \\
10 & 10 & 10 & 10 & 10 & 10 \\
\end{array} \]

\text{Number of nodes}

\text{Speedup}

\( (b) \)

Figure 3.6: Performance of the parabolic solver: a) Explicit algorithm \((\theta = 0)\); b) Semi-implicit algorithm \((\theta = 1)\).

\( \sim 65 \times \) for 1D-problems, \( \sim 60 \times \) for 2D-problems, and \( \sim 50 \times \) for 3D-problems. It is important to point out that, the acceleration observed in Figure 3.6 depends on the size and complexity of the ionic model used.

As mentioned in Section 3.2, the performance of the linear solver in CUDA is highly conditioned by the efficiency of the sparse matrix-vector multiplication, which depends on the sparse matrix format used to store the matrix \( \mathbf{K} \) in (2.42). Figure 3.7 shows the relative computing time with respect to the HYB format for the 3D benchmark problem for both structured and unstructured meshes. For the structured meshes DIA format gave the best performance with a reduction of a 10\% in computing time with respect to the HYB and ELL formats. On the contrary, COO format gave the worst performance. For unstructured meshes HYB gives the best performance although no significant differences were found for the CSR and ELL formats, whereas the DIA format is not suitable for this type of discretisation. These results are in agreement with the benchmark results reported by Bell and Garland [5] for the sparse matrix-vector multiplication.

As part of the benchmark, the performance of the implicit-parabolic solver with different iterative schemes for solving Step II has also been evaluated. Namely: i)
Conjugate Gradient (CG), Bi-Conjugate Gradient (BCG), Stabilised Bi-Conjugate Gradient (BCGSTAB), and iv) Generalised Minimum Residual (GEMRES). Figure 3.8 shows the relative computing time with respect the CG scheme for the 3D benchmark problems with structured meshes. The HYB sparse matrix and ILU preconditioning has been used for the computations.

Figure 3.8 shows that CG gives the best performance for the problem under consideration, whereas the BCG scheme is the most expensive in terms of computing time. The same trend was observed when using different preconditioning offered by the CUSP library, and for different sparse matrix formats. In general it was observed that CG required more iterations to converge to the same tolerance per time increment than the other schemes. However, the lower cost per iteration in terms of matrix-vector multiplications involved in the CG scheme lead to a lower computation time. These findings are in agreement with results reported in [79, 80] for the solution of the electrophysiology problem using the mono domain model.

Figure 3.9 shows the depolarisation front in the voxelised human bi-ventricular model (Figure 3.9a), and a human atria discretised with tetrahedral elements (Figure 3.9b). The speed-up, \(S\), obtained on this realistic models with the semi-implicit scheme was \(50\times\) for the human bi-ventricular heart and \(40\times\) for the human atria.
with respect to a single CPU core, where as for the explicit scheme, the speed-up, $S$ was $\sim 50\times$ for both problems. These results were in excellent agreement with those obtained in the benchmark (see Figure 3.6) for both integration schemes. An additional test was conducted by running the human bi-ventricular heart on 42 cores in our computer cluster. In this case, the GPU still performed better by a factor of $1.8\times$, which implies a slightly better performance as compared to a single processor ($\sim 76\times$), due to the extra cost added by the communication between the different processes during the parallel execution.

In order to gain a better insight of the performance of the GPU solver with non-diagonally block matrices (matrices with large bandwidth), we have solved a voxelised human bi-ventricular heart coupled with a network of one-dimensional linear elements representing the fast conduction system in the heart (see Figure 3.10a). The resulting model involves 1.5 million degrees of freedom, and since the coupling between both meshes occurs only at a discrete number of points, the innervation points (see Figure 3.10a, right panel), the resulting $\mathbf{K}$ matrix in (2.42) has an spar-
sity pattern as shown in Figure 3.10b with a bandwidth of 1418491 and a maximum number of zeros per row of 35. Solving 1600 ms of electric activity with a fixed time step of 0.02 ms took the GPU-solver 9000 s using the semi-implicit scheme and the HYB sparse format for \( \hat{K} \), implying an acceleration of \( 50\times \) with respect a single CPU-core. This result is in agreement with the benchmark results demonstrating the efficiency of the GPU solver to handle ill sparsity patterns as those shown in Figure 3.10b. In fact, when the sparsity pattern was optimised using a CutHill-MacKee algorithm which reduced the bandwidth of the matrix \( \hat{K} \) 80 fold, the performance was not affected in neither the GPU-solver nor the CPU-solver.

### 3.5 Discussion and conclusions

The potential of GPU implementation for solving reaction-diffusion systems describing cardiac electrophysiology has been investigated. Scalability tests were performed
Figure 3.10: A coupled problem with large bandwidth. a) Detailed of the one-dimensional network representing the fast conduction system of the heart (left panel) coupled to the human bi-ventricular model (right panel). Coupling between the two meshes takes place at discrete points only (black markers in the right panel); b) Sparsity pattern of the resulting linear system.

for fully explicit and semi-implicit numeric schemes in 1D-, 2D-, and 3D-model problems with structured and unstructured meshes using novel software entirely developed in C++/CUDA. The performance of the method was ulteriorly demonstrated on realistic models of a biventricular human heart and human atria.

Previous studies [94] have reported speedups of $32 \times$ for the monodomain model using an explicit finite difference scheme along with the phase I Luo-Rudy ionic model (LR1) [60], a much less complex model (8 SV) than the ionic models used in this study (19 SV of the TP06 or 28 SV of the MLK action potential model of atria). In their study Sato et al [94] have established the solution of the PDE equation as the bottle neck of the computation with GPU, in part because they were forced to used single precision for their computations requiring a more stringent step size when solving the Step II than the required for integrating the action potential model in Step I. In addition, Sato et al. were using an adaptive time stepping when integrating the system of ODEs, which introduces latencies in the computations. In our implementation, we integrate with a constant time step size that guarantees both, the stability of the ionic model (most restrictive step size) and the stability of the PDE, similarly as has been proposed in order studies [71]. With this strategy,
along with an efficient data structure for the ionic model and a sparse format for
the matrices, we have been able to make full use of the GPU parallel potential by
reducing latency during GPU execution leading to acceleration of up to $90\times$ on
2D-problems of similar size as those considered in [94] with an explicit scheme. We
must also stress that, Sato et al. used older NVidia GT 8800 and GT 9800 GX2
cards that are faster for single precision floating point operations than the Tesla
GPUs used in our study, but do not support double precision calculations.

For the ODE solver, Lionetti et al. [59] performed 20 ms of a heart beat simulation
on a domain that consisted of 42,240 grid points with no spatial integration (all the
cells were decoupled) since their main interest was to optimise the ODE solver. They
also used different optimization techniques for the different cell models considered.
For the 18 SV Puglisi-Bers model [81], comparable to the TP06 model, they report
a total simulation time of 5.98 s using a time step of 0.0008 ms and single precision
on a NVIDIA GTX295. Our simulations with the TP06 model took 15 s on a the
C2090 using double precision arithmetic. Taking into account the overhead due to
double precision arithmetic, the performance is comparable. In a more recent work
Neic et al. [71] report an execution time of 360 s to simulate 250 ms with the 27 SV
proposed by Mahajan et al. [64] using a time step of 0.025 ms on a domain with
approximately 800000 grid points. The same simulation with the MLK model took
in our case 178 s representing an effective speedup of 2. However, differences in
the mathematical complexity of the two models could account for the difference in
spite of having the same number of state variables. Vigueras et al. [107] has reported
similar speedup values for the ODE solver with the TP06 model, whereas they report
slight lower speedups for the three dimensional problem benchmark with 1.2 million
grid points ($30\times$ against a $70\times$). The reason for this reduction in performance with
respect to our implementation may be related to the implementation of the PDE
solver. While we perform the solution of the linear system entirely in the GPU
memory with the CUSP libraries, they perform make use of a hybrid scheme in
which at each simulation step, the system of PDEs is solved by copying the required
data from CPU memory to GPU memory, and then the result copied back to CPU
memory. This operation introduces a considerable overhead that may reduce the
overall performance of the solver.

In other recent work Neic et al.[71] have used multiple-GPU enhanced simu-
lation code to accelerate cardiac bidomain simulations with the Mahajan 28 SV
rabbit model. In their study Neic et. al. have reported an overall speed-up of $16\times$
for the bi-domain problem, but only moderate speed-up between 8.1×-9.5× for the parabolic PDE equivalent to the monodomain solver. Even though our simulations with the Maleckar model (28 SV) on a 3D problem with 1.5 million grid points showed gave a speedup of the order of 50, a direct comparison is difficult, not only for the different problem size, but also because their implementation considers a computationally more expensive bidomain approximation. In addition, their multi-GPU implementation, communications between GPUs were performed through the host memory and not using a peer-to-peer communication between the GPU cards, which substantially increases communication time, and therefore a penalising overhead. Rocha et al. [85] report a simulation time of 1804 s with the TP06 model for 500 ms on a 2D grid size of 641×641 using single precision arithmetics and an implicit scheme for the PDE. Our implementation takes 867s in double precision, representing an speedup of 2. However, our computations were carried on an Tesla C2090 which is considerably faster than the GTX280 used in their study. The same problem was also solved by Bartocci et al. [4] but using an explicit algorithm and finite differences for the spatial discretisation. Bartocci et al. report a total computation time of 105.4 s on a Tesla C2070 using single precision arithmetic, while our explicit implementation takes 220 s on the same graphic card with double precision arithmetic. Considering that they report an overhead of 1.8 when using double precision arithmetic with respect to their single precision results for the TP06 model, their computation time for this problem would be 198 s which is comparable to our results. It should also be accounted the fact that Bartocci et al., use a finite difference discretisation for the Laplacian term which implies a K matrix with lower bandwidth than the one obtained with the finite element discretisation (bandwidth of 5 for finite differences against a bandwidth of 9 for finite elements), which implicates a reduction in the computing cost.

Results from Figure 3.5 and Figure 3.6 indicate that a minimum problem size is required in order to optimise GPU performance. This is due to the slower clock speed and limited double precision floating point arithmetic of the GPU cores as compare to CPU cores. This lower core performance is, however, compensated by the capability of the GPU to schedule a vast numbers of threads and efficiently reducing latency in this many core architecture. In this regard, for the E2090 our results indicate that for the ODE solver, the GPU is able to accommodate up to 8000 nodes without degradation of computer performance when working with a complex model like TP06. For the monodomain solver, working with the TP06 ionic model,
this threshold was found to be at 800 nodes for the fully explicit scheme and at 8000 nodes for the semi-implicit scheme, independently of problem dimensionality and mesh structure (structured or unstructured mesh). In this regard, we must add that we have used highly optimised unstructured meshes which maximise the bandwidth of the resulting algebraic system of equations. After this threshold is surpassed the scalability of the GPU with the number of degrees of freedom approaches to linearity. Even though these results are dependent on the GPU card used, this threshold appeared to be independent of the problem dimensionality for the cases considered. We must remark that the bandwidth of the linear system did not change much between the structured and unstructured meshes. This explains the similarity between the results obtained for both benchmarks, as well as the speedup found for the ventricle and atria simulations. However, ulterior tests performed on non-diagonally dominant matrices show no-degradation on the computing time with respect to the benchmark problems.

Our study also shows the importance of an efficient sparse matrix format and adequate iterative solver for the computations. We found that using a DIA sparse format is optimal when working with purely structured meshes. However, this format is not suitable for unstructured or semi-structured meshes for which the ELL and HYB formats were found to be the best in agreement with the study conducted by NVIDIA [5]. In this regard, our study points the HYB sparse matrix format as the best alternative in terms of storage flexibility and numerical performance. It is important to point out that complex models of cardiac electrophysiology include the presence of the specialised conduction system, i.e., Purkinje system, (see Figure 3.10a) leading to models with mixed element meshes i.e., 3D elements connected to 1D elements, which result in an ill sparsity pattern as shown in Figure 3.10b. Regarding the iterative solvers for solving the parabolic system, our benchmark confirms (see Figure 3.8) the results from Potse et al. [80] which found that for the monodomain model, an ILU preconditioned in conjunction with a Conjugate Gradient iterative solver were optimum in terms of computing time.

This work confirms other studies that, a significant reduction on computing time can be achieved for solving the cardiac monodomain equations using GPUs. Despite the significant lower performance observed on a single GPU core with respect to a single CPU core, a single GPU card offered excellent performance with speedups of 70× for three-dimensional problems solved explicitly and near 50× for three-dimensional problems solved with a semi-implicit scheme when compared with a
single CPU. However, we must mention that when an adaptive time step is used in the single CPU core (an alternative that is not valid for GPU due to the introduced latencies and dynamic load loss), the overall acceleration of the GPU over the CPU reduces to $42\times$ for 1D problems, and $32\times$ for 2D and 3D problems, implying a reduction of about a 40% with respect the case of constant time step. However, despite this reduction, the performance of the GPU is still outstanding with respect to the CPU. As a final remark, we have seen that a personal workstation is able to perform a simulation of the electric activity of a whole heart in reasonable times and at a reasonable cost due to the significant lower prize of a GPU card as compared to a computer cluster and its outstanding performance. This aspect, enables researchers to interact more easily with their simulations helping them to improve their understanding of the physiology and the synergy between the different scales involved in electrocardiographic simulations. However, working with GPUs poses additional programming challenges over traditional parallel CPU implementations, in particular in aspects related to efficient management of threads and memory, as well as data structure in order to obtains the maximum performance of the GPU.
Modelling and simulation of the electric activity of the heart using GPU

Andrés Mena Tobar
Vulnerability in Regionally Ischemic Human Heart. Effect of the Extracellular Potassium Concentration

Ventricular tachycardia and ventricular fibrillation are two types of cardiac arrhythmias that usually take place during acute ischemia and frequently lead to sudden death. Pro-arrhythmic mechanisms of acute ischemia have been extensively investigated, although often in animal models rather than in human ventricles. In this work, we investigate how hyperkalemia affects the vulnerability window to reentry and the reentry patterns in the heterogeneous substrate caused by acute regional ischemia using an anatomically and biophysically detailed human biventricular model.

Over the last years, mathematical modelling and computer simulations have been a useful tool in analysing electrophysiological phenomena. In this particular, one of the major contributions of computer electrophysiology has been in understanding important relations between electrophysiological parameters [87]. For the ischemic heart, computer models have allowed to address the role of ischemic abnormalities in cardiac electrophysiological behavior [29]. However, most of these simulations have been restricted to 2D [29, 103] or 3D simulations of total ischemic heart [86]. Little work has been carried out in modeling the entire heart subjected to acute ischemic conditions [109, 25]. In the work by Weiss et al., even though they have accounted for heterogeneities caused by ischemia, their characterization of the action poten-
tial under acute ischemic conditions has mostly relied on the model characterized for guinea pigs [28]. Dutta et al. [25] have investigated on how reduced repolarization increase arrhythmic risk in the heterogeneous substrate caused by acute myocardial ischemia. In their work, Dutta et al. [25] developed a human ventricular biophysically-detailed model with acute regional ischemia. Even tough macro-reentries around the ischemic zone were reported, these reentries self-terminated after three complete circuits not being able to establish sustained reentry.

In this chapter we investigate on how hyperkalemia affects the vulnerability window to reentry and the reentry patterns in the heterogeneous substrate caused by acute regional ischemia using an anatomically and biophysically detailed human biventricular model. The proposed mathematical model is based on the mono domain model for simulating the propagation of the electrical signal of the heart. For the biophysical description of the action potential under normal and ischemic condition, the model proposed by ten Tusscher and Panfilov [102], TP06, is used. In this regard, the model has been modified to account for ischemia by incorporating an ATP sensitive Potassium, \( I_{K(\text{ATP})} \), current. By analyzing high spatio-temporal resolution simulation data, we unravel the mechanisms associated with the observed reentrant patterns observed in acutely-ischemic ventricles.

4.1 Methods

4.1.1 Mathematical model

The variation of the transmembrane potential, \( V \), in the heart was modeled by means of the monodomain model [33]

\[
\nabla \cdot (D \nabla V) = C_m \frac{\partial V}{\partial t} + J_{\text{ion}}(V, w) + J_{\text{stim}},
\]

(4.1)

\[
\frac{\partial w}{\partial t} = f(w, V, t).
\]

(4.2)

where \( D \) is the symmetric and positive definite conductivity tensor, \( C_m \) the membrane capacitance, \( J_{\text{ion}}(V, w) \) the transmembrane ionic current, \( J_{\text{stim}} \) the stimulation current, and \( w(w, V, t) \) is a vector of gating variables and concentrations, \( f \) is a vector valued function, and \( t \) refers to time. Both, \( J_{\text{ion}} \) and \( f \) depend on the cellular
model used. The boundary conditions associated with this model are

\[ \mathbf{n} \cdot \nabla (D \nabla V) = 0. \]  

(4.3)

where \( \mathbf{n} \) is the outward normal.

The monodomain model represents an important simplification of the more complex bi-domain model [33] with important advantages for mathematical analysis and computation. Despite its simplicity, this model is adequate for studying a number of electrophysiologic problems as ventricular fibrillation or the onset of ischemia in the electric behaviour of the heart [29, 103, 86, 109, 25, 102].

From a mathematical and computational point of view, the electrophysiology problem is the couple solution of a linear partial differential equation, describing electric conduction, with a nonlinear stiff system of ordinary differential equations describing the cellular ionic currents which leads to a nonlinear reaction-diffusion problem. An efficient way of solving Equations (4.1-4.3) is by applying the Strang based operator splitting scheme [98] in combination with a trapezoidal family of method for time integrations, in conjunction with the finite element method for the spatial discretization [65].

4.2 Model of acute ischemia

Simulation of the ischemic heart requires both, an accurate description of the organ that incorporates its muscular structure and heterogeneity (described in the following sections), and an appropriate model of its electrophysiology. This section is dedicated to the characterization of the mathematical model of action potential used to perform the numerical simulations of the regional acute ischemic heart. All simulations were performed with a modified version of the ten Tusscher and Panfilov (TP06) model of action potential [102].

4.2.1 Action Potential model under ischemic conditions

One of the most important aspects in simulating the ischemic heart is to incorporate ATP sensitive potassium current \( I_{\text{KATP}} \), a dormant depolarization current under physiological conditions that is activated under ischemic condition [28]. \( K_{\text{ATP}} \) ion channels have been investigated in different regions of the heart, i.e., atria
and ventricle, the sinoatrial (SA) and the atrioventricular (AV) node, on different species [31, 74, 108, 58]. However, very little experimental data regarding the $I_{K_{ATP}}$ current for different tissue layers within the ventricle is available. Furukawa et al. [31] has characterized $K_{ATP}$ channels in isolated endocardial and epicardial cells of cats. Experiments by Furukawa et al. suggest that the open probability of $K_{ATP}$ channels reduces with the intracellular ATP concentration, $[ATP]_i$, for both cell types. However, $[ATP]_i$ concentration responsible for a 50% block of $K_{ATP}$ channels is approximately four times less for endocardial cells than for epicardial cells. Similar observations have been done by Nichols et al. [74] and Venkatesh et al. [108] in epicardial cells of guinea pig, and by Light et al. [58] in rabbit.

A modified version of the ten Tusscher cardiac action potential model [102] was used in the simulations. The model describes the principal ionic currents through the cardiac cell membrane with high degree of electrophysiological detail for the three types of cardiac cells. The basic model was completed with the formulation of the ATP-sensitive $K^+$ current ($I_{K_{ATP}}$) described by Ferrero et al. [28].

Figure 4.1: Fraction of open channels for different cell types. Symbols represent experimental values for different cell types in different species. Solid lines represent the model fit Eq. 4.5.

Figure 4.1 shows the open probability of $K_{ATP}$ channels, $f_{ATP}$, measured for different cell types and different animal species. The figure suggests a better cor-
relation between data corresponding to the same cell type rather than the animal specie. This could be interpreted as a low specificity to animal specie but a high specificity to cell type. We have adopted this hypothesis to incorporate the $I_{KATP}$ model by Ferrero et al. [28] originally formulated for guinea pig, to the TP06 model for human. A similar hypothesis has also been adopted by Michailova et al. [66] in developing an heterogeneous model of $I_{KATP}$ for rabbit.

The procedure followed consisted on modifying the original formulation for the $f_{ATP}$ proposed by Ferrero et al. [28] in order to fit the data from Figure 4.1 for different cell types. The maximal conductance for the $I_{KATP}$ was modified to adjust the experimental observations on the AP duration (APD) by Furukawa et al. [31]. In words, a 50% reduction in the APD under conditions of hypoxia and hyperkalemia for the epicardium, and a 10% reduction under the same conditions for the endocardium. For the mid-miocardium, there are no experimental data available regarding the behavior of $K_{ATP}$ channels. For this case, we have adjusted the value of $f_{ATP}$ such that the same reduction in APD as for epicardial cells was obtained without modifying the maximal conductance of the channel, as proposed by others [66, 34]. In addition, for physiological values of $[ATP]_i$ and $[ADP]_i$, the APD and the resting potential should not be affected by the presence of the $I_{KATP}$ current in the AP model. Hence, the $I_{KATP}$ current has been formulated as

$$I_{KATP} = g_0 \left( \frac{[K_+]}{5.4} \right)^{0.24} f_M f_N f_T f_{ATP}(V - E_K),$$

(4.4)

where $g_0$ is the maximum channel conductance in the absence of $Na^+$, $Mg^{2+}$ and $ATP$, $f_M$, $f_N$, $f_T$ are correction factors, $f_{ATP}$ is the fraction of opened channels, $V$ the transmembrane potential, and $E_K$ is the inversion potential of the channel. The maximum channel conductance and the fraction of opened channels, $f_{ATP}$, have been modified with respect to its original formulation

$$f_{ATP} = \frac{1}{1 + ([ATP]_i/K_m)^H},$$

(4.5)

where $[ATP]_i$ is the intracellular concentration of $ATP$, and $K_m$ (in mmol/L) and $H$ (-) are given as

$$K_m = \alpha (35.8 + 17.9[ADP]_i^{0.256})$$

(4.6)

$$H = 1.3 + 4.44 \exp(-0.09[ADP]_i)$$

(4.7)
where $[ADP]_i$ is the intracellular concentration of ADP intracellular en $\mu$mol/L, and $\alpha$, $\beta$ fitting parameters that account for the cellular heterogeneity. Parameters $\alpha$ and $\beta$ where identified by fitting experimental data available for different animal models from Figure 4.1. To adjust $g_0$, 100 stimuli at a basic cycle length (BCL) of 1000ms were applied to an isolated cell. The APD at the last stimulus was measured under physiological and pathological conditions. The variation in APD obtained between the pathological and physiological conditions was used to adjusted $g_0$. Table 4.1 shows the adopted $[K^+]_o$, $[ATP]_i$, $[ADP]_i$ values to define physiologic and ischemic conditions for fitting $g_0$.

<table>
<thead>
<tr>
<th>Condition</th>
<th>$[K^+]_o$</th>
<th>$[ATP]_i$</th>
<th>$[ADP]_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normoxia</td>
<td>5.4</td>
<td>6.8</td>
<td>15.0</td>
</tr>
<tr>
<td>Ischemia</td>
<td>9.9</td>
<td>4.6</td>
<td>99.0</td>
</tr>
</tbody>
</table>

Table 4.1: Extracellular Potassium and Intracellular ADP and ATP concentrations defining physiologic and ischemic conditions for the simulations

Table 4.2 summarizes the parameters identified in the fitting process. The table also shows the obtained APD value under physiologic and ischemic conditions defined in Table 4.1, in addition to the APD value obtained when the $I_{KATP}$ current remains inactive. The APD variation between normal and hypoxic conditions given in Table 4.2 are in good agreement with experimental observations by Furukawa et al. [31].

4.2.2 Heart model

The geometry of the bi-ventricular heart (the atria has not been considered in the model) and the orientation of the muscle fibers were obtained from diffusion tensor magnetic resonance imaging (DT-MRI), from images acquire at John Hopkins University [42]. From the segmented image, a regular volumetric mesh was constructed with hexahedral elements and a resolution of $0.4mm \times 0.4mm \times 0.4mm$, which gave rise to 1.43 million nodes and 1.29 million hexahedra.

Transmural heterogeneity of electrophysiological behavior across the heart is necessary for normal cardiac function, with excessive heterogeneity contributing to ar-
Table 4.2: Value of the parameters for the $I_{KATP}$ current adapted to the ten Tusscher model for different cell types. Also shown, the APD under physiologic and ischemic conditions.

<table>
<thead>
<tr>
<th>Cell Type</th>
<th>Parameter $g_0$ (mS)</th>
<th>$\alpha$</th>
<th>$I_{KATP}$ active $APD_{90}^{Nor}$ (ms)</th>
<th>$APD_{90}^{Isq}$ (ms)</th>
<th>$I_{KATP}$ inactive $APD_{90}^{Nor}$ (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENDO</td>
<td>4.5</td>
<td>0.32</td>
<td>299</td>
<td>265</td>
<td>299</td>
</tr>
<tr>
<td>MID</td>
<td>4.5</td>
<td>0.86</td>
<td>393</td>
<td>178</td>
<td>393</td>
</tr>
<tr>
<td>EPI</td>
<td>4.5</td>
<td>1.0</td>
<td>298</td>
<td>141</td>
<td>298</td>
</tr>
</tbody>
</table>

4.2.3 Electrophysiological heterogeneities under acute ischemia

The ischemic region was located in the inferolateral and posterior side of the left ventricle mimicking the occlusion of the circumflex artery (see Figure 4.2A). The ischemic region was composed of realistically dimensioned transitional border zones (BZ), a normal zone (NZ), and the central zone (CZ) of ischemia in agreement with experimental findings [20] during early stages of ischemia. In the CZ, $[K^+]_o$ was set at three different values: 7.0, 8.0 and 9.0 in order to study three different time frames during acute ischemia. The inward $Na^+$ and L-type $Ca^{2+}$ currents were scaled by a factor of 0.85 to mimic the effect of acidosis [29, 115], whereas $[ATP]_i$ and $[ADP]_i$ concentrations were set to 5 mM and 99 mM respectively [50]. The border zone included a linear variation in electrophysiological properties between
the NZ and the IZ as shown in experiments [20]. In addition, the model includes a 1.0 mm washed zone (not affected by ischemia) in the endocardium as a result of the interaction between the endocardial tissue and the blood in the ventricular cavities as suggested by Wilensky et al. [112] (see right panel in Figure 4.2a). The resulting human ventricular model in acute regional ischemia produced realistic pseudo-ecgs at the six derivations of the standard ECG (see bottom panel in Figure 4.2), exhibiting the ST elevation in V5-V6, with an acute T-wave in V6, and ST depression in V1-V4 consistent with an infarction involving the inferior, lateral and posterior walls caused by the occlusion of the proximal circumflex artery [68].

4.2.4 Stimulation protocol

This work does not incorporate the specialized conduction system to stimulate the heart. However, Purkinje-like stimulation was simulated by stimulating discrete zones of the endocardium according to the work performed by Durrer et al. [24]. According to this work endocardial stimulation in the left ventricle starts at three well defined locations within a time window of 5 ms: i) An area high on the anterior paraseptal wall just below the attachment of the mitral valve; ii) A central area on the left surface of the interventricular septum; and iii) The posterior paraseptal area at about one third of the distance from apex to base. For the right ventricle we have defined an stimulation area near the insertion of the anterior papillary muscle. In the right ventricle stimulation start 5 ms after the onset of the left ventricular potential [24]. Figure 4.2B, shows the location for the four stimulation areas.

The model was preconditioned with endocardial stimulation (S1) consisting on 56 stimulations at a cycle length (CL) of 800 ms (frequency of 1.25Hz). Following the preconditioning, an extra-stimulus, or premature stimulation, (S2) was applied in the subendocardial border zone (see Figure 4.2C) in agreement with the findings by [50]. The coupling intervals (CI), e.g., time differences between S1 and S2 was varied with a resolution of 1ms to determine the vulnerability window of reentry. In this regard, a depolarization pattern was considered as a reentry if at least two cycles were completed around the ischemic zone. Sustained reentrant patterns were studied for three seconds after the extra stimulus. Reentry patterns and vulnerability windows were investigated for different values of extracellular potassium concentration in the CZ.
Chapter 4. Vulnerability in Regionally Ischemic Human Heart. Effect of the Extracellular Potassium Concentration

4.2.5 Numerical simulations

Computations were performed with the GPU-based software TOR [65] using the operator splitting and a semi-implicit scheme with a fixed time step of 0.02 ms.
Simulation of 1 s of electric activity took 1.5 h on a GPU Tesla M2090 (6GRAM DDR5).

4.3 Results

Results show spatial heterogeneities in the propagated action potential, as reported experimentally, throughout the regional ischemic tissue, such as resting membrane potential (-85.2 mV in NZ, and -72.5 mV in the CZ, with potentials varying between these values in the BZ). During a basic beat, activity spreads from three directions into the ischemic area as shown in Figure 4.3a. Premature beat was delivered at the point of first activation in the border zone during the basic beat (see point F in Figure 4.3b).

![Figure 4.3: Depolarization pattern through the ischemic tissue (delimited by the white line) during basic activation (A), and right after the ectopic beat (B). The ischemic tissue is depolarized by three different fronts coming from the normal activation sites shown in Fig 4.2. The ectopic beat was delivered at the first point being depolarized within the ischemic tissue (point F in the right panel).](image)

The ectopic beat was delivered at different coupling intervals to determine the vulnerability window for three levels of extracellular potassium, $[K^+]_o$. Figure 4.4 shows the VW for different levels of $[K^+]_o$, for an early activation initiating at point F in Figure 4.3.

The extracellular potassium was found to have a significant effect on the vulnerability window. For a $[K^+]_o = 7$ mM the vulnerability windows was found to be the largest, and equal to 20 ms, with a CI between 250 and 270 ms. For CI between 250 and 265 ms, sustained reentry (lasting more than 3s, about 7 to 8 beats) was observed, whereas for CIs between 265 and 270, the reentrant activity self-terminated.
Chapter 4. Vulnerability in Regionally Ischemic Human Heart. Effect of the Extracellular Potassium Concentration

Figure 4.4: Vulnerability window for different levels of $[K^+]_o$. Early stimulation starts at point F in Figure 4.3.

after three reentry circuits (3 beats). For $[K^+]_o = 8$ mM the vulnerability window was reduced to 15 ms (CIs between 255 and 270 ms). In this case, sustained reentry was observed for CIs between 255 and 265 ms. On the contrary, for $[K^+]_o = 9.0$ mM no reentrant activity was found. In general, for CIs below the lower bound of the vulnerability window double blocked occurred without generating re-entrant activity. On the contrary, for CIs above the upper limit of the VW, conduction through the ischemic zone occurred.

Results in Figure 4.4 depend on the location of the ectopic stimulation. To explore this fact, the VW was determined for point F1 in Figure 4.3. As in the previous case, no-reentrant activity was obtained for $[K^+]_o = 9.0$ mM. For $[K^+]_o = 7.0$ mM the VW was reduced to 10 ms (CI between 265 and 275 ms) with the reentrant activity self-terminating after two beats (800 ms). A similar result was obtained for the case of $[K^+]_o = 8.0$ mM, with reentrant activity observed for CIs between 290 and 310 ms that self-terminated after two reentrant circuits.

The reentrant activity observed for different levels of $[K^+]_o$, and different ectopic sites, was associated with ventricular tachycardia in all cases, without degeneration in ventricular fibrillation, at least within the observed window of 3 seconds after earlier stimulation (see Figure 4.5). In all cases where sustained reentry was found, the same pattern emerged at the epicardium. After initiating early activity (at point F in Figure 4.3), conduction within the ischemic region slows down (see the
Figure 4.5: Reentrant patterns at endocardium (A) and epicardium (B) for the last recorded beat for the case $[K^+]_o = 7.0$ mM and CI=260 ms. Panel C shows the pseudo-ECG corresponding to the derivation V3, the arrows indicate the the window corresponding to the patterns shown in panels (A) and (B).

...frequency of the pseudo-ECG in Figure 4.5C) and a fragmented wave front with multiple areas of conduction block are observed in the epicardium due to intramural reentry (see bellow). In successive beats this pattern evolves into a circus, or double circus, movement within the CZ around an area of diameter about 4.0 cm and a revolution time in the order of 250 ms as shown in Fig 4.5B for the last recorded beat ($[K^+]_o = 7.0$ mM and CI=260 ms).

Detail observation of the activation activity at the endocardium in Figure 4.5A shows a faster circus movement in the endocardium as compared to the epicardium, with a revolution time in the order of 200 ms. This behavior is due to the presence of the washed zone that allows faster conduction velocity. The washed zone was also found to be related to the fragmentation of the wave front at the epicardium since it favoured transmural reentry. Figure 4.6A shows a snapshot of the transmembrane potential at the epicardium and endocardium at $t=2500$ ms after initiation of earlier stimulation (last recorded beat) for a $[K^+]_o = 7.0$ mM and CI=260 ms. The figure demonstrates how the double circus closes first at the endocardium. The trans-
mural map depicted in Figure 4.6B shows how the faster endocardial conduction allows intramural conduction causing this wavefront to emerge at the epicardium at \( t=2509 \) ms as shown in Figure 4.5B.

![Image of transmembrane potential at \( t=2500 \) ms after initiation of earlier stimulation](image)

Figure 4.6: Transmembrane potential at \( t=2500 \) ms after initiation of earlier stimulation ([\( K^+ \])_o = 7.0 \text{ mM and CI}=260 \text{ ms}). A) Depolarization map at the epicardium and endocardium. B) Transmural depolarization map.

Even though, basically, one circus movement of relatively large dimensions is the responsible for the sustained tachycardia, the pattern of the reentrant wavefront, position and dimensions, changes from beat to beat. Figure 4.7 shows the changes in the pattern from a single circus to a double circus, a figure of eight, for the case of [\( K^+ \])_o = 7.0 \text{ mM and CI}=260 \text{ ms}. The reentry was always observed within the ischemic region, and in all cases where sustained reentry was found, the reentrant front was lead from the endocardial surface where the pattern was clearly identified for all CIs. Reentry patterns were not altered when the basic stimulation was maintained after delivering the extra-stimulus, indicating that tachycardia overrides normal stimulation.

### 4.4 Discussion and conclusions

We have developed a biventricular human model that combines a realistic anatomy and actual muscle fiber orientation obtained from MR-DTI. In addition, the model integrates the biophysical features of membrane kinetics, as well as the electrophysiological alterations induced by acute regional ischemia. In this regard, the size of the ischemic area was defined according to previous studies [20]. In spite of not incorporating the specialized conduction system, simulation results are in agreement with experimental studies and clinical observations. Ventricular activation is in agreement with results from Durrer et al. [24]. Pseudo-ECG at the six deriva-
Figure 4.7: Evolution of the reentry pattern from beat to beat. The sequence indicates an evolution from a single circus to a double circus (figure of eight) to then come back to a single circus pattern ([K$^+$]$_o$ = 7.0 mM and CI=260 ms).

Extracellular potassium [K$^+$]$_o$, the element that most influences the effective refractive period (ERP) in single cell kinetics [28, 29], has also a significant effect on the VW. We found that the size of the VW decreases with [K$^+$]$_o$. Even though we did not perform simulations for lower values of [K$^+$]$_o$ than 7 mM, these results are in agreement with the experimental findings from Smith et al. [96] that reported that the peak of arrhythmic events during the 1a phase of acute regional ischemia occurs between the first five to nine minutes after occlusion. For this time window, Coronel et al. [20] reported that the accumulation of [K$^+$]$_o$ in the CZ goes from 6 mM to 8.5 mM. A possible explanation for this behavior is found on the relationship between the ERP and [K$^+$]$_o$. As [K$^+$]$_o$ increases, the ERP also increases reducing the likelihood of reentry if the size of the ischemic area remains unaltered. In other words, a prolonged ERP in ischemic tissue would require either increased ischemic region size or decreased conduction velocity. The size of the ischemic region is defined by the occluded vessel and remains unchanged while local changes in [K$^+$]$_o$, [ATP]$_i$ and PH occur as demonstrated by a number of studies [50, 112, 96, 20]. In addition, Smith et al. [96] show that there is a relatively small increase in tissue resistivity during the first 10 minutes of ischemia (in the order of 30% to 50% of initial values). However, after 15 minutes of ischemia (1b phase of ischemia), a more prominent rise in tissue resistivity begins due to cell-to-cell electrical uncoupling.
Electrical uncoupling implies a significant reduction in conduction velocity favoring the development of reentry activity.

The macro-reentrant patterns of activation obtained in the regionally ischemic human biventricular model were consistent with those reported by Janse et al. [50] in porcine and canine hearts. In all our simulations, reentrant activity was associated with ventricular tachycardia without registering ventricular fibrillation episodes. This is also in agreement with experiments from Smith et al. [96] which reported the onset of ventricular fibrillation to be cluster between 19 and 30 minutes of ischemia corresponding to the 1b phase. After stabilization of the reentrant activity, one circuit (single rotor) of fairly large dimension was basically observed with the activity circling around an area of block the size of the CZ. For some beats, double circuits were observed but not sustained, and the reentrant activity continued because of one single reentrant circuit. Patterns from our simulations resemble closely those reported Janse et al. [50]. For instance, Figure 8 in [50] shows single circuit and the eight shape reentrant pattern also obtained in our simulations (see Figure 4.7).

Our results indicate that reentrant activity may be in part sustained by the presence of a washed zone in the endocardium as suggested by Wilensky et al. [112]. In order to get a better insight on this, all the simulations were repeated with a model without the washed zone. Figure 4.8 shows the VW for the model without washed zone for an ectopic stimulation delivered at point F in Figure 4.3. The most significant finding was that all reentries self-terminated after two or three complete circuits. While for $[K^+]_o = 7 \text{ mN}$ and $[K^+]_o = 8 \text{ mN}$ the vulnerability window was the same but without sustained reentries, for $[K^+]_o = 9 \text{ mM}$, non sustained reentrant activity was found for CIs between 310 ms and 370 ms. The significant difference in the VW between $[K^+]_o = 9 \text{ mM}$ and the other two lower concentrations is associated with the rapid increase in the ERP with $[K^+]_o$. On the contrary, the presence of a VW for $[K^+]_o = 9 \text{ mM}$ in the model without the washed zone is due to the slower conduction velocity associated with the ischemic tissue in the endocardium which allows to close the reentrant circuit which otherwise will find the tissue within the refractory period. These results suggest that the washed zone could act as a pro-arrhythmic substrate factor helping on establishing sustained ventricular tachycardia.

A number of limitations are associated with this study. The Purkinje system was not present in our model since we were interested in studying the mechanisms associated with reentry rather than in the biophysical mechanisms leading to ec-
Figure 4.8: Vulnerability window for different levels of $[K^+]_o$ in a model without the washed-out zone. Early stimulation starts at point F in Figure 4.3.

topic excitation. In this particular, the study by Janse et al. [50] suggests that earlier activation most likely is due to focal activity localized in the Purkinje fibers close to the ischemic border zone. Future studies, however, are intended to extend our study to evaluate the implication of the Purkinje system in earlier activation during acute regional ischemia and its potential role in contributing to arrhythmia. However, a recent study by Dutta et al. [25] indicates that early activation causing transmural micro-reentry could be generated by electrotonically-triggered EADs at the endocardium. A second important limitation is related to the definition of the ischemic zone itself. We have considered an idealized shape with smooth borders and transitions between the IZ and the NZ. Patient specific acute regional ischemic areas are expected to have tortuous borders that may contribute to modify the evolution of the reentry patterns and the VW. However, we believe that the general patterns that have emerged in this study will not be greatly modified by the actual shape of the ischemic area as long as the general dimensions of the ischemic area are maintained. In addition, we have only monitored the reentrant activity up to three seconds during which we have observed perpetuating activity for a restricted vulnerable window. In some cases, i.e., CI between 265 ms and 270 ms for $[K^+]_o=7.0$ mM, the reentrant activity was spontaneously terminated after completing three reentrant circuits. This type of behavior was also reported in [50] where tachycardia terminated within 30 sec after initiation. Additional studies on larger ischemic zones are required in order to determine if the size of the ischemic area may favor the onset of ventricular fibrillation in the ischemic heart. In addition,
the location of the ectopic activity is important for both, the size of the vulnerable window and the reentrant pattern.

In conclusion, the model predicts the generation of reentry within the ischemic zone due to the heterogeneity in the refractory period between the ischemic affected area and the normal myocardium. The observed patterns obtained with the simulations are in good agreement with experimental studies conducted in porcine and canine hearts subjected to acute regional ischemia. The main results of the simulations can be summarized as follows: i) As a consequence of the applied extra-stimulus that originates an ectopic beat, reentrant activity generated for CIs which range depends on the value of the extracellular potassium concentration $[K^+]_o$; ii) The reentrant activity generated due to the extra-stimulus initiated as a consequence of the interaction between wavefronts emerging from the washed zone into the ischemic zone.
Modelling and simulation of the electric activity of the heart using GPU

Andrés Mena Tobar
Conclusions and future work

The main conclusions and future work of this thesis is summarized in the following.

5.1 Conclusions

5.1.1 GPU implementation

- Graphic Processing Units (GPUS) with their tremendous computation horse power and “low cost” is a potential technology for research and clinic applications in computational electrophysiology. This work confirms that, a significant reduction on computing time can be achieved for solving the monodomain model for the electric activity of the heart using GPUs. Despite the significant lower performance observed on a single GPU core with respect to a single CPU core, a single GPU card offered excellent performance with speedups of $70\times$ for three-dimensional problems solved explicitly and near $50\times$ for three-dimensional problems solved with a semi-implicit scheme when compared with a single CPU. However, we must mention that when an adaptive time step is used in the single CPU core (an alternative that is not valid for GPU due to the introduced latencies and dynamic load loss), the overall acceleration
of the GPU over the CPU reduces to 42× for 1D problems, and 32× for 2D and 3D problems, implying a reduction of about a 40% with respect to the case of constant time step. This aspect, enables researchers to interact more easily with their simulations helping them to improve their understanding of the physiology and the synergy between the different scales involved in electrocardiographic simulations.

• This study also shows the importance of an efficient sparse matrix format and adequate iterative solver for the computations. We found that using a DIA sparse format is optimal when working with purely structured meshes. However, this format is not suitable for unstructured or semi-structured meshes for which the ELL and HYB formats were found to be the best in agreement with the study conducted by NVIDIA [5]. In this regard, our study points the HYB sparse matrix format as the best alternative in terms of storage flexibility and numerical performance. Regarding the iterative solvers for solving the parabolic system, the performed benchmark confirms previous results published in literature which found that for the monodomain model, an ILU preconditioned in conjunction with a Conjugate Gradient iterative solver were optimum in terms of computing time.

• Our results indicate that a minimum problem size is required in order to optimise GPU performance. This is due to the slower clock speed and limited double precision floating point arithmetic of the GPU cores as compared to CPU cores. This lower core performance is, however, compensated by the capability of the GPU to schedule a vast numbers of threads and efficiently reducing latency in this many core architecture.

• The implementation of complex AP models may require splitting the kernel in multiple kernel invocations due to the limitation in the number of registers per thread in the GPU. We implemented a multiple kernel implementation of the TP06 model and compared the performance against a single kernel implementation on a GPU, and a single CPU. Our results indicate that, the computing overhead remains constant for the CPU implementation. However, for the GPU, the computation overhead reduces asymptotically to zero as the number of nodes in the model increases. This indicates that GPUs are able to handle complex ionic models (multi-kernel) much more efficiently than
single CPU-cores. The fact that the overhead reduces with problem size makes the applications with GPU much more profitable in terms of computational resources.

5.1.2 Acute ischemic simulations

- The model predicts the generation of reentry within the ischemic zone due to the heterogeneity in the refractory period between the ischemic affected area and the normal myocardium. The observed patterns obtained with the simulations were consistent with previously reported patterns in the literature in porcine and canine hearts. Further, in all our simulations, reentrant activity was associated with ventricular tachycardia without registering ventricular fibrillation episodes. This was also in agreement with previously reported experiments indicating that the onset of ventricular fibrillation was cluster between 19 and 30 minutes of ischemia corresponding to the 1b phase of ischemia.

- After stabilization of the reentrant activity, one circuit (single rotor) of fairly large dimension was basically observed with the activity circling around an area of block the size of the CZ. For some beats, double circuits were observed but not sustained, and the reentrant activity continued because of one single reentrant circuit. Patterns from our simulations resemble closely those reported by Janse et al. [50].

- Our results indicate that reentrant activity is in part sustained by the presence of a washed zone in the endocardium as suggested by Wilensky et al. [112]. Simulations conducted with and without the washed zone indicated that, for the model without the washed zone all reentries self-terminated after two or three complete circuits. In addition the absence of the washed zone shifted the vulnerability window (but without sustained reentries) toward larger values of $[K^+]_o$ caused by the increase in the ERP and the reduction in conduction velocity associated with the ischemic tissue in the endocardium which allows to close the reentrant circuit which otherwise will find the tissue within the refractory period. In general, these results suggest that the washed zone could act as a pro-arrhythmic substrate factor helping on establishing sustained ventricular tachycardia.
5.2 Conclusiones

5.2.1 Implementación en GPUs

- Las unidades de procesamiento gráfico (GPUS) que poseen una enorme potencia de cálculo y que se pueden considerar de “bajo coste”, es una tecnología muy potente para la investigación y aplicaciones clínicas en electrofisiología computacional. Este trabajo confirma que se puede lograr una reducción significativa en el tiempo computacional para resolver el modelo monodominio para la actividad eléctrica del corazón usando GPU. A pesar del menor rendimiento significativo observado en un solo núcleo de GPU con respecto a un solo núcleo de CPU, una única tarjeta de GPU ofrece un excelente rendimiento con aceleraciones de $70 \times$ para problemas tridimensionales resueltos explícitamente y cerca de $50 \times$ para problemas tridimensionales resueltos con un esquema semi-implícito cuando se compara con una sola CPU. Sin embargo, debemos mencionar que cuando se utiliza un paso de tiempo adaptativo en el núcleo de una sola CPU (una alternativa que no es válida para GPU debido a las latencias introducidas y la pérdida de carga dinámica), la aceleración general de la GPU sobre la CPU se reduce a $42 \times$ para problemas 1D, y $32 \times$ para 2D y 3D, lo que implica una reducción de aproximadamente un 40% con respecto al caso de paso de tiempo constante. Este aspecto, permite a los investigadores interactuar más fácilmente con sus simulaciones, ayudándoles a mejorar su comprensión de la fisiología y la sinergia entre las diferentes escalas implicadas en las simulaciones electrocardiográficas.

- Este estudio también muestra la importancia de un formato de matriz sparse eficiente y un solver iterativo adecuado para los cálculos. Encontramos que usar un formato sparse DIA es óptimo cuando se trabaja con mallas puramente estructuradas. Sin embargo, este formato no es adecuado para las mallas no estructuradas o semiestructuradas para las que los formatos ELL y HYB fueron los mejores de acuerdo con el estudio realizado por NVIDIA [5]. En este sentido, nuestro estudio señala el formato de matriz sparse HYB como la mejor alternativa en términos de flexibilidad de almacenamiento y rendimiento numérico. En cuanto a los solucionadores iterativos para resolver el sistema parabólico, el benchmark realizado confirma resultados publicados previamente en la literatura para el modelo monodominio, una ILU pre-
condicionada en conjunción con un solver iterativo de Gradiente Conjugado fue óptimo en términos de tiempo de cálculo.

- Nuestros resultados indican que se requiere un tamaño de problema mínimo para optimizar el rendimiento de la GPU. Esto se debe a la velocidad de reloj más lenta y la aritmética de punto flotante de doble precisión limitada de los núcleos de la GPU en comparación con los núcleos de la CPU. Este rendimiento inferior de un núcleo en la GPU, es compensado con la capacidad de la GPU para programar un gran número de hilos y reducir eficazmente la latencia en esta arquitectura de muchos núcleos.

- La implementación de modelos complejos de PA (puede requerir la división del kernel en múltiples invocaciones del kernel debido a la limitación en el número de registros por hilo en la GPU. Implementamos una implementación de kernel múltiple del modelo TP06 y comparamos el rendimiento con una única implementación de kernel en una GPU, y una sola CPU. Nuestros resultados indican que, la sobrecarga de computación permanece constante para la implementación de la CPU. Sin embargo, para la GPU, la sobrecarga de cálculo se reduce asintóticamente a cero a medida que aumenta el número de nodos en el modelo. Esto indica que las GPUs son capaces de manejar modelos iónicos complejos (multi-kernel) de manera mucho más eficiente que los núcleos de CPU únicos. El hecho de que la sobrecarga se reduce con el tamaño del problema hace que las aplicaciones con GPU sean mucho más rentables en términos de recursos computacionales.

5.2.2 Simulaciones de la isquémia aguda

- El modelo predice la generación de reentrada dentro de la zona isquémica debido a la heterogeneidad en el período refractario entre el área isquémica afectada y el miocardio normal. Los patrones observados obtenidos con las simulaciones fueron consistentes con los patrones previamente reportados en la literatura en corazones porcinos y caninos. Además, en todas nuestras simulaciones, la actividad reentrante se asoció con taquicardia ventricular sin registrar episodios de fibrilación ventricular. Esto también estuvo de acuerdo con los experimentos previamente informados que indican que el inicio de la
fibrilación ventricular se agrupa entre 19 y 30 minutos de isquemia correspondiente a la fase 1b de la isquemia.

• Después de la estabilización de la actividad reentrante, se observó básicamente un circuito (rotor único) de dimensión bastante grande con la actividad circulando alrededor de un área de bloque del tamaño de la ZC(zona central). Para algunos latidos, se observaron circuitos dobles pero no se mantuvieron, y la actividad reentrante continuó debido a un solo circuito de reentrada. Los patrones de nuestras simulaciones se parecen mucho a los reportados por Janse et al. [50].

• Nuestros resultados indican que la actividad reentrante es en parte sostenida por la presencia de una zona lavada en el endocardio como sugiere Wilensky et al. [112]. Las simulaciones realizadas con y sin la zona lavada indicaron que, para el modelo sin la zona lavada, todas las reentradas se auto terminan después de dos o tres circuitos completos. Además, la ausencia de la zona lavada desplazó la ventana de vulnerabilidad (pero sin reentradas sostenidas) hacia valores mayores de \([K^+]_o\) causados por el aumento en el periodo refractario efectivo y la reducción en la velocidad de conducción asociada con el tejido isquémico en el Endocardio que permite cerrar el circuito reentrante que de otro modo encontrará el tejido dentro del periodo refractario. En general, estos resultados sugieren que la zona lavada podría actuar como un factor de substrato pro-arrítmico ayudando en el establecimiento de taquicardia ventricular sostenida.

5.3 Future work

This thesis opens the development of future research lines in both GPU computing and computational electrophysiology. These future lines are described in the following:

• To extend the implementation of the GPU solver to multi-GPU architectures. Even though high performance GPU cards continue to increase their capabilities in terms of number of cores and memory, a model of the entire heart will require a significant amount of resources that cannot be satisfied by a single GPU card.
• Implement new finite element technologies, in particular, enriched finite elements with residual-free bubbles. Residual-free bubbles are defined in such a way that they are solution of the discretized (in time) homogeneous parabolic equation.

• Incorporate the specialized conduction system in the model of acute ischemia in order to study if the ectopic beat can be generated due to injury currents established in the border zone of the ischemic tissue.

• Consider a realistic geometry of the ischemic lesion obtained from medical images in order to study the impact of the geometry in both the vulnerability window and the reentry patterns.

5.4 Contributions

5.4.1 Projects

This thesis was partially supported by the “VI Plan Nacional de Investigación Científica, Desarrollo e Innovación Tecnológica” from the Ministerio de Economía y Competitividad of Spain (grant numbers TIN2012-37546-C03-01 and TIN2012-37546-C03-03) and the European Commission (European Regional Development Funds - ERDF - FEDER).

5.4.2 Journal papers


Conferences


Modelling and simulation of the electric activity of the heart using GPU
Andrés Mena Tobar


Modelling and simulation of the electric activity of the heart using GPU
Andrés Mena Tobar


108

Modelling and simulation of the electric activity of the heart using GPU

Andrés Mena Tobar


