Asymptotic analysis in cardiac electrophysiology : applications in modeling and in data assimilation

Annabelle Collin

ANALYSE ASYMPTOTIQUE EN ÉLECTROPHYSIOLOGIE CARDIAQUE. APPLICATIONS À LA MODÉLISATION ET À L'ASSIMILATION DE DONNÉES.

THÈSE DE DOCTORAT

Présentée par

Annabelle COLLIN

pour obtenir le grade de

DOCTEUR DE
L' UNIVERSITÉ PIERRE ET MARIE CURIE - Paris VI

Spécialité : MATHÉMATIQUES APPLIQUÉES

Soutenue publiquement le ? ? ? ? ? ? ? ? ? ? ? ? ? devant le jury composé de :

? ? ? ?	Président
Dominique CHAPELLE	Directeur de thèse
Jean-Frédéric GERBEAU	Directeur de thèse
? ? ? ?	Rapporteur
? ? ? ?	Rapporteur
? ? ? ?	Examinateur

Après avis favorables des rapporteurs : ? ? ? ? et ? ? ? ?

ANALYSE ASYMPTOTIQUE EN ÉLECTROPHYSIOLOGIE CARDIAQUE. APPLICATIONS À LA MODÉLISATION ET À L'ASSIMILATION DE DONNÉES.

Resumé :

Cette thèse est dédiée au développement d'outils mathématiques innovants qui ont pour but d'améliorer la modélisation de l'électrophysiologie cardiaque.

Une présentation approfondie du modèle bidomaine – un système d'équations de réaction-diffusion – à domaine fixé est proposée en s'appuyant sur la littérature. De plus, nous donnons une justification mathématique du processus d'homogénéisation à l'aide la convergence «2-scale». Une étude de l'impact des déformations mécaniques dans les lois de conservation à l'aide de la théorie des mélanges est ensuite faite.

Comme les techniques d'imagerie actuelles ne fournissent globalement que des surfaces pour les oreillettes cardiaques dont l'épaisseur est très faible, une réduction dimensionnelle du modèle bidomaine dans une couche mince à une formulation posée sur la surface associée est étudiée. L'enjeu est aussi crucial en termes d'efficacité de calcul. À l'aide de techniques développées pour les modèles de coques, une analyse asymptotique des termes de diffusion est faite sous des hypothèses de gradient d'anisotropie fort à travers l'épaisseur, situation rencontrée dans les oreillettes. Des simulations 2D-3D illustrent les résultats. Puis, une modélisation complète du cœur – avec le modèle asymptotique pour les oreillettes et le modèle volumique pour les ventricules – permet la simulation de cycles complets d'électrocardiogramme. De plus, les méthodes développées dans l'analyse asymptotique sont utilisées pour obtenir des résultats de convergence forte pour les modèles de coque-3D.

Enfin, dans la perspective de «personnaliser» les modèles, une méthode d'estimation des modèles d'électrophysiologie est proposée. Les données médicales intégrées dans notre modèle – au moyen d'un filtre d'état de type *Luenberger* spécialement conçu – sont les cartes d'activation électrique. Ces problématiques apparaissent dans d'autres domaines où les modèles (de réaction-diffusion) et les données (position du front) sont très similaires, comme la propagation de feux ou la croissance tumorale.

Mots-clés : Analyse asymptotique ; Couches minces ; Modélisation ; Assimilation de données ; Électrophysiologie cardiaque

ASYMPTOTIC ANALYSIS IN CARDIAC ELECTROPHYSIOLOGY. APPLICATIONS IN MODELING AND IN DATA ASSIMILATION.

Abstract:

This thesis aims at developing innovative mathematical tools to improve cardiac electrophysiological modeling.

A detailed presentation of the bidomain model – a system of reaction-diffusion equations – with a fixed domain is given based on the literature. Furthermore, we mathematically justify the homogenization process using the 2-scale convergence. Then, a study of the impact of the mechanical deformations in the conservation laws is performed using the mixture theory.

As the atria walls are very thin and generally appear as thick surfaces in medical imaging, a dimensional reduction of the bidomain model in a thin domain to a surface-based formulation is studied. The challenge is also crucial in terms of computational efficiency. Following similar strategies used in shell mechanical modeling, an asymptotic analysis of the diffusion terms is done with assumptions of strong anisotropy through the thickness, as observed in the atria. Simulations in 2D and 3D illustrate these results. Then, a complete modeling of the heart – with the asymptotic model for the atria and the volume model for the ventricles – allow the simulation of full electrocardiogram cycles. Furthermore, the methods developed in the asymptotic analysis are used to obtain strong convergence results for the 3D-shell models.

Finally, a specific data assimilation method is proposed in order to «personalize» the electrophysiological models. The medical data assimilated in the model – using a *Luenberger*-like state filter specially designed – are the maps of electrical activation. The proposed methods can be used in other application fields where models (reaction-diffusion) and data (front position) are very similar, as for fire propagation or tumor growth.

Keywords: Asymptotic analysis; Thin domains; Modeling; Data assimilation; Cardiac electrophysiology

Contents

Introduction

Cette thèse s'inscrit dans un vaste et ambitieux projet de modélisation du cœur humain entrepris par les équipes MΞDISIM (anciennement MACS) et REO chez Inria depuis une dizaine d'années environ. Ma thèse s'est effectuée sous la direction de Dominique Chapelle, MΞDISIM et Jean-Frédéric Gerbeau, REO. La dernière année, j'ai aussi été encadrée par Philippe Moireau, MΞDISIM. Ma thèse a été financée par Inria sur une bourse CORDI-S.

Contexte de la thèse

L'objectif de cette thèse est de développer des outils mathématiques performants afin d'améliorer la modélisation des phénomènes électriques ayant lieu dans le cœur humain. Ainsi, cette thèse mélange plusieurs disciplines et communautés scientifiques que sont la médecine, la biologie et les mathématiques. La médecine est devenue un champ d'application très présent dans le domaine des mathématiques appliquées. De nombreuses simulations de systèmes biologiques variés sont aujourd'hui proposées aux médecins avec l'idée que ces simulations – adaptées le plus possible à chaque patient – aident les médecins dans leur diagnostic ainsi que dans leur stratégie thérapeutique. L'objectif n'est pas de remplacer les médecins mais de mettre à leur disposition de nouveaux outils innovants qui puissent les aider à planifier une opération, à conseiller ainsi qu'à optimiser les traitements et aussi bien sûr à établir un diagnostic rapidement et de façon non invasive de préférence. Trouver des modèles mathématiques qui permettent une première compréhension de systèmes – même très complexes – est la première étape vers cet objectif. En plus de leur intérêt descriptif, ces modèles – et les méthodes mathématiques développées autour d'eux – peuvent être utilisés dans des situations médicales très diverses. Depuis longtemps, par exemple, ils permettent d'anticiper des événements médicaux. En effet, en épidémiologie des maladies transmissibles, ils peuvent prévoir une épidémie ainsi que sa période d'incubation, comme celle de l'infection par le VIH [3]. Par la suite, ils ont aussi été utilisés dans la modélisation de situations cliniques variées, comme par exemple la prédiction de la rupture d'un anévrisme cérébral [60] ou le taux de survie après un cancer [59].

Dans cette thèse nous nous sommes intéressés à la modélisation de l'électrophysiologie cardiaque qui étudie les phénomènes électrochimiques qui se produisent dans le tissu cardiaque. Il y a un réel enjeu économique et sociétal autour des maladies cardio-vasculaires (MCV) qui sont une des principales causes de mortalité en Europe [52], étant à l'origine de 47% des décès en Europe (40% dans l'Union Européenne). Les MCV ont aussi un coût économique majeur estimé à près de 196 milliards d'euros par an dans l'Union Européenne. Dans la plupart des pays européens, les chiffres s'améliorent. Cependant, certains facteurs favorisant ces maladies continuent d'augmenter, comme par exemple le tabagisme, les mauvaises habitudes alimentaires ou encore le manque d'activité physique. Ces facteurs sont communs à d'autres pays développés ou en voie de développement qui connaissent donc les mêmes problématiques [25, 41]. Un nombre important de phénomènes physiques complexes a lieu dans le cœur, ce qui est en fait un véritable trésor pour la modélisation. Cet organe a pour

rôle d'assurer la circulation du sang en le pompant par des contractions rythmées. Une onde électrique, mesurée par les cardiologues à l'aide d'un électrocardiogramme, est à l'origine de chacun de ses battements. Les cellules cardiaques qui sont impactées par cette onde de dépolarisation répondent en se contractant. L'analyse mathématique des phénomènes biologiques qui ont lieu dans le cœur couplent alors trois domaines d'étude : l'électrophysiologie, la mécanique tissulaire ainsi que la mécanique des fluides. Afin d'obtenir une simulation complète du cœur, une étude approfondie de ces 3 phénomènes est nécessaire ainsi qu'une étude de leur impact les uns sur les autres. La littérature à ce sujet est abondante. Nous pouvons citer par exemple les travaux pionniers de *Peskin* [56] en 1977 et plus récemment *Watanabe et al.* [68] sur la modélisation de l'écoulement du sang dans le cœur avec un couplage fluide-structure. Des méthodes complexes d'interaction fluide-structure pour l'écoulement sanguin en général ont été développées comme par exemple dans [22] où un couplage de l'équation de Navier-Stokes incompressible sur la partie fluide avec des équations mécaniques régissant la dynamique de la structure est proposé. En ce qui concerne la mécanique cardiaque, nous pouvons citer les travaux de [32] (côté mécanique du solide) et [57] (côté fluide) à qui nous devons les premières contributions majeures. Plus récemment, un modèle électromécanique du cœur a été proposé par [64]. Dans cette étude, la loi de comportement du matériau est décomposée en une partie passive (propriétés visco-élastiques du myocarde) et une partie active (réactive à l'activation électrique). Des premiers modèles cellulaires apparaissent dès le début des années 50 avec les travaux de *Hodgkin* et *Huxley* [30, 31] (prix Nobel de médecine en 1963) qui proposent la première modélisation du potentiel d'action pour l'axone géant. Ces travaux seront ensuite appliqués en électrophysiologie cardiaque par *Noble* [53] et de nombreux modèles de plus en plus complexes suivront. Nous pouvons citer entre autres le modèle *Luo-Rudy* [44] adapté aux ventricules, les deux chambres inférieures du cœur et le modèle *Courtemanche-Ramirez-Nattel* [18] pour les oreillettes, les deux chambres supérieures du cœur. Se couplant à ces modèles de l'échelle cellulaire représentés par des équations différentielles ordinaires (EDOs), des modèles de propagation du signal dans le tissu se sont développés. Ces équations aux dérivées partielles (EDPs) qui modélisent la propagation de l'onde électrique sont basées sur des modèles de réaction diffusion comme le très classique modèle bidomaine proposé par *Tung* en 1978 [65]. Une fois ces modèles bien établis, la communauté mathématique de l'analyse numérique s'est intéressée de près à ces modèles d'électrophysiologie. On peut citer par exemple les travaux précurseurs de *Colli-Franzone* à partir de 1979 [15, 14] ainsi que ceux de *Panfilov* et *Winfree* [55] en 1985. Comme on pourra le constater dans la suite, la littérature autour des études d'existence, d'unicité ainsi que sur le développement de schémas numériques adaptés est abondante. Comme de nombreuses pathologies cardio-vasculaires sont dues à des troubles électro-physiologiques qui perturbent le rythme cardiaque, l'étude de cette dernière est un véritable enjeu. De nombreux progrès ont été faits que ce soit au niveau de la modélisation et de l'étude mathématique et numérique des EDOs et des EDPs qui apparaissent. Cependant, nous ne sommes pas encore en mesure de produire des simulations – dans un temps raisonnable et adaptées à chaque patient – pouvant aider les médecins par exemple dans leur diagnostic ainsi qu'à prédire les effets d'un traitement. Cette thèse apporte des contributions sur ce sujet et permet de se rapprocher de cet objectif.

Objectifs de la thèse et enjeux théoriques

De nombreuses questions se posent encore en modélisation sur la nature et l'impact des divers mécanismes de couplage entre phénomènes mécaniques et électriques. Nous nous sommes intéressés ici au modèle bidomaine, un modèle très connu et très utilisé par la communauté. L'étude de l'impact des déformations du cœur par exemple sur la convection des charges ioniques demande une excellente compréhension de ce modèle. La compréhension des hypothèses qui sont utilisées pour l'établir sur un domaine supposé fixe est primordiale afin de vérifier si elles sont toujours acceptables en domaine mobile. C'est une des contributions de cette thèse qui peut se résumer par :

Étudier l'impact des déformations mécaniques sur les lois de conservation du modèle bidomaine.

La réduction dimensionnelle d'un modèle décrit par des EDPs dans une couche mince à une formulation posée sur la surface associée présente des enjeux importants en termes d'efficacité de calcul et prend un éclairage particulier dans les applications cliniques, avec par exemple la perspective de réaliser des simulations au cours d'une intervention afin de guider celle-ci (comme l'ablation par radio-fréquences en cardiologie). Dans ce contexte, nous souhaitons proposer un modèle surfacique d'électrophysiologie – sous des hypothèses de gradient d'anisotropie fort à travers l'épaisseur du domaine mince considéré – à l'aide d'une analyse asymptotique. Cette situation de forte anisotropie dans l'épaisseur malgré des parois très fines se rencontre dans les oreillettes, les deux chambres supérieures du cœur. Il s'agit de la principale contribution de cette thèse :

Établir un modèle réduit du modèle bidomaine mathématiquement justifié par une analyse asymptotique et permettant de prendre en compte des fortes anisotropies dans l'épaisseur.

Enfin, une personnalisation des modèles d'électrophysiologie sera proposée. Les nombreuses données médicales non-invasives qui sont à notre disposition peuvent être utilisées afin de produire des simulations prédictives d'un patient donné et/ou de retrouver les paramètres d'un modèle. Les données que nous souhaiterons intégrer dans notre modèle – à l'aide de méthodes séquentielles développées initialement par *Kalman* (filtre optimal) [35] et *Luenberger* [43] – sont les cartes d'activation électrique. Ces problématiques apparaissent aussi dans d'autres domaines où les modèles (modèles de réaction-diffusion) et les données (position du front) sont très similaires. On peut citer par exemple l'étude de propagation de feux de forêt ou de croissance tumorale. Cette dernière contribution peut se résumer par :

Proposer un estimateur d'état d'un modèle de réaction-diffusion assimilant des observations de type propagation de front.

Tous ces objectifs sont l'occasion d'utiliser de nombreux outils mathématiques performants dont le rôle est de justifier par une analyse mathématique les résultats. Un outil revient très régulièrement dans cette thèse et sert de fil conducteur à celle-ci. Il s'agit de l'analyse asymptotique. Elle peut être vue comme un outil de modélisation mathématique

car elle permet de déduire des modèles plus simples à partir d'autres et de hiérarchiser les modèles décrivant un phénomène donné. L'analyse asymptotique joue aussi un rôle fondamental dans la réduction du temps de calcul permettant d'utiliser des modèles qui n'ont pas des temps de simulation prohibitifs mais qui sont justifiés dans un certain cadre à l'aide de théorèmes de convergence. L'analyse asymptotique est aussi utilisée ici dans le cadre des problèmes inverses.

Analyse asymptotique

Modélisation

Pour modéliser le signal électrique se propageant dans le cœur, on utilise classiquement le modèle bidomaine [65, 58], un système d'équations de réaction diffusion couplé à un modèle ionique cellulaire modélisé par une ou plusieurs équations différentielles. Si nous considérons le modèle ionique de *Mitchell-Schaeffer* [46], un modèle phénoménologique très apprécié pour sa performance et surtout sa simplicité, le modèle bidomaine s'écrit :

$$\begin{cases} A_m\Big(C_m\dfrac{\partial V_m}{\partial t}+I_{ion}(V_m,w)\Big)-\vec{\nabla}\cdot\big(\vec{\vec{\sigma}}_i\cdot\vec{\nabla}V_m\big)=\vec{\nabla}\cdot\big(\vec{\vec{\sigma}}_i\cdot\vec{\nabla}u_e\big)+A_mI_{app} & \text{sur } \Omega\times(0,T),\\ \vec{\nabla}\cdot\Big(\big(\vec{\vec{\sigma}}_i+\vec{\vec{\sigma}}_e\big)\cdot\vec{\nabla}u_e\Big)=-\vec{\nabla}\cdot\big(\vec{\vec{\sigma}}_i\cdot\vec{\nabla}V_m\big), & \text{sur } \Omega\times(0,T),\\ (\vec{\vec{\sigma}}_i\cdot\vec{\nabla}V_m)\cdot\vec{n}_{\partial\Omega}=-(\vec{\vec{\sigma}}_i\cdot\vec{\nabla}u_e)\cdot\vec{n}_{\partial\Omega}, & \text{sur } \partial\Omega\times(0,T),\\ (\vec{\vec{\sigma}}_e\cdot\vec{\nabla}u_e)\cdot\vec{n}_{\partial\Omega}=0, & \text{sur } \partial\Omega\times(0,T),\\ \partial_t w+g(V_m,w)=0, & \text{sur } \Omega\times(0,T), \end{cases} \tag{1}$$

où

$$I_{ion}(V_m,w) = -\frac{w(V_m-V_{min})^2(V_{max}-V_m)}{\tau_{in}(V_{max}-V_{min})}+\frac{V_m-V_{min}}{\tau_{out}(V_{max}-V_{min})}, \tag{2}$$

$$g(V_m,w) = \begin{cases} \dfrac{w}{\tau_{open}}-\dfrac{1}{\tau_{open}(V_{max}-V_{min})^2} & \text{if } V_m<V_{gate},\\ \dfrac{w}{\tau_{close}} & \text{if } V_m\geq V_{gate}, \end{cases} \tag{3}$$

Tous les paramètres seront présentés dans la suite mais on peut rapidement donner quelques éléments pour la lecture de ces équations. Les deux inconnues sont le potentiel extracellulaire u_e et le potentiel transmembranaire $V_m = u_i - u_e$ où u_i est le potentiel intracellulaire. Nous pouvons remarquer qu'il s'agit d'un système réaction-diffusion couplé avec une ODE qui représente le modèle ionique choisi, ici le modèle de *Mitchell-Schaeffer*. Ce modèle ne prend pas en compte les déformations mécaniques du cœur. Pour pouvoir intégrer l'impact de ces déformations, il faut avoir une vision très précise de la mise en équation du modèle bidomaine afin de comprendre ce qui doit être adapté.

Nous proposons donc dans cette thèse une étude détaillée des différents phénomènes qui mènent aux équations bidomaines. Utilisant la littérature abondante sur ce sujet, nous avons

regroupé l'ensemble des étapes et apporté une justification mathématique à certaines parties. Tout d'abord un modèle microscopique de l'activité électrique peut être écrit [40, 49] à l'aide des lois de conservation et de l'équation de Nernst-Planck. À l'aide d'une analyse asymptotique présentée dans [48, 62], une hypothèse d'électroneutralité est validée. Un modèle bidomaine microscopique couplant les phénomènes intra-cellulaires et extra-cellulaires – d'où l'origine du nom «bidomaine» – est alors écrit. Pour obtenir le modèle bidomaine, une homogénéisation est nécessaire et a été initialement présentée dans [51]. La méthode pour obtenir la convergence n'est pas explicitée mais dans l'appendice de [16], une justification formelle est donnée. Cependant des hypothèses sur la cellule semblent nécessaires et la convergence du terme surfacique n'est pas explicitée. Nous proposons ici une étape supplémentaire vers la justification mathématique complète de l'homogénéisation, un processus asymptotique, en utilisant la convergence à 2 échelles proposée par [1, 6], et qui permet d'arriver à un modèle bidomaine macroscopique. En particulier le passage du terme surfacique écrit sur la membrane du domaine à un terme volumique sur tout le cœur est étudié dans un cas régulier. Une homogénéisation sur un cas d'étude proche a été proposée très récemment dans [2].

Dans un second temps, les questions qui se posent en modélisation à propos de l'impact des déformations mécaniques sur les phénomènes électriques ainsi qu'électrochimiques ayant lieu dans la cellule sont étudiées. L'influence des déformations sur la propagation de l'onde électrique dans le tissu cardiaque – connue dans la littérature sous le nom de «mechano-electrical feedback» [50, 37] – est aujourd'hui acceptée. Cependant, à notre connaissance, les modèles proposés dans la littérature modifient uniquement les lois de comportement. Nous prenons en compte dans ce travail les phénomènes mécaniques directement dans les lois de conservation. Afin d'éviter l'étape d'homogénéisation, précédemment évoquée, la théorie des mélanges, dans laquelle on suppose qu'à tout instant toutes les phases sont présentes à chaque point matériel, est utilisée. Cette théorie des mélanges – où le milieu considéré n'a pas besoin d'avoir de structure cellulaire – est comparée avec le processus d'homogénéisation. Des lois de comportement viennent fermer le système. Nous nous sommes attachés dans ce travail à mettre en évidence les phénomènes prépondérants. Le modèle que nous proposons s'écrit sous forme variationnelle : $\forall \vec{v}_1^*,\, v_2^*,\, v_3^*$,

$$\begin{cases} \displaystyle\int_{\Omega(t)} \big(\vec{\vec{\lambda}} : \vec{\vec{\varepsilon}}(\vec{v}_1^*) - \rho_s(\vec{f} - \vec{\gamma}_s)\cdot \vec{v}_1^*\big) d\Omega - \int_{\partial\Omega(t)} \vec{t}\cdot\vec{v}_1^*\, dS = 0, \\ \displaystyle\int_{\Omega(t)} \Big(\frac{\partial(\phi_i\rho_i^0 + \phi_e\rho_e^0)}{\partial t} - \vec{\nabla}_{\vec{x}}\cdot(\phi_i\rho_i^0 + \phi_e\rho_e^0)\vec{v}_s - \vec{\nabla}_{\vec{x}}\cdot(\sigma_i\vec{\nabla}_{\vec{x}}\, u_i + \sigma_e\vec{\nabla}_{\vec{x}}\, u_e)\Big) v_2^*\, d\Omega = 0, \\ \displaystyle\int_{\Omega(t)} \Big(\phi_i\frac{\partial\rho_i^0}{\partial t} - \phi_i\vec{\nabla}_{\vec{x}}\cdot\rho_i^0\vec{v}_s - \vec{\nabla}_{\vec{x}}\cdot\sigma_i\vec{\nabla}_{\vec{x}}\, u_i\Big) v_3^*\, d\Omega + \int_{\Omega(t)} A_m C_m\Big(\frac{\partial V_m}{\partial t}\Big|_{\vec{\xi}} + I_{ion}^{tot}\Big) v_3^*\, d\Omega = 0. \end{cases} \tag{4}$$

De nouveaux termes apparaissent, comme par exemple les termes prenant en compte la vitesse du solide $\vec{v}_s$, et les équations sont écrites sur des domaines mobiles. Une discussion sur ce nouveau modèle et des simulations numériques dans le cas où le cœur est considéré comme incompressible sont proposées. Ce travail sera la seule partie dans cette thèse où

les phénomènes mécaniques seront pris en compte. Dans le reste, le tissu cardiaque sera considéré comme un tissu non contractile isolé mécaniquement et électriquement du reste du corps. Bien que ces hypothèses de modélisation puissent paraître restrictives, elles sont justifiées dans un certain cadre. Ces travaux seront présentés dans le Chapitre 2.

Modèles asymptotiques

Origine du problème de l'utilisation d'un modèle 3D en électrophysiologie atriale Une grande partie de cette thèse a pour but d'établir un modèle d'électrophysiologie réduit obtenu par une analyse asymptotique, de le valider et de l'utiliser dans des cas réalistes. Les oreillettes qui forment la partie haute du cœur sont très fines. De plus, elles apparaissent comme des surfaces dans la plupart des données d'imagerie médicales. Même dans le cas où un maillage 3D des oreillettes est accessible, les temps de calcul d'un modèle bidomaine sur ce domaine mince sont prohibitifs. Le contexte asymptotique – permettant de proposer un modèle dont les équations sont posées sur la surface moyenne – semble alors tout à fait adapté. La principale difficulté réside dans la forte anisotropie présente à travers l'épaisseur des oreillettes. Afin de prendre en compte – dans les équations du modèle bidomaine (1) – les directions privilégiées que sont les fibres musculaires pour le signal électrique, les tenseurs de diffusion $\vec{\vec{\sigma}}^{i,e}$ sont décomposés comme suit

$$\vec{\vec{\sigma}}^{i,e} = \sigma_{i,e}^{t}\vec{\vec{I}} + (\sigma_{i,e}^{l} - \sigma_{i,e}^{t})\,\vec{\tau}\otimes\vec{\tau}, \tag{5}$$

où le vecteur unitaire $\vec{\tau}$ est parallèle à la direction locale de la fibre. La propagation du signal se fait donc pour une partie de façon homogène et pour une autre en suivant les directions des fibres musculaires qui composent le coeur. Dans les oreillettes, les directions des fibres ne restent pas constantes dans l'épaisseur. Dans la littérature, les directions principales des fibres à l'endocarde (surface interne) ainsi qu'à l'épicarde (surface externe) sont données [29, 28]. Les changements d'orientation des fibres dans les parois des oreillettes sont mal connus et nous avons fait le choix d'une modélisation par une variation linéaire des fibres dans l'épaisseur. Une autre hypothèse de modélisation peut être faite comme par exemple les modèles double-couche proposés par [33, 66, 39] qui semblent bien adaptés aux zones proches des veines pulmonaires. Notre analyse asymptotique peut s'adapter à une autre hypothèse. Dans notre cas où les fibres sont supposées varier linéairement dans l'épaisseur, un vecteur $\vec{\tau}$ de l'épaisseur se décompose sous la forme suivante :

$$\vec{\tau}(\xi^1,\xi^2,\xi^3) = \underline{\tau}_0(\xi^1,\xi^2)\cos\left(\frac{2\theta(\xi^1,\xi^2)\xi^3}{d}\right) + \underline{\tau}_0^{\perp}(\xi^1,\xi^2)\sin\left(\frac{2\theta(\xi^1,\xi^2)\xi^3}{d}\right), \tag{6}$$

où $\underline{\tau}_0$ et $\underline{\tau}_0^{\perp}$ sont des vecteurs unitaires tels que

- $\underline{\tau}_0$ appartient au plan tangent de la surface en chaque point
- $\underline{\tau}_0^{\perp}$ est un vecteur unitaire orthogonal à $\underline{\tau}_0$ qui appartient au plan tangent,

et 2θ correspond à l'angle total de rotation. La Figure 1 illustre cette variation des fibres dans l'épaisseur des oreillettes.

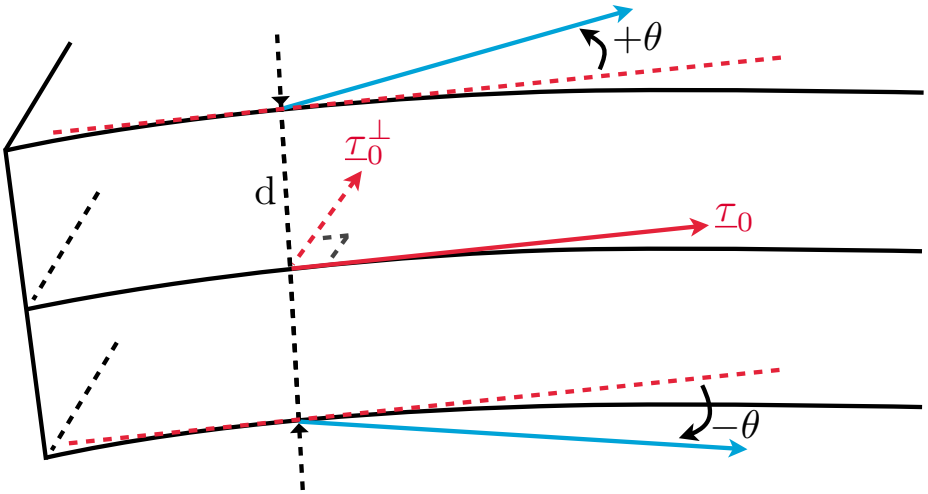

FIGURE 1: Rotations des fibres à travers l'épaisseur

Convergences faible et forte du modèle de diffusion Comme la difficulté réside dans l'anisotropie des tenseurs de diffusion, l'analyse asymptotique a été effectuée seulement sur le modèle de diffusion. Les préliminaires géométriques qui permettent de considérer un nouveau système de coordonnées dans laquelle l'épaisseur correspond à la troisième variable notée ξ^3 sont donnés. Nous avons suivi les stratégies utilisées en modélisation mécanique des coques [12, 9]. Nous considérons alors une décomposition de Galerkin du modèle de diffusion posé sur un sous-espace formé par des variations polynomiales – quadratiques dans notre cas – dans la variable correspondant à l'épaisseur de la grandeur qui nous intéresse. Nous définissons ε comme le ratio entre l'épaisseur et le diamètre de la surface moyenne. Pour chaque ε, une unique solution au problème posé dans le sous-espace existe et on la note $u^\varepsilon = u_0^\varepsilon + \xi^3 u_1^\varepsilon + (\xi^3)^2 u_2^\varepsilon$. En ne gardant que les premiers termes du développement asymptotique, nous trouvons un problème limite

$$\sigma^t \underline{\nabla} \cdot (\underline{\nabla} u_0^l) + (\sigma^l - \sigma^t)\, \underline{\nabla} \cdot \Big(\big(I_0(\theta)\, \underline{\tau}_0 \otimes \underline{\tau}_0 + J_0(\theta)\, \underline{\tau}_0^{\perp} \otimes \underline{\tau}_0^{\perp} \big) \cdot \underline{\nabla} u_0^l \Big) = f_0, \tag{7}$$

avec $I_0(\theta) = \dfrac{1}{2} + \dfrac{1}{4\theta}\sin(2\theta)$ et $J_0(\theta) = \dfrac{1}{2} - \dfrac{1}{4\theta}\sin(2\theta) = 1 - I_0(\theta)$. Ce problème admet une unique solution que nous notons u_0^l. La convergence faible ainsi que la convergence forte ont été prouvées en deux étapes différentes qui sont résumées dans le théorème suivant dont la preuve est donnée dans le Chapitre 3.

Proposition 1. *Les suites $(u_0^\varepsilon, u_1^\varepsilon)$ convergent fortement vers $(u_0^l, 0)$ dans $H^1(\mathcal{S})$ quand ε tend vers 0. Les suites $\varepsilon u_1^\varepsilon$ et $\varepsilon^2 u_2^\varepsilon$ convergent fortement vers 0 dans $H^1(\mathcal{S})$ quand ε tend vers 0. De plus $\varepsilon u_2^\varepsilon$ converge fortement vers 0 dans $L^2(\mathcal{S})$ quand ε tend vers 0.*

Application pour la convergence forte dans les modèles de coques-3D La preuve de la Proposition 1 se fait en 2 temps. La démonstration d'une version faible s'inspire fortement de ce qui est fait dans [12, 9] sur des modèles dits de coques-3D. Cependant, pour montrer

la convergence forte, de nouvelles méthodes sont introduites. Nous les avons ensuite appliquées sur les modèles de coques-3D dont la convergence forte n'avait pas encore été établie. Dans ce même travail, nous proposons aussi une preuve simplifiée de la convergence faible pour les modèles de coques. Les preuves sont données dans le Chapitre 4. La formulation variationnelle des modèles de coques s'écrit : chercher $\vec{U} \in \mathcal{V}_{\mathrm{m}}^{3D}$ tel que

$$A^{3D}(\vec{U}, \vec{V}) = F^{3D}(\vec{V}), \forall \vec{V} \in \mathcal{V}^{3D}. \tag{8}$$

Les définitions de A^{3D} et F^{3D} seront données dans la suite. Nous utilisons les mêmes décompositions de Galerkin, c'est-à-dire des variations polynomiales quadratiques, et le problème dans le nouvel espace devient : trouver $(\vec{u}^{\varepsilon}, \vec{\theta}^{\varepsilon}, \vec{\varrho}^{\varepsilon}) \in \mathcal{V}^{3Ds}$ tel que

$$A^{3Ds}(\vec{u}^{\varepsilon}, \vec{\theta}^{\varepsilon}, \vec{\varrho}^{\varepsilon}; \vec{v}, \vec{\eta}, \vec{\varsigma}) = F^{3Ds}(\vec{v}, \vec{\eta}, \vec{\varsigma}), \quad \forall (\vec{v}, \vec{\eta}, \vec{\varsigma}) \in \mathcal{V}^{3Ds}. \tag{9}$$

Nous isolons les deux premiers termes du développement asymptotique avec lesquels nous définissons les deux formes linéaires suivantes

$$A_{\mathrm{m}}^{3Ds} \text{ et } A_{\mathrm{b}}^{3Ds}.$$

La première est de degré 1 dans l'épaisseur et la deuxième de degré 3. Les termes pairs s'annulent lors de l'intégration. Contrairement au cas du modèle de diffusion, la semi-norme $\|\cdot\|_m$ construite à partir du premier terme A_{m}^{3Ds} n'est pas nécessairement une norme, c'est-à-dire que des éléments non nuls peuvent l'annuler. Le cas où $\|\cdot\|_m$ n'est pas une norme correspond à de la «flexion pure non-inhibée». Un espace $\mathcal{V}_{0\mathrm{b}}^{3Ds}$ complet pour la norme $\|\cdot\|_b$ (construite à partir de A_{b}^{3Ds}), dans lequel les éléments qui annulent la semi-norme $\|\cdot\|_m$ sont contenus, est construit. Avec des hypothèses de modélisation sur le terme source, le problème limite dans le cas non-inhibé est défini par : trouver $\vec{u}^0, \vec{\theta}^0, \vec{\varrho}^0 \in \mathcal{V}_{0\mathrm{b}}^{3Ds}$,

$$A_{\mathrm{b}}^{3Ds}(\vec{u}^0, \vec{\theta}^0, \vec{\varrho}^0; \vec{v}, \vec{\eta}, \vec{\varsigma}) = G^{3Ds}(\vec{v}), \quad \forall (\vec{v}, \vec{\eta}, \vec{\varsigma}) \in \mathcal{V}_{0\mathrm{b}}^{3Ds}, \tag{10}$$

où $G^{3Ds} = \varepsilon^2 F^{3Ds}$. La preuve du théorème de convergence forte qui suit est donnée dans cette thèse.

Proposition 2. *La solution $(\vec{u}^{\varepsilon}, \vec{\theta}^{\varepsilon}, \vec{\varrho}^{\varepsilon})$ du problème (9) converge fortement vers $(\vec{u}^0, \vec{\theta}^0, \vec{\varrho}^0)$ solution de (10) quand ε tend vers 0. De plus, la suite $\frac{1}{\varepsilon}\|\vec{u}^{\varepsilon}, \vec{\theta}^{\varepsilon}\|_{\mathrm{m}}$ tend vers 0.*

Le second cas correspond à une situation de «flexion pure inhibée» et avec $G^{3Ds} = F^{3Ds}$ le problème limite devient

$$A_{\mathrm{m}}^{3Ds}(\vec{u}^{\mathrm{m}}, \vec{\theta}^{\mathrm{m}}; \vec{v}, \vec{\eta}) = G^{3Ds}(\vec{v}), \quad \forall (\vec{v}, \vec{\eta}) \in \mathcal{V}_{\mathrm{m}}^{3Ds}. \tag{11}$$

Sous certaines hypothèses, le résultat suivant a été démontré

Proposition 3. *La suite $\left(\vec{u}^{\varepsilon}, \vec{\theta}^{\varepsilon}\right)$ converge fortement quand ε tend vers 0 vers $(\vec{u}^{\mathrm{m}}, \vec{\theta}^{\mathrm{m}})$ la solution de (11). De plus, la suite $\varepsilon\|\vec{u}^{\varepsilon}, \vec{\theta}^{\varepsilon}, \vec{\varrho}^{\varepsilon}\|_{\mathrm{b}}$ tend vers 0.*

Modèle asymptotique bidomaine Nous revenons à présent à l'électrophysiologie cardiaque. Utilisant la Proposition 1, nous sommes en mesure de proposer un modèle bidomaine asymptotique dont les équations sont posées sur la surface moyenne du domaine d'étude. Sous forme variationnelle, ce modèle s'écrit : pour tout $t > 0$, on cherche $V_m(\cdot,t) \in H^1(\mathcal{S})$, $u_e(\cdot,t) \in H^1(\mathcal{S})$ et $w(\cdot,t) \in L^\infty(\mathcal{S})$ tels que $\int_\mathcal{S} u_e \, dS = 0$ vérifiant

$$\begin{cases} A_m \displaystyle\int_\mathcal{S} \Big(C_m \frac{\partial V_m}{\partial t} + I_{ion}(V_m, w) \Big) \phi \, dS + \int_\mathcal{S} \Big(\underline{\underline{\sigma}}_i \cdot \big(\underline{\nabla} V_m + \underline{\nabla} u_e \big) \Big) \cdot \underline{\nabla} \phi \, dS \\ \qquad\qquad = A_m \displaystyle\int_\mathcal{S} I_{app}(V_m, w) \phi \, dS, \\ \displaystyle\int_\mathcal{S} \Big((\underline{\underline{\sigma}}_i + \underline{\underline{\sigma}}_e) \cdot \underline{\nabla} u_e \Big) \cdot \underline{\nabla} \psi \, dS + \int_\mathcal{S} \Big(\underline{\underline{\sigma}}_i \cdot \underline{\nabla} V_m \Big) \cdot \underline{\nabla} \psi \, dS = 0, \\ \partial_t w + g(V_m, w) = 0, \ \text{in } \mathcal{S} \times (0,T), \end{cases} \tag{12}$$

pour tout $\phi, \psi \in H^1(\mathcal{S})$ tels que $\int_\mathcal{S} \psi \, dS = 0$. Les tenseurs surfaciques sont définis par

$$\underline{\underline{\sigma}}_{i,e} = \sigma^t_{i,e} \underline{\underline{I}} + (\sigma^l_{i,e} - \sigma^t_{i,e})(I_0(\theta) \underline{\tau}_0 \otimes \underline{\tau}_0 + J_0(\theta) \underline{\tau}_0^\perp \otimes \underline{\tau}_0^\perp),$$

avec

$$I_0(\theta) = \frac{1}{2} + \frac{1}{4\theta} \sin(2\theta) \text{ et } J_0(\theta) = \frac{1}{2} - \frac{1}{4\theta} \sin(2\theta).$$

Nous pouvons remarquer que le cas $\theta = 0$ correspond à l'absence de variation des fibres dans l'épaisseur est nulle, et toute la partie anisotrope des tenseurs surfaciques est alors portée par $\underline{\tau}_0$. Plus θ est important, plus la direction orthogonale à la direction moyenne des fibres intervient. Des comparaisons de simulation par éléments finis entre ce modèle bidomaine asymptotique et le modèle bidomaine classique sont réalisées dans des cas de forte anisotropie dans l'épaisseur. Les résultats obtenus sont excellents, avec moins de 2% d'erreur en norme l^2. De plus, les temps de simulation sont très réduits avec le modèle asymptotique (432 min. avec un modèle 3D pour 26 min. avec un modèle surfacique pour 1500 ms de simulation – i.e. 2 cycles cardiaques – avec 15,000 pas de temps). Ces résultats sont aussi présentés dans le Chapitre 4.

Afin de valider ce modèle surfacique, nous avons proposé des simulations réalistes d'électrophysiologie atriale. À partir du Zygote[1] – un modèle de cœur basé sur des données anatomiques réelles – et en utilisant les logiciels de maillage `3-matic`[2] et `Yams` [24], un maillage surfacique valide pour des calculs de type éléments finis des deux oreillettes a été obtenu. En utilisant la littérature [29, 28], des directions de fibres réalistes ont été identifiées et prescrites. Nous avons ensuite implémenté le modèle ionique complexe car spécialement adapté pour les oreillettes, développé par Courtemanche, Ramirez et Nattel [18], pour le coupler au modèle bidomaine surfacique. La simulation obtenue est très réaliste et peut être réalisée – grâce à ces temps de simulation raisonnables (3-4 minutes pour un cycle cardiaque) – dans le cadre d'une consultation médicale par exemple. Ces résultats validés par des cardiologues sont présentés dans le Chapitre 5. Ils sont comparés avec des simulations 3D trouvées dans la littérature [20, 26, 38].

[1] `www.3dscience.com`
[2] `www.materialise.com`

Un électrocardiogramme est un examen médical non invasif qui donne une représentation graphique de l'activité électrique du cœur en mesurant le potentiel par différentes électrodes posées sur la surface du corps [67]. La simulation d'électrocardiogrammes est un véritable enjeu puisque c'est l'outil médical le plus utilisé par les médecins pour la détection de pathologies cardiaque. L'article [7] est le point de départ de ce travail. Des simulations du complexe QRST d'un électrocardiogramme correspondant à la dépolarisation et à la repolarisation des ventricules sont données dans cet article. Notre objectif est d'obtenir des simulations d'électrocardiogrammes complets c'est-à-dire qui prennent aussi en compte les oreillettes. Nous avons pour cela proposé un modèle bidomaine couplé, avec une partie volumique pour les ventricules et une partie asymptotique pour les oreillettes dans le Chapitre 6. Ce modèle a été couplé avec deux modèles ioniques complexes. Sur la surface des oreillettes, le modèle physiologique présenté par Courtemanche, Ramirez et Nattel [18] est utilisé. Dans la partie ventriculaire, le «*Minimal model for human Ventricular action potentials*», un modèle phénoménologique efficace, est appliqué (voir [8]) alors que pour les simulations de [7] le modèle plus simple de Mitchell-Schaeffer [46] a été utilisé. Ces modèles complexes nous permettent de proposer des électrocardiogrammes très réalistes dans des cas sains et pathologiques. Ce travail a été effectué en collaboration avec E. Schenone (doctorante dans l'équipe REO chez Inria).

Assimilation de données

Ces modèles complexes doivent alors être adaptés à chaque état physiologique afin de produire des simulations prédictives d'un patient donné. Dans cette optique, on peut s'appuyer naturellement sur les nombreuses données médicales disponibles, en particulier les cartes d'activation électrique d'un patient, [61]. Ces cartes donnent la position des fronts de dépolarisation à différents instants. Dans cette dernière partie, nous proposons d'assimiler ces données médicales à l'aide d'un observateur [43], fondé sur un modèle bidomaine et corrigeant sa dynamique en fonction des écarts aux mesures observées. Pour des questions de simplicité – la théorie reste cependant valide pour le modèle bidomaine – mais aussi de généralisation, nous considérons dans cette partie le modèle de réaction diffusion suivant

$$\left\{\begin{array}{rcll} \partial_t u - \vec{\nabla}\cdot(\vec{\vec{\sigma}}\cdot\vec{\nabla}u) & = & kf(u), & \mathcal{B}\times(0,T), \\ (\vec{\vec{\sigma}}\cdot\vec{\nabla}u)\cdot\vec{n} & = & 0, & \partial\mathcal{B}\times(0,T), \\ u(\vec{x},0) & = & u_0(\vec{x}), & \mathcal{B}. \end{array}\right. \tag{13}$$

Il peut par exemple s'agir d'un modèle monodomaine, un cas simplifié du modèle bidomaine avec $u = V_m$. Nous avons pris cette notation plus générale car ces travaux peuvent être utilisés dans d'autres domaines – comme par exemple les feux de forêt ou encore les propagations de tumeur – où des modèles similaires sont utilisés [4, 23]. Des données de même type sont disponibles – comme par exemple la position d'un feu de forêt à différents instants – et des méthodes d'assimilation existent dans la littérature [45, 63] basées sur des filtres de type Kalman non linéaires qui permettent de prédire la position du feu. L'idée ici est de proposer un observateur de type Luenberger [43] afin de contrer le grand défaut des filtres de type

Kalman [35, 34, 21, 47] : des temps de simulations qui sont prohibitifs dès que la dimension de l'espace où ils sont appliqués est grande.

Les cartes d'activation sont des données de type ensembles de niveaux. Or des *level sets* similaires apparaissent naturellement en traitement d'images, notamment dans la détection d'objets (*i.e.* la segmentation) où l'on résout une équation dite eikonale caractérisant l'évolution du contour de l'objet poursuivant celui qui est observé [69, 54]. Or par une analyse asymptotique exploitant la «raideur» du front, le modèle bidomaine peut être lui-même relié à une équation eikonale [36]. La méthode consiste à se placer dans un nouveau système de coordonnées dans lequel la première direction ξ_1 correspond à la direction normale au front. À l'aide d'une analyse asymptotique dans les autres coordonnées, l'équation eikonale de courbure est obtenue :

$$\partial_t \phi_u = |\vec{\nabla}\,\phi_u|\Big(\sigma\vec{\nabla}\cdot\Big(\frac{\vec{\nabla}\,\phi_u}{|\vec{\nabla}\,\phi_u|}\Big)+\sqrt{\sigma k}\,c_0\Big),\ \mathcal{B}\times(0,T), \tag{14}$$

où ϕ_u est une ligne de niveau associée à u. Elle est définie telle que

$$\phi_u > 0 \text{ si } u > c_{\text{th}},\ \phi_u < 0 \text{ si } u < c_{\text{th}} \text{ et } \phi_u = 0 \text{ si } u = c_{\text{th}},$$

où c_{th} est la valeur du potentiel à partir de laquelle une zone est considérée comme dépolarisée.

Nous définissons donc notre observateur d'état du modèle bidomaine de telle sorte qu'asymptotiquement il corresponde à un observateur de l'équation eikonale résultante, s'inspirant alors des propriétés de suivi de contours obtenues en traitement d'images. Nous utilisons des méthodes séquentielles d'assimilation de données, soit de type Kalman [35, 34, 21, 47] – c'est-à-dire optimal au sens d'un certain critère [5] – ou soit de type Luenberger [42, 43]. On suppose alors que la solution cible vérifie le modèle considéré (13) mais que la condition initiale $u_0 = \hat{u}_0 + \zeta^u$ est mal connue, c'est-à-dire que l'on a un *a priori* $\hat{u}_0$ sur celle-ci mais aussi une incertitude représentée par ζ^u. Cette solution n'est pas connue mais des données obtenues grâce à des observations sont disponibles. Les données considérées ici correspondent aux cartes d'activation qui nous donnent la position du contour correspondant à la position du front. À partir du contour à chaque instant, un objet $z_u(t)$ est défini avec une valeur à l'intérieur du contour et une autre à l'extérieur.

La méthode d'assimilation de données que nous utilisons consiste alors à modifier la dynamique d'un nouveau système vérifié par l'observateur $\hat{u}$ tel que

$$\hat{u} \to u.$$

La figure 2 expose dans un cas 2D la situation à un instant t. L'objet $\mathrm{z}_u(t)$ défini à partir de l'observation du front de la cible u à l'instant t est représenté en rouge. En bleu, il s'agit de la position du front au même instant de la solution de l'observateur $\hat{u}$. L'objectif est de modifier le modèle vérifié par l'observateur afin que le contour bleu $\Gamma_{\hat{u}}$ converge vers le contour rouge.

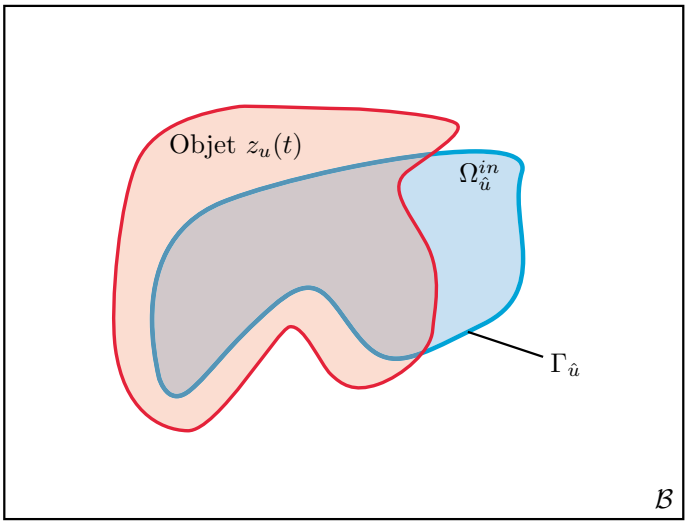

FIGURE 2: Observation et observateur à un instant t dans un cas 2D

Le modèle vérifié par l'observateur que nous proposons dans cette thèse est

$$\begin{cases} \partial_t \hat{u} - \vec{\nabla} \cdot (\vec{\hat{\sigma}} \cdot \vec{\nabla}\, \hat{u}) &= kf(\hat{u}) & \\ + \lambda \hat{\delta}(\Gamma_{\hat{u}}, \vec{x}) \alpha(|\vec{\nabla}\, \hat{u}|) \Big(-\big(z_u - C_1(\Omega^{\text{in}}_{\hat{u}})\big)^2 + \big(z_u - C_2(\Omega^{\text{in}}_{\hat{u}})\big)^2 \Big), & \mathcal{B} \times (0,T), \\ (\vec{\hat{\sigma}} \cdot \vec{\nabla}\, \hat{u}) \cdot \vec{n} &= 0, & \partial\mathcal{B} \times (0,T), \\ \hat{u}(\vec{x}, 0) &= \hat{u}_0(\vec{x}), & \mathcal{B}, \end{cases} \tag{15}$$

avec $\Omega^{\text{in}}_{\hat{u}}$, un espace – évoluant en temps – qui correspond à la zone déjà dépolarisée de l'observateur à chaque instant. Les fonctions C_1 et C_2 correspondent à la moyenne de l'objet z_u à l'intérieur et à l'extérieur de $\Omega^{\text{in}}_{\hat{u}}$:

$$C_1(\Omega^{\text{in}}_{\hat{u}}) = \frac{\int_{\Omega^{\text{in}}_{\hat{u}}} z_u d\vec{x}}{\int_{\Omega^{\text{in}}_{\hat{u}}} d\vec{x}} \text{ et } C_2(\Omega^{\text{in}}_{\hat{u}}) = \frac{\int_{\mathcal{B}\setminus\Omega^{\text{in}}_{\hat{u}}} z_u d\vec{x}}{\int_{\mathcal{B}\setminus\Omega^{\text{in}}_{\hat{u}}} d\vec{x}}.$$

Ces quantités qui permettent de comparer les deux fronts sont issues de la théorie du traitement d'images [69, 54].

Cet observateur est justifié mathématiquement. À l'aide de la théorie des dérivées de forme [19, 27], le terme de gain de l'observateur est linéarisé et le théorème suivant est alors montré.

Proposition 4. *Si la raideur du front de l'objet z_u est suffisamment élevée (la condition exacte sera donnée)*

$$\lambda \hat{\delta}(\Gamma_{\hat{u}}, \vec{x}) \alpha(|\vec{\nabla}\, \hat{u}|) \Big(-\big(z_u - C_1(\Omega^{in}_{\hat{u}})\big)^2 + \big(z_u - C_2(\Omega^{in}_{\hat{u}})\big)^2 \Big)$$

est un terme stabilisant pour le modèle de l'erreur définie par $\tilde{u} = u - \hat{u}$.

Cet observateur d'état se couple naturellement avec un observateur paramétrique afin d'identifier par exemple la vitesse du front électrique à partir des cartes d'activation en utilisant l'estimation conjointe état-paramètre proposé par *Chapelle*, *Moireau* et *Le Tallec* dans

[13]. Pour l'estimation de paramètres, nous utilisons le filtre de Kalman RoUKF (Reduced-order Unscented Kalman Filtering) [47]. Des résultats sur données synthétiques réalistes montrent tout le potentiel de cette approche et nous rapprochent de la personnalisation de modèles d'électrophysiologie adaptés à des patients. Ce travail sera présenté dans le Chapitre 8.

Organisation de la thèse

Organisation de la thèse

La thèse est décomposée en 4 parties qui contiennent chacune 2 chapitres.

Partie I (Chapitres 1 et 2) - Modélisation cardiaque Le Chapitre 1 donne une brève présentation du rôle du cœur et de son fonctionnement. L'objectif est de présenter – ainsi que d'illustrer – en quelques pages toutes les notions physiologiques qui sont nécessaires pour la lecture de cette thèse.
Le Chapitre 2 présente les modèles disponibles en électrophysiologie. Une analyse très détaillée du modèle bidomaine est donnée en s'appuyant sur la littérature. De plus, nous proposons une justification mathématique du processus d'homogénéisation qui mène au modèle bidomaine. Dans ce deuxième chapitre, nous avons étudié l'influence des phénomènes mécaniques directement dans les lois de conservation du modèle bidomaine à l'aide de la théorie des mélanges alors que la littérature l'étudie uniquement à travers les lois de comportement.

Partie II (Chapitres 3 et 4) - Modèles 3D et surfaciques Le Chapitre 3 prend la forme d'un article [11] publié dans *M3AS* avec D. Chapelle et J.-F. Gerbeau et a pour titre :

A surface-based electrophysiology model relying on asymptotic analysis and motivated by cardiac atria modeling.

Ce chapitre propose – à l'aide d'une analyse asymptotique détaillée – et valide – avec des comparaisons de simulations 2D et 3D – le modèle bidomaine asymptotique (12). En particulier, la démonstration de la Proposition 1 est établie dans ce chapitre.
Le Chapitre 4 est un article [10] publié dans *Journal of Elasticity* avec D. Chapelle qui s'intitule :

Strong convergence results for the asymptotic behavior of the 3D-shell model.

Ce chapitre propose les preuves des Propositions 2 et 3. Numerical simulations of full electrocardiograms

Partie III (Chapitres 5 et 6) - Applications du modèle bidomaine surfacique Dans le Chapitre 5, le modèle surfacique (12) proposé et validé dans le Chapitre 3 est appliqué sur une anatomie réelle des oreillettes. Ce chapitre a fait l'objet d'un court article [17] publié dans les actes de la conférence *Functional Imaging et Modeling of the Heart 2013* (FIMH) et qui s'intitule :

Surface-based electrophysiology modeling and assessment of physiological simulations in atria.

Cet article a été écrit en collaboration avec J.-F. Gerbeau et D. Chapelle ainsi qu'avec deux cardiologues de l'Hôpital Cardiologique du Haut-Lévêque à Bordeaux, M. Haïssaguerre et M. Hocini.
Nous proposons dans le Chapitre 6 une méthode pour la simulation numérique d'électrocardiogrammes complets. Un modèle bidomaine couplé – pour les oreillettes le modèle surfacique (12) et pour les ventricules le modèle bidomaine classique – est introduit. Couplant ce modèle complexe avec deux modèles ioniques adaptés, nous obtenons des simulations très réalistes du cœur complet. Ces simulations sont utilisées afin de produire des électrocardiogrammes contenant toutes les ondes caractéristiques (PQRST). Ce chapitre prend la forme d'un article pre-print dont le titre est

Numerical simulations of full electrocardiogram cycles.

Ce travail est en collaboration avec J.-F. Gerbeau et E. Schenone.

Partie IV (Chapitres 7 et 8) - Problèmes inverses Cette dernière partie concerne l'assimilation de données sur le modèle de réaction-diffusion (13). Le Chapitre 7 est une brève introduction des méthodes d'assimilation de données qui existent. L'objectif n'est pas de proposer une présentation complète et mathématiquement justifiée des outils d'assimilation de données, mais de donner les clés principales ainsi que les références essentielles pour la compréhension du chapitre suivant.
Le Chapitre 8 donne l'origine de l'observateur d'état (15) et le justifie mathématiquement en donnant la preuve de la Proposition 4. Un observateur RoUKF [47] est ajouté sur la dynamique des paramètres et des simulations 1D et 2D d'estimation conjointe sont proposées avec d'excellents résultats. Ce chapitre prend aussi la forme d'un article pre-print intitulé

A joint state and parameter observer for the reaction diffusion model. Application in cardiac electrophysiology

et a été écrit en collaboration avec D. Chapelle et P. Moireau.

Principales contributions

Articles et conférences

Pour conclure cette introduction, un bilan des articles qui illustrent le travail produit dans cette thèse dans leur ordre d'apparition est effectué. Des conférences ont aussi été l'occasion de présenter ces travaux.

Articles publiés et pre-print

- D. Chapelle, A. Collin, et J.-F. Gerbeau. A surface-based electrophysiology model relying on asymptotic analysis and motivated by cardiac atria modeling. *M3AS*, 23(14) :2749-2776, 2013.

- D. Chapelle et A. Collin. Strong convergence results for the asymptotic behavior of the 3D-shell model. *Journal of Elasticity*, 115(2) :173-192, 2014.
- A. Collin, J.-F. Gerbeau, M. Hocini, M. Haïssaguerre, et D. Chapelle. Surface-based electrophysiology modeling and assessment of physiological simulations in atria. *FIMH 2013*, 7945 :352-359, 2013.
- A. Collin, J.-F. Gerbeau, et E. Schenone. Numerical simulations of full electrocardiogram cycles. Pre-print.
- D. Chapelle, A. Collin, et P. Moireau. A joint state and parameter observer for the reaction diffusion model. Application in cardiac electrophysiology. Pre-print.

Conférences

- 41ème Congrès National d'Analyse Numérique – CANUM12. Session poster : «A surface-based electrophysiology model motivated by cardiac atria modeling and relying on asymptotic considerations» - Prix du meilleur poster.
- V International Symposium on modeling of Physiological Flows - MPF2013. Session orale : «Atrial electrophysiology study using a surface-based model relying on asymptotic analysis».
- 7th International Conference on Functional Imaging et Modeling of the Heart - FIMH2013. Session poster : «Surface-based electrophysiology modeling and assessment of physiological simulations in atria».
- 42ème Congrès National d'Analyse Numérique – CANUM14. Session orale : «Inverse problems in cardiac electrophysiology using front observers».
- 4th International Conference on Engineering Frontiers in Pediatric and Congenital Heart Disease 2014. Session poster (en collaboration avec Elisa Schenone) : «Numerical simulations of full electrocardiogram cycles».

Développement Logiciel Pour les simulations des Chapitres 2, 3, 5, 6, de nombreuses contributions (environ 10000 lignes) ont été développées dans la librairie C++ d'éléments finis FELiScE[3] développée chez Inria par les équipes MΞDISIM et REO. Le modèle bidomaine classique couplé avec plusieurs modèles ioniques, a été implémenté (en collaboration avec Elisa Schenone, doctorante de l'équipe REO). Une partie de la version mobile du modèle a été ajoutée. Les outils nécessaires aux éléments finis curvilignes ont été développés et ont permis l'implémentation du modèle bidomaine surfacique. Le modèle de Courtemanche-Ramirez-Nattel a été ajouté afin de proposer les simulations réalistes d'oreillettes. Finalement, l'ajout de l'observateur d'état permettant de corriger les conditions initiales a complété ce travail.

[3] felisce.gforge.inria.fr

Bibliographie

[1] G. Allaire. Homogenization and two-scale convergence. *SIAM Journal of Mathematical Analysis*, 23(6) :1482–1518, 1992.

[2] H. Ammari, J. Garnier, L. Giovangigli, W. Jing, and J.-K. Seo. Spectroscopic imaging of a dilute cell suspension. Submitted.

[3] R.M. Anderson. Mathematical models of the potential demographic impact of AIDS in Africa. *AIDS*, 5(Suppl 1) :37–4, 1991.

[4] M.I. Asensio and L. Ferragut. On a wildland fire model with radiation. *International Journal for Numerical Methods in Engineering*, 54(1) :137–157, 2002.

[5] A. Bensoussan. *Filtrage optimal des systèmes linéaires*. Dunod, 1971.

[6] A. Bensoussan, J.-L. Lions, and G. Papanicolaou. *Asymptotic Analysis for Periodic Structures*. Elsevier, 1978.

[7] M. Boulakia, S. Cazeau, M.A. Fernández, J.-F. Gerbeau, and N. Zemzemi. Mathematical modeling of electrocardiograms : a numerical study. *Annals of Biomedical Engineering*, 38(3) :1071–1097, 2010.

[8] A. Bueno-Orovio, E. M. Cherry, and F. H. Fenton. Minimal model for human ventricular action potentials in tissue. *Journal of Theretical Biology*, 253 :544–560, 2008.

[9] D. Chapelle and K.J. Bathe. The mathematical shell model underlying general shell elements. *Internat. J. Numer. Methods Engrg.*, 48(2) :289–313, 2000.

[10] D. Chapelle and A. Collin. Strong convergence results for the asymptotic behavior of the 3D-shell model. *Journal of Elasticity*, 115(2) :173–192, 2014.

[11] D. Chapelle, A. Collin, and J.-F. Gerbeau. A surface-based electrophysiology model relying on asymptotic analysis and motivated by cardiac atria modeling. *M3AS*, 23(14) :2749–2776, 2013.

[12] D. Chapelle, A. Ferent, and K.J. Bathe. 3D-shell elements and their underlying mathematical model. *M3AS*, 14(1) :105–142, 2004.

[13] D. Chapelle, P. Moireau, and P. Le Tallec. Robust filtering for joint state-parameter estimation in distributed mechanical systems. *Discrete and Continuous Dynamical Systems*, (1–2) :65–84, 2009.

[14] P. Colli Franzone, L. Guerri, C. Viganotti, E. Macchi, S. Baruffi, S. Spaggiari, and B. Taccardi. Potential fields generated by oblique dipole layers modeling excitation wavefronts in the anisotropic myocardium. Comparison with potential fields elicited by paced dog hearts in a volume conductor. *Circulation Research*, 51(3) :330–346, 1982.

[15] P. Colli Franzone and E. Magenes. On the inverse potential problem of electrocardiology. *Calcolo*, 16(4) :459–538, 1979.

[16] P. Colli Franzone and G. Savaré. Degenerate evolution systems modeling the cardiac electric field at micro and macroscopic level. *Progress in Nonlinear Differential Equations and Their Applications*, 50 :49–78, 2002.

[17] A. Collin, J.-F. Gerbeau, M. Hocini, M. Haïssaguerre, and D. Chapelle. Surface-based electrophysiology modeling and assessment of physiological simulations in atria. *FIMH 2013*, 7945 :352–359, 2013.

[18] M. Courtemanche, R.J. Ramirez, and S. Nattel. Ionic mechanisms underlying human atrial action potential properties : insights from a mathematical model. *American Journal of Physiology*, (275) :H301–H321, 1998.

[19] M.C. Delfour and J.-P. Zolésio. *Shapes and Geometries : Analysis, Differential Calculus, and Optimization (Advances in Design and Control)*. SIAM, second edition, 2011.

[20] D. Deng, Y. Gong, G. Shou, and P. Jiao. Simulation of biatrial conduction via different pathways during sinus rhythm with a detailed human atrial model. *Journal of Zheijiang University-SCIENCE B (Biomedicine and Biotechnology)*, pages 1862–1783, 2012.

[21] G. Evensen. *Data Assimilation : The Ensemble Kalman Filter*. Springer, 2009.

[22] M.A. Fernández, J.-F. Gerbeau, and C. Grandmont. A projection semi-implicit scheme for the coupling of an elastic structure with an incompressible fluid. *International Journal for Numerical Methods in Engineering*, 69(4) :794–821, 2007.

[23] S.C. Ferreira, Jr. M. L. Martins, and M. J. Vilela. Reaction-diffusion model for the growth of avascular tumor. *Physical Review E*, 65 :021907 (8pp), 2002.

[24] P. Frey. Yams : A fully automatic adaptive isotropic surface remeshing procedure. Technical report 0252, Inria, Rocquencourt, France, Nov. 2001.

[25] A.S. Go, D. Mozaffarian, V. L. Roger, E. J. Benjamin, J. D Berry, M. J. Blaha, S. Dai, E. S. Ford, C. S. Fox, S. Franco, et al. Heart disease and stroke statistics—2014 update a report from the American Heart Association. *Circulation*, 129(3) :e28–e292, 2014.

[26] D.M. Harrild and S.H. Craig. A computer model of normal conduction in the human atria. *Circulation Research*, (87) :e25–e36, 2000.

[27] M. Hintermüller and W. Ring. An inexact Newton-C-G-type active contour approach for the minimization of the Mumford-Shah functional. *Journal of Mathematical Imaging and Vision*, 20 :19–42, 2004.

[28] S.Y. Ho, R.H. Anderson, and D. Sánchez-Quintana. Atrial structure and fibres : morphologic bases of atrial conduction. *Cardiovascular Research*, (54) :325–336, 2002.

[29] S.Y. Ho and D. Sanchez-Quintana. The Importance of Atrial Structure and Fibers. *Clinical Anatomy*, 22(1) :52–63, 2009.

[30] A.L. Hodgkin and A. F. Huxley. Currents carried by sodium and potassium ions through the membrane of the giant axon of loligo. *Journal of Physiology*, 116(4) :449–472, 1952.

[31] A.L. Hodgkin and A. F. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *Journal of Physiology*, 117(4) :500–544, 1952.

[32] P.J. Hunter. *Finite element analysis of cardiac muscle mechanics*. PhD thesis, University of Oxford, 1975.

[33] V. Jacquemet. *A biophysical model of atrial fibrillation and electrograms : formulation, validation and applications*. PhD thesis, EPFL, 2004.

[34] S.J. Julier and J.K. Uhlmann. A new extension of the Kalman filter to nonlinear systems. In *Proc. of AeroSense : The 11th Int. Symp. on Aerospace/Defence Sensing, Simulation and Controls*, 1997.

[35] R. Kalman and R. Bucy. New results in linear filtering and prediction theory. *Trans. ASME J. Basic. Eng.*, 83 :95—108, 1961.

[36] J.P. Keener. An eikonal-curvature equation for action potential propagation in myocardium. *Journal of Mathematical Biology*, 29 :629–651, 1991.

[37] R.H. Keldermann, M.P. Nash, and A.V. Panfilov. Modeling cardiac mechano-electrical feedback using reaction-diffusion-mechanics systems. *Physica D*, 238 :1000–1007, 2008.

[38] M. Krueger, V. Schmidt, C. Tobón, F. Weber, C. Lorenz, D. Keller, H. Barschdorf, M. Burdumy, P. Neher, G. Plank, K. Rhode, G. Seemann, D. Sanchez-Quintana, J. Saiz, R. Razavi, and O. Dössel. Modeling atrial fiber orientation in patient-specific geometries : a semi-automatic rule-based approach. *Functional Imaging and Modeling of the Heart*, pages 223–232, 2011.

[39] S. Labarthe, E. Vigmond, Y. Coudière, J. Henry, H. Cochet, and P. Jais. A computational bilayer surface model of human atria. *FIMH 2013 - 7th International Conference on Fonctional Imaging and Modeling of the Heart*, pages 1530s–1533, 2013.

[40] C. Liu, J.M. Zachara, A. Felmy, and Y. Gorby. An electrodynamics-based model for ion diffusion in microbial polysaccharides. *Colloids and Surfaces B : Biointerfaces*, 38 :55–65, 2004.

[41] L. Liu. Cardiovascular diseases in China. *Biochemistry and cell biology*, 85(2) :157–163, 2007.

[42] D.G. Luenberger. *Determining the State of a Linear with Observers of Low Dynamic Order*. PhD thesis, Stanford University, 1963.

[43] D.G. Luenberger. An introduction to observers. *IEEE Transactions on Automatic Control*, 16 :596–602, 1971.

[44] C.H. Luo and Y. Rudy. A dynamic model of the cardiac ventricular action potential. I. simulations of ionic currents and concentration changes. *Circulation Research*, 74 :1071–1096, 1994.

[45] J. Mandel, L.S. Bennethum, J.D. Beezley, J.L. Coen, C.C Douglas, L.P. Franca, M. Kim, and A. Vodacek. A wildland fire model with data assimilation. *Mathematics and Computers in Simulation*, 79(3) :584–606, 2008.

[46] C.C. Mitchell and D.G. Schaeffer. A two-current model for the dynamics of cardiac membrane. *Bulletin Math. Bio.*, 65 :767–793, 2003.

[47] P. Moireau and D. Chapelle. Reduced-order Unscented Kalman Filtering with application to parameter identification in large-dimensional systems. *ESAIM : Control, Optimisation and Calculus of Variations*, 17(2) :380–405, 2011.

[48] Y. Mori. From three-dimensional electrophysiology to the cable model : an asymptotic study. *arXiv preprint arXiv :0901.3914*, 2009.

[49] Y. Mori, J.W. Jerome, and C.S. Peskin. A three-dimensional model of cellular electrical activity. *Bulletin of the Institute of Mathematics, Academia Sinica*, 2(2) :367–390, 2007.

[50] M.P. Nash and A.V. Panfilov. Electromechanical model of excitable tissue to study reentrant cardiac arrhythmias. *Progress in Biophysics and Molecular Biology*, 85(2-3) :501–522, 2004.

[51] J.C. Neu and W. Krassowska. Homogenization of syncytial tissues. *Critical Reviews in Biomedical Engineering*, 21(2) :137–199, 1993.

[52] M. Nichols, N. Townsend, R. Luengo-Fernandez, J. Leal, A. Gray, P. Scarborough, and M. Rayner. European cardiovascular disease statistics. *European Heart Neatwork, Brussels, European Society of Cardiology, Sophia Antipolis*, 2012.

[53] D. Noble. A modification of the Hodgkin-Huxley equation applicable to Purkinje fiber action and pacemaker potentials. *Journal of Physiology*, 160 :317–352, 1962.

[54] S. Osher and R. Fedkiw. *Level Set Methods and Dynamic Implicit Surfaces*. Applied Mathematical Sciences, 2002.

[55] A.V. Panfilov and A.T. Winfree. Dynamical simulations of twisted scroll rings in three-dimensional excitable media. *Physica D : Nonlinear Phenomena*, 17(3) :323–330, 1985.

[56] C. S. Peskin. Numerical analysis of blood flow in the heart. *Journal of Computational Physics*, 25(3) :220–252, 1977.

[57] C.S. Peskin. The fluid dynamics of heart valves : experimental, theoretical, and computational methods. *Annual Review of Fluid Mechanics*, 14 :235–259, 1982.

[58] A.J. Pullan, M.L. Buist, and L.K. Cheng. *Mathematically Modeling the Electrical Activity of the Heart*. World Scientific, 2005.

[59] V. Quaranta, A.M. Weaver, P.T. Cummings, and A.R. Anderson. Mathematical modeling of cancer : the future of prognosis and treatment. *Clin Chim Acta*, 357(2) :137–139, 2005.

[60] Berguer R., J.L. Bull, and Khanafer K. Refinements in mathematical models to predict aneurysm growth and rupture. *Ann N Y Acad Sci.*, 1085 :110–116, 2006.

[61] C. Ramanathan, R.J. Ghanem, P. Jia, K. Ryu, and Y. Rudy. Noninvasive electrocardiographic imaging for cardiac electrophysiology and arrhythmia. *Nature Medicine*, 10 :422–428, 2004.

[62] G. Richardson. A multiscale approach to modelling electrochemical processes occurring across the cell membrane with application to transmission of action potentials. *Mathematical Medicine and Biology*, 26(3) :201–224, 2009.

[63] C.M. Rochoux, B. Cuenot, S. Ricci, A. Trouvé, B. Delmotte, Se. Massart, R Paoli, and R. Paugam. Data assimilation applied to combustion. *Comptes Rendus Mécanique*, 341 :266–276, 2013.

[64] J. Sainte-Marie, D. Chapelle, R. Cimrman, and M. Sorine. Modeling and estimation of the cardiac electromechanical activity. *Computers and Structures*, 84 :1743–1759, 2006.

[65] L. Tung. *A bi-domain model for describing ischemic myocardial d-c potentials.* PhD thesis, Massachusetts Institute of Technology. Dept. of Electrical Engineering and Computer Science., 1978.

[66] E. Vigmond, S. Labarthe, H. Cochet, Y Coudière, J. Henry, and P. Jais. A bilayer representation of the human atria. *EMBC - 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 1530s–1533, 2013.

[67] F. Wartak. *Electrocardiogram interpretation*. Medical Education Systems, 1975.

[68] H. Watanabe, S. Sugiura, H. Kafuku, and T. Hisada. Multiphysics simulation of left ventricular filling dynamics using fluid–structure interaction finite element method. *Biophys. J.*, 87(3) :2074–2085, 2004.

[69] H.K. Zhao, T. Chan, B. Merriman, and S. Osher. A variational level set approach to multiphase motion. *Journal of Computational Physics*, 127(1) :179–195, 1996.

Part I

Cardiac Modeling

Introduction of Part I

The first part of this thesis is decomposed into two chapters. Chapter 1 *A brief overview of the role of the heart and its behavior* gives a presentation of the human heart. This knowledge is necessary in order to understand the complexity of the heart and to determine the challenges of cardiac modeling, and in particular in the cardiac electrophysiological framework. This means that the first chapter of this part describes the impact of the electrical activity in the behavior of the heart and justifies our interest for the improvement of the modeling of cardiac electrophysiology.

Chapter 2 *Mathematical models in cardiac electrophysiology* gives a detailed presentation of a very classical electrophysiological model. The electric wave propagating in the cardiac tissue can be represented by a nonlinear reaction-diffusion partial differential equation (PDE), coupled with an ordinary differential equation (ODE) representing cellular activity. We consider the bidomain model. This model represents the diffusion in the intra- and the extra-cellular domains hence the name. We start with the model of cellular activity in the intra- and the extra-cellular domains and we present **a mathematical-justified homogenization**, an asymptotic process which leads to the equations of the bidomain model. We use this model in the rest of this thesis. In the second part of this chapter, we study the **impact of the mechanical deformation** of the heart on the electrical activity. We use for this part a mixture theory. We establish the **conservation laws**: the mass conservation and the principle of virtual work for the solid and the charge conservation for the ionic species and we also give the **constitutive laws**. The last section of the second chapter presents simulations of the bidomain model using finite elements. Realistic simulations of both ventricles are given with a fixed domain and also simulations in moving domain in simpler cases are given.

CHAPTER 1
A brief overview of the role of the heart and its behavior

Abstract The heart contracts regularly and the continuity of its beating is essential to life. Indeed, an arrest of the heartbeat is one of the most obvious signs of death. These pulses allow to irrigate organs with oxygenated blood and cannot stop even for a short period. In fact, some organs cannot survive a brief cessation of heartbeats. This is the case of the brain which is extremely sensitive to any circulatory abnormality. A deep knowledge of the behavior of the heart is necessary in order to formulate adapted models and obtain realistic simulations. These tools can then be provided to doctors for example to predict diseases or tailoring treatments. We present in this chapter a brief overview of the role of the heart and its behavior. The heart is a muscle that pumps blood by repeated, rhythmic contractions. A representation of the heart and of the circulatory system is provided in the first section of this chapter. More specifically, at each beating of the heart, an electrical wave crosses over the heart and triggers the muscle contraction. The second section provides a presentation of this coupled action – electrical and mechanical – which allows the blood circulation during a cardiac cycle. Cardiac electrophysiology is the study of the electrical wave which precedes the cardiac contraction. In this thesis we are interested in cardiac electrophysiology (models, numerical simulations, data assimilation ...). Hence, the last part of this chapter is dedicated to the electrical function of the heart, in particular we give a presentation of the electrocardiograms.

1.1 The human heart

1.1.1 The heart anatomy

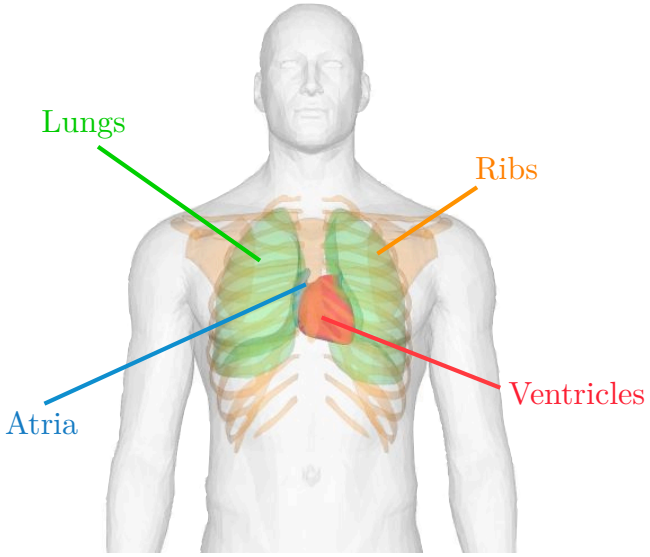

Figure 1.1: Location and orientation of the heart in the thorax

The human heart is an organ which functions as a pump. This pump provides a continuous circulation of the blood through the complete body. The heart is located in the thoracic cavity between the two lungs. In Figure 1.1, the location and the orientation of the heart in the thoracic cavity is shown. Its size is about that of a fist, and its weight is about $250 - 300$g. A protective sac named the *pericardium* surrounds the heart. The wall of the heart is decomposed into three parts: the *endocardium* (in the interior of the muscle), the *myocardium* (in the middle) and the *epicardium* (in the exterior).

Figure 1.2 gives a detailed representation of the heart. It is composed of four main chambers, the two superior left and right *atria* and the two inferior left and right *ventricles*. The inferior part of the ventricles is named the *apex*. The atria are very thin compared to the ventricles. The *septum* separates the left ventricle from the right ventricle. The orientation of the heart in the body can vary significantly from one human to another and its major axis is often defined as the smallest principal inertia axis of the left ventricle.

1.1.2 The circulatory system

Figure 1.3 represents the circulatory system. At each cardiac cycle, the deoxygenated blood arrives in the right atrium by the *vena cava* and the right atrium ejects the blood into the right ventricle after the opening of the *tricuspid valve*. The right ventricle pumps out the blood by the *pulmonary artery* towards the lungs after the opening of the *pulmonary valve*. The lungs oxygenate the blood and this oxygenated blood returns to the heart by the *pulmonary vein* and arrives in the left atrium. The left atrium ejects the blood into the

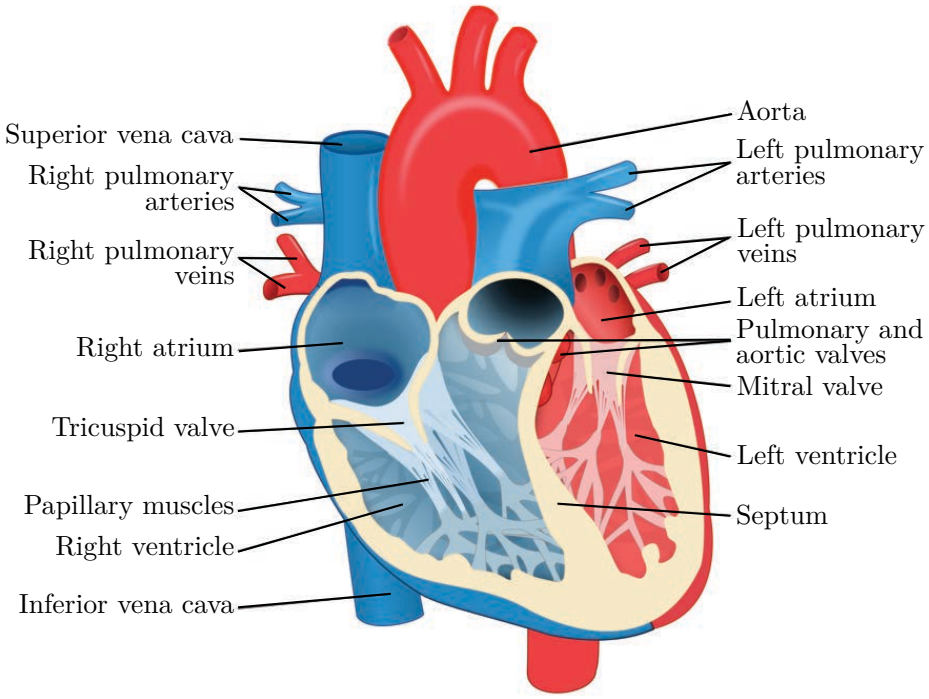

Figure 1.2: Representation of the heart anatomy (adapted from wikipedia) - In blue, the right atrium and the right ventricle and in red, the left atrium and the left ventricle

left ventricle after the opening of the *mitral valve*. Then the left ventricle can pump out the blood towards the whole body by the *aorta* after the opening of the *aortic valve*. All these new notions are also represented in Figure 1.2. As we can see in Figure 1.3, the deoxygenated blood is classically represented in blue and the oxygenated blood in red. The same colors are classically used for the representation of the left (which contains the oxygenated blood) and the right (which contains the deoxygenated blood) parts of the heart. For more information or more illustrations, see [1].

1.1.3 The muscle fibers

Cardiac muscle has a fiber architecture and we will see in what follows that the fibers orientation is very important in cardiac electrophysiology. Many anatomical studies have been devoted to describing the ventricular geometric organization of the fibers. It has been shown, for example in [6], that three different myocardial layers – superficial (subepicardial), middle, and deep (subendocardial) – have been distinguished. The cardiac muscle fibers of the ventricles are oriented in a helicoidal shape in these three layers. The position of the atrial fibers is more complicated to determine because there is a high heterogeneity from one human to another. Still, main directions have been detected, see [2, 3] for more details.

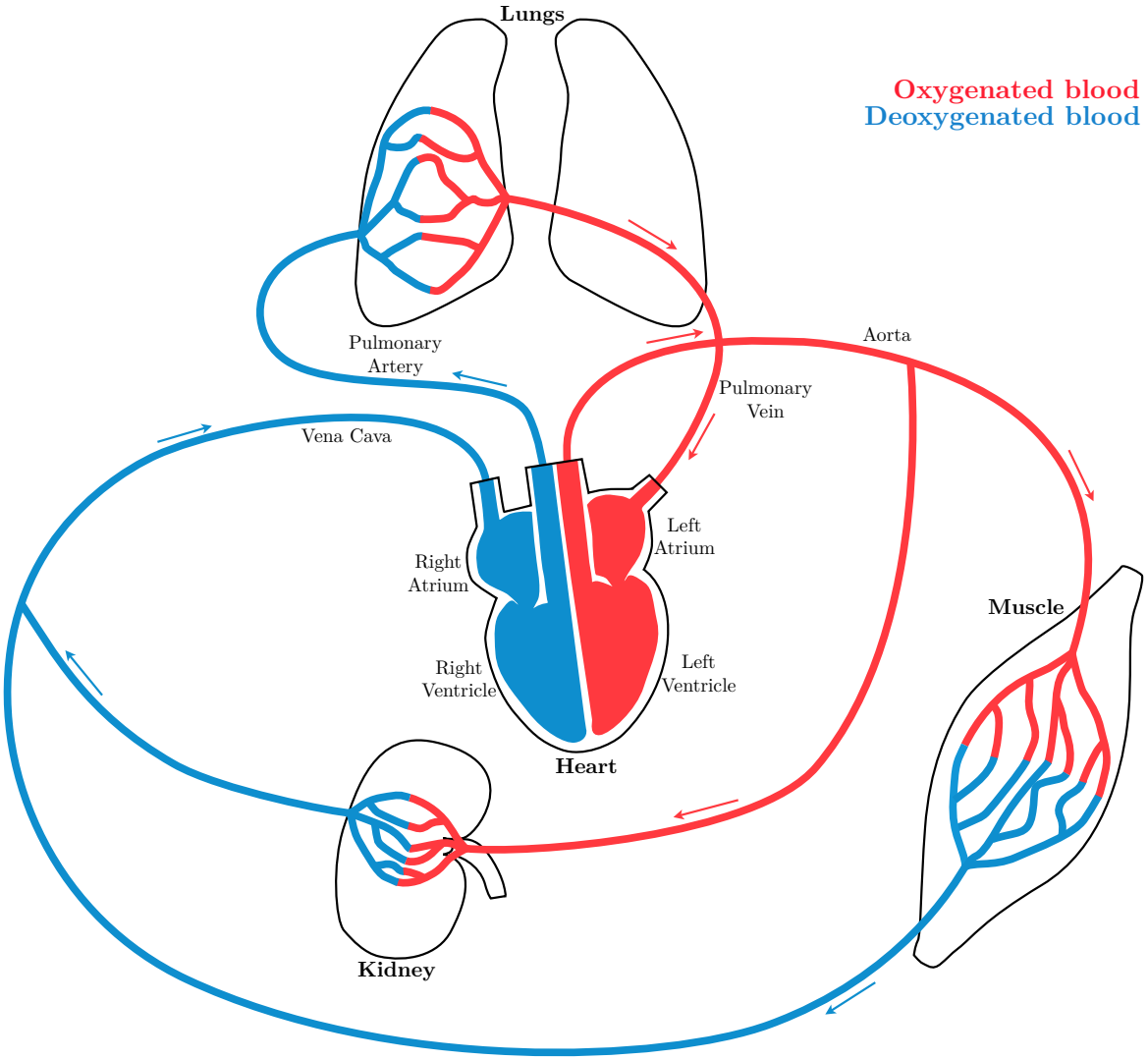

Figure 1.3: Circulatory System

1.2 The heart behavior - Presentation of a cardiac cycle

In order to pump the blood, the atria and the ventricles contract at each cardiac cycle. An electrical signal is at the origin of the contraction. Indeed, at each cardiac beat, an electrical signal crosses the heart and depolarizes the cardiac cells thus triggering their contraction.

At the cellular scale, we are interested in the action potential, see [5] for more details. It corresponds to the evolution of the transmembrane potential, the difference between the extra- and the intra-cellular potentials, in the cell over time. As can be seen in Figure 1.4, at the beginning, the cell is at the *resting potential*. Due to a stimulus, the cell becomes depolarized very quickly. The *depolarization* is at the origin of the mechanical activity of the cell. After a *plateau phase* in which the cell is depolarized and then contracted, the *repolarization* of the cell begins. During this phase, the cell is in *refractory phase*, which

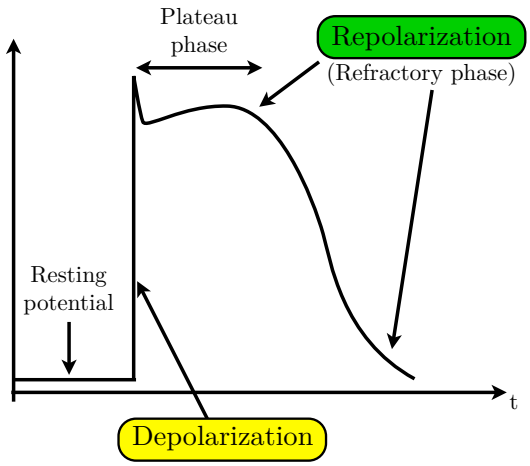

Figure 1.4: Action potential corresponding to the different states of a cell

means that a new stimulus cannot trigger a new depolarization.

At the heart scale, we are interested in the propagation of the electrical signal in the heart. Figure 1.5 represents the sucessive steps of the cardiac cycle. At the top of the figure, we have a simplified description of the heart. For each step of the cardiac cycle, the electrical phenomenon is described at the left and the right represents the consequence of the electrical activation on the mechanical activity of the heart and how this implies the blood circulation.

1. The cardiac cycle starts with the activation of the natural pacemaker of the heart, the *sinoatrial node* (also named *sinus node*) represented by a yellow star. A wave of excitation spreads out through the atria along specialized conduction channels triggering the atrial contraction. Due to the contraction of the atria, the blood flows through the tricuspid/mitral valve to the right/left ventricle.

2. This wave also activates the *atrio-ventricular (AV) node* represented by a yellow star which is the only electrical connection between the atria and the ventricles. Otherwise, the atria and the ventricles are electrically isolated. The atrioventricular node delays impulses. This delay in the cardiac pulse is very important because it ensures that the atria have ejected the blood into the ventricles first before the ventricles contract. The signal follows the *Purkinje fibers* – a fast conduction channel – also represented in yellow. When the atria have ejected all the blood, the tricuspid and mitral valves close.

3. The wave spreads throughout the ventricles from the endocardium to the epicardium triggering the ventricular contraction. It implies the opening of the aortic and pulmonary valves due to the rise in ventricular pressure and the blood flows to the aorta and pulmonary artery respectively. During this time, the repolarization of the atria takes place and it allows atrial relaxation.

4. Close to the end of the ventricular contraction, the ventricular repolarization begins. As we can see, the direction of the repolarization is inverted compared to the depo-

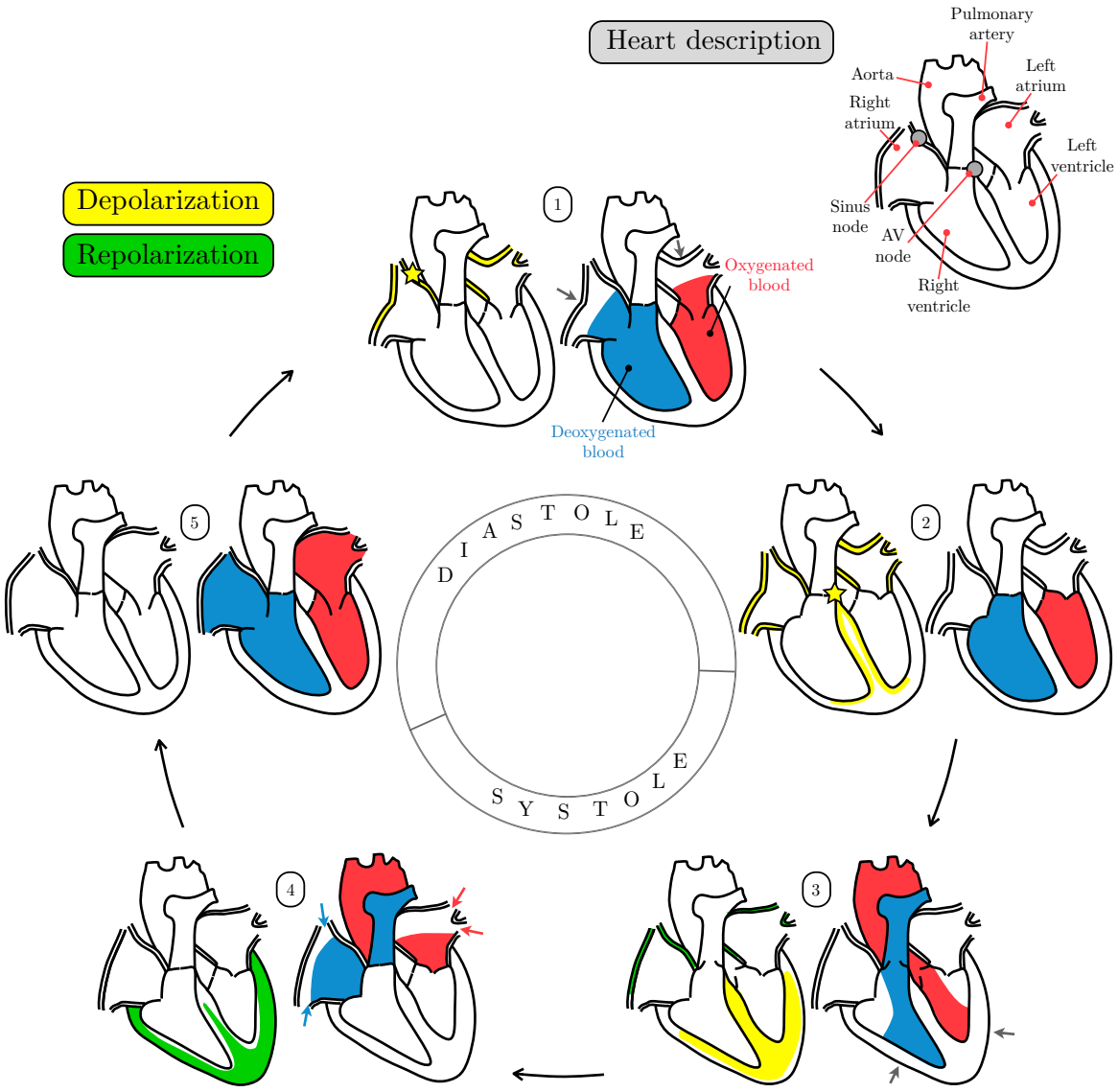

Figure 1.5: Cardiac cycle with deoxygenated blood in blue and oxygenated blood in red

larization and starts from the epicardium to the endocardium. The blood comes in the right and the left atria by the vena cava and the pulmonary veins respectively.

5. The whole heart is at the resting potential. Ventricular relaxation induces a fast decrease in pressure that triggers the opening of the tricuspid and mitral valves. Then the blood enters in the ventricles.

In Figure 1.6 the left ventricular and atrial volumes are represented during a cardiac cycle. We include the five steps presented in Figure 1.5 and we represent the different phases of the cardiac cycle.

The importance of the orientation of the muscle fibers for the electrical activity of the heart is due to the fact that the electrical conductivity is higher along the fiber direction than

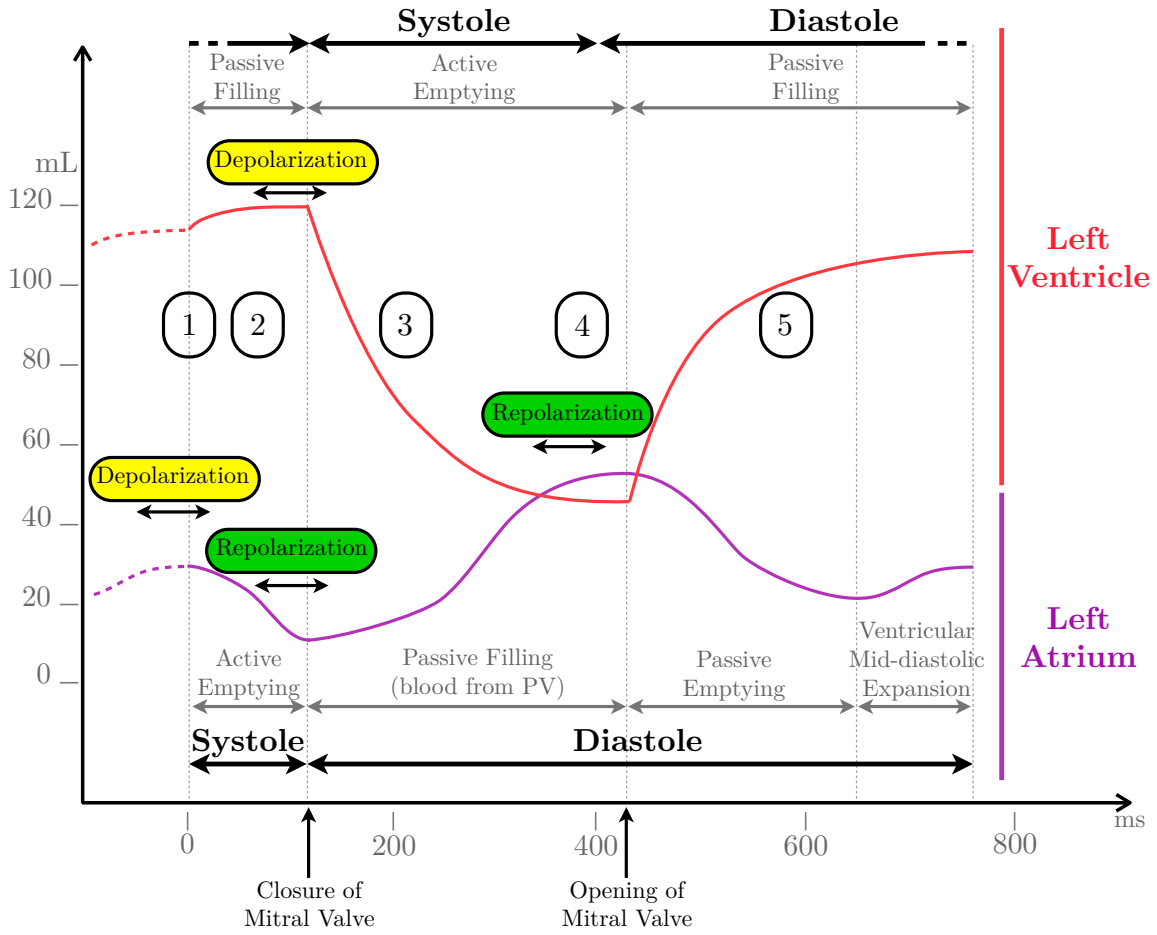

Figure 1.6: Left atrial (purple) and ventricular (red) volumes during a cardiac cycle

across. As we will see in the following chapter, the model that we use is able to take into account the fibers variation.

1.3 Electrophysiology and Electrocardiograms

The electrocardiogram (ECG) is an external measurement of the electrical activity of the heart, see [4]. The electrocardiograph – which produces the ECGs – measures the difference of potential between different positions of the body surface. This allows to obtain a non-invasive image of the electrophysiological state of the patient. Figure 1.7 represents the first derivation of the ECG (a classical ECG has twelve derivations) in a healthy case. We can see in the figure:

- the P wave which represents the atrial depolarization,
- the QRS complex which represents the ventricular depolarization,
- the QT segment which corresponds to the plateau of ventricular action potential,

- the T wave which represents the ventricular repolarization.

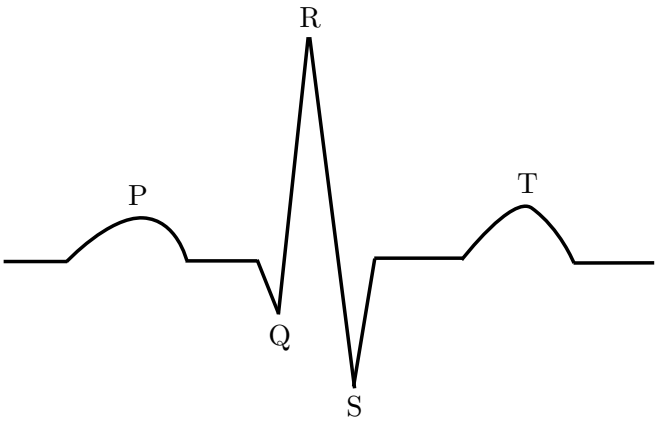

Figure 1.7: First derivation of an ECG during a cardiac cycle

Figure 1.8 gives a rapid physical explanation of the ECG behavior. The first derivation is the potential difference between the left and the right arms. It means that if the projection of the propagation vector on the axis left-right is positive, the wave is positive, and inversely. As previously explained, at each cardiac beat, the sinoatrial node situated in the top of the right atrium, is the natural pacemaker of the heart and triggers an excitation wave every 750ms. This wave propagates in the whole atria. As the sinus node is in the right atrium and as the heart axis is slightly oriented on the right, the P wave is positive for the first derivation. The only connection between the atria and the ventricles is the atrio-ventricular node which delays the signal and this implies a non zero PQ segment. Following the rapid conduction bundle called the Purkinje fibers, most of the ventricular endocardium is depolarized at the same time. The propagation goes from the left ventricular endocardium to the myocardium and then the Q wave is negative. Indeed, due to the smaller thickness of the right ventricle, the right part of the epicardium is yet depolarized. Rapidly the signal propagates in the ventricles from the top to the bottom thus implying this positivity of the R wave. As can be seen in Figure 1.8, the Q wave corresponds to the end of the ventricular depolarization. During the ventricular depolarization, the atrial repolarization occurs. There is no wave corresponding to the atrial repolarization because this wave is hidden by the ventricular depolarization. In few pathological cases, the atrial repolarization is visible due to a desynchronization between the atria and the ventricles and we know that the atrial repolarization is negative, see [7]. It means that the repolarization is in the same direction of depolarization as can be seen in the figure. However, the T wave is positive like the R wave. It means that the repolarization is in the opposite direction of the depolarization. Indeed, the plateau phase varies across the thickness of ventricular cardiac tissue as explained in [8].

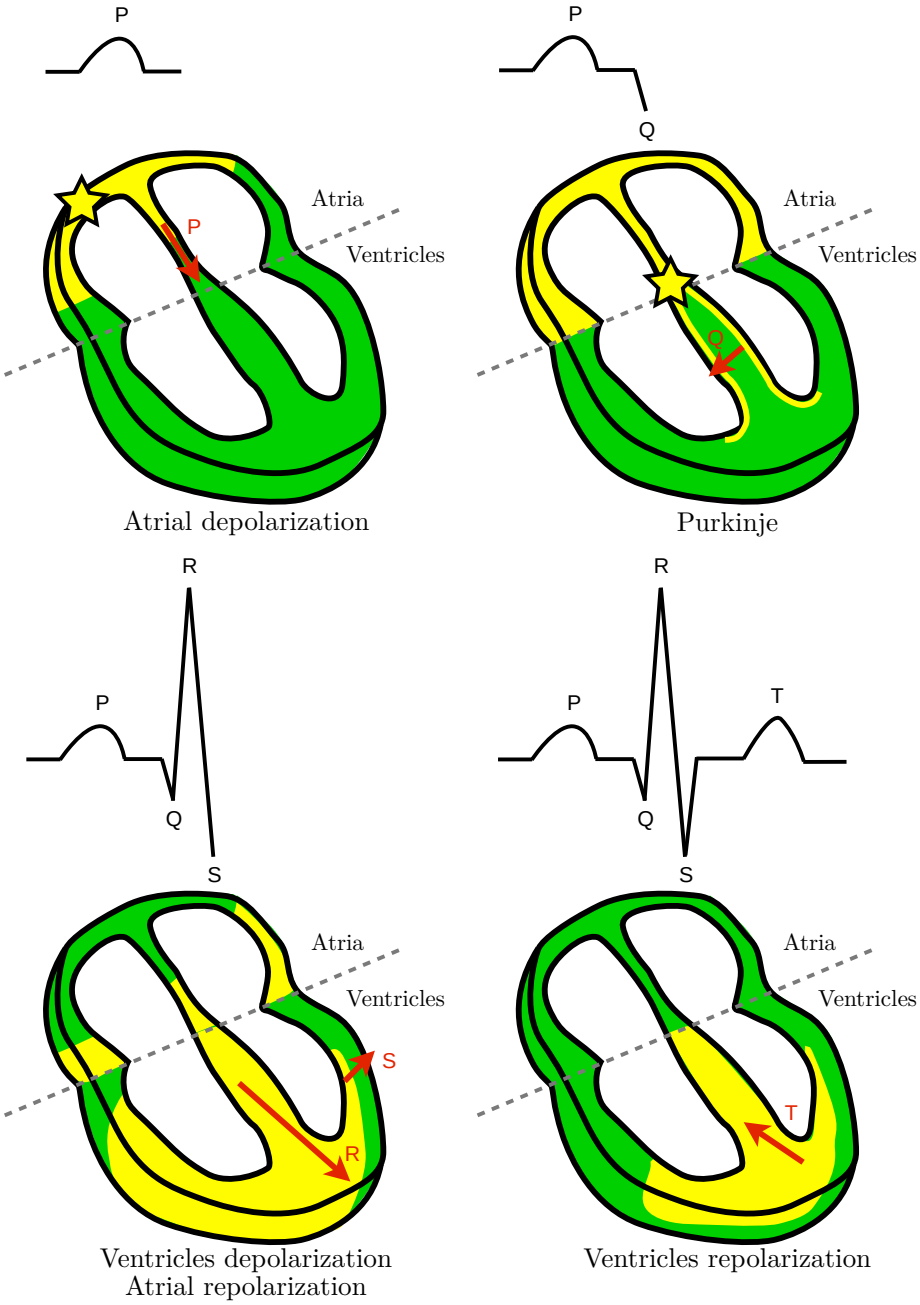

Figure 1.8: Cardiac electrical cycle

Bibliography

[1] Netter F. H. *The Heart - The Netter collection of medical illustrations*, volume 5. Saunders Elsevier, 1969.

[2] S.Y. Ho, R.H. Anderson, and D. Sánchez-Quintana. Atrial structure and fibres: morphologic bases of atrial conduction. *Cardiovascular Research*, (54):325–336, 2002.

[3] M. Krueger, V. Schmidt, C. Tobón, F. Weber, C. Lorenz, D. Keller, H. Barschdorf, M. Burdumy, P. Neher, G. Plank, K. Rhode, G. Seemann, D. Sanchez-Quintana, J. Saiz, R. Razavi, and O. Dössel. Modeling atrial fiber orientation in patient-specific geometries: a semi-automatic rule-based approach. *Functional Imaging and Modeling of the Heart*, pages 223–232, 2011.

[4] J. Malmivuo and R. Plonsey. *Bioelectromagnetism - Principles and Applications of Bioelectric and Biomagnetic Fields*. Oxford University Press, 1995.

[5] F.B. Sachse. *Computational Cardiology: Modeling of Anatomy, Electrophysiology and Mechanics*. Springer-Verlag, 2004.

[6] D.D. Streeter. Gross morphology and fiber geometry of the heart. *Handbook Physiology, The cardiovascular system*, 1:61–112, 1979.

[7] J. Tranchesi, V. Adelardi, and J. Martins de Oliveira. Atrial repolarization—its importance in clinical electrocardiography. *Circulation*, 22:635–644, 1960.

[8] G.-X. Yan and C. Antzelevitch. Cellular basis for the normal T wave and the electrocardiographic manifestations of the long-QT syndrome. *Circulation*, 98:1928–1936, 1998.

CHAPTER 2
Mathematical models in cardiac electrophysiology

Abstract As explained in the previous chapter, cardiac electrophysiology is the study of the electrical wave which triggers each contraction of the heart. In this chapter, we are interested in the modeling of cardiac electrophysiology. Starting with a study of the motion of electric charges at the cellular scale and proceeding with an homogenization process, we present a derivation of the bidomain model which is a very classical electrophysiological model. We also present a study of the impact of the mechanical deformation of the heart on the electrical wave using the mixture theory. We finish this chapter by a numerical part which presents the space and time discretization used in this thesis and gives some simulations, in particular a simulation of a realistic ventricular case.

Introduction

Cardiac electrophysiology studies the electrical activity of the heart. Knowledge concerning cardiac electrophysiology is necessary for the understanding of the complete behavior of the organ. At each beat, an electrical wave crosses the heart and triggers the muscle cells contraction which is at the origin of the heart contraction. The study of this electrical current requires a thorough understanding of the chemical activity triggering the motion of electric charges.

There are two modeling scales in cardiac electrophysiology. The microscopic scale aims at producing a detailed description of the origin of the electric wave in the cells. The macroscopic scale describes the propagation of the electrical wave in the heart. One of the most popular mathematical model in cardiac electrophysiology is the bidomain model, see for example [41, 44, 36, 7]. Introduced in the late 70's [45], the bidomain model is a macroscopic model coupled with one (or more) ordinary differential equation(s) representing the cellular activity. The bidomain model is based upon the assumption that the cardiac muscle is segmented into the intra- and extracellular domains – hence its name. Section 2.1 gives a detailed description of the bidomain model using the literature. Following the work of [24, 28], the microscopic model of cellular activity is presented in Section 2.1.2 and allows to obtain the microscopic bidomain model (2.12). A homogenization procedure – see for more details [32, 35, 39] – is used to link the microscopic and macroscopic behaviors and leads to the equations of the bidomain model presented in (2.34). Section 2.1.4 gives a brief presentation of the principal models concerning the cellular activity. The literature is abundant on this subject, see for example [41, 44, 26].

In what follows, we are interested in the effects of the mechanical deformation on the heart electro-chemical activity. As previously explained, every beat is triggered by an electrical wave traveling across the heart, inducing heart muscle contractions. The influence of the electrical function of the heart in the cardiac mechanical mechanism has been well investigated, see for example [42]. It has also been shown in experiments that the stretch produced by the heart deformation leads to significant electro-physiological changes. The influence of deformations on the electrical wave propagation in the cardiac tissue – the so-called mechano-electrical feedback [31, 22] – is becoming an accepted concept. However, the models given in the literature mainly consider the alteration of the constitutive laws of the bidomain model caused by the deformation. The objective in Section 2.2 is to suggest a mechano-electrical model able to take into account the influence of the heart deformation in the conservation laws. Using the detailed analysis of the equations of the bidomain model presented in Section 2.1, we rewrite the conservation laws in moving domain in Section 2.2.2. Next, in Section 2.2.3, the constitutive laws which complete the model formulation are presented. We summarize the model in (2.64).

We finish with a numerical part in Section 2.3. In this section, a presentation of the space and time discretizations of the bidomain model is given. We present a simulation in a realistic case of the ventricle depolarization and repolarization in a fixed domain. We also give some simulations in moving domain where we suppose that the mesh deformation is known in time. Hence, the complete coupled model is not considered. The obtained results are discussed in Section 2.3.5.3.

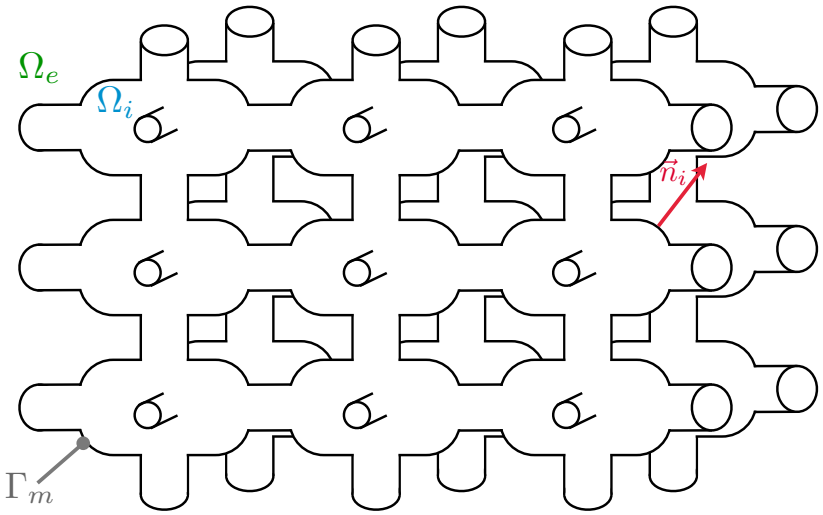

Figure 2.1: Intracellular and extracellular regions of the heart

2.1 Bidomain model

This section begins with the heart description used in electrophysiology. We continue with the model of cellular activity and we finish after a homogenization process by the presentation of the equations of the bidomain model.

2.1.1 Heart description

As explained in the introduction, the cardiac muscle is decomposed into two parts. We denote by Ω the volume of the heart, Ω_i the intracellular region and Ω_e the extracellular region. The cells are connected by many gap junctions, see Figure 2.1. Geometrically, Ω_i and Ω_e are two connected domains verifying

$$\bar{\Omega} = \bar{\Omega}_e \cup \bar{\Omega}_i \text{ and } \Omega_e \cap \Omega_i = \emptyset.$$

We suppose that the membrane $\Gamma_m = \partial\Omega_e \cap \partial\Omega_i = \partial\Omega_i \setminus \partial\Omega$ is regular and we define $\vec{n}_i$ as the unit normal pointing from Ω_i to Ω_e.

In Section 2.1.2.1 we give the equations that hold in the intracellular and extracellular regions.

2.1.2 Model of cellular activity

This section is inspired from [24, 28].

2.1.2.1 Ion motion

General model We consider ion motion in an isothermal medium with negligible solvent flows. Here, this medium corresponds to the cell interior or the extracellular space. We sup-

pose that we have N ionic species in the system. The individual conservation gives

$$\frac{\partial c_k}{\partial t} + \vec{\nabla}_{\vec{x}} \cdot \vec{j}_k = 0,\ k = 1, 2 \dots N, \tag{2.1}$$

where c_k, $\vec{j}_k = c_k \vec{v}_k$ are the concentration and the flux of k-th ionic species, respectively, with $\vec{v}_k$ the velocity corresponding. The charge density $\rho = \rho_0 + \mathrm{F} \sum_{k=1}^{N} z_k c_k$ – with F the Faraday constant – is related to an electrostatic force $\vec{E}$ by the Gauss law of electrostatics

$$\vec{\nabla}_{\vec{x}} \cdot \vec{E} = \frac{\rho}{\varepsilon_0 \varepsilon_s} = \frac{1}{\varepsilon_0 \varepsilon_s}(\rho_0 + \mathrm{F} \sum_{k=1}^{N} z_k c_k), \tag{2.2}$$

with ε_s the relative permittivity of the medium and ε_0, the permittivity of free space.
In order to determine the flux $\vec{j}_k = c_k \vec{v}_k$ of the species k, we want to apply the Nernst-Planck equation. The Nernst-Planck equation results from Newton's second law. The forces for each species k are given by

$$\begin{array}{rcll} \vec{F}_k^{Fick} & = & -\dfrac{RTq}{\mathrm{F}c_k} \vec{\nabla}_{\vec{x}}\, c_k, & \text{Fick law,} \\ \vec{F}_k^{frict} & = & -\mu_k \vec{v}_k, & \text{friction,} \\ \vec{F}_k^{elec} & = & z_k q \vec{E}, & \text{electromagnetism,} \end{array} \tag{2.3}$$

with R is the gas constant, T the temperature, q the elementary charge and μ_k the friction coefficient of the species k. We apply Newton's second law and we neglect the inertia

$$\vec{F}_k^{Fick} + \vec{F}_k^{frict} + \vec{F}_k^{elec} = 0.$$

We infer

$$-\frac{RTq}{\mathrm{F}c_k} \vec{\nabla}_{\vec{x}}\, c_k + z_k q \vec{E} = \mu_k \vec{v}_k.$$

We use it in order to find the flux $\vec{j}_k = c_k \vec{v}_k$ of the species k and we find the Nernst-Planck equation

$$\vec{j}_k = -\frac{RTq}{\mathrm{F}\mu_k} \vec{\nabla}_{\vec{x}}\, c_k + \frac{z_k q c_k}{\mu_k} \vec{E}.$$

This equation becomes

$$\vec{j}_k = -D_k \Big(\vec{\nabla}_{\vec{x}}\, c_k - \frac{\mathrm{F} z_k c_k}{RT} \vec{E} \Big), \tag{2.4}$$

with $D_k = \frac{RTq}{\mathrm{F}\mu_k}$, the diffusion coefficient of the species k. The electrostatic force is related to an electrostatic potential u by $\vec{E} = -\vec{\nabla}_{\vec{x}}\, u$. Using (2.1), (2.4) and (2.2) we finally obtain

$$\left\{ \begin{array}{rcll} \dfrac{\partial c_k}{\partial t} & = & -\vec{\nabla}_{\vec{x}} \cdot \vec{j}_k, & k = 1, 2 \dots N, \quad \text{ionic conservation} \\ \vec{j}_k & = & -D_k \left(\vec{\nabla}_{\vec{x}}\, c_k + \dfrac{F z_k c_k}{RT} \vec{\nabla}_{\vec{x}}\, u \right), & k = 1, 2 \dots N, \quad \text{drift-diffusion flux} \\ \Delta_{\vec{x}}\, u & = & -\dfrac{1}{\varepsilon_0 \varepsilon_s}(\rho_0 + F \sum_{k=1}^{N} z_k c_k). & \quad \text{Gauss law} \end{array} \right. \tag{2.5}$$

This system of equations is named the Nernst-Planck-Poisson equation.

Assumptions: electroneutrality and negligible diffusion current In order to obtain the first equations of the microscopic bidomain model, we make some assumptions. The first assumption is the **electroneutrality** of the solution away from the cell membranes. We make a dimensional analysis of (2.5), see [14, 27, 28]. Firstly we normalize the potential by the *thermal voltage* $\frac{RT}{F}$ and the concentration by a characteristic concentration denoted by c_0. We define $\tilde{u}$, $\tilde{c}_k$, $\tilde{\rho}_0$ by

$$u = \frac{RT}{F}\tilde{u},\ c_k = c_0\tilde{c}_k \text{ and } \rho_0 = Fc_0\tilde{\rho}_0.$$

We also define $\vec{\tilde{j}}_k$ as $\vec{j}_k = c_0\vec{\tilde{j}}_k$. Let L_0 be the typical length of the system over which the flux and the potential vary. The constant D_0 is the typical diffusion coefficient for the ionic species. Given L_0 we define a typical time T_0 by $T_0 = \frac{L_0^2}{D_0}$, *i.e.* the diffusion time scale. We define $x = L_0\tilde{x}$, $t = T_0\tilde{t}$, $D_k = D_0\tilde{D}_k$ and $\vec{\tilde{j}}_k = \frac{D_0}{L_0}\vec{\tilde{j}}_k$ and we obtain

$$\begin{cases} \dfrac{\partial\tilde{c}_k}{\partial\tilde{t}} = -\vec{\nabla}_{\vec{\tilde{x}}}\cdot\vec{\tilde{j}}_k, & k=1,2\dots N, \quad \text{charge conservation} \\ \vec{\tilde{j}}_k = -\tilde{D}_k\left(\vec{\nabla}_{\tilde{x}}\tilde{c}_k + z_k\tilde{c}_k\vec{\nabla}_{\tilde{x}}\tilde{u}\right), & k=1,2\dots N, \quad \text{drift-diffusion flux} \\ \beta^2\Delta_{\vec{\tilde{x}}}\tilde{u} = -(\tilde{\rho}_0 + \sum_{k=1}^{N} z_k\tilde{c}_k), & \text{Gauss law} \end{cases} \tag{2.6}$$

with

$$\beta = \frac{\varepsilon_s\varepsilon_0 RT}{F^2c_0L_0^2} = \frac{r_d^2}{L_0^2} \text{ with } r_d = \sqrt{\frac{\varepsilon_s\varepsilon_0 RT}{F^2c_0}},$$

the Debye length. The Debye length is the length over which mobile charge carriers screen out electric fields in an electrolytic solution. It means that the Debye length is the distance over which significant charge separation can occur. The length over which the electric field can vary for a given degree of charge separation, and hence over which $\Delta_{\tilde{x}}\tilde{u}$ can take a large value, is blocked by the Debye length.

An asymptotic analysis on the parameter β is presented in [27, 38]. When this parameter is small, we obtain the electroneutrality which holds away from the cell membranes, where $r_d \approx 1nm$ and L_0 is of the order of microns or more. Making tend β to zero we obtain the electroneutrality assumption

$$\tilde{\rho}_0 + \sum_{k=1}^{N} z_k\tilde{c}_k = 0.$$

Formally we can obtain this limit by making tend $\varepsilon_0\varepsilon_s$ to zero in (2.2). It does not imply that $\vec{\nabla}_{\vec{x}}\cdot\vec{E} = 0$, since (2.2) reduces to $\Delta_{\vec{x}}\,u = \frac{0}{0}$. We infer

$$\sum_{k=1}^{N} z_k\frac{\partial\tilde{c}_k}{\partial\tilde{t}} = 0 \text{ and } \sum_{k=1}^{N} z_k\frac{\partial c_k}{\partial t} = 0.$$

We finally obtain

$$\vec{\nabla}_{\vec{x}} \cdot \left(\sum_{k=1}^{N} z_k \vec{j}_k \right) = \vec{\nabla}_{\vec{x}} \cdot \left(\sum_{k=1}^{N} z_k D_k \left(\vec{\nabla}_{\vec{x}}\, c_k + \frac{F z_k c_k}{RT} \vec{\nabla}_{\vec{x}}\, u \right) \right) = 0. \tag{2.7}$$

The drift current is larger than the diffusion current and dominates, hence (2.7) becomes

$$\vec{\nabla}_{\vec{x}} \cdot \left(\sum_{k=1}^{N} D_k \frac{F^2 (z_k)^2 c_k}{RT} \vec{\nabla}_{\vec{x}}\, u \right) = \vec{\nabla}_{\vec{x}} \cdot (\sigma \vec{\nabla}_{\vec{x}}\, u) = 0, \tag{2.8}$$

where

$$\sigma = \frac{F^2}{RT} \sum_{k=1}^{N} D_k (z_k)^2 c_k$$

is the total conductivity.

Application to cardiac electrophysiology We denote by u_i and u_e, the intracellular and extracellular potentials respectively with the associated conductivities σ_i and σ_e. Using (2.8), we have

$$\begin{cases} \vec{\nabla}_{\vec{x}} \cdot (\sigma_i \vec{\nabla}_{\vec{x}}\, u_i) = 0, \ \Omega_i, \\ \vec{\nabla}_{\vec{x}} \cdot (\sigma_e \vec{\nabla}_{\vec{x}}\, u_e) = 0, \ \Omega_e. \end{cases} \tag{2.9}$$

We need to introduce some definitions for the following sections. Then we define $\vec{j}_i = -\sigma_i \vec{\nabla}_{\vec{x}}\, u_i = \sum_k z_k \vec{j}_{i,k}$ and $\vec{j}_e = -\sigma_e \vec{\nabla}_{\vec{x}}\, u_e = \sum_k z_k \vec{j}_{e,k}$. The dynamics of the bidomain model only depend on the transmembrane electrical phenomena. Section 2.1.2.2 describes them.

2.1.2.2 Electrical transmembrane conditions

General model Near the membrane, there is an accumulation of electric charge, whose thickness is of the order of the Debye length r_d. In deriving the electroneutrality condition, we considered the limit when $\frac{r_d}{L}$ tends to zero. In agreement with this limit, we consider that this electrical charge forms a layer of infinitesimal thickness concentrated on both sides of the membrane surface, see [28]. The electric current $\vec{j}$ that hits the membrane can be blocked by the membrane – and then changes the surface charge μ – or can pass through the membrane with ionic current I_{ion}^{tot} measured from Ω_i to Ω_e. The charge conservation gives

$$\begin{cases} \dfrac{\partial \mu_i}{\partial t} + I_{ion}^{tot} &= \vec{j}_i \cdot \vec{n}_i, \\ \dfrac{\partial \mu_e}{\partial t} - I_{ion}^{tot} &= -\vec{j}_e \cdot \vec{n}_i \end{cases} \tag{2.10}$$

Assumptions In order to obtain the last equations of the bidomain model, we make two assumptions about the surface charge. Any charge accumulation on one side of the membrane is automatically balanced by a charge on the other side

$$\mu_i + \mu_e = 0,$$

and we suppose that the membrane behaves as a capacitance, *i.e.* the surface charge is linearly proportional to the transmembrane potential

$$\mu_i = C_m V_m,$$

where $V_m = u_i - u_e$ denotes the transmembrane potential and C_m the capacitance per unit area of the membrane. We obtain

$$\left\{ \begin{array}{rcl} C_m \dfrac{\partial V_m}{\partial t} + I_{ion}^{tot} & = & \vec{j}_i \cdot \vec{n}_i, \\ \vec{j}_i \cdot \vec{n}_i & = & \vec{j}_e \cdot \vec{n}_i. \end{array} \right. \tag{2.11}$$

2.1.2.3 Microscopic bidomain model

Combining (2.9) and (2.11), the complete microscopic bidomain model is

$$\left\{ \begin{array}{rcll} \vec{\nabla}_{\vec{x}} \cdot (\sigma_i \vec{\nabla}_{\vec{x}} u_i) & = & 0, & \Omega_i \\ \vec{\nabla}_{\vec{x}} \cdot (\sigma_e \vec{\nabla}_{\vec{x}} u_e) & = & 0, & \Omega_e \\ \sigma_i \vec{\nabla}_{\vec{x}} u_i \cdot \vec{n}_i & = & \sigma_e \vec{\nabla}_{\vec{x}} u_e \cdot \vec{n}_i, & \Gamma_m \\ -\sigma_i \vec{\nabla}_{\vec{x}} u_i \cdot \vec{n}_i & = & C_m \dfrac{\partial V_m}{\partial t} + I_{ion}^{tot}, & \Gamma_m. \end{array} \right. \tag{2.12}$$

We define $I_m = C_m \dfrac{\partial V_m}{\partial t} + I_{ion}^{tot}$, the current surface density that arrives on the membrane from Ω_i. We recall that the conductivities $\sigma_{i,e}$ are defined by

$$\sigma_i = \frac{F^2}{RT} \sum_{k=1}^{N} D_k (z_k)^2 c_{i,k} \text{ and } \sigma_e = \frac{F^2}{RT} \sum_{k=1}^{N} D_k (z_k)^2 c_{e,k}.$$

This implies that if we have more than one species, these conductivities are not constant. However, for the homogenization process, we assume that they are constant. One perspective of this work could be to do the homogenization process in a more general case. Note that this assumption is generally used in the literature, explicitly of implicitly.

2.1.3 Macroscopic model

2.1.3.1 Homogenization

The microscopic model is unusable for the whole heart. At the macroscopic scale, the heart appears as a continuous material with a fiber-based structure. At this scale, the intracellular and extracellular media are undistinguishable and we consider that the cardiac volume is "$\Omega = \Omega_i = \Omega_e$". We use an homogenization of the microscopic bidomain model in order to obtain a bidomain model where all the unknowns are defined everywhere. The

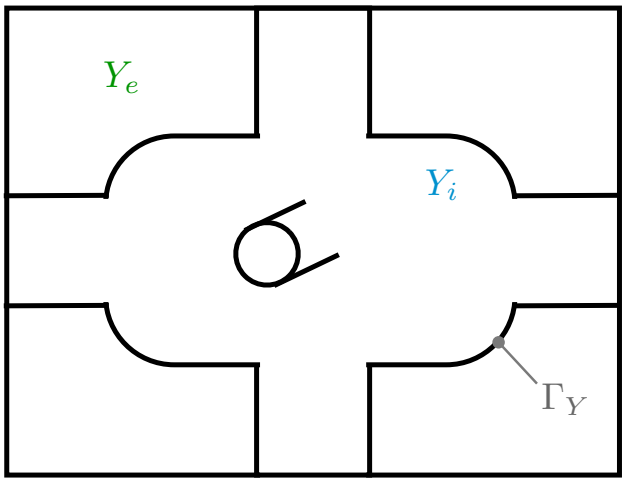

Figure 2.2: A 3D cell (corresponding to the domain Y)

homogenization process has been investigated and written in collaboration with Sebastien Imperiale (researcher in MΞDISIM team).

A normalization of the system is necessary in order to understand the relative amplitudes of the different terms involved. Let L_0^h be a characteristic length of the heart and T_0^h be a characteristic time of a cardiac cycle. In the same idea, we denote by σ_0 a characteristic conductivity, C_m^0 a characteristic membrane capacitance, and u_0 a characteristic potential. We set

$$u_{i,e}(\vec{x},t) = u_0\tilde{u}_{i,e}\Big(\frac{\vec{x}}{L_0^h},\frac{t}{T_0^h}\Big),\quad C_m = C_{m,0}\tilde{C}_m,\ \text{and}\ \sigma_{i,e} = \sigma_0\tilde{\sigma}_{i,e},$$

We assume that $I_{ion}^{tot}(u_i-u_e)$ can be written as

$$I_{ion}^{tot}(u_i-u_e) = I_{ion,0}^{tot}(u_0)\tilde{I}_{ion}^{tot}(\tilde{u}_i-\tilde{u}_e).$$

All quantities denoted by a tilde are dimensionless quantities. We obtain from (2.12), the dimensionless system

$$\begin{cases} \vec{\nabla}_{\vec{x}}\cdot(\tilde{\sigma}_i\vec{\nabla}_{\vec{x}}\tilde{u}_i) &= 0, & \Omega_i^{L_0^h}\\ \vec{\nabla}_{\vec{x}}\cdot(\tilde{\sigma}_e\vec{\nabla}_{\vec{x}}\tilde{u}_e) &= 0, & \Omega_e^{L_0^h}\\ \tilde{\sigma}_i\vec{\nabla}_{\vec{x}}\tilde{u}_i\cdot\vec{n}_i &= \tilde{\sigma}_e\vec{\nabla}_{\vec{x}}\tilde{u}_e\cdot\vec{n}_i, & \Gamma_m^{L_0^h}\\ -\tilde{\sigma}_i\vec{\nabla}_{\vec{x}}\tilde{u}_i\cdot\vec{n}_i &= \dfrac{L_0^hC_{m,0}}{\sigma_0T_0^h}\tilde{C}_m\dfrac{\partial\tilde{u}_i-\tilde{u}_e}{\partial t}+\dfrac{L_0^hI_{ion,0}^{tot}(u_0)}{\sigma_0u_0}\tilde{I}_{ion}^{tot}(\tilde{u}_i-\tilde{u}_e), & \Gamma_m^{L_0^h}, \end{cases} \tag{2.13}$$

where $\Omega_{i,e}^{L_0^h}$ and $\Gamma_m^{L_0^h}$ are rescaled by L_0^h. We define ε – the parameter which tends to zero in the homogenization process – as the ratio between the length of a cell denoted by $l_c = 10^{-2}$cm

and $L_0^h = 10$cm. This implies that ε is the order of 10^{-3}. We study the dimensionless quantity

$$\frac{L_0^h C_{m,0}}{\sigma_0 T_0^h}$$

with T_0^h of the order of 1s, $C_{m,0}$ of 10^{-6}mF.cm^{-2} and σ_0 of 10^{-2}S.cm^{-1} and we see that the quantity is of the same order of ε and can be set to ε by a small modification of the reference quantities. We assume that the term $I_{ion,0}^{tot}(u_0)$ allows to have

$$\frac{L_0^h I_{ion,0}^{tot}(u_0)}{\sigma_0 u_0}$$

of the order of ε and to set it to ε. To study the mathematical properties of this problem, we consider a family of such problems where ε tends to and characterizes the limit equation. The fact that ε is small means that the microscopic scale and the macroscopic scale are well separated. For the sake of clarity, we do not keep the tilde notation but we write the dependance in ε.

We use the results of the 2-scale convergence, see [2, 4] for the homogenization process. This classical method has been used in many fields of science and engineering.

We denote by Y the micro-structure corresponding to a cardiac cell, see Figure 2.2. This micro-structure is decomposed into two parts: the intracellular part Y_i and the extracellular part Y_e. The intra and the extra-cellular domains are separated by Γ_Y. We recall that ε is a characteristic length of this cell. The global position vector is denoted by $\vec{x}$ and the local position vector by $\vec{y}$. We have

$$\vec{y} = \frac{1}{\varepsilon}\,\vec{x}.$$

We define the domain $\Omega = \Omega_i^\varepsilon \cup \Omega_e^\varepsilon$ by ε-periodicity and we denote by Γ^ε the boundary between the intra- and the extra-cellular domains Ω_i^ε and Ω_e^ε. We assume that Ω_i^ε and Ω_i^ε are a union of cells and we have

$$\Omega_{i,e}^\varepsilon = \bigcup_k (\varepsilon Y_{i,e} + \varepsilon \vec{w}_k) = \varepsilon \bigcup_k (Y_{i,e} + \vec{w}_k) \tag{2.14}$$

and then

$$\Gamma^\varepsilon = \bigcup_k (\varepsilon \Gamma_Y + \varepsilon \vec{w}_k) = \varepsilon \bigcup_k (\Gamma_Y + \vec{w}_k), \tag{2.15}$$

with $\vec{w}_k$ is the vector corresponding to the translation between the cell considered and the reference cell. The objective is to homogenize the following problem, *i.e.* study the convergence when ε tends to zero

$$\left\{\begin{array}{rcll} \vec{\nabla}_{\vec{x}} \cdot (\sigma_{i,e} \vec{\nabla}_{\vec{x}}\, u_{i,e}^\varepsilon) & = & 0, & \Omega_{i,e}^\varepsilon \\ \sigma_i \vec{\nabla}_{\vec{x}}\, u_i^\varepsilon \cdot \vec{n}_i & = & \sigma_e \vec{\nabla}_{\vec{x}}\, u_e^\varepsilon \cdot \vec{n}_i, & \Gamma^\varepsilon \\ -\sigma_i \vec{\nabla}_{\vec{x}}\, u_i^\varepsilon \cdot \vec{n}_i & = & \varepsilon\Big(C_m \dfrac{\partial (u_i^\varepsilon - u_e^\varepsilon)}{\partial t} + I_{ion}^{tot}(u_i^\varepsilon - u_e^\varepsilon)\Big), & \Gamma^\varepsilon. \end{array}\right. \tag{2.16}$$

In order to have a unique solution, we add the conditions

$$\int_{\Omega^\varepsilon_{i,e}} u^\varepsilon_{i,e} = 0.$$

We study the simplified case $I^{tot}_{ion} = \alpha(u^\varepsilon_i - u^\varepsilon_e) - \beta(\vec{x},t)$, with $\alpha > 0$ and assuming β regular sufficiently. The associated variational problem is find $(u^\varepsilon_i, u^\varepsilon_e) \in C^1(0,T;H^1(\Omega^\varepsilon_i)) \times C^1(0,T;H^1(\Omega^\varepsilon_e))$ such that $\forall (v^\varepsilon_i, v^\varepsilon_e) \in H^1(\Omega^\varepsilon_i) \times H^1(\Omega^\varepsilon_e)$,

$$\Big(\sigma_i \vec{\nabla}_{\vec{x}}\, u^\varepsilon_i, \vec{\nabla}_{\vec{x}}\, v^\varepsilon_i\Big)_{\Omega^\varepsilon_i} + \Big(\sigma_e \vec{\nabla}_{\vec{x}}\, u^\varepsilon_e, \vec{\nabla}_{\vec{x}}\, v^\varepsilon_e\Big)_{\Omega^\varepsilon_e} + \varepsilon C_m \Big\langle \frac{\partial (u^\varepsilon_i - u^\varepsilon_e)}{\partial t}, v^\varepsilon_i - v^\varepsilon_e \Big\rangle_{\Gamma^\varepsilon} + \varepsilon\alpha \Big\langle u^\varepsilon_i - u^\varepsilon_e, v^\varepsilon_i - v^\varepsilon_e \Big\rangle_{\Gamma^\varepsilon} = \varepsilon \Big\langle \beta, v^\varepsilon_i - v^\varepsilon_e \Big\rangle_{\Gamma^\varepsilon}. \quad (2.17)$$

The regularity that we assume here is strong compared to the classical results of existence and unicity. We make this choice for the sake of simplicity. We take $v^\varepsilon_{i,e} = \partial_t u^\varepsilon_{i,e}$ in (2.17) in order to derive an a priori estimate and we obtain

$$\frac{1}{2}\partial t \Big[\Big(\sigma_i \vec{\nabla}_{\vec{x}}\, u^\varepsilon_i, \vec{\nabla}_{\vec{x}}\, u^\varepsilon_i\Big)_{\Omega^\varepsilon_i} + \Big(\sigma_e \vec{\nabla}_{\vec{x}}\, u^\varepsilon_e, \vec{\nabla}_{\vec{x}}\, u^\varepsilon_e\Big)_{\Omega^\varepsilon_e} + \varepsilon\alpha \Big\langle u^\varepsilon_i - u^\varepsilon_e, u^\varepsilon_i - u^\varepsilon_e \Big\rangle_{\Gamma^\varepsilon} \Big] = -\varepsilon C_m \Big\langle \frac{\partial (u^\varepsilon_i - u^\varepsilon_e)}{\partial t}, \frac{\partial (u^\varepsilon_i - u^\varepsilon_e)}{\partial t} \Big\rangle_{\Gamma^\varepsilon} + \varepsilon \Big\langle \beta, \frac{\partial (u^\varepsilon_i - u^\varepsilon_e)}{\partial t} \Big\rangle_{\Gamma^\varepsilon}. \quad (2.18)$$

We define

$$\mathcal{E}^\varepsilon(t) = \Big(\sigma_i \vec{\nabla}_{\vec{x}}\, u^\varepsilon_i, \vec{\nabla}_{\vec{x}}\, u^\varepsilon_i\Big)_{\Omega^\varepsilon_i} + \Big(\sigma_e \vec{\nabla}_{\vec{x}}\, u^\varepsilon_e, \vec{\nabla}_{\vec{x}}\, u^\varepsilon_e\Big)_{\Omega^\varepsilon_e} + \varepsilon\alpha \|u^\varepsilon_i - u^\varepsilon_e\|^2_{L^2(\Gamma^\varepsilon)}.$$

By integrating (2.17) between 0 and t, we obtain

$$\begin{aligned}
\frac{1}{2}\mathcal{E}^\varepsilon(t) - \frac{1}{2}\mathcal{E}^\varepsilon(0) &= -\varepsilon C_m \int_0^t \Big\langle \frac{\partial (u^\varepsilon_i - u^\varepsilon_e)}{\partial s}, \frac{\partial (u^\varepsilon_i - u^\varepsilon_e)}{\partial s} \Big\rangle_{\Gamma^\varepsilon} ds + \varepsilon \int_0^t \Big\langle \beta, \frac{\partial (u^\varepsilon_i - u^\varepsilon_e)}{\partial s} \Big\rangle_{\Gamma^\varepsilon} ds, \\
&\leq \varepsilon \int_0^t \Big\langle \beta(\vec{x},s), \frac{\partial (u^\varepsilon_i - u^\varepsilon_e)}{\partial s}(\vec{x},s) \Big\rangle_{\Gamma^\varepsilon}, \\
&\leq \varepsilon \Big\langle \beta(\vec{x},t), (u^\varepsilon_i - u^\varepsilon_e)(\vec{x},t) \Big\rangle_{\Gamma^\varepsilon} - \varepsilon \Big\langle \beta(\vec{x},0), (u^\varepsilon_i - u^\varepsilon_e)(\vec{x},0) \Big\rangle_{\Gamma^\varepsilon} \\
&\qquad -\varepsilon \int_0^t \Big\langle \frac{\partial \beta}{\partial s}(\vec{x},s), (u^\varepsilon_i - u^\varepsilon_e)(\vec{x},s) \Big\rangle_{\Gamma^\varepsilon} ds.
\end{aligned}$$

We study the first and the third terms of this inequality separately. We start with

$$\begin{aligned}
\varepsilon \Big\langle \beta(\vec{x},t), (u^\varepsilon_i - u^\varepsilon_e)(\vec{x},t) \Big\rangle_{\Gamma^\varepsilon} &\leq \varepsilon \|\beta(\vec{x},t)\|_{L^2(\Gamma^\varepsilon)} \|(u^\varepsilon_i - u^\varepsilon_e)(\vec{x},t)\|_{L^2(\Gamma^\varepsilon)} \\
&\leq \varepsilon^{\frac{1}{2}} \|\beta(\vec{x},t)\|_{L^2(\Gamma^\varepsilon)} \alpha^{-\frac{1}{2}} \sqrt{\mathcal{E}^\varepsilon(t)} \\
&\leq \frac{1}{2}\varepsilon \|\beta(\vec{x},t)\|^2_{L^2(\Gamma^\varepsilon)} \alpha^{-1} r_1 + \frac{1}{2r_1}\mathcal{E}^\varepsilon(t),\ \forall r_1 > 0.
\end{aligned}$$

The third term becomes

$$\begin{aligned}
-\varepsilon\int_0^t\Big\langle\frac{\partial\beta(\vec{x},s)}{\partial s},(u_i^\varepsilon-u_e^\varepsilon)(\vec{x},s)\Big\rangle_{\Gamma^\varepsilon}ds &\leq \varepsilon\int_0^t\|\partial_s\beta(\vec{x},s)\|_{L^2(\Gamma^\varepsilon)}\|(u_i^\varepsilon-u_e^\varepsilon)(\vec{x},s)\|_{L^2(\Gamma^\varepsilon)}ds\\
&\leq \sup_{s\in[0,t]}\Big(\varepsilon^{\frac{1}{2}}t\|\partial_s\beta(\vec{x},s)\|_{L^2(\Gamma^\varepsilon)}\alpha^{-\frac{1}{2}}\sqrt{\mathcal{E}^\varepsilon(\vec{x},s)}\Big)\\
&\leq \sup_{s\in[0,t]}\Big(\frac{1}{2}\varepsilon t^2\|\partial_s\beta(\vec{x},s)\|^2_{L^2(\Gamma^\varepsilon)}\alpha^{-1}r_2+\frac{1}{2r_2}\mathcal{E}^\varepsilon(s)\Big),\ \forall r_2>0,\\
&\leq \frac{1}{2}\varepsilon t^2\alpha^{-1}r_2\sup_{[0,t]}\|\partial_s\beta(\vec{x},s)\|^2_{L^2(\Gamma^\varepsilon)}+\frac{1}{2r_2}\sup_{[0,t]}\mathcal{E}^\varepsilon(s),\ \forall r_2>0.
\end{aligned}$$

This implies

$$\begin{aligned}
\frac{1}{2}\mathcal{E}^\varepsilon(t) &\leq \frac{1}{2}\mathcal{E}^\varepsilon(0)-\varepsilon\Big\langle\beta(0),(u_i^\varepsilon-u_e^\varepsilon)(0)\Big\rangle_{\Gamma^\varepsilon}+\frac{1}{2}\varepsilon\|\beta(\vec{x},t)\|^2_{L^2(\Gamma^\varepsilon)}\alpha^{-1}r_1+\frac{1}{2r_1}\mathcal{E}^\varepsilon(t)\\
&\quad+\frac{1}{2}\varepsilon t^2\alpha^{-1}r_2\sup_{[0,t]}\|\partial_s\beta(\vec{x},s)\|^2_{L^2(\Gamma^\varepsilon)}+\frac{1}{2r_2}\sup_{[0,t]}\mathcal{E}^\varepsilon(s),\ \forall r_1,r_2>0.
\end{aligned}$$

With $r_1=r_2=4$, we have

$$\begin{aligned}
\frac{1}{4}\sup_{[0,T]}\mathcal{E}^\varepsilon(t) &\leq \frac{1}{2}\mathcal{E}^\varepsilon(0)-\varepsilon\Big\langle\beta(0),(u_i^\varepsilon-u_e^\varepsilon)(0)\Big\rangle_{\Gamma^\varepsilon}+2\varepsilon\sup_{[0,T]}\|\beta(\vec{x},t)\|^2_{L^2(\Gamma^\varepsilon)}\alpha^{-1}\\
&\quad+2\varepsilon T^2\alpha^{-1}\sup_{[0,T]}\|\partial_t\beta(\vec{x},t)\|^2_{L^2(\Gamma^\varepsilon)}.
\end{aligned}$$

We suppose that β is regular enough to have

$$\varepsilon\sup_{[0,T]}\|\beta(\vec{x},t)\|^2_{L^2(\Gamma^\varepsilon)}\text{ and }\varepsilon\sup_{[0,T]}\|\partial_t\beta(\vec{x},t)\|^2_{L^2(\Gamma^\varepsilon)}$$

bounded. Note that $|\Gamma^\varepsilon|=O(\varepsilon^{-1})$. With regular initial conditions, we finally obtain

$$\|\vec{\nabla}_{\vec{x}}\,u_i^\varepsilon\|^2_{L^2(\Omega_i^\varepsilon)}+\|\vec{\nabla}_{\vec{x}}\,u_e^\varepsilon\|^2_{L^2(\Omega_e^\varepsilon)}+\varepsilon\|u_i^\varepsilon-u_e^\varepsilon\|^2_{L^2(\Gamma^\varepsilon)}\leq C, \tag{2.19}$$

where C is independent of ε. According to the Poincaré inequality and (2.19), $(u_{i,e}^\varepsilon)_\varepsilon$ are uniformly bounded in $H^1(\Omega_{i,e}^\varepsilon)$. The same method can be applied with time derivatives, as soon as β is regular enough, and allows to obtain that $(\partial_t u_{i,e}^\varepsilon)_\varepsilon$ is uniformly bounded in the same norm. We denote by $\tilde{\cdot}$ the extension by zero in Ω of the following functions

$$\tilde{u}_i^\varepsilon=\begin{vmatrix} u_i^\varepsilon & \Omega_i^\varepsilon,\\ 0 & \Omega_e^\varepsilon,\end{vmatrix}\qquad \tilde{u}_e^\varepsilon=\begin{vmatrix} 0 & \Omega_i^\varepsilon,\\ u_e^\varepsilon & \Omega_e^\varepsilon,\end{vmatrix}$$

$$\widetilde{\vec{\nabla}_{\vec{x}}\,u_i^\varepsilon}=\begin{vmatrix} \vec{\nabla}_{\vec{x}}\,u_i^\varepsilon & \Omega_i^\varepsilon,\\ 0 & \Omega_e^\varepsilon,\end{vmatrix}\qquad \widetilde{\vec{\nabla}_{\vec{x}}\,u_e^\varepsilon}=\begin{vmatrix} 0 & \Omega_i^\varepsilon,\\ \vec{\nabla}_{\vec{x}}\,u_e^\varepsilon & \Omega_e^\varepsilon,\end{vmatrix}$$

$$\sigma=\begin{vmatrix} \sigma_i & \Omega_i^\varepsilon,\\ \sigma_e & \Omega_e^\varepsilon.\end{vmatrix}$$

According to the variational form (2.17), we search $(\tilde{u}_i^\varepsilon,\ \tilde{u}_e^\varepsilon,\ \widetilde{\vec{\nabla}_{\vec{x}}\, u_i^\varepsilon},\ \widetilde{\vec{\nabla}_{\vec{x}}\, u_e^\varepsilon})$ with $(u_i^\varepsilon, u_e^\varepsilon) \in C^1(0,T;H^1(\Omega_i^\varepsilon)) \times C^1(0,T;H^1(\Omega_e^\varepsilon))$ verifying $\forall v_i^\varepsilon,\ v_e^\varepsilon \in H^1(\Omega)$

$$\Big(\sigma_i \widetilde{\vec{\nabla}_{\vec{x}}\, u_i^\varepsilon}, \vec{\nabla}_{\vec{x}}\, v_i^\varepsilon\Big)_\Omega + \Big(\sigma_e \widetilde{\vec{\nabla}_{\vec{x}}\, u_e^\varepsilon}, \vec{\nabla}_{\vec{x}}\, v_e^\varepsilon\Big)_\Omega + \varepsilon C_m \Big\langle \frac{\partial(\tilde{u}_i^\varepsilon - \tilde{u}_e^\varepsilon)}{\partial t}, v_i^\varepsilon - v_e^\varepsilon \Big\rangle_{\Gamma^\varepsilon} + \varepsilon\alpha \Big\langle \tilde{u}_i^\varepsilon - \tilde{u}_e^\varepsilon, v_i^\varepsilon - v_e^\varepsilon \Big\rangle_{\Gamma^\varepsilon} = \varepsilon \Big\langle \beta, v_i^\varepsilon - v_e^\varepsilon \Big\rangle_{\Gamma^\varepsilon}. \quad (2.20)$$

The a priori estimation (2.19) allows to apply the 2-scale convergence – see [2, 4]. We fix t and when ε tends to zero, we have

$$\begin{cases} \tilde{u}_i^\varepsilon & \to & u_i^0(\vec{x},t)\chi_{Y_i}(\vec{y}),\ \forall t, \\ \tilde{u}_e^\varepsilon & \to & u_e^0(\vec{x},t)\chi_{Y_e}(\vec{y}),\ \forall t, \\ \widetilde{\vec{\nabla}_{\vec{x}}\, u_i^\varepsilon} & \to & (\vec{\nabla}_{\vec{x}}\, u_i^0 + \vec{\nabla}_{\vec{y}}\, u_i^1)\chi_{Y_i}(\vec{y}),\ \forall t, \\ \widetilde{\vec{\nabla}_{\vec{x}}\, u_e^\varepsilon} & \to & (\vec{\nabla}_{\vec{x}}\, u_e^0 + \vec{\nabla}_{\vec{y}}\, u_e^1)\chi_{Y_e}(\vec{y}),\ \forall t, \end{cases} \quad (2.21)$$

where χ_{Y_i} and χ_{Y_e} are the characteristic functions of Y_i and Y_e respectively and with $u_i^0,\ u_e^0 \in C^0(0,T;H^1(\Omega))$ and $u_i^1,\ u_e^1 \in C^0(0,T;L^2(\Omega;H^1_\sharp(Y)/\mathbb{R}))$, where $H^1_\sharp(Y)$ is the completion for the norm $H^1(Y)$ of $C^\infty_\sharp(Y)$, the space of infinitely differentiable functions that are periodic of period Y.

Remark 1. *The 2-scale convergence $\tilde{u}_i^\varepsilon \in C^1(0,T;L^2(\Omega)) \to u_i^0(\vec{x})\chi_{Y_i}(\vec{y}),\ \forall t$ means that there exists a function f^ε such that*

$$\int_\Omega \tilde{u}_i^\varepsilon(\vec{x})\, v\Big(\vec{x}, \frac{1}{\varepsilon}\vec{x}\Big)\, d\vec{x} - \int_\Omega \int_{Y_i} u_i^0(\vec{x})\, v(\vec{x},\vec{y})\, d\vec{y}d\vec{x} = f^\varepsilon(t),\ \forall t,\ \forall v \in L^2(\Omega \times Y), \quad (2.22)$$

with $\sup_t |f^\varepsilon(t)|$ tends to zero.

The theory of the 2-scale homogenization is applied in a static case, see [2], but we also need the convergences of the following proposition.

Proposition 1. *The functions $u_i^0,\ u_e^0 \in C^1(0,T;H^1(\Omega))$ and we have*

$$\begin{cases} \partial_t \tilde{u}_i^\varepsilon & \to & \partial_t u_i^0(\vec{x},t)\chi_{Y_i}(\vec{y}),\ \forall t, \\ \partial_t \tilde{u}_e^\varepsilon & \to & \partial_t u_e^0(\vec{x},t)\chi_{Y_e}(\vec{y}),\ \forall t. \end{cases} \quad (2.23)$$

Proof. We have $(u_i^\varepsilon, u_e^\varepsilon) \in C^1(0,T;H^1(\Omega_i^\varepsilon)) \times C^1(0,T;H^1(\Omega_e^\varepsilon))$ and we can apply the 2-scale convergence to $\partial_t \tilde{u}_i^\varepsilon$ and $\partial_t \tilde{u}_e^\varepsilon$. We obtain

$$\begin{cases} \partial_t \tilde{u}_i^\varepsilon & \to & w_i^0(\vec{x},t)\chi_{Y_i}(\vec{y}),\ \forall t, \\ \partial_t \tilde{u}_e^\varepsilon & \to & w_e^0(\vec{x},t)\chi_{Y_e}(\vec{y}),\ \forall t. \end{cases} \quad (2.24)$$

We want to show that $w_i^0 = \partial_t u_i^0$. There exists a function g^ε such that

$$\Big(\partial_t \tilde{u}_i^\varepsilon,\ v\Big)_\Omega - \Big(w_i^0,\ v\Big)_{Y_i \times \Omega} = g^\varepsilon(t),\ \forall t,\ \forall v \in H^1(\Omega)$$

with $\sup_t |g^\varepsilon(t)|$ tends to zero. We integrate this equation in time between 0 and T and using an integration by parts, we have

$$\int_0^T \int_\Omega \partial_t \tilde{u}_i^\varepsilon \, v - \int_0^T \int_{Y_i \times \Omega} w_i^0 \, v = \int_\Omega \tilde{u}_i^\varepsilon(T) \, v - \int_\Omega \tilde{u}_i^\varepsilon(0) \, v - \int_{Y_i \times \Omega} \int_0^T w_i^0 \, v = \int_0^T g^\varepsilon(t), \, \forall T.$$

Making tend ε to zero, we obtain

$$\int_{Y_i \times \Omega} u_i^0(T) \, v - \int_{Y_i \times \Omega} u_i^0(0) \, v = \int_{Y_i \times \Omega} \int_0^T w_i^0 \, v, \, \forall T$$

and then

$$u_i^0(T) - u_i^0(0) = \int_0^T w_i^0, \, \forall T.$$

This last equation implies that u_i^0 is time-differentiable and $w_i^0 = \partial_t u_i^0$. We can use the same method for u_e^0. □

We use test functions of the form $v_{i,e}^\varepsilon = v_{i,e}^0(\vec{x}) + \varepsilon v_{i,e}^1(\vec{x}, \frac{\vec{x}}{\varepsilon})$ and (2.20) becomes

$$\begin{aligned} \Big(\sigma_i \widetilde{\vec{\nabla}_{\vec{x}} u_i^\varepsilon}, \vec{\nabla}_{\vec{x}} v_i^0 + \vec{\nabla}_{\vec{y}} v_i^1 + \varepsilon \vec{\nabla}_{\vec{x}} v_i^1 \Big)_\Omega + \Big(\sigma_e \widetilde{\vec{\nabla}_{\vec{x}} u_e^\varepsilon}, \vec{\nabla}_{\vec{x}} v_i^0 + \vec{\nabla}_{\vec{y}} v_i^1 + \varepsilon \vec{\nabla}_{\vec{x}} v_e^1 \Big)_\Omega \\ + \varepsilon C_m \Big\langle \frac{\partial(\tilde{u}_i^\varepsilon - \tilde{u}_e^\varepsilon)}{\partial t}, v_i^0 + \varepsilon v_i^1 - v_e^0 - \varepsilon v_e^1 \Big\rangle_{\Gamma^\varepsilon} \\ + \varepsilon \alpha \Big\langle \tilde{u}_i^\varepsilon - \tilde{u}_e^\varepsilon, v_i^0 + \varepsilon v_i^1 - v_e^0 - \varepsilon v_e^1 \Big\rangle_{\Gamma^\varepsilon} = \varepsilon \Big\langle \beta, v_i^0 + \varepsilon v_i^1 - v_e^0 - \varepsilon v_e^1 \Big\rangle_{\Gamma^\varepsilon}. \end{aligned} \tag{2.25}$$

We want to apply the results of the 2-scale convergence but the surface terms need a more detailed analysis. The following proposition gives the convergence of the surface term.

Proposition 2. *Let $(u^\varepsilon)_\varepsilon$ be a bounded sequence in $H^1(\Omega_{i,e}^\varepsilon)$ and $\tilde{u}^\varepsilon$ its extension in $H^1(\Omega)$. We define u^0 the two scale limit of the sequence $(\tilde{u}^\varepsilon)_\varepsilon$,* i.e.

$$\tilde{u}^\varepsilon \underset{\varepsilon \to 0}{\to} u^0(\vec{x}) \, \chi_{Y_{i,e}}(\vec{y}).$$

Then we have, for all $v : (\vec{x}, \vec{y}) \mapsto v(\vec{x}, \vec{y}) \in \mathcal{D}(\Omega, C_\sharp^\infty(Y))$, where $\mathcal{D}(\Omega, C_\sharp^\infty(Y))$ is the space of continuous functions on Ω with values in $C_\sharp^\infty(Y)$ and having compact supports.

$$\varepsilon \int_{\Gamma^\varepsilon} u^\varepsilon(\vec{x}) \, v\Big(\vec{x}, \frac{\vec{x}}{\varepsilon}\Big) d\Gamma^\varepsilon \;\longrightarrow\; \int_\Omega u^0(\vec{x}) \int_{\Gamma_Y} v(\vec{x}, \vec{y}) \, d\Gamma_Y \, d\vec{x}. \tag{2.26}$$

Proof. We consider here the intra-cellular case and the same arguments can be applied for the extra-cellular case. We define $w : (\vec{x}, \vec{y}) \mapsto w(\vec{x}, \vec{y})$ periodic in Y as satisfying, for each $\vec{x}$ (seen here as a parameter),

$$\left\{ \begin{array}{rcll} \Delta_{\vec{y}} \, w(\vec{x}, \cdot) & = & \dfrac{1}{|Y_i|} \displaystyle\int_{\Gamma_Y} v(\vec{x}, \vec{y}) d\Gamma_Y & \text{in } Y_i, \\ \vec{\nabla}_{\vec{y}} \, w(\vec{x}, \vec{y}) \cdot \vec{n}_{\Gamma_Y}(\vec{y}) & = & v(\vec{x}, \vec{y}) & \text{on } \Gamma_Y, \end{array} \right. \tag{2.27}$$

and

$$\int_{Y_i} w(\vec{x},\vec{y})\, dY_i = 0.$$

We can see that $w(\vec{x},\cdot) \in H^1(Y_i)$ is well defined, *i.e.* the compatibility assumptions are satisfied. Substituting v in the left term of (2.26) we find

$$\varepsilon \int_{\Gamma^\varepsilon} u^\varepsilon(\vec{x})\, v\Big(\vec{x},\frac{\vec{x}}{\varepsilon}\Big) d\,\Gamma^\varepsilon = \varepsilon \int_{\Gamma^\varepsilon} u^\varepsilon(\vec{x})\, \vec{\nabla}_{\vec{y}}\, w\Big(\vec{x},\frac{\vec{x}}{\varepsilon}\Big)\cdot \vec{n}_{\Gamma_Y}\Big(\frac{\vec{x}}{\varepsilon}\Big)\, d\,\Gamma^\varepsilon. \tag{2.28}$$

Without limiting the generality, Ω_i^ε is assumed to be a union of cells εY_i (see (2.14)). Then using (2.15), we can show that

$$\vec{n}_{\Gamma_Y}\Big(\frac{\vec{x}}{\varepsilon}\Big) = \vec{n}(\vec{x}),\ \vec{x}\in\Gamma^\varepsilon.$$

Applying Green's formula in (2.28), we have

$$\begin{aligned}\varepsilon \int_{\Gamma^\varepsilon} u^\varepsilon(\vec{x})\, v\Big(\vec{x},\frac{\vec{x}}{\varepsilon}\Big) d\,\Gamma^\varepsilon &= \varepsilon \int_{\Gamma^\varepsilon} u^\varepsilon(\vec{x})\, \vec{\nabla}_{\vec{y}}\, w\Big(\vec{x},\frac{\vec{x}}{\varepsilon}\Big)\cdot \vec{n}\, d\,\Gamma^\varepsilon \\ &= \varepsilon \int_{\Omega_i^\varepsilon} u^\varepsilon(\vec{x})\vec{\nabla}\cdot\Big[\vec{\nabla}_{\vec{y}}\, w\Big(\vec{x},\frac{\vec{x}}{\varepsilon}\Big)\Big] d\vec{x} + \varepsilon \int_{\Omega_i^\varepsilon} \vec{\nabla} u^\varepsilon(\vec{x})\cdot \vec{\nabla}_{\vec{y}}\, w\Big(\vec{x},\frac{\vec{x}}{\varepsilon}\Big) d\vec{x}.\end{aligned}$$

The last term in the RHS vanishes in the limit. Indeed, $\vec{\nabla} u^\varepsilon(\vec{x})$ is uniformly bounded in $L^2(\Omega_i^\varepsilon)$ by assumption and $\vec{\nabla}_{\vec{y}}\, w$ is also bounded due to the regularity assumption of v. Then, we remark that

$$\varepsilon \int_{\Omega_i^\varepsilon} u^\varepsilon(\vec{x})\vec{\nabla}\cdot\Big[\vec{\nabla}_{\vec{y}}\, w\Big(\vec{x},\frac{\vec{x}}{\varepsilon}\Big)\Big] d\vec{x} = \int_{\Omega_i^\varepsilon} u^\varepsilon(\vec{x})\Delta_{\vec{y}}\, w\Big(\vec{x},\frac{\vec{x}}{\varepsilon}\Big) d\vec{x} + \varepsilon \int_{\Omega_i^\varepsilon} u^\varepsilon(\vec{x})\vec{\nabla}_{\vec{x}}\cdot \vec{\nabla}_{\vec{y}}\, w\Big(\vec{x},\frac{\vec{x}}{\varepsilon}\Big) d\vec{x}.$$

The last term also vanishes in the limit as $w(\vec{x},\vec{y})$ is regular in $\vec{x}$, still due to the regularity property of v. Finally we have

$$\int_{\Omega_i^\varepsilon} u^\varepsilon(\vec{x})\Delta_{\vec{y}}\, w\Big(\vec{x},\frac{\vec{x}}{\varepsilon}\Big) d\vec{x} = \int_{\Omega} \tilde{u}^\varepsilon(\vec{x})\Delta_{\vec{y}}\, w\Big(\vec{x},\frac{\vec{x}}{\varepsilon}\Big) d\vec{x} = \frac{1}{|Y_i|}\int_{\Omega} \tilde{u}^\varepsilon(\vec{x}) \int_{\Gamma_Y} v(\vec{x},\vec{y}) d\,\Gamma_Y d\vec{x}.$$

Using (2.22) with v corresponds here to $\int_{\Gamma_Y} v(\vec{x},\vec{y}) d\,\Gamma_Y$, the two scale convergence of $(\tilde{u}^\varepsilon)_\varepsilon$ implies

$$\frac{1}{|Y_i|}\int_{\Omega_i^\varepsilon} \tilde{u}^\varepsilon(\vec{x}) \int_{\Gamma_Y} v(\vec{x},\vec{y})\, d\,\Gamma_Y\, d\vec{x} \longrightarrow \frac{1}{|Y_i|}\int_{\Omega}\int_{Y_i} u^0(\vec{x}) \int_{\Gamma_Y} v(\vec{x},\vec{y})\, d\,\Gamma_Y\, d\vec{y}\, d\vec{x}.$$

As

$$u^0(\vec{x}) \int_{\Gamma_Y} v(\vec{x},\vec{y})\, d\,\Gamma_Y$$

is independent of $\vec{y}$, we obtain

$$\frac{1}{|Y_i|}\int_{\Omega_i^\varepsilon} \tilde{u}^\varepsilon(\vec{x}) \int_{\Gamma_Y} v(\vec{x},\vec{y})\, d\,\Gamma_Y\, d\vec{x} \longrightarrow \int_{\Omega} u^0(\vec{x}) \int_{\Gamma_Y} v(\vec{x},\vec{y})\, d\,\Gamma_Y\, d\vec{x}.$$

This last convergence completes the proof. □

With the test functions $v^{\varepsilon}_{i,e} = v^0_{i,e}(\vec{x}) + \varepsilon v^1_{i,e}(\vec{x}, \frac{\vec{x}}{\varepsilon})$, Proposition 2 directly gives

$$\varepsilon \int_{\Gamma^\varepsilon} u^{\varepsilon}_{i,e}\, v^{\varepsilon}_{i,e} d\Gamma^\varepsilon \longrightarrow |\Gamma_Y| \int_\Omega u^0_{i,e} v^0_{i,e} d\vec{x}. \tag{2.29}$$

Using (2.21) and (2.29) to $\tilde{u}^{\varepsilon}_i$, $\tilde{u}^{\varepsilon}_e$, $\partial_t \tilde{u}^{\varepsilon}_i$ and $\partial_t \tilde{u}^{\varepsilon}_e$ we obtain

$$\begin{aligned}&\Big(\sigma_i(\vec{\nabla}_{\vec{x}}\, u^0_i + \vec{\nabla}_{\vec{y}}\, u^1_i), \vec{\nabla}_{\vec{x}}\, v^0_i + \vec{\nabla}_{\vec{y}}\, v^1_i\Big)_{Y_i\times\Omega} + \Big(\sigma_e(\vec{\nabla}_{\vec{x}}\, u^0_e + \vec{\nabla}_{\vec{y}}\, u^1_e), \vec{\nabla}_{\vec{x}}\, v^0_e + \vec{\nabla}_{\vec{y}}\, v^1_e\Big)_{Y_e\times\Omega}\\ &+ |\Gamma_Y| C_m \Big(\frac{\partial(u^0_i - u^0_e)}{\partial t}, v^0_i - v^0_e\Big)_\Omega + |\Gamma_Y|\alpha\Big(u^0_i - u^0_e, v^0_i - v^0_e\Big)_\Omega = |\Gamma_Y|\Big(\beta, v^0_i - v^0_e\Big)_\Omega.\end{aligned} \tag{2.30}$$

In order to obtain the cell problem in Y_i, we take $v^0_i = v^0_e = v^1_e = 0$ and (2.30) becomes

$$\Big(\sigma_i(\vec{\nabla}_{\vec{x}}\, u^0_i + \vec{\nabla}_{\vec{y}}\, u^1_i), \vec{\nabla}_{\vec{y}}\, v^1_i\Big)_{Y_i\times\Omega} = 0.$$

Using the classical method of the cell problem, see [2, 4] for the method, we decompose u^1_i in the form

$$u^1_i = \sum_{j=1}^{3} \mathcal{X}^j_i \partial_{x^j} u^0_i,$$

and the variational form of the cell problem is find $\mathcal{X}^j_i$, $j = 1..3$ such that

$$\Big(\vec{\nabla}_{\vec{y}}\, \mathcal{X}^j_i,\, \vec{\nabla}_{\vec{y}}\, w\Big)_{Y_i} = -\Big(\vec{e}^{\,j},\, \vec{\nabla}_{\vec{y}}\, w\Big)_{Y_i},\ \forall w \in H^1_\sharp. \tag{2.31}$$

The associated effective medium is the tensor $\vec{\vec{T}}_i$ defined by

$$(\vec{\vec{T}}_i)_{j,k} = \int_{Y_i} (\vec{\nabla}_{\vec{y}}\, \mathcal{X}^j_i + \vec{e}^{\,j}) \cdot (\vec{\nabla}_{\vec{y}}\, \mathcal{X}^k_i + \vec{e}^{\,k}) dY_i.$$

We define $\vec{\vec{J}}^{\mathcal{X}}_i$ the Jacobian matrix of the function $\mathcal{X}_i = (\mathcal{X}^1_i,\, \mathcal{X}^2_i,\, \mathcal{X}^3_i)$ and we can show that

$$\vec{\vec{T}}_i = \int_{Y_i} (\vec{\vec{I}} + \big(\vec{\vec{J}}^{\mathcal{X}}_i\big)^{\intercal}) dY_i = |Y_i|\vec{\vec{I}} + \int_{Y_i} \big(\vec{\vec{J}}^{\mathcal{X}}_i\big)^{\intercal} dY_i.$$

We use exactly the same method in order to find the cell problem in the extracellular domain Y_e and the effective medium $\vec{\vec{T}}_e$. We take $v^1_i = v^1_e = v^0_e = 0$ in (2.30) and we obtain

$$\begin{aligned}&\Big(\sigma_i(\vec{\nabla}_{\vec{x}}\, u^0_i + \vec{\nabla}_{\vec{y}}\, u^1_i), \vec{\nabla}_{\vec{\vec{x}}}\, v^0_i\Big)_{Y_i\times\Omega} + |\Gamma_Y| C_m \Big(\frac{\partial(u^0_i - u^0_e)}{\partial t}, v^0_i\Big)_\Omega + |\Gamma_Y|\alpha\Big(u^0_i - u^0_e, v^0_i\Big)_\Omega\\ &= |\Gamma_Y|\Big(\beta, v^0_i\Big)_\Omega.\end{aligned} \tag{2.32}$$

Using the decomposition of $u_i^1 = \sum_{j=1}^3 \mathcal{X}_i^j \partial_{x^j} u_i^0$ and $\vec{\nabla}_{\vec{x}} u_i^0 = \sum_{j=1}^3 \partial_{x^j} u_i^0 \vec{e}^j$, we obtain

$$\begin{aligned}\Big(\sigma_i(\vec{\nabla}_{\vec{x}} u_i^0 + \vec{\nabla}_{\vec{y}} u_i^1), \vec{\nabla}_{\vec{x}} v_i^0\Big)_{Y_i\times\Omega} &= \Big(\sigma_i \sum_{j=1}^3 (\vec{e}^j + \vec{\nabla}_{\vec{y}} \mathcal{X}_i^j)\partial_{x^j} u_i^0, \vec{\nabla}_{\vec{x}} v_i^0\Big)_{Y_i\times\Omega} \\ &= \Big(\sigma_i \vec{\vec{T}}_i \vec{\nabla}_{\vec{x}} u_i^0, \vec{\nabla}_{\vec{x}} v_i^0\Big)_{\Omega},\end{aligned}$$

and this implies that

$$\Big(\sigma_i \vec{\vec{T}}_i \vec{\nabla}_{\vec{x}} u_i^0, \vec{\nabla}_{\vec{x}} v_i^0\Big)_{\Omega} + |\Gamma_Y| C_m \Big(\frac{\partial(u_i^0 - u_e^0)}{\partial t}, v_i^0\Big)_{\Omega} + |\Gamma_Y| \alpha \Big(u_i^0 - u_e^0, v_i^0\Big)_{\Omega} = |\Gamma_Y| \Big(\beta, v_i^0\Big)_{\Omega},$$

and for the extracellular potential

$$\Big(\sigma_e \vec{\vec{T}}_e \vec{\nabla}_{\vec{x}} u_e^0, \vec{\nabla}_{\vec{x}} v_e^0\Big)_{\Omega} - |\Gamma_Y| C_m \Big(\frac{\partial(u_i^0 - u_e^0)}{\partial t}, v_e^0\Big)_{\Omega} - |\Gamma_Y| \alpha \Big(u_i^0 - u_e^0, v_e^0\Big)_{\Omega} = -|\Gamma_Y| \Big(\beta, v_e^0\Big)_{\Omega}.$$

Denoting by $A_m = \frac{|\Gamma_Y|}{|Y|}$ – the ratio of membrane area per unit volume – the homogenization of the microscopic bidomain model gives

$$\begin{cases} -\vec{\nabla}_{\vec{x}} \cdot \dfrac{\sigma_i}{|Y|} \vec{\vec{T}}_i \vec{\nabla}_{\vec{x}} u_i^0 = A_m \Big(C_m \dfrac{\partial(u_i^0 - u_e^0)}{\partial t} + (\alpha(u_i^0 - u_e^0) - \beta)\Big), & \text{on } \Omega \times (0,T), \\ \\ -\vec{\nabla}_{\vec{x}} \cdot \dfrac{\sigma_e}{|Y|} \vec{\vec{T}}_e \vec{\nabla}_{\vec{x}} u_e^0 = -A_m \Big(C_m \dfrac{\partial(u_i^0 - u_e^0)}{\partial t} + (\alpha(u_i^0 - u_e^0) - \beta)\Big), & \text{on } \Omega \times (0,T), \end{cases}$$

where $\vec{\vec{T}}_{i,e} = |Y_{i,e}| \vec{\vec{I}} + \int_{Y_{i,e}} \big(\vec{\vec{J}}_{i,e}^{\mathcal{X}}\big)^{\intercal} dY_{i,e}$ with $\mathcal{X}_{i,e}$ the solutions of the cell problems. We define

$$\begin{aligned}\vec{\vec{\sigma}}_{i,e} &= \frac{\sigma_{i,e}}{|Y|} \vec{\vec{T}}_{i,e} = \sigma_{i,e} \frac{|Y_i|}{|Y|} \vec{\vec{I}} + \sigma_{i,e} \frac{1}{|Y|} \int_{Y_{i,e}} \big(\vec{\vec{J}}_{i,e}^{\mathcal{X}}\big)^{\intercal} dY_{i,e} \\ &= \frac{|Y_i|}{|Y|} \frac{F^2}{RT} \sum_{k=1}^N (z_k)^2 D_k c_i^k \vec{\vec{I}} + \frac{1}{|Y|} \frac{F^2}{RT} \sum_{k=1}^N (z_k)^2 D_k c_i^k \int_{Y_{i,e}} \big(\vec{\vec{J}}_{i,e}^{\mathcal{X}}\big)^{\intercal} dY_{i,e} \qquad (2.33)\end{aligned}$$

A classical article for the homogenization of the microscopic bidomain model is [32]. In this article, another epsilon $\tilde{\varepsilon}$ is defined as $\tilde{\varepsilon} = \sqrt{\frac{d_c}{R_m \sigma_i}}$ with d_c the diameter of a unit cell and R_m the surface resistivity of the membrane. The convergence is not proven in this article but in the appendix of [12] a formal proof using the 2-scale convergence is given. In [12], the $\tilde{\varepsilon}$ parameter is considered. As $\tilde{\varepsilon}$ does not correspond to a characteristic length of a cell, we cannot write

$$\Omega_{i,e}^{\tilde{\varepsilon}} = \bigcup_k (\tilde{\varepsilon} Y_{i,e} + \tilde{\varepsilon} \vec{w}_k) \text{ and } \Gamma^{\tilde{\varepsilon}} = \bigcup_k (\tilde{\varepsilon} \Gamma_Y + \tilde{\varepsilon} \vec{w}_k),$$

but another type of decomposition is possible. Indeed, in [12], $\nu = \frac{\sigma_e}{\sigma_i}$ is introduced and the cell is defined by

$$Y_{i,e}^{\tilde{\varepsilon}} = \Omega_{i,e} \cap Y^{\tilde{\varepsilon}} \text{ with } Y^{\tilde{\varepsilon}} = [0, \tilde{\varepsilon}]^2 \times [0, \nu\tilde{\varepsilon}],$$

i.e. Y has the following form $Y = [0,1]^2 \times [0,\nu]$. The quantity ν is interpreted as a ratio between the length and the diameter of an elongated cardiac cell, see Figure 2.2 for example. It seems surprising to use a conductivity parameter as a geometric parameter. The results obtained in [12] are the same as ours, but compared to our analysis, the proof is very formal and in particular, the study of the surface term is not detailed.

We suppose that this homogenization can be extended to any ionic current I_{ion}^{tot}. In practice we do not solve the cell equations given in (2.31) in Y_i and its equivalent in Y_e. However, we consider an anisotropy due to the fact that the electrical conductivity is higher along than across the cardiac fiber directions, see Section 2.3.2 for more details.

Remark 2. *In [18], an homogenization of the gap junctions is given. In these articles, two separate domains corresponding to two different cells are considered. The bidomain model, and not only the intra-cellular equation, is written in each cell. A complete homogenization of the intra-cellular domain with the gap junctions and the extra-cellular domain could be an interesting perspective for future works.*

2.1.3.2 Bidomain model

These equations can be rewritten as

$$\begin{cases} A_m\Big(C_m\dfrac{\partial V_m}{\partial t} + I_{ion}^{tot}\Big) - \vec{\nabla}_{\vec{x}} \cdot \big(\vec{\vec{\sigma}}_i \cdot \vec{\nabla}_{\vec{x}} V_m\big) \\ \qquad\qquad = \vec{\nabla}_{\vec{x}} \cdot \big(\vec{\vec{\sigma}}_i \cdot \vec{\nabla}_{\vec{x}} u_e\big), \qquad \text{on } \Omega \times (0,T), \\ \vec{\nabla}_{\vec{x}} \cdot \Big(\big(\vec{\vec{\sigma}}_i + \vec{\vec{\sigma}}_e\big) \cdot \vec{\nabla}_{\vec{x}} u_e\Big) = -\vec{\nabla}_{\vec{x}} \cdot \big(\vec{\vec{\sigma}}_i \cdot \vec{\nabla}_{\vec{x}} V_m\big), \quad \text{on } \Omega \times (0,T). \end{cases}$$

We add boundary conditions on $\partial\Omega$. At the cell level, we observe that the intra-cellular current $\vec{j}_i$ does not spread outside the heart. Then we have the first following boundary condition,

$$(\vec{\vec{\sigma}}_i \cdot \vec{\nabla}_{\vec{x}} u_i) \cdot \vec{n}_{\partial\Omega} = 0, \text{ on } \partial\Omega,$$

with $\vec{n}_{\partial\Omega}$, the exterior normal of $\partial\Omega$. In terms of V_m and u_e, we obtain

$$(\vec{\vec{\sigma}}_i \cdot \vec{\nabla}_{\vec{x}} V_m) \cdot \vec{n}_{\partial\Omega} = -(\vec{\vec{\sigma}}_i \cdot \vec{\nabla}_{\vec{x}} u_e) \cdot \vec{n}_{\partial\Omega}, \text{ on } \partial\Omega.$$

We suppose in this study that the heart is electrically isolated, *i.e.* the heart is not coupled with the thorax and we have

$$(\vec{\vec{\sigma}}_e \cdot \vec{\nabla}_{\vec{x}} u_e) \cdot \vec{n}_{\partial\Omega} = 0, \text{ on } \partial\Omega.$$

We finally obtain

$$\begin{cases} A_m\Big(C_m\dfrac{\partial V_m}{\partial t}+I_{ion}^{tot}\Big)-\vec{\nabla}_{\vec{x}}\cdot\big(\vec{\vec{\sigma}}_i\cdot\vec{\nabla}_{\vec{x}}V_m\big) \\ \qquad\qquad =\vec{\nabla}_{\vec{x}}\cdot\big(\vec{\vec{\sigma}}_i\cdot\vec{\nabla}_{\vec{x}}u_e\big), & \text{on } \Omega\times(0,T), \\ \vec{\nabla}_{\vec{x}}\cdot\Big(\big(\vec{\vec{\sigma}}_i+\vec{\vec{\sigma}}_e\big)\cdot\vec{\nabla}_{\vec{x}}u_e\Big)=-\vec{\nabla}_{\vec{x}}\cdot\big(\vec{\vec{\sigma}}_i\cdot\vec{\nabla}_{\vec{x}}V_m\big), & \text{on } \Omega\times(0,T), \\ \vec{\vec{\sigma}}_i\cdot\vec{\nabla}_{\vec{x}}V_m\cdot\vec{n}_{\partial\Omega}=-\vec{\vec{\sigma}}_i\cdot\vec{\nabla}_{\vec{x}}u_e\cdot\vec{n}_{\partial\Omega}, & \text{on } \partial\Omega\times(0,T), \\ \vec{\vec{\sigma}}_e\cdot\vec{\nabla}_{\vec{x}}u_e\cdot\vec{n}_{\partial\Omega}=0, & \text{on } \partial\Omega\times(0,T). \end{cases} \tag{2.34}$$

This model is commonly used in simulations of cardiac electrophysiology. The bidomain model is able to take into account different electrical conductivities of the intracellular and extracellular spaces.

2.1.4 Ionic models representing the cell activity

This part concerns the description of the term I_{ion}^{tot} which appears in (2.34). Ionic models represent the transmembrane currents and other cellular ionic processes which give rise to action potentials. We can see in Figure 2.3 the different states of a cell. Initially the cell is at the resting potential near -80mV and is polarized. Due to a stimulus, the cell becomes depolarized very quickly. The depolarization is at the origin of the mechanical activity of the cell. After this depolarization, the repolarization of the cell begins. During this phase, the cell is in refractory phase which means that a new stimulus cannot trigger a new depolarization. The form of the action potential varies in the heart. Figure 2.4 gives a presentation of these various forms. The term I_{ion}^{tot} aims at describing these different states of the cell with various levels of details according to the choice of the ionic model. In terms of modeling, action potentials are produced as a result of ionic currents that pass across the cell membrane, triggering a depolarization or repolarization of the membrane over time. The currents are produced by the displacement of ionic species across the membrane through ionic channels. The channels open and close in response to various stimuli that regulate the transport of ions across the membrane. Each type of channel is highly selective for a specific ion, the most important of which being Na^+, K^+, Ca^{2+}, and Cl^-, and different types of channels sensitive to a given ion can also exist with different kinetics governing their opening and closing. Many physiological models exist in the literature, see [41] for a survey of various models. Historically, the first action potential model is the Hodgkin-Huxley model [19]. In order to understand the complexity of physiological models, we present a brief description of the Hodgkin-Huxley model, the most important model in all of the physiological theory, see [21], originally formulated for neurons. In fact, the Hodgkin-Huxley theory had a remarkable influence on electrophysiology and is the basis of all action potential models. As seen before, the cell membrane can be modeled as a combined resistor and capacitor,

$$C_m\frac{\partial V_m}{\partial t}+I_{ion}^{tot},$$

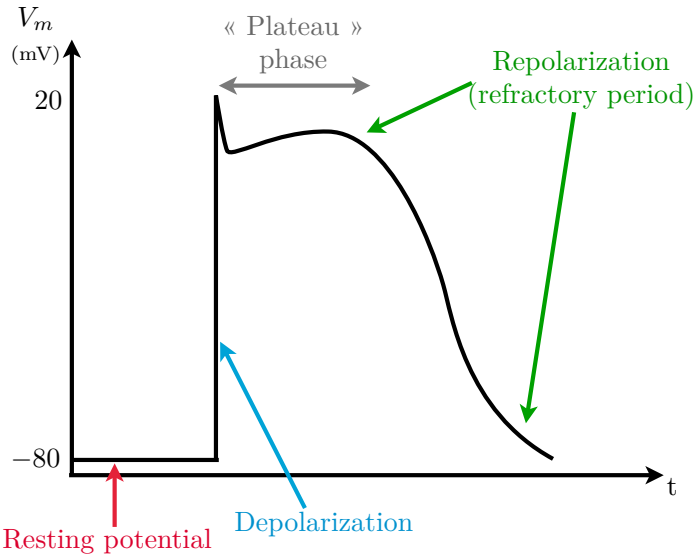

Figure 2.3: Action potential *i.e.* the different states of a cell

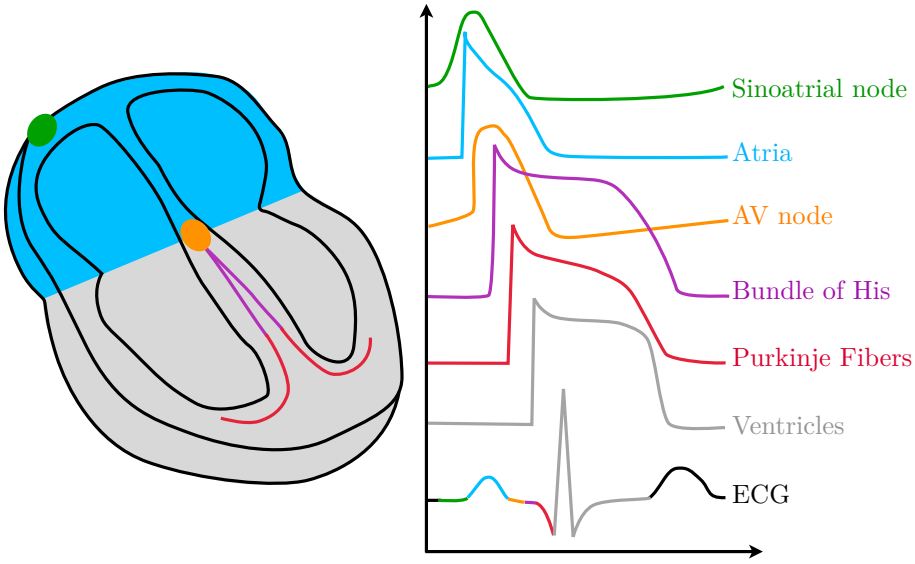

Figure 2.4: Electrical response of each type of cardiac cell, inspired from [17]

with the ionic current I_{ion}^{tot} decomposed into two parts,

$$I_{ion}^{tot} = I_{ion} - I_{app}. \tag{2.35}$$

The term I_{app} corresponds to the external stimulus current. The transmembrane current I_{ion} proposed by the Hodgkin-Huxley model consists of

$$I_{ion} = I_{Na} + I_K + I_l,$$

with the sodium current I_{Na}, the potassium current I_K, and the leakage current I_l which concerns various ions and primarily chloride ions. The currents are determined by

$$\begin{aligned} I_{Na} &= g_{Na}(V_m - E_{Na}), \\ I_K &= g_K(V_m - E_K), \\ I_l &= g_K(V_m - E_l), \end{aligned}$$

where g_{Na}, g_K and g_l are the conductances and E_{Na}, E_K and E_l, the equilibrium voltages of the currents I_{Na}, I_K and I_l respectively. The conductance g_l is supposed to be constant and the other conductances are defined by

$$\begin{aligned} g_{Na} &= m^3 h \bar{g}_{Na}, \\ g_K &= m^4 \bar{g}_K, \end{aligned}$$

with $\bar{g}_{Na}$ and $\bar{g}_K$ are the maximal conductances of the sodium and potassium currents, respectively. The dimensionless state variables m, n and the inactivation variable h verify

$$\frac{dw}{dt} = \alpha_w(1-w) - \beta_w w, \ w = m, \, n, \, h,$$

where α_w and β_w, $w = m, \, n, \, h$, are the voltage-dependent rate constants and control the activation and inactivation of the variables m, n and h. An adaptation of this model to the cardiac action potential is given with the Noble model, see [33]. More complex models – with many ionic currents – were based on the Noble model. We can cite the Luo-Rudy model [25] for the ventricles (19 ionic currents and 30 other variables) and the Courtemanche-Ramirez-Nattel model [13] for the atria (12 ionic currents and 20 other variables, see Figure 2.5), for example. The number of variables of these complex models increase the number of ordinary differential equations coupled with the bidomain model and can induce significant computation times.

All the above-described models are physiological models. As already explained, these models describe the detailed ionic exchanges. Other models – named phenomenological models – are an approximation of the ionic channels. These models are intended to describe the excitability process with a lower complexity. With only one (or a few) additional variable(s) w named the state variable(s) – and then only one (or a few) ordinary differential equation(s) – these models are able to reproduce the depolarization or repolarization of the membrane. The first phenomenological model is the FitzHugh-Nagumo model described in

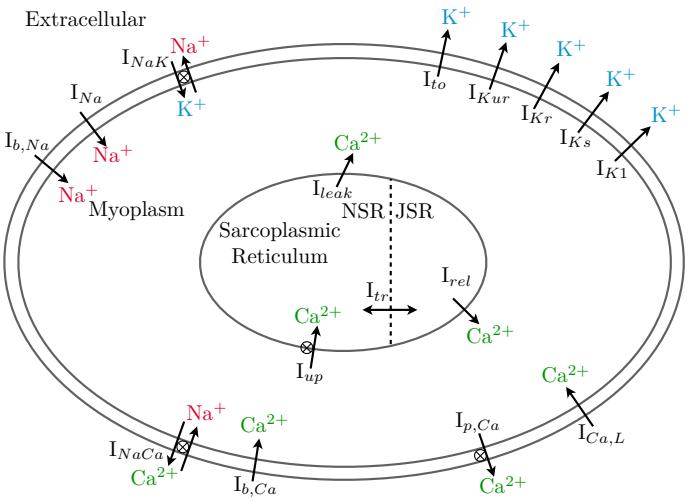

Figure 2.5: A presentation of the Courtemanche model, a classical ionic model for the atria

[15, 29],

$$I_{ion} = kV_m(V_m - a)(V_m - 1) + w, \qquad \frac{dw}{dt} + g(V_m, w) = 0, \tag{2.36}$$

with

$$g(V_m, w) = -\varepsilon(\gamma V_m - w),$$

and with k and a, positive constants. A more widely phenomenological accepted model for ventricular action potential is the Mitchell-Schaeffer model presented in [26],

$$I_{ion} = \frac{w}{\tau_{in}}(V_m)^2(V_m - 1) - \frac{V_m}{\tau_{out}}, \qquad \frac{dw}{dt} + g(V_m, w) = 0, \tag{2.37}$$

with

$$g(V_m, w) = \begin{cases} \dfrac{w-1}{\tau_{open}} \text{ if } V_m \leq v_{gate}, \\ \dfrac{w}{\tau_{close}} \text{ if } V_m > v_{gate} \end{cases} \tag{2.38}$$

and with τ_{open}, τ_{close}, τ_{in}, τ_{out} and v_{gate}, positive constants. We use this phenomenological model for the simulations in Section 2.3. Figure 2.6 gives the form of an action potential with the Mitchell-Schaeffer model. We can also see in this figure the area of influence of each parameter.

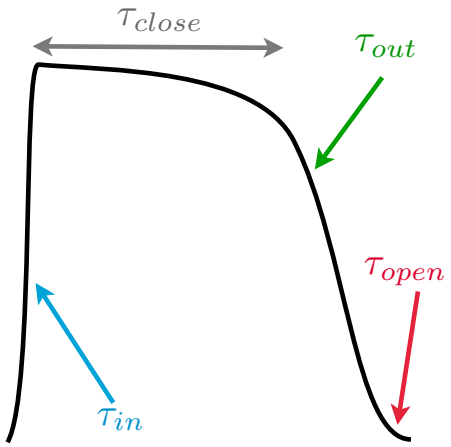

Figure 2.6: An action potential with Mitchell-Schaeffer model (and influence of parameters)

2.1.5 Existence and uniqueness of the bidomain model

We recall that the bidomain model can be written in the following form

$$\begin{cases} A_m\Big(C_m\dfrac{\partial V_m}{\partial t}+I_{ion}^{tot}\Big)-\vec{\nabla}_{\vec{x}}\cdot\big(\vec{\vec{\sigma}}_i\cdot\vec{\nabla}_{\vec{x}}V_m\big) \\ \qquad\qquad =\vec{\nabla}_{\vec{x}}\cdot\big(\vec{\vec{\sigma}}_i\cdot\vec{\nabla}_{\vec{x}}u_e\big), & \text{on } \Omega\times(0,T), \\ \vec{\nabla}_{\vec{x}}\cdot\Big(\big(\vec{\vec{\sigma}}_i+\vec{\vec{\sigma}}_e\big)\vec{\nabla}_{\vec{x}}u_e\Big)=-\vec{\nabla}_{\vec{x}}\cdot\big(\vec{\vec{\sigma}}_i\cdot\vec{\nabla}_{\vec{x}}V_m\big), & \text{on } \Omega\times(0,T), \\ (\vec{\vec{\sigma}}_i\cdot\vec{\nabla}_{\vec{x}}V_m)\cdot\vec{n}_{\partial\Omega}=-(\vec{\vec{\sigma}}_i\cdot\vec{\nabla}_{\vec{x}}u_e)\cdot\vec{n}_{\partial\Omega}, & \text{on } \partial\Omega\times(0,T), \\ (\vec{\vec{\sigma}}_e\cdot\vec{\nabla}_{\vec{x}}u_e)\cdot\vec{n}_{\partial\Omega}=0, & \text{on } \partial\Omega\times(0,T). \end{cases} \tag{2.39}$$

The problem is complemented with initial conditions

$$V_m(0,\vec{x})=V_m^0(\vec{x}),\ \forall x\in\Omega.$$

We notice that u_e is defined up to a constant. We fix this constant by enforcing the following condition

$$\int_\Omega u_e=0$$

on the extra-cellular potential. These equations are coupled with one or more ordinary differential equations governing I_{ion}^{tot}, see Section 2.1.4 for more details.

The existence and uniqueness of the bidomain model has been studied for different ionic models in the literature. The first analysis of the bidomain model has been reported in [11]. The proof is valid for the FitzHugh-Nagumo ionic model (2.36) and is based on a reformulation of the equations in terms of an abstract evolutionary variational inequality. The analysis

for a simplified ionic model, *i.e.* when $I_{ion}(V_m, \cdots) = I_{ion}(V_m)$, has been presented in [3]. For example in this article, it is shown that – under assumptions on the ionic term and on the diffusion tensors and if $V_m(0) \in L^2(\Omega)$ – a unique solution $(u_i, u_e, V_m) \in L^2\big(0, T, H_0^1(\Omega)\big)$ with $V_m = u_i - u_e$ exists and verifies $\partial_t V_m \in L^2\big(0, T, (H_0^1(\Omega))'\big)$. In a more recent work [7], existence and uniqueness of a solution are proved for the bidomain model with a class of general ionic models (including FitzHugh-Nagumo, Aliev-Panfilov [1] and MacCulloch [6]), using a semi-group approach. Finally, in [46], existence, uniqueness and some regularity with a generalized phase-I Luo-Rudy ionic model [25] are proved.

Remark 3. *In practice, we only represent the transmembrane potential in the results of the simulations. Indeed, the diffusion of the front propagation and also the evolution of action potential is represented by the transmembrane potential. A simplified version of the bidomain model named the monodomain model exists, see [10] for more details,*

$$\begin{cases} A_m\Big(C_m \dfrac{\partial V_m}{\partial t} + I_{ion,m}^{tot}\Big) - \vec{\nabla}_{\vec{x}} \cdot \big(\vec{\vec{\sigma}}_m \cdot \vec{\nabla}_{\vec{x}} V_m\big) = 0, & \text{on } \Omega \times (0,T), \\ (\vec{\vec{\sigma}}_m \cdot \vec{\nabla}_{\vec{x}} V_m) \cdot \vec{n}_{\partial\Omega} = 0, & \text{on } \partial\Omega \times (0,T). \end{cases} \tag{2.40}$$

This model can be obtained in the case where the intra- and the extra-cellular domains have the same anisotropic ratio, i.e.

$$\vec{\vec{\sigma}}_i = \lambda \vec{\vec{\sigma}}_e.$$

We use the bidomain model because the monodomain model is a restrictive case and we need to have the knowledge of the extra-cellular potential in order to compute electrocardiograms, see Chapter 6.

2.2 Mechanical bidomain model

In this part, we are interested in the effects of mechanical deformation on the heart electrical activity. In Section 2.2.1, we give the notations and the kinematical definitions. In Section 2.2.2, we establish the conservation laws: the mass conservation and the principle of virtual work for the solid and the charge conservation for the ionic species. In Section 2.2.3, we give the constitutive laws and we finish in Section 2.2.4 by the presentation of the modified bidomain model.

2.2.1 Classical definitions

We consider the heart as a deformable porous medium. However, in a first approximation, we will neglect the fluid flows hence consider an equivalent solid medium. This medium occupies at time t the space domain $\Omega(t)$ with boundary $\partial\Omega(t)$. We denote by $(\Omega_0, \partial\Omega_0)$ the reference configuration, not necessarily corresponding to $(\Omega(0), \partial\Omega(0))$. We consider a material point M_0 at position $\vec{\xi}$ in Ω_0. The trajectory of the point M_0 is defined by the curve $\vec{x}(t) = \vec{\phi}(\vec{\xi}, t)$. The total Lagrangian formulation consists in considering the physical and mechanical quantities of interest in the variable $\vec{\xi} \in \Omega_0$. The deformation is then a unique

mapping from the reference configuration to the current configuration giving the position of each material point over time,

$$\vec{\phi} : \begin{cases} \bar{\Omega}_0 \times [0,T] & \rightarrow & \mathbb{R}^3 \\ (\vec{\xi},\, t) & \mapsto & \vec{x}(t) = \vec{\phi}(\vec{\xi},t). \end{cases}$$

We define the displacement $\vec{y}$ by $\vec{y}(\vec{\xi},t) = \vec{x} - \vec{\xi} = \vec{\phi}(\vec{\xi},t) - \vec{\xi}$ and the gradient by

$$\vec{\vec{F}}(\vec{\xi},t) = \vec{\nabla}_{\vec{\xi}}\vec{\phi} = \vec{\vec{1}} + \vec{\nabla}_{\vec{\xi}}\vec{y}.$$

We denote by $d\Omega$ the volume measure in the reference configuration. The deformed volume is then given by $Jd\Omega$ with $J = \det \vec{\vec{F}}$ and the deformed area vector by $J\vec{\vec{F}}^{-T} \cdot d\vec{S}$.
The local changes of geometry are given by the Green-Lagrange strain tensor denoted by $\vec{\vec{e}}$ and defined by

$$\vec{\vec{e}} = \frac{1}{2}(\vec{\nabla}_{\vec{\xi}}\vec{y} + (\vec{\nabla}_{\vec{\xi}}\vec{y})^{\intercal} + (\vec{\nabla}_{\vec{\xi}}\vec{y})^{\intercal} \cdot \vec{\nabla}_{\vec{\xi}}\vec{y}).$$

The linearized expression of $\vec{\vec{e}}$ is

$$\vec{\vec{\varepsilon}} = \frac{1}{2}(\vec{\nabla}_{\vec{\xi}}\vec{y} + (\vec{\nabla}_{\vec{\xi}}\vec{y})^{\intercal}).$$

2.2.2 Conservation laws

2.2.2.1 Principle of virtual work for the solid

We start with the mass conservation law of the solid and we obtain

$$\frac{\partial \rho_s}{\partial t} + \vec{\nabla}_{\vec{x}} \cdot (\rho_s\, \vec{v}_s) = 0,\ \Omega(t), \tag{2.41}$$

with ρ_s the solid density and $\vec{v}_s$ the solid velocity. Using the mass conservation and the conservation of momentum, we can derive Newton's second law of motion,

$$\vec{\nabla}_{\vec{x}} \cdot \vec{\vec{\lambda}} + \rho_s \vec{f} - \rho_s \vec{\gamma}_s = 0,\ \Omega(t), \tag{2.42}$$

with $\vec{\gamma}_s$ the acceleration of the solid and $\vec{\vec{\lambda}}$ the Cauchy stress tensor. The weak formulation of this fundamental law of dynamics is given on the space of admissible displacements $V(\Omega(t))$. We obtain in the Eulerian and Lagrangian formulations

$$\int_{\Omega(t)} \big(\vec{\vec{\lambda}} : \vec{\vec{\varepsilon}}(\vec{v}_1^*) - \rho_s(\vec{f} - \vec{\gamma}_s) \cdot \vec{v}_1^*\big) d\Omega - \int_{\partial\Omega(t)} \vec{t} \cdot \vec{v}_1^*\, dS = 0,\ \forall \vec{v}_1^* \in V(\Omega(t)), \tag{2.43}$$

$$\int_{\Omega_0} \big(\vec{\vec{\Lambda}} : d_{\vec{y}}\vec{\vec{e}} \cdot \vec{v}_1^* - \rho_0(\vec{f} - \vec{\gamma}_s) \cdot \vec{v}_1^*\big) d\Omega - \int_{\partial\Omega_0} \vec{t}_0 \cdot \vec{v}_1^*\, dS = 0,\ \forall \vec{v}_1^* \in V(0), \tag{2.44}$$

with

$$d_{\vec{y}}\,\vec{\vec{e}} = \frac{1}{2}\big(\vec{\vec{F}}^T \cdot \vec{\nabla}_{\vec{\xi}}\vec{v}^* + (\vec{\nabla}_{\vec{\xi}}\vec{v}^*)^\intercal \cdot \vec{\vec{F}}\big)$$
$$\rho_0(\xi) = J\rho_s(x(\xi,t),t),$$
$$\vec{\vec{\Lambda}} = J\vec{\vec{F}}^{-1}\vec{\vec{\lambda}}\,\vec{\vec{F}}^{-T},$$
$$\text{and } \vec{t}_0 = J\|\vec{\vec{F}}^{-T}\cdot\vec{n}_0\|\vec{t}.$$

2.2.2.2 Charge conservation

We denote by $\vec{v}_{i,k,r}$ and $\vec{v}_{e,k,s}$, the relative velocities of the species k in the intra and extracellular domains respectively. These relative velocities correspond to the ionic velocities in the classical bidomain model. In this part, the species also moves due to the transport effect of the solid with velocity $\vec{v}_s$. Then the flux of the species k in the intra and extracellular domains is respectively defined by

$$\vec{\jmath}_{i,k} = \vec{\jmath}_{i,k,r} + \vec{\jmath}_{i,k,s} = c_{i,k}(\vec{v}_{i,k,r} + \vec{v}_s),$$

$$\vec{\jmath}_{e,k} = \vec{\jmath}_{e,k,r} + \vec{\jmath}_{e,k,s} = c_{e,k}(\vec{v}_{e,k,r} + \vec{v}_s),$$

with $c_{i,k}$, $c_{e,k}$ the concentrations of the species k in the intra and extracellular domains, respectively. We denote by ϕ_i and ϕ_e, the volume fractions of the intra- and extra-cellular domains respectively. We suppose that ϕ_i and ϕ_e are constant in space and time for the sake of simplicity. Based on the previous analysis of the homogenization process in fixed domain, the conservation matter is directly extended at the macroscopic level and the equations are posed in $\Omega(t)$. Indeed, the conservation of matter gives

$$\begin{cases} \phi_i \dfrac{\partial c_{i,k}}{\partial t} + \phi_i \vec{\nabla}_{\vec{x}} \cdot \vec{\jmath}_{i,k} + A_m\,\mu_{i,k} &= 0,\ \Omega(t), \\ \phi_e \dfrac{\partial c_{e,k}}{\partial t} + \phi_e \vec{\nabla}_{\vec{x}} \cdot \vec{\jmath}_{e,k} + A_m\,\mu_{e,k} &= 0,\ \Omega(t). \end{cases} \tag{2.45}$$

where $\mu_{e,k}$ and $\mu_{i,k}$ correspond to the membrane interactions. The fact – that any charge accumulation on one side of the membrane is automatically balanced by a charge on the other side – allows to have

$$p_i + p_e = \sum_{k=1}^{N} z_k p_{i,k} + \sum_{k=1}^{N} z_k p_{e,k} = 0,\ \Omega(t).$$

The formulation (2.45) is a mixture theory like formulation. The mixture theory is a purely macroscopic framework in which the cardiac medium is considered to be made of a superposition of intra- and extra-cellular phases at each point. This theory can be applied in many situations, even if the studied domain does not have a periodic or a quasi-periodic structure. Here, the mixture consists of N ionic species distributed in the intra- and extra-cellular domains and directly gives (2.45).

This method can be compared with the classical method which consists of a microscopic model followed by a homogenization presented in Section 2.1. In order to justify this conservation law, we present the relation with the first part. The homogenization needs a periodic domain but this hypothesis is not necessary in the mixture theory. We can compare (2.45) with (2.1) posed in each subdomain Ω_i and Ω_e. We have with the microscopic model,

$$\left\{\begin{array}{rcll} \partial_t c^\varepsilon_{i,k} + \vec{\nabla}_{\vec{x}} \cdot \vec{j}^\varepsilon_{i,k} & = & 0, & \Omega^\varepsilon_i, \\ \partial_t c^\varepsilon_{e,k} + \vec{\nabla}_{\vec{x}} \cdot \vec{j}^\varepsilon_{e,k} & = & 0, & \Omega^\varepsilon_e, \\ \vec{j}^\varepsilon_{i,k} \cdot \vec{n}_i - \vec{j}^\varepsilon_{e,k} \cdot \vec{n}_i & = & 0, & \Gamma^\varepsilon, \end{array}\right. \tag{2.46}$$

with ε defined as Section 2.1.3.1. Defining

$$\vec{j}^\varepsilon_{i,k} \cdot \vec{n}_i = \varepsilon\, p^\varepsilon_i,$$

we have the following variational formulation, $\forall v_i,\ v_e$,

$$\Big(\partial_t c^\varepsilon_{i,k}, v_i\Big)_{\Omega^\varepsilon_i} - \Big(\vec{j}^\varepsilon_{i,k}, \vec{\nabla}_{\vec{x}}\, v_i\Big)_{\Omega^\varepsilon_i} + \Big(\partial_t c^\varepsilon_{e,k}, v_e\Big)_{\Omega^\varepsilon_e} - \Big(\vec{j}^\varepsilon_{e,k}, \vec{\nabla}_{\vec{x}}\, v_e\Big)_{\Omega^\varepsilon_e} + \varepsilon\Big\langle p^\varepsilon_i, v_i - v_e\Big\rangle_{\Gamma^\varepsilon} = 0. \tag{2.47}$$

We suppose that we have enough regularity so that the surface term converges as in Section 2.1.3.1, *i.e.* there exists p_i such that

$$\varepsilon\Big\langle p^\varepsilon_i, v_i - v_e\Big\rangle_{\Gamma^\varepsilon} \to |\Gamma_Y|\Big(p_i, v_i - v_e\Big)_\Omega.$$

The 2-scale convergence implies that there exist $c_{i,k},\ c_{e,k},\ \vec{j}_{i,k}, \vec{j}_{e,k}$ such that $\forall v_i,\ v_e$,

$$\begin{aligned} |Y_i|\Big(\partial_t c_{i,k}, v_i\Big)_\Omega - |Y_i|\Big(\vec{j}_{i,k}, \vec{\nabla}_{\vec{x}}\, v_i\Big)_\Omega + |Y_e|\Big(\partial_t c_{e,k}, v_e\Big)_\Omega \\ - |Y_e|\Big(\vec{j}_{e,k}, \vec{\nabla}_{\vec{x}}\, v_e\Big)_\Omega + |\Gamma_Y|\Big(p_i, v_i - v_e\Big)_\Omega = 0. \end{aligned} \tag{2.48}$$

We then obtain

$$\left\{\begin{array}{rcll} \dfrac{|Y_i|}{|Y|}\partial_t c_{i,k} + \dfrac{|Y_i|}{|Y|}\vec{\nabla}_{\vec{x}} \cdot \vec{j}_{i,k} + \dfrac{|\Gamma_Y|}{|Y|} p_i & = & 0, & \Omega, \\ \dfrac{|Y_e|}{|Y|}\partial_t c_{e,k} + \dfrac{|Y_e|}{|Y|}\vec{\nabla}_{\vec{x}} \cdot \vec{j}_{e,k} - \dfrac{|\Gamma_Y|}{|Y|} p_i & = & 0, & \Omega. \end{array}\right. \tag{2.49}$$

We see that this homogenization process in the periodic case gives exactly (2.45) presented in the mixture theory.

Using the electroneutrality assumption, we obtain

$$\rho_i = \mathrm{F}\sum_{k=1}^N z_k c_{i,k} + \rho^0_i = 0, \text{ and } \rho_e = \mathrm{F}\sum_{k=1}^N z_k c_{e,k} + \rho^0_e = 0,$$

with ρ^0_i, ρ^0_e the background charge densities, F the Faraday constant and z_k the valence of

the ionic species k. The background density corresponds to the ionic species which are in the fluid but do not move through the membrane because of the very weak permeability of the membrane for these ionic species. For example, in most ionic models, we neglect the variation of the chloride ions and we only consider the potassium, the sodium and the calcium. The electroneutrality assumption implies

$$\mathrm{F}\sum_{k=1}^{N} z_k c_{i,k} = -\rho_i^0, \text{ and } \mathrm{F}\sum_{k=1}^{N} z_k c_{e,k} = -\rho_e^0. \tag{2.50}$$

2.2.2.3 Generalized bidomain model

The total flux $\vec{j}_i$ in the intracellular domain is defined by $\vec{j}_i = \vec{j}_{i,r} + \vec{j}_{i,s}$ with

$$\vec{j}_{i,r} = \mathrm{F}\sum_{k=1}^{n} z_k \vec{j}_{i,k,r} = \mathrm{F}\sum_{k=1}^{n} z_k c_{i,k} \vec{v}_{i,k,r}$$

and

$$\vec{j}_{i,s} = \mathrm{F}\sum_{k=1}^{n} z_k c_{i,k} \vec{v}_s = -\rho_i^0\, \vec{v}_s,$$

We have the same definition for $\vec{j}_e$,

$$\vec{j}_e = \vec{j}_{e,r} + \vec{j}_{e,s} = \mathrm{F}\sum_{k=1}^{n} z_k c_{e,k} \vec{v}_{e,k,r} - \rho_e^0\, \vec{v}_s. \tag{2.51}$$

Using (2.45) and the definition of the flux, we then obtain a generalization of the bidomain model

$$\begin{cases} \dfrac{\partial(\phi_i\rho_i^0 + \phi_e\rho_e^0)}{\partial t} + \phi_i \vec{\nabla}_{\vec{x}} \cdot \vec{j}_i + \phi_e \vec{\nabla}_{\vec{x}} \cdot \vec{j}_e = 0, \\ \phi_i \dfrac{\partial \rho_i^0}{\partial t} + \phi_i \vec{\nabla}_{\vec{x}} \cdot \vec{j}_i + A_m p_i = 0, \end{cases} \tag{2.52}$$

We obtain for the first equation of (2.52) in the Eulerian and Lagrangian formulations,

$$\begin{aligned} &\int_{\Omega(t)} \left(\frac{\partial(\phi_i\rho_i^0 + \phi_e\rho_e^0)}{\partial t} v_2^* + \vec{\nabla}_{\vec{x}} \cdot (\phi_i \vec{j}_i + \phi_e \vec{j}_e) v_2^* \right) d\Omega \\ &\qquad = \int_{\Omega_0} \left(\frac{\partial(\phi_i \mathrm{R}_i^0 + \phi_e \mathrm{R}_e^0)}{\partial t} v_2^* + \vec{\nabla}_{\vec{\xi}} \cdot (\phi_i \vec{J}_i + \phi_e \vec{J}_e) v_2^* \right) d\Omega = 0, \ \forall v_2^*, \end{aligned} \tag{2.53}$$

with

$$\vec{J}_i = J\vec{\vec{F}}^{-1}\vec{j}_i, \ \vec{J}_e = J\vec{\vec{F}}^{-1}\vec{j}_e,$$
$$\mathrm{R}_i^0 = J\rho_i^0, \text{ and } \mathrm{R}_e^0 = J\rho_e^0.$$

2.2.2.4 Conservation laws

We finally obtain

$$\begin{cases} \vec{\nabla}_{\vec{x}} \cdot \vec{\vec{\lambda}} + \rho_s \vec{f} - \rho_s \vec{\gamma}_s & = & 0, \quad \Omega(t), & \text{Newton's second law of motion} \\ \dfrac{\partial(\phi_i \rho_i^0 + \phi_e \rho_e^0)}{\partial t}_{|_{\vec{x}}} + \phi_i \vec{\nabla}_{\vec{x}} \cdot \vec{j}_i + \phi_e \vec{\nabla}_{\vec{x}} \cdot \vec{j}_e & = & 0, \quad \Omega(t), & \text{Bidomain model} \\ \phi_i \dfrac{\partial \rho_i^0}{\partial t}_{|_{\vec{x}}} + \phi_i \vec{\nabla}_{\vec{x}} \cdot \vec{j}_i + A_m p_i & = & 0, \quad \Omega(t), & \text{Bidomain model.} \end{cases} \tag{2.54}$$

We need constitutive laws in order for this model to be complete. The constitutive laws will concern $\vec{\vec{\lambda}}$ for the mechanical part and $\vec{j}_{i,e}$, p_i for the bidomain part.

2.2.3 Constitutive laws

This section gives the constitutive laws.

2.2.3.1 Stress tensor

We suppose that the stress tensor is the sum of different contributions. We use the classical decomposition of the stress tensor in cardiac mechanical model, for more details see for example [9].

2.2.3.2 Nernst-Planck equation

We use the detailed presentation of the Nernst-Planck equation given in Section 2.1.2.1 of the classical bidomain model. Necessarily, the study of the Nernst-Planck equation must be done at the cellular level and the epsilon notation is used in order to distinguish the microscopic scale from the macroscopic scale. The presentation of the Nernst-Planck equation is essential to understand how velocity appears in the friction force and the result is the relative velocity and not the complete velocity. In fact, the friction force is independent of the solid displacement. Then the Nernst-Planck equation – when the domain moves – gives the flux of the species $\vec{j}^{\varepsilon}_{i,k,r} = c^{\varepsilon}_{i,k} \vec{v}^{\varepsilon}_{i,k,r}$ in the intracellular domain and we have,

$$\begin{aligned} \vec{j}^{\varepsilon}_{i,k,r} &= -\frac{RTq}{\mathrm{F}\mu_k} \vec{\nabla}_{\vec{x}}\, c^{\varepsilon}_{i,k} + \frac{z_k q c^{\varepsilon}_{i,k}}{\mu_k} \vec{E}^{\varepsilon}_i \\ &= -D_k \Big(\vec{\nabla}_{\vec{x}}\, c^{\varepsilon}_{i,k} - \frac{\mathrm{F} z_k c^{\varepsilon}_{i,k}}{RT} \vec{E}^{\varepsilon}_i \Big), \end{aligned} \tag{2.55}$$

with $D_k = \frac{RTq}{\mathrm{F}\mu_k}$, the diffusion coefficient of the species k and the electric field denoted by $\vec{E}^{\varepsilon}_i$. The drift current is larger than the diffusion current and dominates. Then (2.55) becomes

$$\vec{j}^{\varepsilon}_{i,k,r} = D_k \frac{\mathrm{F} z_k c^{\varepsilon}_{i,k}}{RT} \vec{E}^{\varepsilon}_i, \tag{2.56}$$

and we have

$$\vec{\jmath}^{\,\varepsilon}_{i,r} = \mathrm{F}\sum_{k=1}^{n} z_k \vec{\jmath}^{\,\varepsilon}_{i,k,r} = \sigma_i \vec{E}^{\varepsilon}_i,\ \Omega(t), \tag{2.57}$$

with

$$\sigma_i = \frac{\mathrm{F}^2}{RT}\sum_{k=1}^{n}(z_k)^2 D_k c^{\varepsilon}_{i,k}, \tag{2.58}$$

the conductivity coefficient that we have supposed to be constant even if it depends on $c^{\varepsilon}_{i,k}$. The electrostatic force $\vec{E}_i$ is related to an electrostatic potential u^{ε}_i by $\vec{E}^{\varepsilon}_i = -\vec{\nabla}_{\vec{x}}\, u^{\varepsilon}_i$. We finally obtain

$$\vec{\jmath}^{\,\varepsilon}_{i,r} = -\sigma_i \vec{\nabla}_{\vec{x}}\, u^{\varepsilon}_i. \tag{2.59}$$

With the same analysis in the extracellular domain, we obtain

$$\vec{\jmath}^{\,\varepsilon}_{e,r} = -\sigma_e \vec{\nabla}_{\vec{x}}\, u^{\varepsilon}_e. \tag{2.60}$$

Using the previous homogenization process, we know that at the macroscopic scale there exist two diffusion tensors defined from a cell problem such that

$$\phi_i \vec{\jmath}_{i,r} = -\vec{\vec{\sigma}}_i \cdot \vec{\nabla}_{\vec{x}}\, u_i,\ \Omega(t) \tag{2.61}$$

and

$$\phi_e \vec{\jmath}_{e,r} = -\vec{\vec{\sigma}}_e \cdot \vec{\nabla}_{\vec{x}}\, u_e,\ \Omega(t). \tag{2.62}$$

We recall that $\vec{\vec{\sigma}}_i$ (and also $\vec{\vec{\sigma}}_e$) has the following form

$$\vec{\vec{\sigma}}_i = \phi_i \sigma_i \vec{\vec{I}} + \text{ cell terms},$$

but that in practice, we do not compute the solution of the cell problems. However we take into account the fibers of the cardiac muscle which correspond to preferred directions for the signal. We can remark that we need to use the microscopic scale and the information issued from the homogenization process in order to establish these constitutive laws.

2.2.3.3 Membrane behavior

As previously explained, the membrane is composed of a combined resistor and capacitor. We assume that the ionic species of the membrane layer are fixed to the membrane and we then infer

$$p_i = C_m \left.\frac{\partial V_m}{\partial t}\right|_{\vec{\xi}} + I^{tot}_{ion}(V_m, \cdots),\ \Omega(t). \tag{2.63}$$

We use an ionic model – as explained in Section 2.1.4 – for the term I_{ion}.

2.2.4 Modified bidomain model

We finally obtain $\forall \vec{v}_1^*,\, v_2^*,\, v_3^*$,

$$\begin{cases} \displaystyle\int_{\Omega(t)} \Big(\vec{\vec{\lambda}} : \vec{\vec{\varepsilon}}(\vec{v}_1^*) - \rho_s(\vec{f} - \vec{\gamma}_s)\cdot \vec{v}_1^*\Big) d\Omega - \int_{\partial\Omega(t)} \vec{t}\cdot \vec{v}_1^*\, dS = 0, \\ \displaystyle\int_{\Omega(t)} \Bigg(\frac{\partial(\phi_i\rho_i^0 + \phi_e\rho_e^0)}{\partial t} - \vec{\nabla}_{\vec{x}}\cdot\Big((\phi_i\rho_i^0 + \phi_e\rho_e^0)\vec{v}_s\Big) \\ \qquad\qquad -\vec{\nabla}_{\vec{x}}\cdot\Big(\vec{\vec{\sigma}}_i\cdot\vec{\nabla}_{\vec{x}}\, u_i + \vec{\vec{\sigma}}_e\cdot\vec{\nabla}_{\vec{x}}\, u_e\Big)\Bigg) v_2^*\, d\Omega = 0, \\ \displaystyle\int_{\Omega(t)} \Bigg(\phi_i\frac{\partial \rho_i^0}{\partial t} - \phi_i\vec{\nabla}_{\vec{x}}\cdot\Big(\rho_i^0\vec{v}_s\Big) - \vec{\nabla}_{\vec{x}}\cdot\Big(\vec{\vec{\sigma}}_i\cdot\vec{\nabla}_{\vec{x}}\, u_i\Big) + A_m\Big(C_m\frac{\partial V_m}{\partial t}\Big|_{\vec{\xi}} + I_{ion}^{tot}\Big)\Bigg) v_3^*\, d\Omega = 0. \end{cases} \tag{2.64}$$

This model can be rewritten in the Lagrangian framework

$$\begin{cases} \displaystyle\int_{\Omega_0} \Big(\vec{\vec{\Lambda}} : d_{\vec{y}}\,\vec{\vec{e}}\cdot\vec{v}_1^* - \rho_0(\vec{f} - \vec{\gamma}_s)\cdot\vec{v}_1^*\Big) d\Omega - \int_{\partial\Omega_0} \vec{t}_0\cdot\vec{v}_1^*\, dS = 0, \\ \displaystyle\int_{\Omega_0} \Bigg(\frac{\partial(\phi_i \mathrm{R}_i^0 + \phi_e \mathrm{R}_e^0)}{\partial t} - \vec{\nabla}_{\vec{\xi}}\cdot\Big((\phi_i \mathrm{R}_i^0 + \phi_e \mathrm{R}_e^0)\vec{\vec{F}}^{-1}\vec{v}_s\Big) \\ \qquad\qquad -\vec{\nabla}_{\vec{\xi}}\cdot\Big(\vec{\vec{\Sigma}}_i\cdot\vec{\nabla}_{\vec{\xi}} u_i + \vec{\vec{\Sigma}}_e\cdot\vec{\nabla}_{\vec{\xi}} u_e\Big)\Bigg) v_2^* d\Omega = 0, \\ \displaystyle\int_{\Omega_0} \Bigg(\phi_i\frac{\partial \mathrm{R}_i^0}{\partial t} - \phi_i\vec{\nabla}_{\vec{\xi}}\cdot\Big(\mathrm{R}_i^0\vec{\vec{F}}^{-1}\vec{v}_s\Big) - \vec{\nabla}_{\vec{\xi}}\cdot\Big(\vec{\vec{\Sigma}}_i\cdot\vec{\nabla}_{\vec{\xi}} u_i\Big) + A_m\Big(C_m\frac{\partial V_m}{\partial t}\Big|_{\vec{\xi}} + I_{ion}^{tot}\Big)\Bigg) v_3^*\, d\Omega = 0, \end{cases} \tag{2.65}$$

where

$$\vec{\vec{\Sigma}}_{i,e} = J\vec{\vec{F}}^{-1}\cdot\vec{\vec{\sigma}}_{i,e}\cdot\vec{\vec{F}}^{-T}. \tag{2.66}$$

Comparing to the bidomain model, new terms appear which contain the background densities and the solid velocity. The equations are written in moving domains. Some remarks about this model are presented in the following section.

2.2.5 Remarks about the modified bidomain model

2.2.5.1 Background density and incompressibility

In the classical bidomain model, we assume that ρ_i^0 and ρ_e^0 are constant in time but it is not necessarily a correct assumption in moving domain. In fact, when the material is incompressible, we have $\partial_t\rho_i^0 = \partial_t\rho_e^0 = 0$ but if the material is compressible, we must consider these new source terms.

We can also consider the assumption of ρ_i^0 and ρ_e^0 constant in space and in which case, we have

$$\vec{\nabla}_{\vec{x}}\cdot(\phi_i\rho_i^0 + \phi_e\rho_e^0)\vec{v}_s = (\phi_i\rho_i^0 + \phi_e\rho_e^0)\vec{\nabla}_{\vec{x}}\cdot\vec{v}_s.$$

In the case of incompressibility of the material we have $\vec{\nabla}_{\vec{x}} \cdot \vec{v}_s = 0$, *i.e.* $J = 1$ and we finally obtain

$$\vec{\nabla}_{\vec{x}} \cdot (\phi_i \rho_i^0 + \phi_e \rho_e^0) \vec{v}_s = 0.$$

In the intracellular domain the fluid cannot move, hence the intracellular medium is an incompressible material. We cannot suppose this concerning the extracellular domain. In fact, during systole, the blood fluid cannot move but it is not verified during the diastole. Depending on the incompressibility hypothesis, we have real source terms or not. The terms will probably not have a large impact.

2.2.5.2 Conductivity

As previously said, the conductivities parameters depend on $c_{e,k}$ and $c_{i,k}$ but are considered constant as generally done in the literature and in particular in the homogenization process. This particularity has been also noted in [34]. The influence of this dependence is unknown and should be studied in order to quantify it.

2.2.5.3 Boundary conditions

The modified bidomain model requires boundary conditions. Concerning the two unknowns u_i and u_e, we can use the same boundary conditions of the classical bidomain model,

$$(\vec{\vec{\sigma}}_i \cdot \vec{\nabla}_{\vec{x}}\, u_i) \cdot \vec{n}_{\partial\Omega} = 0, \text{ on } \partial\Omega(t)$$

and

$$(\vec{\vec{\sigma}}_e \cdot \vec{\nabla}_{\vec{x}}\, u_e) \cdot \vec{n}_{\partial\Omega} = 0, \text{ on } \partial\Omega(t),$$

where $\vec{n}_{\partial\Omega}$ is the exterior normal of $\partial\Omega(t)$.

2.2.5.4 Comparison with Mechano-Electrical Feedback

We want to compare our results with the literature and in particular, with the Mechano-Electrical Feedback (MEF) [40, 31, 22, 23]. In this work, we study the impact of the displacements in the conservation laws whereas in MEF, it is generally done through the constitutive laws. Indeed, in MEF, various changes are considered as for example through the diffusion tensors, the membrane capacitance C_m and/or the ionic current I_{ion} – with in particular, the consideration of the stretch-activated channels in [22, 23]. In theses studies, $\vec{\nabla}_{\vec{\xi}} u_{i,e}$ is used instead of $\vec{\nabla}_{\vec{x}}\, u_{i,e}$ in our work. For the intra-cellular tensor, it could be justified by the fact that the number of the gap junctions is maybe the most important cause of the ions displacement and not the actual distance. A full homogenization process with the consideration of the gap junctions is necessary – for more detail see Remark 2 – if we want to validate this hypothesis. However, this hypothesis cannot be applied in the extra-cellular domain.

As explained in Section 2.2.5.1, we do not expect an important influence of the new source terms and this implies that the main consideration of the displacements impact in our work is $\vec{\nabla}_{\vec{x}}\, u_{i,e}$. As detailed in what follows, in order to take into account the directions

of the fibers in the cardiac muscle, we assume in this work, that the diffusion tensors are decomposed into two parts

$$\vec{\vec{\sigma}}_{i,e} = \sigma^t_{i,e}\,\vec{\vec{I}} + (\sigma^l_{i,e} - \sigma^t_{i,e})\,\vec{\tau}\otimes\vec{\tau},$$

where the vector $\vec{\tau}$ is of unit length and parallel to the local fiber direction, and $\sigma^l_{i,e}$ and $\sigma^t_{i,e}$ are the conductivity coefficients in the intra- and extra-cellular mediums measured along and across – respectively – the fiber direction. In this work $\sigma^{l,t}_{i,e}$ are constant and the impact of the displacements on the diffusion tensors is only represented in the fact that the fibers are displaced with the heart. A consideration of the impact also in $\sigma^{l,t}_{i,e}$ using the MEF could be very interesting. However, if we compare the mechano-electrical model presented in [31, 22] with our model in Lagrangian framework (2.65), we remark that the results appear to be very close for the diffusion tensors. Indeed, in [31, 22], the diffusion tensors are defined by

$$\vec{\vec{\Sigma}}_{i,e} = J\vec{\vec{F}}^{-1}\cdot\vec{\vec{F}}^{-T}\cdot\vec{\vec{\sigma}}_{i,e},$$

to be compared with (2.66) above. As the method which justifies this choice is not detailed, we do not know if it means that the diffusion tensors are changed in order to replace $\vec{\nabla}_{\vec{x}}\,u_{i,e}$ in the case where $\vec{\vec{F}}^{-1}$ and $\vec{\vec{\sigma}}_{i,e}$ commute (e.g. in 1D). It is one of the perspectives of this work to better understand this difference. Then the impact in the ionic current, still using the MEF, is also an other very interesting perspective.

2.3 Numerical part

In this section we give a brief presentation of the numerical method of the bidomain model that we have used and we present some simulations in a fixed and then in a moving domain. The space and time discretizations used in the simulations are presented in Section 2.3.1 and correspond to the discretizations presented in [5]. In collaboration with Elisa Schenone, PhD student in REO Inria team, we implemented these procedures in the finite element library named *FELiScE*[1] developed at Inria.

2.3.1 Space and time discretization

The discretization in space is performed by applying the finite element method to an appropriate weak formulation of the bidomain model. The bidomain model (2.39) coupled with one ODE for example can be rewritten in a weak form as follows. For all $t > 0$, find

[1] http://felisce.gforge.inria.fr

$V_m(\cdot,t) \in H^1$, $u_e()\cdot,t) \in H^1$ and $w(\cdot,t) \in L^\infty$ with $\int_\Omega u_e = 0$, such that

$$\begin{cases} A_m \int_\Omega \Big(C_m \dfrac{\partial V_m}{\partial t} + I_{ion}(V_m,\, w)\Big)\phi + \displaystyle\int_\Omega \Big[\vec{\vec{\sigma}}_i \cdot \big(\vec{\nabla} V_m + \vec{\nabla} u_e\big)\Big] \cdot \vec{\nabla}\phi \\ \qquad\qquad = A_m \displaystyle\int_\Omega I_{app}\phi, \\ \displaystyle\int_\Omega \Big[(\vec{\vec{\sigma}}_i + \vec{\vec{\sigma}}_e) \cdot \vec{\nabla} u_e\Big] \cdot \vec{\nabla}\psi + \int_\Omega \Big[\vec{\vec{\sigma}}_i \cdot \vec{\nabla} V_m\Big] \cdot \vec{\nabla}\psi = 0, \\ \partial_t w + g(V_m, w) = 0, \ \text{ on } \Omega, \end{cases} \tag{2.67}$$

for all $\phi, \psi \in H^1$ such that $\int_\Omega \psi = 0$. The current I_{app} is a given applied current stimulus as seen in (2.35).

All the simulations in this thesis except the last part are performed by applying the finite element method to the previous weak formulation of the bidomain model.

The resulting system is discretized in time by combining a second-order implicit scheme (backward differentiation formulae, see [37]) with an explicit treatment of the ionic current. Let $N \in \mathbb{N}^*$ be a given integer and consider a uniform partition $\{t_n, t_{n+1}\}_{0\le n\le N-1}$ with $t_n = nT/N = n\Delta t$ of the time interval $[0,T]$. Denote by $(V_m^n,\, u_e^n,\, w^n)$ the approximation of $(V_m,\, u_e,\, w)$ obtained at time t_n.

Then, $(V_m^{n+1},\, u_e^{n+1},\, w^{n+1})$ is computed as follows. For $0 \le n \le N-1$,

---

Algorithm

1. We compute the second order extrapolation of V_m^{n+1} by $\tilde{u}_m^{n+1} = 2V_m^n - V_m^{n-1}$.
2. We solve for w^{n+1},
$$\frac{1}{\Delta t}\Big(\frac{3}{2}w^{n+1} - 2w^n + \frac{1}{2}w_{n-1}\Big) + g(\tilde{u}_m^{n+1}, w^n) = 0.$$
3. We solve for $(V_m^{n+1},\, u_e^{n+1})$ with $\int_\Omega u_e^{n+1} = 0$,
$$\begin{cases} A_m \displaystyle\int_\Omega \frac{C_m}{\Delta t}\Big(\frac{3}{2}V_m^{n+1} - 2V_m^n + \frac{1}{2}V_m^{n-1}\Big)\phi + \int_\Omega \Big[\vec{\vec{\sigma}}_i \cdot \big(\vec{\nabla} V_m^{n+1} + \vec{\nabla} u_e^{n+1}\big)\Big] \cdot \vec{\nabla}\phi \\ \qquad\qquad = A_m \displaystyle\int_\Omega \Big[I_{app} - I_{ion}(\tilde{u}_m^{n+1},\, w^{n+1})\Big]\phi, \\ \displaystyle\int_\Omega \Big[(\vec{\vec{\sigma}}_i + \vec{\vec{\sigma}}_e) \cdot \vec{\nabla} u_e^{n+1}\Big] \cdot \vec{\nabla}\psi + \int_\Omega \Big[\vec{\vec{\sigma}}_i \cdot \vec{\nabla} V_m^{n+1}\Big] \cdot \vec{\nabla}\psi = 0. \end{cases}$$

2.3.2 Heart conductivity

Cardiac muscle has a fiber architecture. The electrical conductivity is higher along than across the fiber direction. As previously discussed, the anisotropy of the intracellular and

extracellular media is included in our model through the conductivity tensors

$$\vec{\vec{\sigma}}_{i,e} = \sigma_{i,e}^{t}\,\vec{\vec{g}} + (\sigma_{i,e}^{l} - \sigma_{i,e}^{t})\,\vec{\tau}\otimes\vec{\tau},$$

where $\vec{\vec{g}}$ denotes the metric tensor – the components of which are given by the identity matrix in an orthonormal coordinate system – the vector $\vec{\tau}$ is of unit length and parallel to the local fiber direction, and σ_i^l and σ_i^t are the conductivity coefficients in the intra-cellular medium measured along and across – respectively – the fiber direction, and likewise for σ_e^l and σ_e^t in the extra-cellular medium.

2.3.3 External stimulus

The cells need an activation to become depolarized. Physiologically, the activation wave begins at the sinus node in the right atrium and the signal then propagates throughout the heart. In fact, the sinus node is the heart's natural pacemaker. The sinus node generates an electrical impulse during a small period of time denoted by t_{stim} at each cardiac cycle. Then for each simulation (with a realistic mesh or not) the applied current I_{app} is a space - time dependent function of the following form

$$I_{app}(\vec{x},t) = \begin{cases} I_{app}^{0} & \forall \vec{x}\in\Omega_{stim},\ \forall t\in[0,t_{stim}], \\ 0 & \text{otherwise.} \end{cases}$$

where Ω_{stim} corresponds to the stimulation area.

2.3.4 Simulations of the ventricles in realistic case

We present a simulation of the bidomain model in a ventricle realistic mesh in a healthy case. This simulation is very much inspired by what was done in [5] that can be considered as the starting point of this thesis from the numerical viewpoint. There were some differences between the simulation of this section and the simulations of [5]. In fact, in collaboration with Elisa Schenone when implementing the bidomain model in the finite element library *FELiScE*, we devised a more realistic ventricle mesh and we adapted the parameters of the ionic model and the applied stimulus for this new geometry. In Chapter 6, a coupled model of the whole heart with a classical 3D model for the ventricles and an asymptotic surface-based – presented in Chapter 3 – for the very fine atria is presented in order to obtain full simulated electrocardiograms. In Chapter 6, we use a more realistic phenomenological model called the *Minimal model for human Ventricular action potentials* (MV) for the ventricles introduced by [8].

Realistic mesh and fibers position of the ventricles We start with the Zygote[2] heart model, a geometric model derived from anatomical data. We used the `3-matic` software to obtain a computationally-correct surface mesh and the `Yams` meshing software to refine the surface mesh, see [16]. Then, we meshed the volume of the two ventricles using `Ghs3d` and

[2] `www.3dscience.com`

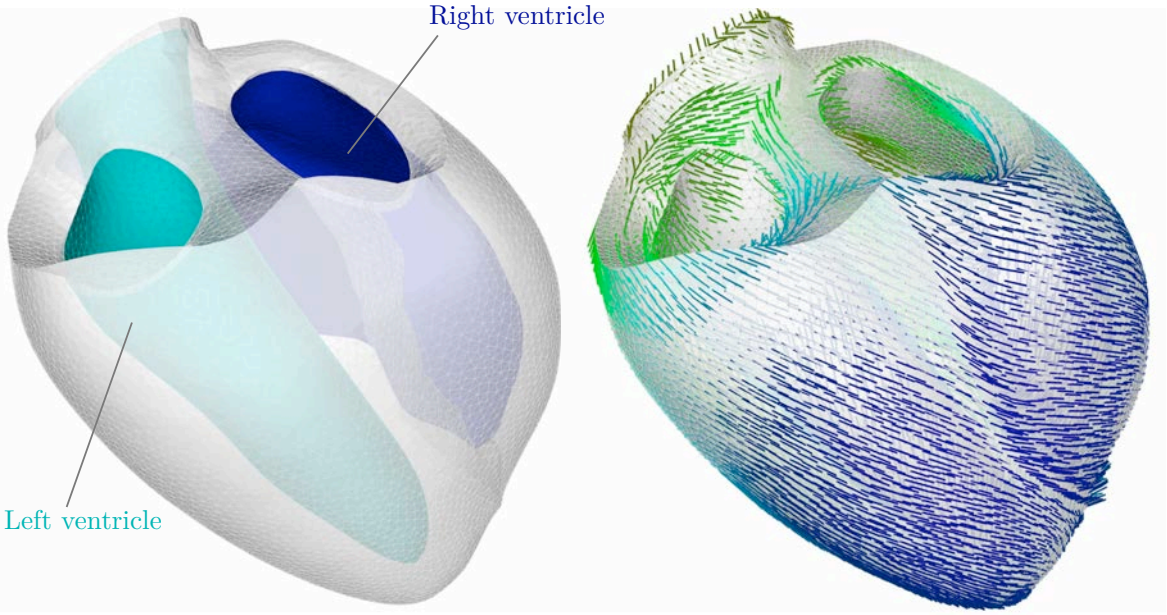

Figure 2.7: Mesh (left) and fibers directions (right) of the ventricles

σ_e^t	σ_e^l	σ_i^t	σ_i^l
$6.0\,10^{-4}$	$2.0\,10^{-3}$	$2.0\,10^{-4}$	$2.0\,10^{-3}$

Table 2.1: Conductivity parameters (all in S.cm^{-1})

`Gmsh`. In Chapter 6 we compare the volume, the mass and the valves diameters of the left ventricle of this mesh with literature reports of normal human heart dimension indicators, and the anatomy corresponds to ventricular end-systole, namely, when the ventricle chamber is smallest. This configuration corresponds to the configuration in the heart cycle that has the smallest amount of internal stresses. That is why this configuration is used in some mechanical modeling studies [20]. We can see in Figure 2.7 (left) different views of the whole heart mesh, which contains about 230,600 tetrahedrons. We recall that the intra and extra -cellular tensors are defined by

$$\vec{\vec{\sigma}}_{i,e} = \sigma_{i,e}^t \, \vec{\vec{g}} + (\sigma_{i,e}^l - \sigma_{i,e}^t) \, \vec{\tau} \otimes \vec{\tau},$$

in order to take into account the privileged direction of the fiber $\vec{\tau}$ in the diffusion. We use [43, 30] to prescribe the fibers in the ventricles, see Figure 2.7 (right). Table 2.1 gives the value of the tensor parameters.

τ_{in}	τ_{out}	τ_{open}	$\tau_{close}^{\text{endo}}$	$\tau_{close}^{\text{Mcell}}$	$\tau_{close}^{\text{epi}}$	τ_{close}^{RV}	V_{min}	V_{max}	v_{gate}
($\text{cm}^2.\text{mA}^{-1}$)				(ms)				(mV)	
4.0	90.0	300.0	120.0	100.0	80.0	90.0	-80.0	20.	-67.0

Table 2.2: Mitchell and Schaeffer parameters

Ionic model and parameters of the bidomain model We use the Mitchell-Schaeffer model (2.37), (2.38) presented in [26] with a normalization of the model

$$\begin{cases} I_{ion}^{tot} = I_{app} + \dfrac{w}{\tau_{in}}(V_m - V_{min})^2(V_m - V_{max}) - \dfrac{V_m - V_{min}}{\tau_{out}(V_{max} - V_{min})}, \\ \\ \dfrac{dw}{dt} + g(V_m, w) = 0, \end{cases} \tag{2.68}$$

with

$$g(V_m, w) = \begin{cases} \dfrac{w}{\tau_{open}} - \dfrac{1}{\tau_{open}(V_{max} - V_{min})} & \text{if } V_m \leq v_{gate}, \\ \\ \dfrac{w}{\tau_{close}} & \text{if } V_m > v_{gate} \end{cases}$$

and with τ_{open}, τ_{close}, τ_{in}, τ_{out} and v_{gate}, positive constants. The value of V_{min} corresponds to the value of the resting potential, *i.e.* when the cell is polarized, and the value of V_{max} corresponds to the maximal potential, *i.e.* when the cell just finishes to become depolarized. Table 2.2 gives the parameter values that we have used. Different values of τ_{close} are given in this table because we consider an heterogeneous tissue. As previously shown in Figure 2.6, this parameter corresponds to the length of the plateau phase which varies through the thickness of the ventricles, see [47]. Here, we consider three layers in the left ventricle τ_{close}^{endo} near the endocardium, τ_{close}^{Mcells} in the mid-myocardium (M-cells) and τ_{close}^{epi} near the epicardium. For the sake of simplicity, we take a constant value τ_{close}^{RV} in the right ventricle. Other ionic models are adapted for the ventricles as for example the *Minimal model for human Ventricular action potentials* [8] used in Chapter 6.

The values of the membrane parameters are $A_m = 200.0\,\text{cm}^{-1}$ and $C_m = 10^{-3}$ mF.cm^{-2}.

External stimulus As previously explained and in Chapter 1, the activation wave begins at the sinus node in the right atrium and it propagates to the ventricles through the atrioventricular node and reach the His-bundle, which activates the septum and the Purkinje network to stimulate all the endocardium. Our model does not include the atria and in order to model the electric activation wave in an appropriate and simple way, we stimulate a part of the endocardium. We apply a given volume current density to a thin subendocardial layer of the ventricles during a small period of time t_{stim}. See [5] for more details about the external ventricular stimulus.

Simulations The results are presented in Figures 2.8 and 2.10. Figure 2.8 displays the repolarization and depolarization of the ventricles during a cardiac cycle. The colors represent

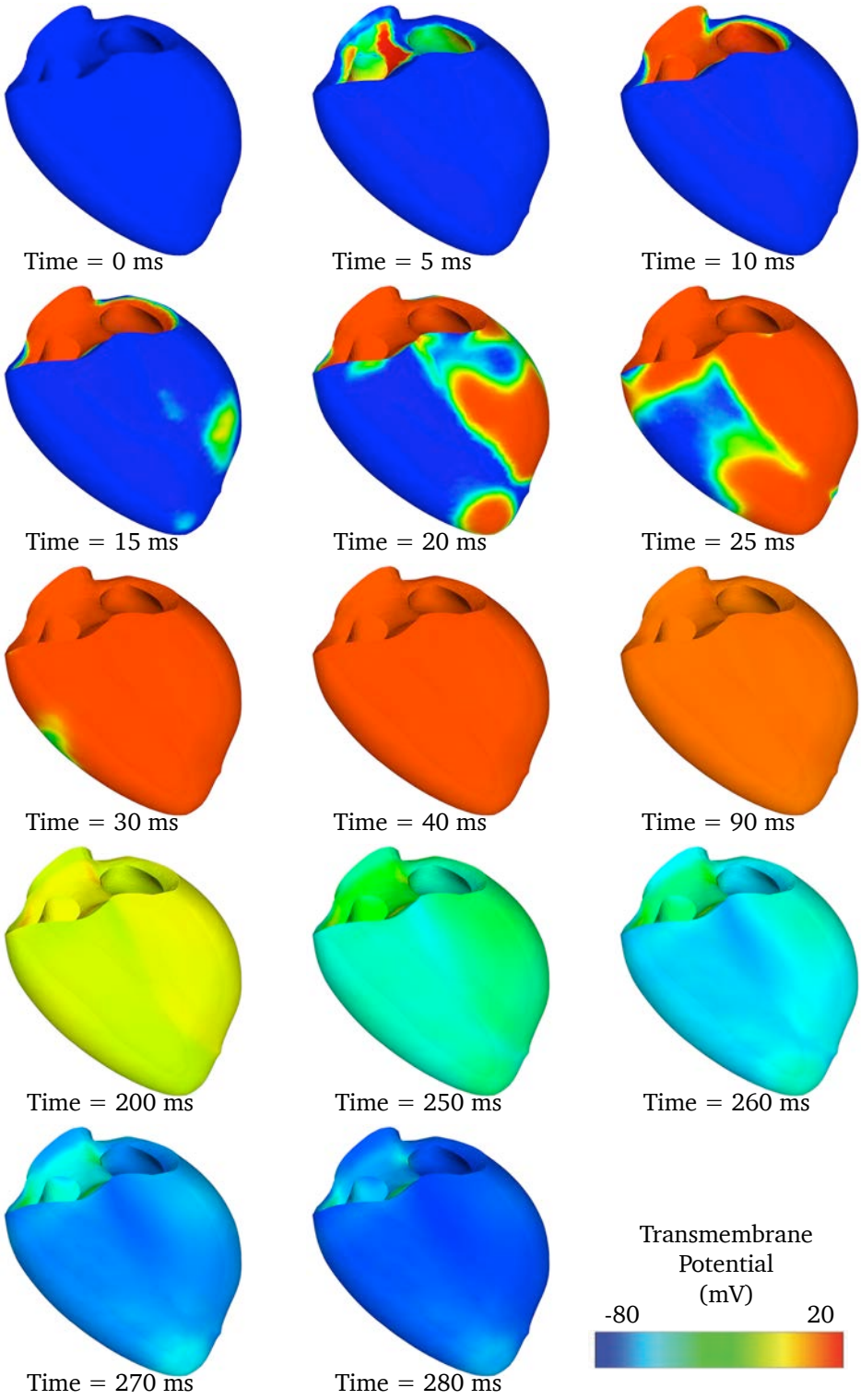

Figure 2.8: Depolarization and repolarization of the ventricles during a cardiac cycle.

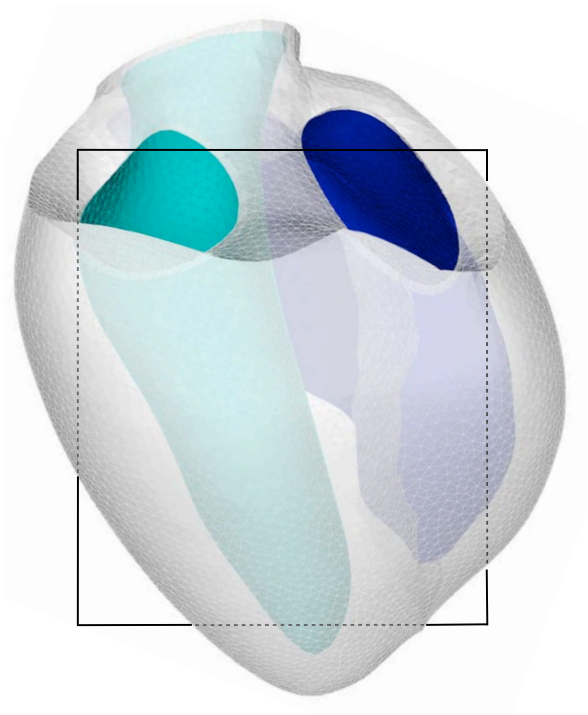

Figure 2.9: Vertical section of the heart using for Figure 2.10.

the transmembrane potential over time. At the beginning, *i.e.* at $t = 0$ms the value of the transmembrane potential is -80mV in all the ventricles. This value corresponds to the resting potential *i.e.* when the ventricles are polarized. The period $t = 0$ms to 35ms corresponds to the depolarizing phase. The ventricles become depolarized starting with the endocardium and finishing at the epicardium. Thereafter we can see that the whole heart remains depolarized for some time, and at $t = 200$ms, we see that the repolarization of the heart has started. In Figure 2.8, we only display a ventricular section of this simulation. Figure 2.9 shows the selected section. This ventricular section allows to observe more details about the depolarization and the repolarization of the ventricles. We can see that the endocardium is depolarized first. Then the signal propagates through the thickness and reaches the pericardium. Figure 2.8 also allows to understand why the ventricles are considered as a heterogeneous tissue. As the plateau phase varies across the left ventricle wall, the repolarization is not in the same direction than the depolarization.

2.3.5 Simulations in moving domain

In this section, we present some simulations with the mechanical bidomain model studied in Section 2.2. We simplify the coupled problem presented in (2.64) by neglecting the new source terms. We suppose that we have a mesh deformation over time and we use it to find

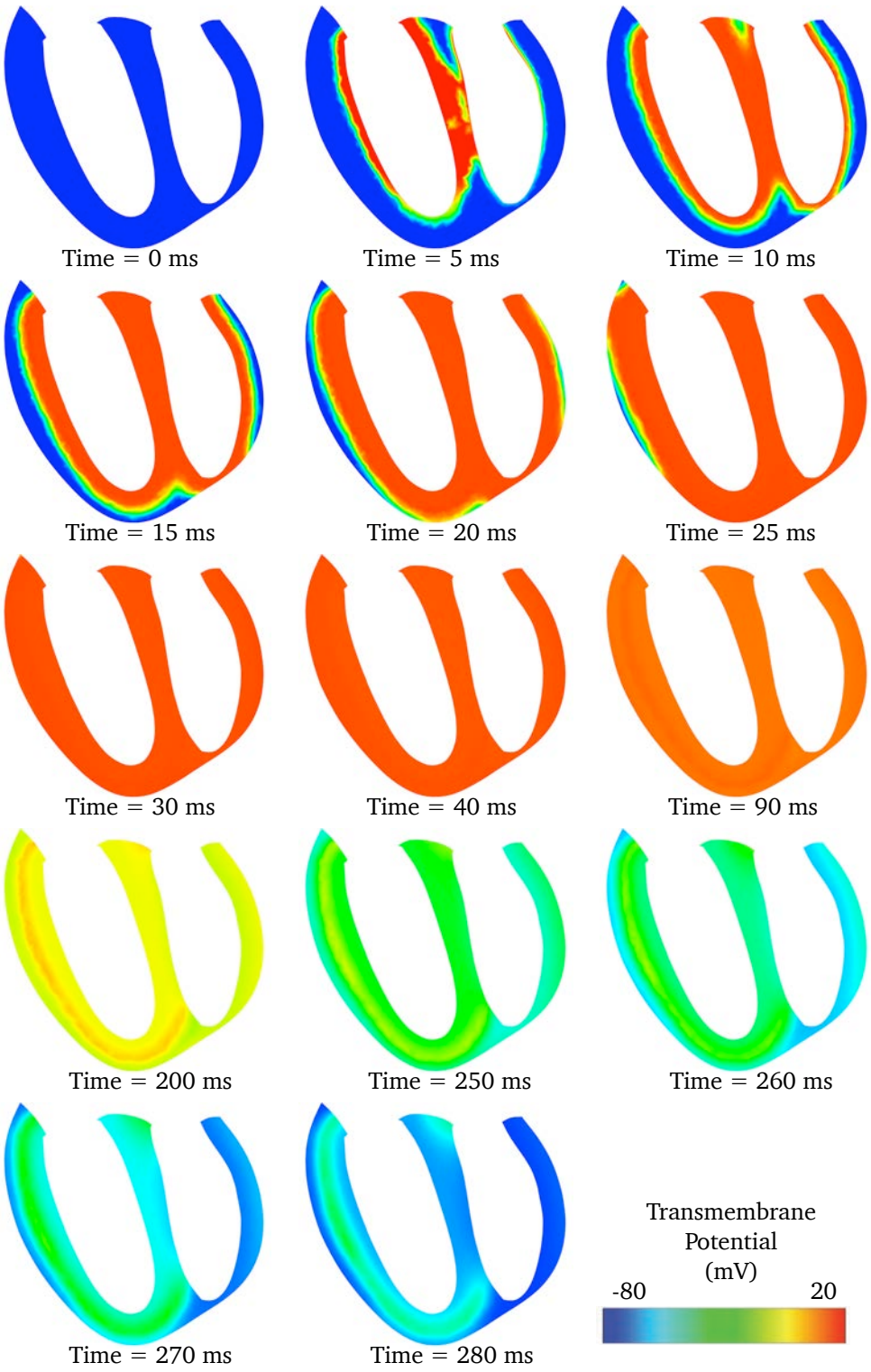

Figure 2.10: Depolarization and repolarization of a ventricular section (see Figure 2.9) during a cardiac cycle.

for all $t > 0$, $u_e(\cdot, t)$, $V_m(\cdot, t) \in H^1(\Omega(t))$ with $\int_{\Omega(t)} u_e = 0$ such that

$$\left\{ \begin{array}{l} \displaystyle\int_{\Omega(t)} \Big(\vec{\vec{\sigma}}_i \cdot \big(\vec{\nabla}_{\vec{x}} V_m + \vec{\nabla}_{\vec{x}} u_e\big)\Big) \cdot \vec{\nabla}_{\vec{x}} \phi \, d\Omega + \int_{\Omega(t)} A_m C_m \Big(\frac{\partial V_m}{\partial t}\Big|_{\vec{\xi}} + I_{ion}^{tot]}\Big)\phi = 0 \, d\Omega, \\ \displaystyle\int_{\Omega(t)} \Big(\vec{\vec{\sigma}}_i \cdot \vec{\nabla}_{\vec{x}} V_m + (\vec{\vec{\sigma}}_e + \vec{\vec{\sigma}}_i) \cdot \vec{\nabla}_{\vec{x}} u_e\Big) \cdot \vec{\nabla}_{\vec{x}} \psi \, d\Omega = 0, \end{array} \right. \tag{2.69}$$

for all ϕ, $\psi \in H^1(\Omega(t))$ with $\int_{\Omega(t)} \psi = 0$. The condition $\int_{\Omega(t)} u_e = 0$ is necessary to have the uniqueness of the solution. We consider here the heart as an incompressible material – at least a quasi-incompressible material see [48] for more details – which justifies that we do not take into account the new source terms obtained in Section 2.2.5.1.

In a first stage, we present a test case in order to verify the finite library *FELiScE* with a moving domain.

2.3.5.1 Test case in heat equation

The deformable structure occupies at time t the space domain $\Omega(t)$ with boundary $\partial\Omega(t)$. We denote by $(\Omega_0, \partial\Omega_0)$ the reference configuration, equal here to $(\Omega(0), \partial\Omega(0))$. The domain Ω_0 is defined by $[-0.5, 0.5] \times [-0.5, 0.5] \times [0, 1]$ and the deformation is defined by $\vec{\phi}(\vec{\xi}, t) = \Big(\frac{1}{k(t)}\xi_1, \frac{1}{k(t)}\xi_2, \xi_3\Big)$, with $k(t)$ a time dependent function. It corresponds to a scaling in the directions x and y. We define

$$\vec{\vec{F}}(\vec{\xi}, t) = \vec{\nabla}_{\vec{\xi}} \vec{\phi}.$$

We study the problem, find $u \in H_0^1(\Omega(t))$,

$$\int_{\Omega(t)} \partial_t u \, v + \int_{\Omega(t)} \vec{\nabla}_{\vec{x}} u \cdot \vec{\nabla}_{\vec{x}} v = \int_{\Omega(t)} f v, \, \forall v \in H_0^1(\Omega(t)). \tag{2.70}$$

The Lagrangian representation of (2.70) is, find $\tilde{u} \in H_0^1(\Omega_0)$

$$\int_{\Omega_0} \partial_t \tilde{u} \, \tilde{v} \, J + \int_{\Omega_0} \vec{\vec{F}}^{-1} \vec{\nabla}_{\vec{\xi}} \tilde{u} \cdot \vec{\vec{F}}^{-1} \vec{\nabla}_{\vec{\xi}} \tilde{v} \, J = \int_{\Omega_0} \tilde{f} \tilde{v} \, J, \, \forall \tilde{v} \in H_0^1(\Omega_0), \tag{2.71}$$

with $\tilde{u} = u \circ \vec{\phi}$ and $\tilde{f} = f \circ \vec{\phi}$.

We want to solve the following problem

$$\left\{ \begin{array}{cc} \partial_t \tilde{u} - (\vec{\vec{F}}^{-1} \vec{\nabla}_{\vec{\xi}} \tilde{u})^2 = \tilde{f}, & \Omega_0 \\ \tilde{u} = 0, & \partial\Omega_0 \end{array} \right. \tag{2.72}$$

The function $\tilde{u}(\vec{\xi}, t) = (\xi_1 - 0.5)(\xi_1 + 0.5)(\xi_2 - 0.5)(\xi_2 + 0.5)\xi_3(\xi_3 - 1.0)t$ is solution of

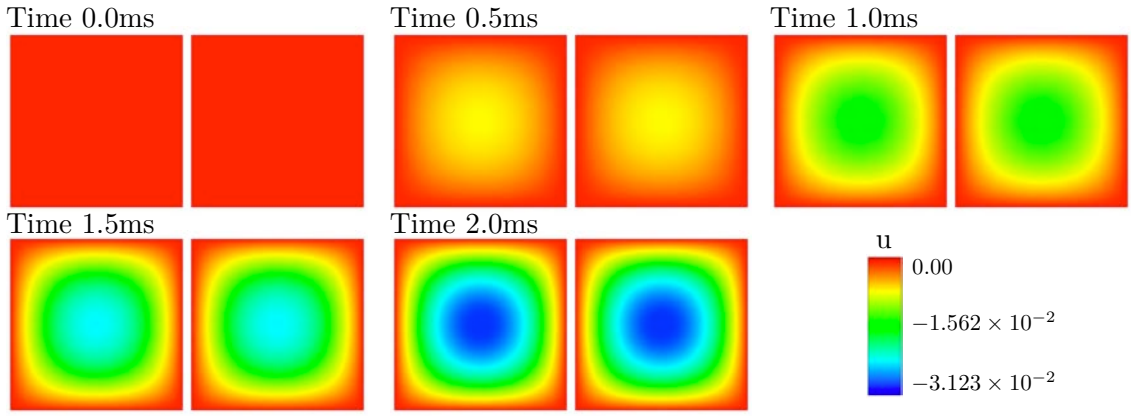

Figure 2.11: Exact solution (left) and Computed solution (right)

Time (ms)	0	0.5	1.0	1.5	2.0
Norm l^∞	0	1.359×10^{-5}	2.742×10^{-5}	4.146×10^{-5}	5.571×10^{-5}

Table 2.3: Error norms

(2.72) with

$$\begin{aligned}\tilde{f}(\vec{\xi},t) = (\xi_1-0.5)(\xi_1+0.5)(\xi_2-0.5)(\xi_2+0.5)\xi_3(\xi_3-1.0)\\ -2k^2(t)(\xi_2-0.5)(\xi_2+0.5)\xi_3(\xi_3-1.0)t\\ -2k^2(t)(\xi_1-0.5)(\xi_1+0.5)\xi_3(\xi_3-1.0)t\\ -2(\xi_1-0.5)(\xi_1+0.5)(\xi_2-0.5)(\xi_2+0.5)t.\end{aligned}$$

We compute the solution of (2.70) using *FELiScE* with

$$\bar{\Omega}(t) = [-0.5k(t), 0.5k(t)] \times [-0.5k(t), 0.5k(t)] \times [0,1],$$

$$\begin{aligned}f(\vec{x},t) = (k(t)x-0.5)(k(t)x+0.5)(k(t)y-0.5)(k(t)y+0.5)z(z-1.0)\\ -2k^2(t)(k(t)y-0.5)(k(t)y+0.5)z(z-1.0)t\\ -2k^2(t)(k(t)x-0.5)(k(t)x+0.5)z(z-1.0)t\\ -2(k(t)x-0.5)(k(t)x+0.5)(k(t)y-0.5)(k(t)y+0.5)t.\end{aligned}$$

and $k(t) = \frac{1.0}{1.0+0.1t}$.

In Figure 2.11, we compare the exact solution $\tilde{u}$ and the solution obtained with *FELiScE*. We observe a good agreement. We also report in Table 2.3 the l^∞ difference between the true and the computed solutions. These results allow to validate the test case in a deformable domain.

τ_{in} ($cm^2.mA^{-1}$)	τ_{out} ($cm^2.mA^{-1}$)	τ_{open} (ms)	τ_{close} (ms)	v_{gate} (mV)	V_{min} (mV)	V_{max} (mV)
4.0	90.0	500.0	100.0	-67.0	-80.0	20.0

Table 2.4: Parameters of the Mitchell-Schaeffer model

A_m (cm^{-1})	C_m ($mF.cm^{-2}$)	σ_i ($S.cm^{-1}$)	σ_e ($S.cm^{-1}$)
200	10^{-3}	2.5×10^{-4}	9.0×10^{-4}

Table 2.5: Parameters of the bidomain model

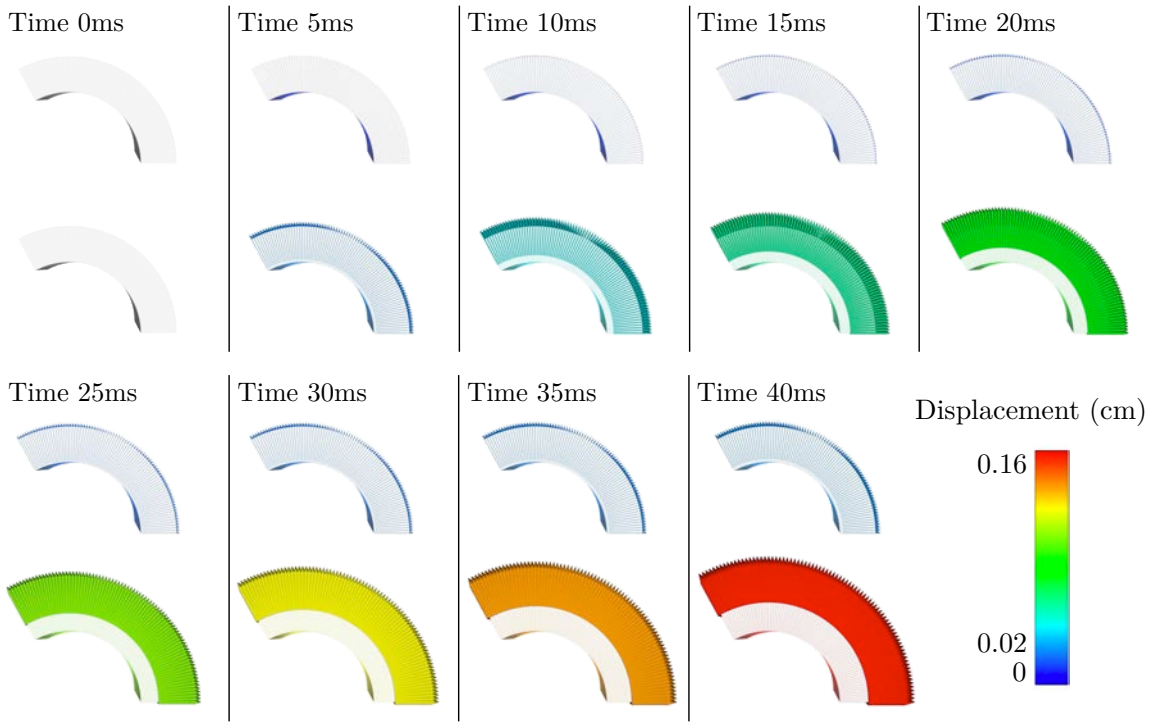

Figure 2.12: Displacements of the small (top) and large (bottom) deformations

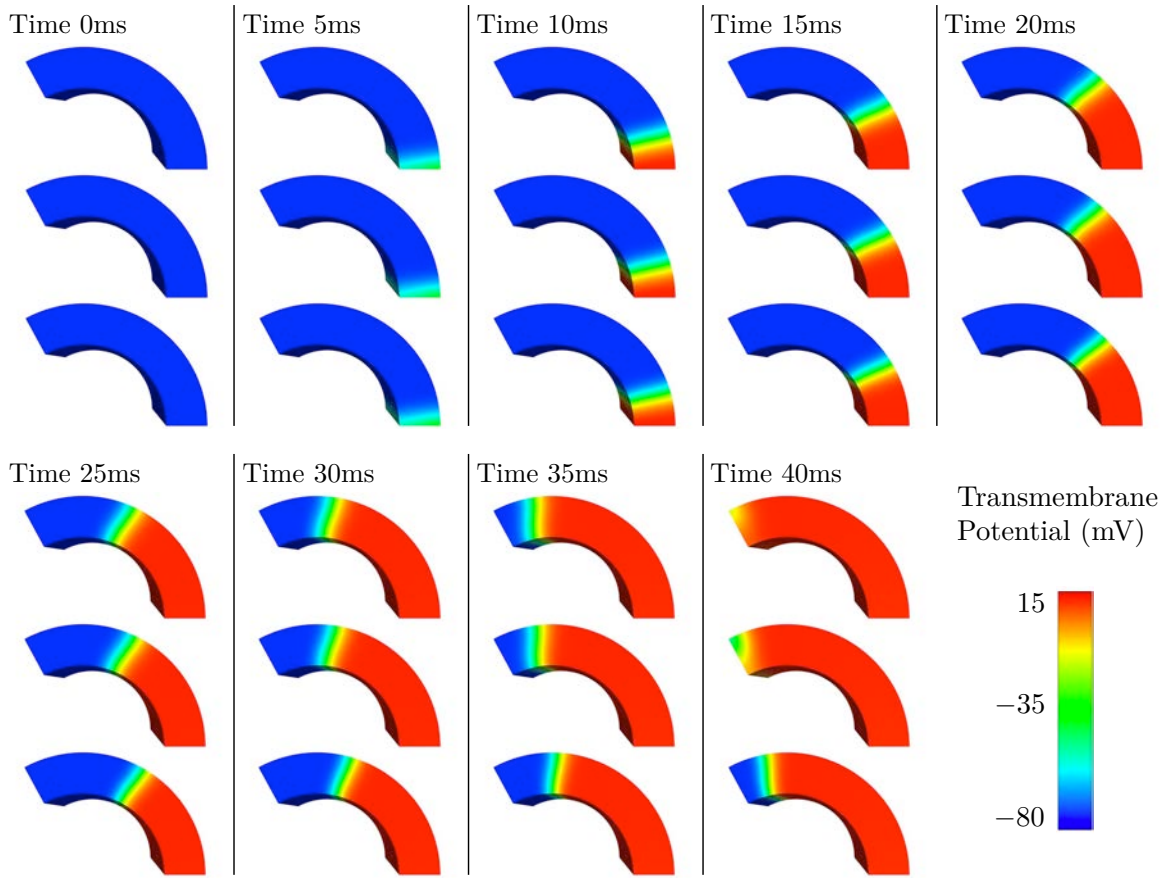

Figure 2.13: Results with the fixed domain (top), with the small deformation (middle) and with the large deformation (bottom)

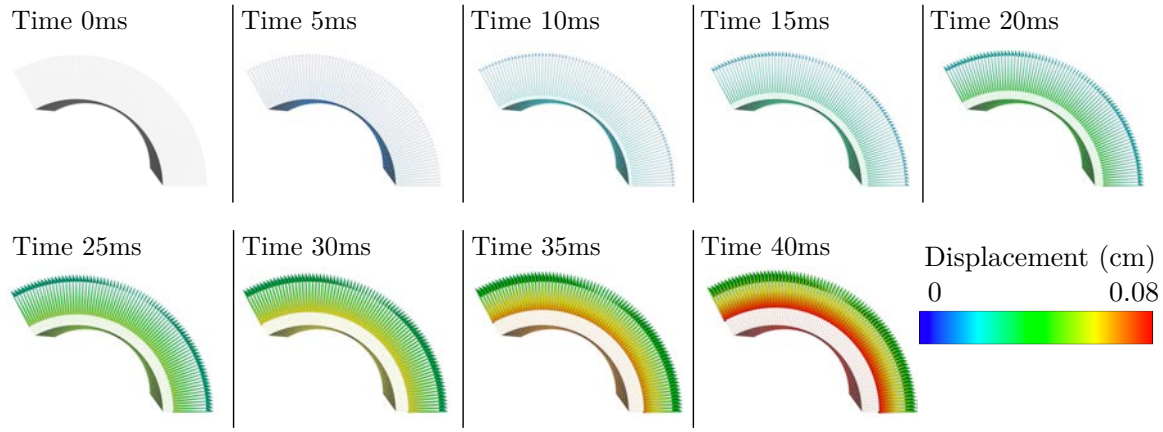

Figure 2.14: Displacements of the space-dependent deformation

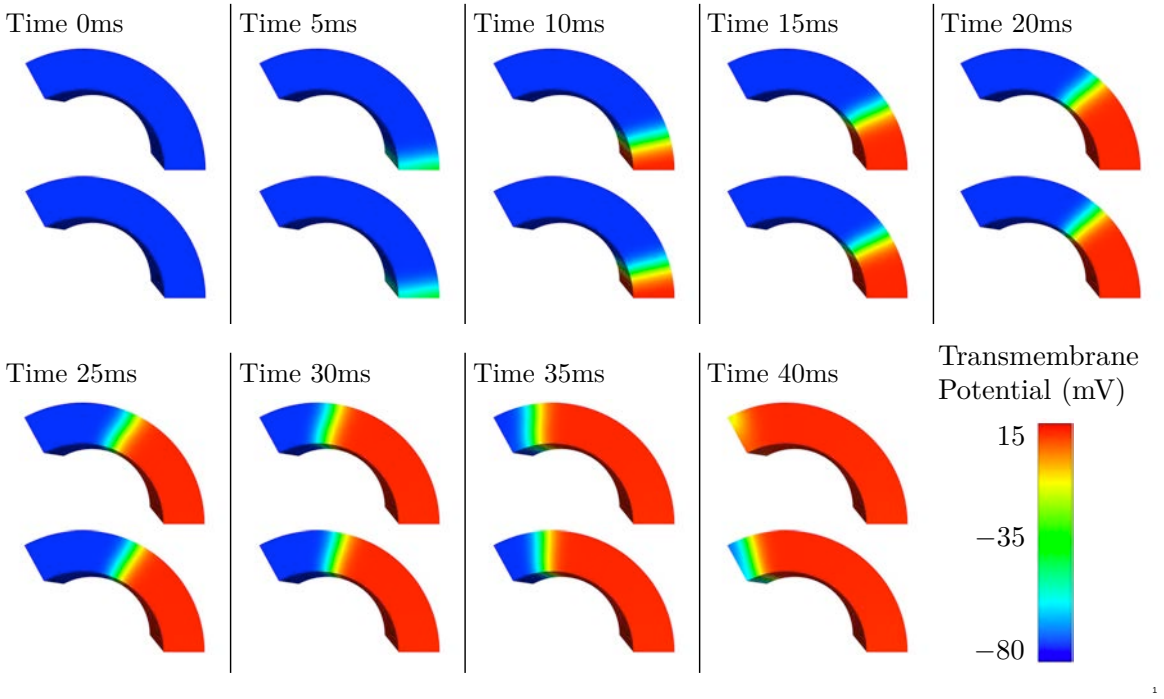

Figure 2.15: Results with the fixed domain (top) and with the deformation (bottom)

2.3.5.2 Simulations of the bidomain model in moving domain

As previously explained, we want to compute the following problem. Find for all $t > 0$, $u_e(\cdot, t)$, $V_m(\cdot, t) \in H^1(\Omega(t))$ with $\int_{\Omega(t)} u_e = 0$ such that

$$\begin{cases} \displaystyle\int_{\Omega(t)} \Big(\vec{\vec{\sigma}}_i \cdot \big(\vec{\nabla}_{\vec{x}}\, V_m + \vec{\nabla}_{\vec{x}}\, u_e\big)\Big) \cdot \vec{\nabla}_{\vec{x}}\, \phi \, d\Omega + \int_{\Omega(t)} A_m C_m \Big(\frac{\partial V_m}{\partial t}\Big|_{\vec{\xi}} + I_{ion}^{tot}\Big)\phi = 0 \, d\Omega, \\ \displaystyle\int_{\Omega(t)} \Big(\vec{\vec{\sigma}}_i \cdot \vec{\nabla}_{\vec{x}}\, V_m + (\vec{\vec{\sigma}}_e + \vec{\vec{\sigma}}_i) \cdot \vec{\nabla}_{\vec{x}}\, u_e\Big) \cdot \vec{\nabla}_{\vec{x}}\, \psi \, d\Omega = 0, \end{cases} \tag{2.73}$$

for all ϕ, $\psi \in H^1(\Omega(t))$ with $\int_{\Omega(t)} \psi = 0$.

We use the space and time discretization presented in Section 2.3.1. The displacement of the mesh is prescribed. We also use the Mitchell-Schaeffer model. The parameters of the Mitchell-Schaeffer model are listed in Table 2.4. Table 2.5 gives the other parameters of the bidomain model. The reference domain – corresponding here to $\Omega(0)$ – is a 3D third-cylinder, with internal radius $r_{min} = 4$, external radius $r_{max} = 6$ and height $h = 0.5$ (all dimensions are given in cm). In the following two cases (simple and more complex), we suppose that we only have an homogeneous diffusion.

First simulation - simple case To start with, we present two different deformations given by Figure 2.12. In these first two cases, the meshes move following the exterior normal direction. The associated displacements are linear in time and the norm of the displacements is constant in space. The second deformation is larger than the first deformation. In this

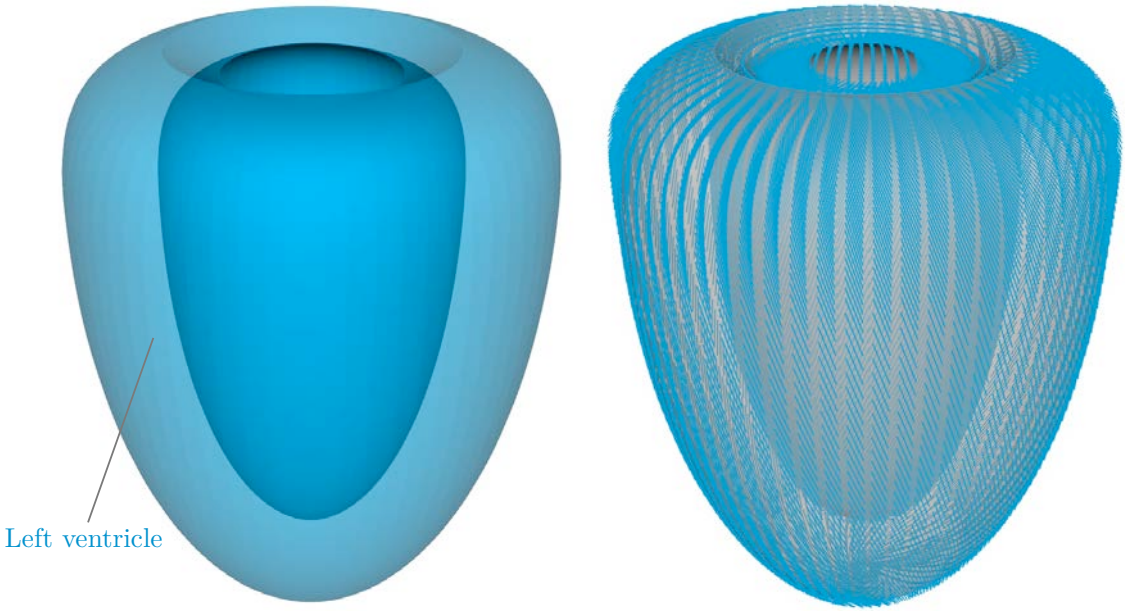

Figure 2.16: Mesh (left) and fibers directions (right) of the left ventricle

case, we do not take into account the fibers.
The results are presented in Figure 2.13. The top corresponds to the solution without deformation. The middle corresponds to the solution with the small deformation and we can remark that the electrical signal is slowed. The solution with the large deformation is given at the bottom and we can see that the deceleration of the signal is more important. The two deformations increase the length of the domain and increase the distance that the wave must cross.

Second simulation - more complex case During the second deformation, the mesh always moves following the exterior normal direction but the norm of the displacement is not constant in space as we can see in Figure 2.14. The external surface of the mesh moves less than the internal surface.
The results are presented in Figure 2.15. The solution with the classical bidomain model is given at the top and the bottom corresponds to the mechanical bidomain model with the naive heart deformation. The displacement near the internal surface is larger and the wave near this surface decelerates more than the wave next to the external surface.

Simple ventricle deformation In this paragraph, we present a more realistic case. We use a simplified one-ventricle mesh, see Figure 2.16, in which the left ventricle is considered as symmetric with respect to two orthogonal planes. The anatomy corresponds to ventricular end-systole, namely, when the ventricle chamber is smallest and is the reference configuration for the mechanical modeling. The displacements of this left ventricle are the results of a bio-mechanical simulation which has been performed by colleagues of the Inria team

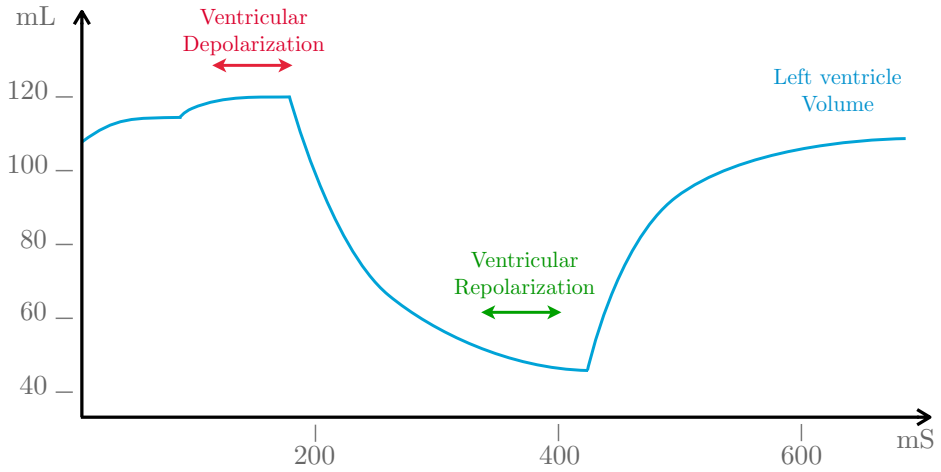

Figure 2.17: Volume of the left ventricle during a cardiac cycle

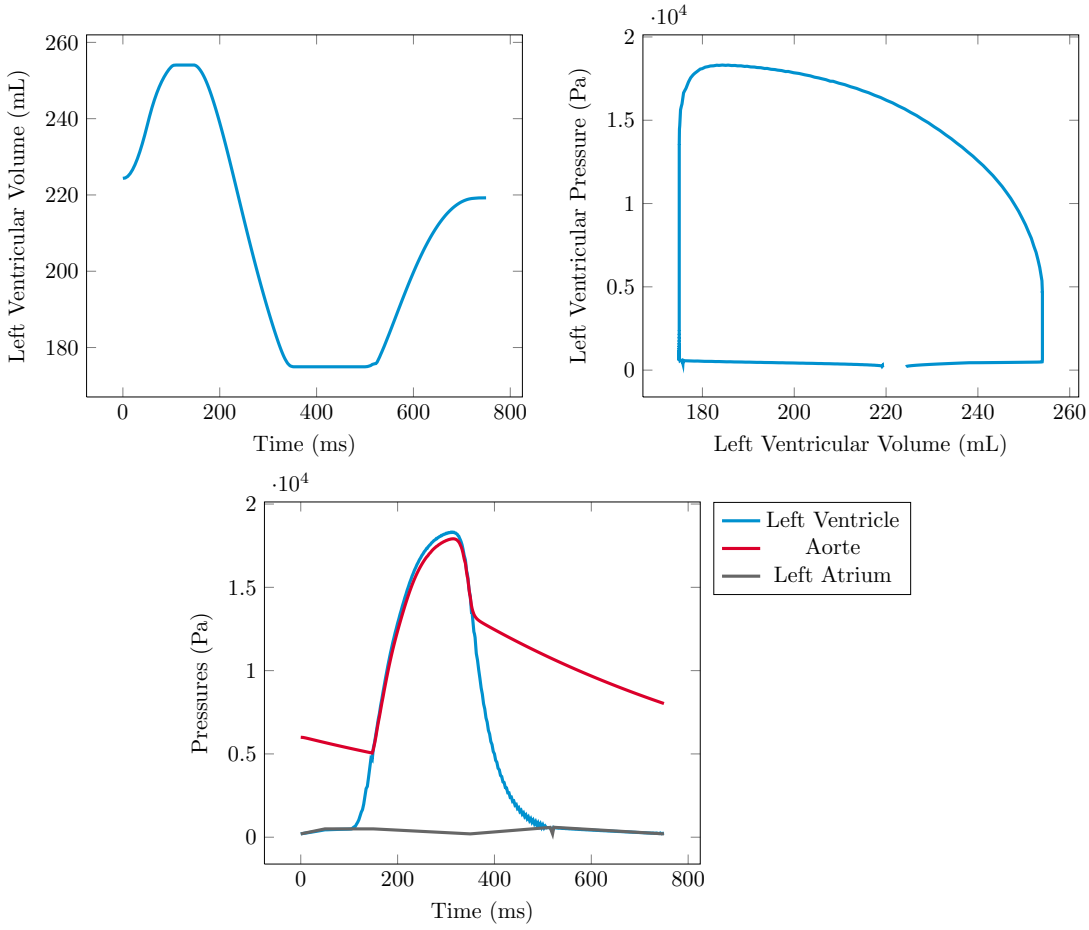

Figure 2.18: Volume of the left cavity (top-left), PV loop of the left cavity (top-right) and Pressures of the left cavity (bottom)

σ_e^t	σ_e^l	σ_i^t	σ_i^l
$4.0\,10^{-3}$	$1.2\,10^{-2}$	$1.2\,10^{-3}$	$1.2\,10^{-2}$

Table 2.6: Conductivity parameters (all in S.cm^{-1}) for the left ventricle case

MΞDISIM. The bio-mechanical model is presented for example in [9]. In this model, the electrical stimulation is an input which is represented by a very simplified action potential (close to a crenel function). The electrical signal starts at the bottom of the left ventricle – named the apex – and the wave rises along the ventricle at constant speed. We adapt our electrical signal at this activation pattern.

Figure 2.17 recalls the time evolution of the left ventricle volume during a cardiac cycle. The depolarization occurs when the ventricle chamber is largest and the repolarization when the chamber is smallest. Figure 2.18 gives the evolution of various indicators of the left ventricle during one cardiac cycle obtained by the bio-mechanical simulation. We see that the volume of the left ventricle is higher than in a normal case. Indeed, the left ventricle mesh is a bit large compared to a realistic case.

As previously, we present two simulations one in fixed domain and one in moving domain. The mesh used for the simulation in fixed domain corresponds to the mechanical reference configuration. The conductivity parameters used for the bidomain model are given in Table 2.6. The other parameters are the same as in Section 2.3.4 expect the value of τ_{close} which is constant and equals 90ms. Due to the large size of the left ventricle mesh, the conductivity parameters are high compared to the realistic case presented in Section 2.3.4. The results of a cardiac cycle are presented in Figure 2.19. At each time, the left corresponds to the case in fixed domain and the right the results in moving domain. We apply the deformation of the ventricles using the displacements. Boundary conditions are prescribed in the apex. The left ventricle is depolarized (between 80 and 160ms) when the left atrium ejects the blood in the ventricles, see Figure 1.5 in Chapter 1. This implies that the ventricular chamber is greater and as can be seen in the figure, the duration of depolarization is then longer. During the repolarization, the volume of the ventricles is smaller and the delay of the signal is slightly caught. We also remark that the signal does not remains horizontal, see for example $t = 150$ms and $t = 300$ms. The origin of this phenomenon is that the ventricle contracts following the directions of the fibers, see Figure 2.16 and that the ventricle is larger in one direction. Figure 2.20 allows to better understand these differences between the two simulations. Indeed, the transmembrane potential for three nodes of the mesh is represented. The green curves correspond to the fixed case and the purple to the moving case. For the node near the apex, we do not see differences due to the boundary conditions. Furthermore, this node is one of the first to be depolarized, *i.e.* the delay due to the filling of the ventricle is not yet significant. In the other two nodes, the delay during the depolarization is significant and decreases a little during the repolarization. Indeed during the repolarization, the ventricles are contracted. This more realistic case gives just an example of the impact of the mechanical deformation of the heart on the electrical activation but remains very limited.

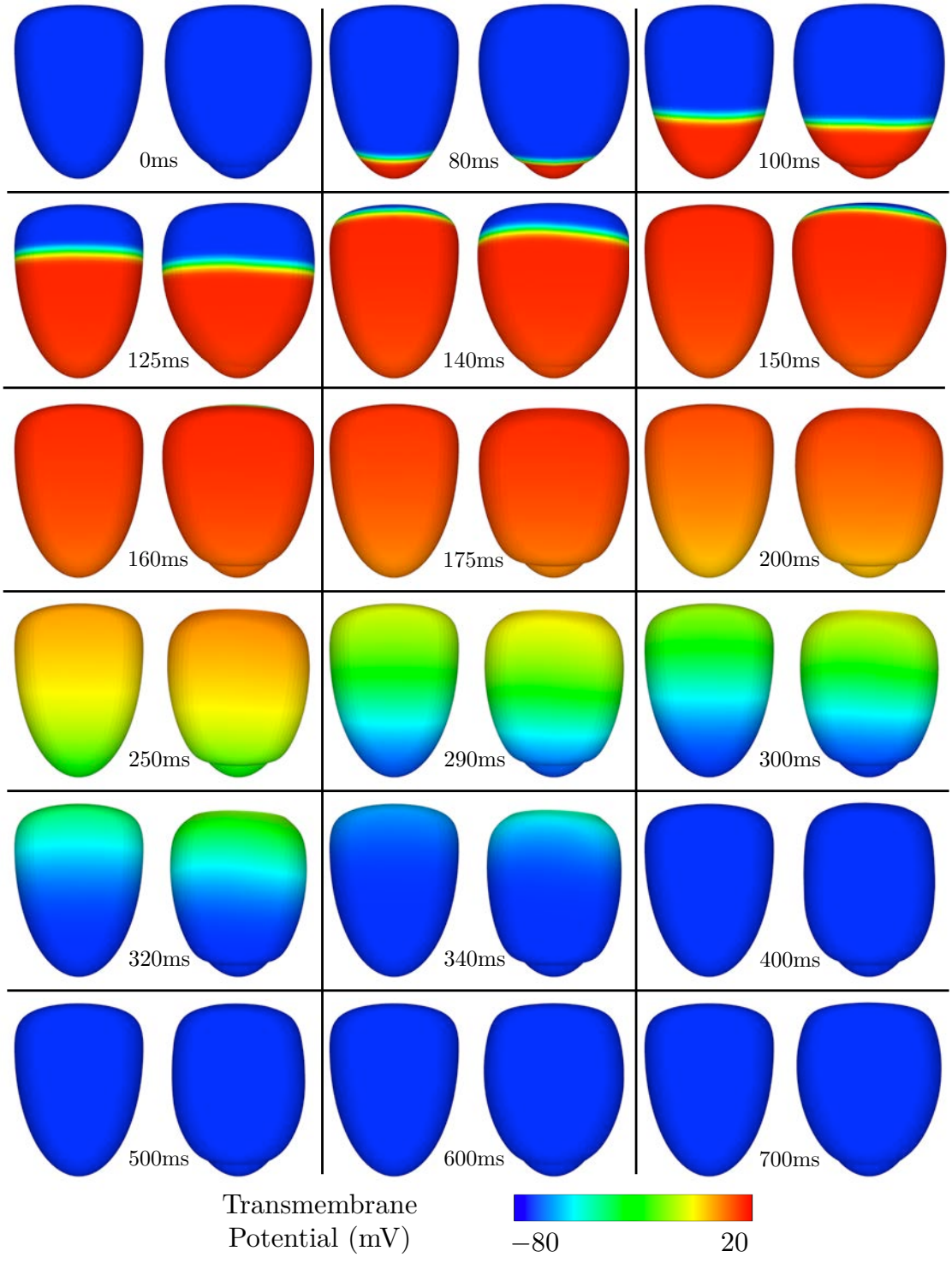

Figure 2.19: Cardiac cycle of the left ventricle (in fixed domain at the left and in moving domain at the right)

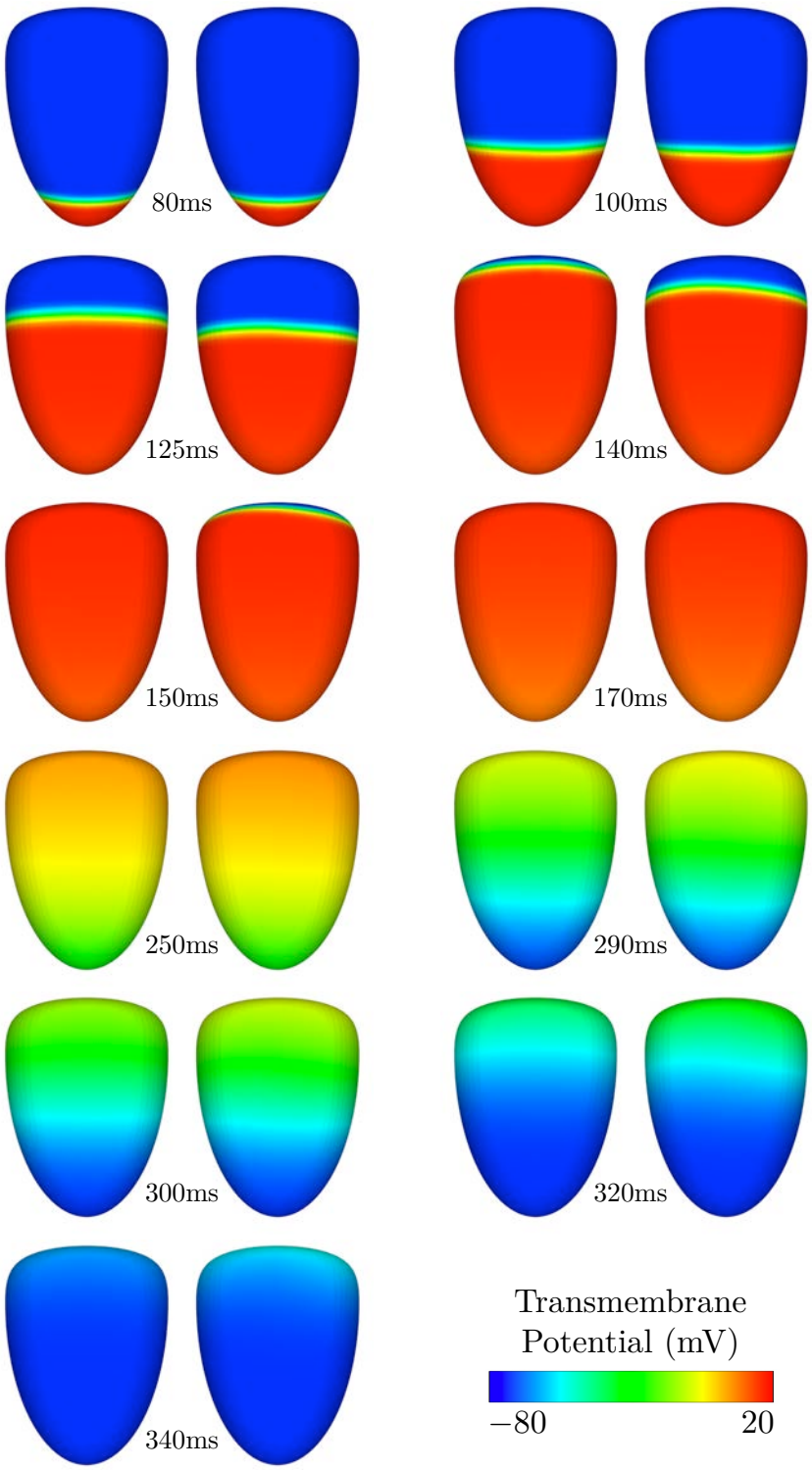

Figure 2.20: Cardiac cycle of the left ventricle (in fixed domain at the left and in moving domain without the mesh deformations at the right)

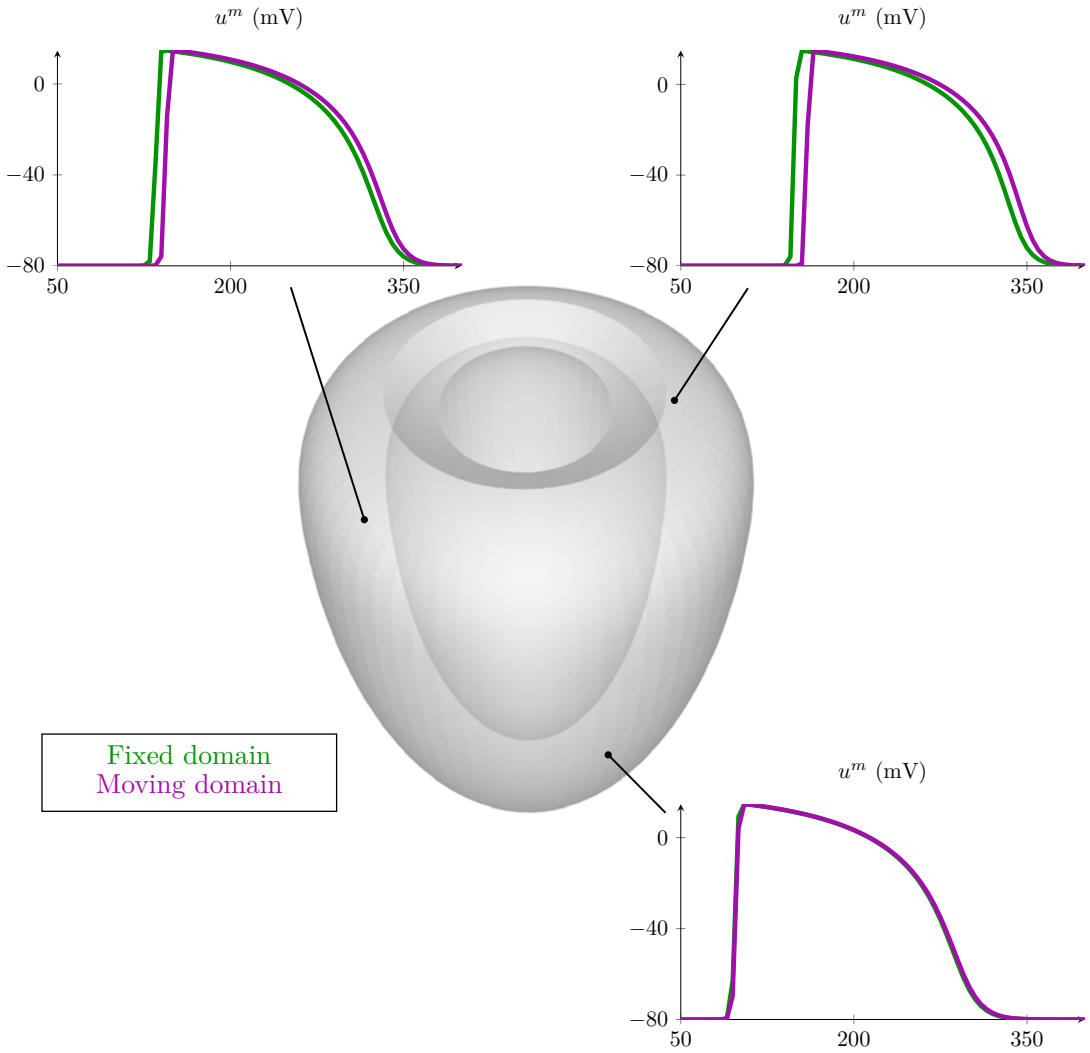

Figure 2.21: Comparison of transmembrane potential V_m in three nodes of the mesh between the fixed case (green) and the moving case (purple)

2.3.5.3 Discussion

We obtain expected results, *i.e.* the effects obtained with a deformation are of the same order than the deformation. The impact in the results seems to be comparable to these obtained with the classical mechano-electrical models, which only consider changes in the constitutive laws. One perspective of this work is to compare our mechanically deformed bidomain model – which takes into account the heart deformation in conservation laws – with the classical mechano-electrical model in a real case. We can also consider the new source terms – see Section 2.2.5.1 – if we suppose that the heart is not incompressible.

Conclusion

In this chapter, we have presented the bidomain model, a classical electrophysiological model. This description was divided into two main parts. The first part consisted of a complete presentation of the equations of the bidomain model, with the model of cellular activity, the considered electrical transmembrane conditions and the homogenization process. The second part studied the effects of mechanical deformation on the heart electrical signal based on the mixture theory. We concluded this chapter with a numerical part composed of a simulation in realistic case of the ventricles with the classical bidomain model and simulations in moving domain. We have also provided details on the discretization algorithm.

Bibliography

[1] R. Aliev and A. Panfilov. A simple two-variable model of cardiac excitation. *Chaos, Solitons Fractals*, 3(7):293–301, 1996.

[2] G. Allaire. Homogenization and two-scale convergence. *SIAM Journal of Mathematical Analysis*, 23(6):1482–1518, 1992.

[3] M. Bendahmane and H.K. Karlsen. Analysis of a class of degenerate reaction-diffusion systems and the bidomain model of cardiac tissue. *Networks and Heterogeneous Media*, 1:185–218, 2006.

[4] A. Bensoussan, J.-L. Lions, and G. Papanicolaou. *Asymptotic Analysis for Periodic Structures*. Elsevier, 1978.

[5] M. Boulakia, S. Cazeau, M.A. Fernández, J.-F. Gerbeau, and N. Zemzemi. Mathematical modeling of electrocardiograms: a numerical study. *Annals of Biomedical Engineering*, 38(3):1071–1097, 2010.

[6] Y. Bourgault, Y. Coudière, and C. Pierre. A collocation-galerkin finite element model of cardiac action potential propagation. *IEEE Transactions on Biomedical Engineering*, 41(8):743–757, 1994.

[7] Y. Bourgault, Y. Coudière, and C. Pierre. Existence and uniqueness of the solution for the bidomain model used in cardiac electrophysiology. *Nonlinear Analysis-Real World Applications*, 10(1):458–482, 2009.

[8] A. Bueno-Orovio, E. M. Cherry, and F. H. Fenton. Minimal model for human ventricular action potentials in tissue. *Journal of Theretical Biology*, 253:544–560, 2008.

[9] D. Chapelle, P. Le Tallec, P. Moireau, and M. Sorine. Energy-preserving muscle tissue model: formulation and compatible discretizations. *International Journal for Multiscale Computational Engineering*, 10(2):189–211, 2012.

[10] P. Colli Franzone, L.F. Pavarino, and B. Taccardi. Simulating patterns of excitation, repolarization and action potential duration with cardiac Bidomain and Monodomain models. *Mathematical Biosciences*, 197(1):35–66, 2005.

[11] P. Colli Franzone and G. Savaré. Degenerate evolution systems modeling the cardiac electric field at micro- and macroscopic level. In *Evolution equations, semigroups and functional analysis (Milano, 2000)*, volume 50 of *Progr. Nonlinear Differential Equations Appl.*, pages 49–78. Birkhäuser, 2002.

[12] P. Colli Franzone and G. Savaré. Degenerate evolution systems modeling the cardiac electric field at micro and macroscopic level. *Progress in Nonlinear Differential Equations and Their Applications*, 50:49–78, 2002.

[13] M. Courtemanche, R.J. Ramirez, and S. Nattel. Ionic mechanisms underlying human atrial action potential properties: insights from a mathematical model. *American Journal of Physiology*, (275):H301–H321, 1998.

[14] J.F.E. Dickinson, J.G. Limon-Petersen, and Compton R.G. The electroneutrality approximation in electrochemistry. *Journal of Solid State Electrochemistry*, 15(7-8):1335–1345, 2011.

[15] R. Fitzhugh. Impulses and physiological states in theoretical models of nerve membrane. *Biophysical Journal*, 1(6):445–466, 1961.

[16] P. Frey. Yams: A fully automatic adaptive isotropic surface remeshing procedure. Technical report 0252, Inria, Rocquencourt, France, Nov. 2001.

[17] Netter F. H. *The Heart - The Netter collection of medical illustrations*, volume 5. Saunders Elsevier, 1969.

[18] P. E. Hand and C. S. Peskin. Homogenization of an electrophysiological model for a strand of cardiac myocytes with gap-junctional and electric-field coupling. *Bulletin of Mathematical Biology*, 72:1408–1424, 2010.

[19] A.L. Hodgkin and A. F. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *Journal of Physiology*, 117(4):500–544, 1952.

[20] A. Imperiale. *Méthodes d'assimilation de la donnée image pour la personnalisation de modèles mécaniques*. PhD thesis, University Pierre and Marie Curie, Paris, 2013.

[21] J. Keener and J. Sneyd. *Mathematical Physiology*. Springer, 2004.

[22] R.H. Keldermann, M.P. Nash, and A.V. Panfilov. Modeling cardiac mechano-electrical feedback using reaction-diffusion-mechanics systems. *Physica D*, 238:1000–1007, 2008.

[23] P. Kohl and F. Sachs. Mechanoelectric feedback in cardiac cells. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 359(1783):1173–1185, 2001.

[24] C. Liu, J.M. Zachara, A. Felmy, and Y. Gorby. An electrodynamics-based model for ion diffusion in microbial polysaccharides. *Colloids and Surfaces B: Biointerfaces*, 38:55–65, 2004.

[25] C.H. Luo and Y. Rudy. A dynamic model of the cardiac ventricular action potential. I. simulations of ionic currents and concentration changes. *Circulation Research*, 74:1071–1096, 1994.

[26] C.C. Mitchell and D.G. Schaeffer. A two-current model for the dynamics of cardiac membrane. *Bulletin Math. Bio.*, 65:767–793, 2003.

[27] Y. Mori. From three-dimensional electrophysiology to the cable model: an asymptotic study. *arXiv preprint arXiv:0901.3914*, 2009.

[28] Y. Mori, J.W. Jerome, and C.S. Peskin. A three-dimensional model of cellular electrical activity. *Bulletin of the Institute of Mathematics, Academia Sinica*, 2(2):367–390, 2007.

[29] J.S. Nagumo, S. Arimoto, and S. Yoshizawa. An active pulse transmission line stimulating nerve axon. *Proceedings of the IEEE*, 50:2061–2071, 1962.

[30] M.P. Nash and P.J. Hunter. Computational mechanics of the heart, from tissue structure to ventricular function. *Journal of elasticity and the physical science of solids*, 61:113–141, 2000.

[31] M.P. Nash and A.V. Panfilov. Electromechanical model of excitable tissue to study reentrant cardiac arrhythmias. *Progress in Biophysics and Molecular Biology*, 85(2-3):501–522, 2004.

[32] J.C. Neu and W. Krassowska. Homogenization of syncytial tissues. *Critical Reviews in Biomedical Engineering*, 21(2):137–199, 1993.

[33] D. Noble. A modification of the Hodgkin-Huxley equation applicable to Purkinje fiber action and pacemaker potentials. *Journal of Physiology*, 160:317–352, 1962.

[34] J. Okada, S. Sugiura, and T. Hisada. Modeling for cardiac excitation propagation based on the Nernst-Planck equation and homogenization. *Physical Review E*, 87(6):062701, 2013.

[35] C. Pierre. *Modélisation et simulation de l'activité électrique du coeur dans le thorax, analyse numérique et méthodes de volumes finis*. PhD thesis, Université Nantes, 2005.

[36] A.J. Pullan, M.L. Buist, and L.K. Cheng. *Mathematically Modeling the Electrical Activity of the Heart*. World Scientific, 2005.

[37] A. Quarteroni, R. Sacco, and F. Saleri. *Numerical mathematics*, volume 37. Texts in Applied Mathematics. Springer-Verlag, Berlin, second edition., 2007.

[38] G. Richardson. A multiscale approach to modelling electrochemical processes occurring across the cell membrane with application to transmission of action potentials. *Mathematical Medicine and Biology*, 26(3):201–224, 2009.

[39] G. Richardson and J. Chapman. Derivation of the bidomain equations for a beating heart with a general microstructure. *SIAM Journal on Applied Mathematics*, 71(3):657–675, 2011.

[40] F. B. Sachse, G. Seemann, C. Riedel, C. D. Werner, and Dössel O. Modeling of the cardiac mechano-electrical feedback. *International Journal of Bioelectromagnetism*, 2(2), 2000.

[41] F.B. Sachse. *Computational Cardiology: Modeling of Anatomy, Electrophysiology and Mechanics*. Springer-Verlag, 2004.

[42] J. Sainte-Marie, D. Chapelle, R. Cimrman, and M. Sorine. Modeling and estimation of the cardiac electromechanical activity. *Computers and Structures*, 84:1743–1759, 2006.

[43] D.D. Streeter. Gross morphology and fiber geometry of the heart. *Handbook Physiology, The cardiovascular system*, 1:61–112, 1979.

[44] J. Sundnes, G.T. Lines, X. Cai, B.F. Nielsen, K.A. Mardal, and A. Tveito. *Computing the Electrical Activity in the Heart*, volume 1 of *Monographs in Computational Science and Engineering*. Springer-Verlag, 2006.

[45] L. Tung. *A bi-domain model for describing ischemic myocardial d-c potentials.* PhD thesis, Massachusetts Institute of Technology. Dept. of Electrical Engineering and Computer Science., 1978.

[46] M. Veneroni. Reaction–diffusion systems for the macroscopic bidomain model of the cardiac electric field. *Nonlinear Analysis: Real World Applications*, 10(2):849–868, 2009.

[47] G.-X. Yan and C. Antzelevitch. Cellular basis for the normal T wave and the electrocardiographic manifestations of the long-QT syndrome. *Circulation*, 98:1928–1936, 1998.

[48] F. C. P. Yin, C. C. H. Chan, and R. M. R. Judd. Compressibility of perfused passive myocardium. *American Journal of Physiology*, 271:1864–1870, 1996.

Part II

3D and Surface-based Models

Introduction of Part II

In the following part of this thesis, we are interested in the atrial electrophysiology. Indeed, as you have seen in the previous chapter, we are able to provide realistic electrophysiological numerical simulations but only for the ventricles. The specificity of the atria is that they have very thin walls and this implies prohibitive computational times. Moreover the atria mostly appear as thick surfaces in medical imaging. For all these reasons, we want to derive a 2D electrophysiology model from a 3D model. This surface-based model should be defined over the midsurface of the thin region and should take into account the **strong anisotropy** variations across the thickness. It is the subject of Chapter 3 – the first chapter of this part – which takes the form of an article [3] gathering the work carried out in collaboration with D. Chapelle and J.-F. Gerbeau

A surface-based electrophysiology model relying on asymptotic analysis and motivated by cardiac atria modeling.

Due to the strong anisotropy, we present in this article **an asymptotic analysis** of the diffusion term. Following similar strategies used in shell mechanical modeling, we consider a Galerkin reduction of the problem posed in a subspace given by polynomial variations of the quantity of interest in the thickness variable - in our case quadratic polynomials. An asymptotic analysis then consists in studying the behavior of the diffusion model when varying the thickness parameter ε which is the ratio between between the size of the thickness and the diameter of the midsurface while keeping the midsurface fixed. We find a limit problem in keeping only the first term of the development after an integration along the thickness variable. In this article, we give a proof of the **weak convergence** of the quadratic decomposition of the 3D quantity to the solution of limit problem which is a surface-based model. The proof of this weak convergence is inspired from methods using in shell mechanical modeling. We also give in this article the **strong convergence**. Using the asymptotic analysis of the diffusion problem, we propose a **bidomain asymptotic model** valid for thin cardiac structure as atria. We validate this model by comparing simulations of the bidomain asymptotic model with the classical bidomain model.

The strategy used for the passage between the weak and the strong convergence is also adapted for the proof of the **strong convergence of the 3D-shell model**. This strong convergence is given in Chapter 4 – the second chapter of this part – which is an article [2] in collaboration with D. Chapelle

Strong convergence results for the asymptotic behavior of the 3D-shell model.

This article also contains an improved version of the proof of the weak convergence previously presented in [4, 1].

Bibliography

[1] D. Chapelle and K.J. Bathe. *The Finite Element Analysis of Shells - Fundamentals*. Springer, second edition, 2011.

[2] D. Chapelle and A. Collin. Strong convergence results for the asymptotic behavior of the 3D-shell model. *Journal of Elasticity*, 115(2):173–192, 2014.

[3] D. Chapelle, A. Collin, and J.-F. Gerbeau. A surface-based electrophysiology model relying on asymptotic analysis and motivated by cardiac atria modeling. *M3AS*, 23(14):2749–2776, 2013.

[4] D. Chapelle, A. Ferent, and K.J. Bathe. 3D-shell elements and their underlying mathematical model. *M3AS*, 14(1):105–142, 2004.

CHAPTER 3

A surface-based electrophysiology model relying on asymptotic analysis and motivated by cardiac atria modeling

Abstract Computational electrophysiology is a very active field with tremendous potential in medical applications, albeit it leads to highly intensive simulations. We here propose a surface-based electrophysiology formulation, motivated by the modeling of thin structures such as cardiac atria, which greatly reduces the size of the computational models. Moreover, our model is specifically devised to retain the key features associated with the anisotropy in the diffusion effects induced by the fiber architecture, with rapid variations across the thickness that cannot be adequately represented by naive averaging strategies. Our proposed model relies on a detailed asymptotic analysis in which we identify a limit model and establish strong convergence results. We also provide detailed numerical assessments that confirm an excellent accuracy of the surface-based model – compared with the reference 3D model – including in the representation of a complex phenomenon, namely, spiral waves.

A surface-based electrophysiology model relying on asymptotic analysis and motivated by cardiac atria modeling

Dominique Chapelle, Annabelle Collin,
Inria Saclay Ile-de-France, Palaiseau, France,
Jean-Frédéric Gerbeau,
Inria Paris-Rocquencourt, France,

M3AS, DOI:10.1142/S0218202513500450

Abstract

Computational electrophysiology is a very active field with tremendous potential in medical applications, albeit it leads to highly intensive simulations. We here propose a surface-based electrophysiology formulation, motivated by the modeling of thin structures such as cardiac atria, which greatly reduces the size of the computational models. Moreover, our model is specifically devised to retain the key features associated with the anisotropy in the diffusion effects induced by the fiber architecture, with rapid variations across the thickness that cannot be adequately represented by naive averaging strategies. Our proposed model relies on a detailed asymptotic analysis in which we identify a limit model and establish strong convergence results. We also provide detailed numerical assessments that confirm an excellent accuracy of the surface-based model – compared with the reference 3D model – including in the representation of a complex phenomenon, namely, spiral waves.

Keywords: Computational electrophysiology; Asymptotic analysis; Thin domains; Cardiac modeling
Mathematics Subject Classification (2010): 22E46, 53C35, 57S20

1 Introduction

Cardiac electrophysiology purports to describe and model chemical and electrical phenomena taking place in the cardiac tissue, and which are responsible for activating the mechanical contraction in the myocytes, namely, the cardiac muscle cells, see e.g. Refs. [16, 15, 19, 3] and references therein. Given the frequent occurrence of pathologies – such as atrial fibrillation and ventricular tachycardia – directly affecting electrophysiology, hence impairing cardiac function, the detailed understanding of the associated underlying mechanisms is particularly important. In this context, there is a tremendous potential for modeling, in particular with models adequately "personalized" to given patients, in order to guide medical decision in selecting the best-adapted therapeutic strategy [18, 17]. For instance, personalized models may be used to optimize the procedure called radio-frequency ablation that aims at treating cardiac arrythmias – e.g., atrial fibrillation – by blocking some abnormal conduction pathways.

Computational electrophysiology modeling – when aiming at representing the behavior of the global heart or whole subparts thereof, which of course is crucial when considering conduction phenomena and their incidence on cardiac function – is known to give rise to highly intensive simulations [7, 9]. This is due, in particular, to the substantial refinements required in the meshes in order to accurately capture the conduction waves. As personalized modeling generally involves the solution of inverse problems in order to estimate anatomical and biophysical

parameters characterizing a given patient-specific model – in order to enable predictive modeling – based on clinical data, this may lead to prohibitive computational times, indeed. The same difficulty holds as regards optimization loops needed to adapt the therapeutic strategy.

Nevertheless, when considering the thin-ness of cardiac structures such as the atria, there is clearly a case for modeling simplifications based on considering surface-based descriptions – hence, in essence two-dimensional (2D) – instead of 3D models, very much in the spirit of structural models – such as shells – in mechanics. Such mechanical models have been extensively analyzed and provide extremely computationally-effective formulations, with very limited loss of accuracy when the structure is thin. As regards electrophysiology modeling, this path has been scarcely explored. A frequently-argued explanation for that lies in the complexity of the 3D anatomical details that appear to be needed to accurately capture the relevant phenomena. In particular, the fiber architecture of the muscle gives rise to strong anisotropy in the diffusion phenomena, and the fiber directions are known to be very rapidly varying across the thickness of the cardiac walls, including in the atria [10]. This makes naive surface reductions of the models based on isotropy or "average anisotropy" assumptions very inaccurate, and in fact altogether unable to represent some complex conduction patterns [12].

The objective in this paper is to propose a surface-based electrophysiology model relying on a detailed asymptotic analysis. This type of derivation – also available for shell structures in mechanics, see e.g. Refs. [6, 4, 5] – provides a "mathematically justified" surface model designed for thin structures. In our case, in order to incorporate the above-mentioned key anatomical details, we include the rapid variations of the fiber direction within our asymptotic analysis assumptions. We point out that some asymptotic derivations have also been performed in earlier works,[14] albeit without taking into account the anisotropy and variations thereof.

An outline of the paper is as follows. In Sections 2 and 3 we recall some relevant elements of cardiac electrophysiology modeling, and differential geometry, respectively. Next, in Section 4 we formulate an asymptotic framework for an anisotropic diffusion model with anisotropy directions rapidly varying across the thickness, and we establish strong convergence results to a limit model. This allows us to propose a surface-based electrophysiology model in Section 5, with detailed numerical assessments.

2 Elements of cardiac electrophysiology

The electric wave propagating in the cardiac tissue can be represented by a nonlinear reaction-diffusion partial differential equation (PDE), coupled with an ordinary differential equation (ODE) representing cellular activity. Considering in particular the bidomain model – see for example Refs. [16, 15, 19, 2] – equations can be written in terms of the extracellular potential u_e, the transmembrane potential $V_m = u_i - u_e$, with u_i the intracellular potential, and the ionic variable w, as

$$\left\{\begin{array}{ll} A_m\Big(C_m\dfrac{\partial V_m}{\partial t} + I_{ion}(V_m,w)\Big) - \operatorname{div}\big(\vec{\vec{\sigma}}_i\cdot\vec{\nabla}V_m\big) & \\ \qquad\qquad = \operatorname{div}\big(\vec{\vec{\sigma}}_i\cdot\vec{\nabla}u_e\big) + A_m I_{app}, & \text{in } \mathcal{B}\times(0,T),\\ \operatorname{div}\Big(\big(\vec{\vec{\sigma}}_i+\vec{\vec{\sigma}}_e\big)\cdot\vec{\nabla}u_e\Big) = -\operatorname{div}\big(\vec{\vec{\sigma}}_i\cdot\vec{\nabla}V_m\big), & \text{in } \mathcal{B}\times(0,T),\\ \partial_t w + g(V_m,w) = 0, & \text{in } \mathcal{B}\times(0,T), \end{array}\right. \tag{1}$$

where $\mathcal{B}$ denotes the 3D domain of interest, and with appropriate boundary conditions

$$\left\{\begin{array}{rcll} \big(\vec{\vec{\sigma}}_i\cdot\vec{\nabla}u_e\big)\cdot\vec{n} &=& -\big(\vec{\vec{\sigma}}_i\cdot\vec{\nabla}V_m\big)\cdot\vec{n}, & \text{in } \partial\mathcal{B}\times(0,T),\\ \big(\vec{\vec{\sigma}}_e\cdot\vec{\nabla}u_e\big)\cdot\vec{n} &=& 0, & \text{in } \partial\mathcal{B}\times(0,T), \end{array}\right. \tag{2}$$

where A_m is a positive constant denoting the ratio of membrane area per unit volume, C_m the membrane capacitance per unit surface, $I_{ion}(V_m, w)$ a reaction term representing the ionic current across the membrane, and I_{app} a given applied stimulus current.

The current I_{ion} can be described by a phenomenological model. In this study, the model proposed by Mitchell and Schaeffer in Ref. [13] is considered. The functions g and I_{ion} are then given by

$$\begin{aligned} I_{ion}(V_m, w) &= -\frac{w(V_m - V_{min})^2(V_{max} - V_m)}{\tau_{in}(V_{max} - V_{min})} + \frac{V_m - V_{min}}{\tau_{out}(V_{max} - V_{min})}, \\ g(V_m, w) &= \begin{cases} \dfrac{w}{\tau_{open}} - \dfrac{1}{\tau_{open}(V_{max} - V_{min})^2} & \text{if } V_m < V_{gate}, \\ \dfrac{w}{\tau_{close}} & \text{if } V_m \geq V_{gate}, \end{cases} \end{aligned}$$

where $\tau_{in}, \tau_{out}, \tau_{open}, \tau_{close}$ are given parameters and V_{min}, V_{max}, $V_{min} < V_{gate} < V_{max}$ auxiliary constants.

Cardiac muscle has a fiber architecture. The electrical conductivity is higher along than across the fiber direction. The intracellular and extracellular media are therefore anisotropic. This anisotropy is included in our model defining the conductivity tensors $\vec{\vec{\sigma}}_i$ and $\vec{\vec{\sigma}}_e$ by

$$\vec{\vec{\sigma}}_{i,e} = \sigma_{i,e}^t \, \vec{\vec{g}} + (\sigma_{i,e}^l - \sigma_{i,e}^t) \, \vec{\tau} \otimes \vec{\tau},$$

where $\vec{\vec{g}}$ denotes the 3D metric tensor – the components of which are given by the identity matrix in an orthonormal coordinate system, see e.g. Ref. [4] – the vector $\vec{\tau}$ is of unit length and parallel to the local fiber direction, and σ_i^l and σ_i^t are the conductivity coefficients in the intra-cellular medium measured along and across – respectively – the fiber direction, and likewise for σ_e^l and σ_e^t in the extra-cellular medium.

The bidomain model can be rewritten in weak form as follows. For all $t > 0$, find $V_m(\cdot, t) \in H^1$, $u_e(\cdot, t) \in H^1$ and $w(\cdot, t) \in L^\infty$ with $\int_{\mathcal{B}} u_e = 0$, such that

$$\begin{cases} A_m \displaystyle\int_{\mathcal{B}} \Big(C_m \frac{\partial V_m}{\partial t} + I_{ion}(V_m, w) \Big) \phi + \int_{\mathcal{B}} \Big[\vec{\vec{\sigma}}_i \cdot \big(\vec{\nabla} V_m + \vec{\nabla} u_e \big) \Big] \cdot \vec{\nabla} \phi \\ \qquad\qquad = A_m \displaystyle\int_{\mathcal{B}} I_{app}(V_m, w) \phi, \\ \displaystyle\int_{\mathcal{B}} \Big[(\vec{\vec{\sigma}}_i + \vec{\vec{\sigma}}_e) \cdot \vec{\nabla} u_e \Big] \cdot \vec{\nabla} \psi + \int_{\mathcal{B}} \Big[\vec{\vec{\sigma}}_i \cdot \vec{\nabla} V_m \Big] \cdot \vec{\nabla} \psi = 0, \\ \partial_t w + g(V_m, w) = 0, \text{ in } \mathcal{B}, \end{cases}$$

for all $\phi, \psi \in H^1$ such that $\int_{\mathcal{B}} \psi = 0$.

With our motivation of deriving a 2D model valid for thin cardiac structures – the atrial walls, in particular – and defined over the midsurface of the thin region, we observe that the main difficulty in representing the anisotropy resulting from the preferred conduction direction along the muscle fibers – which may rapidly vary across the thickness – arises from the diffusion term in the equations. Hence, in our mathematical analysis we will focus on this term, with the primary concern of taking into account both major features, namely, anisotropic behavior and heterogeneous distribution across the thickness.

3 Geometric preliminaries

The midsurface $\mathcal{S}$ of the body is described by a mapping $\vec{\phi}$ defined over $\overline{\omega}$, with ω a domain of $\mathbb{R}^2$, and with values in the three-dimensional Euclidean space $\mathcal{E}$, see Figure 1. We assume that $\vec{\phi}$

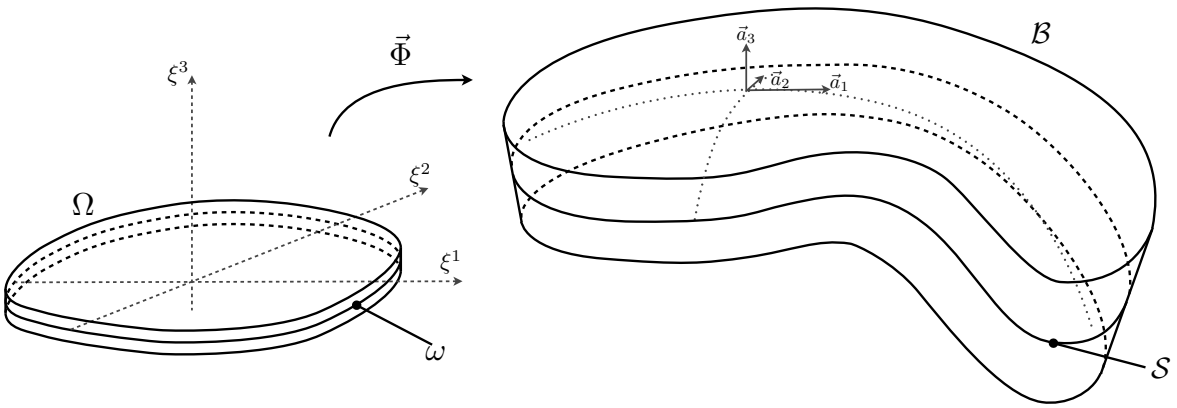

Figure 1: Geometric domain and midsurface $\mathcal{S}$ descriptions

is as regular as needed. For the sake of simplicity in our asymptotic analysis we suppose that the thickness of the body, denoted by d, is constant over $\mathcal{S}$. We introduce $L = \mathrm{diam}(\mathcal{S})$ and $\varepsilon = \frac{d}{L}$ a dimensionless thickness parameter. Note that this parameter is defined to be geometrically *intrinsic*, namely, it is invariant with respect to changes in the geometric mapping $\vec{\phi}$ (hence, in the coordinate system and its domain ω) that preserve the surface.

Denoting by (ξ^1, ξ^2) the coordinates used in $\mathbb{R}^2$ (hence in ω), we define

$$\vec{a}_\alpha(\xi^1, \xi^2) = \partial_\alpha \vec{\phi}(\xi^1, \xi^2), \quad \alpha = 1, 2.$$

We suppose that $\vec{\phi}$ is such that the vectors $(\vec{a}_\alpha)_{\alpha=1,2}$ form a basis – called the covariant basis – for the tangential plane to the midsurface $\mathcal{S}$ at any point with coordinates in $\overline{\omega}$. We also introduce the contravariant basis $(\vec{a}^1, \vec{a}^2)$ such that

$$\vec{a}^\alpha(\xi^1, \xi^2) \cdot \vec{a}_\beta(\xi^1, \xi^2) = \delta^\beta_\alpha, \quad \alpha, \beta = 1, 2, \quad \forall (\xi^1, \xi^2) \in \omega,$$

and we denote by

$$a_{\alpha\beta} = \vec{a}_\alpha \cdot \vec{a}_\beta,$$

the covariant components of the surface-attached metric tensor, see e.g. Ref. [4].

Remark 1 *Introducing the* 2×2 *matrix* $A = (a_{\alpha\beta})_{\alpha,\beta=1,2}$*, we have* $A^{-1} = (a^{\alpha\beta})_{\alpha,\beta=1,2} = (\vec{a}^\alpha \cdot \vec{a}^\beta)_{\alpha,\beta=1,2}$.

We denote by a the quantity $a = \|\vec{a}_1 \times \vec{a}_2\|^2$, and the unit vector normal to the tangential plane at any point with coordinates in $\overline{\omega}$ is given by

$$\vec{a}_3 = \frac{\vec{a}_1 \times \vec{a}_2}{\sqrt{a}}.$$

Defining now $\Omega = \omega \times \left]-\frac{d}{2}, +\frac{d}{2}\right[$, the 3D geometry is described by the mapping

$$\vec{\Phi} : \left\{ \begin{array}{ccc} \overline{\Omega} & \longrightarrow & \mathcal{E} \\ (\xi^1, \xi^2, \xi^3) & \longmapsto & \vec{\phi}(\xi^1, \xi^2) + \xi^3 \vec{a}_3(\xi^1, \xi^2) \end{array} \right.$$

and the 3D geometric domain of interest is given by $\mathcal{B} = \vec{\Phi}\big(\overline{\Omega}\big)$, see Figure 1. We also introduce the 3D covariant basis

$$\vec{g}_i(\xi^1, \xi^2, \xi^3) = \partial_i \vec{\Phi}(\xi^1, \xi^2, \xi^3), \quad i = 1, 2, 3, \quad \forall (\xi^1, \xi^2, \xi^3) \in \Omega,$$

and the corresponding contravariant basis $(\vec{g}^1, \vec{g}^2, \vec{g}^3)$

$$\vec{g}^i(\xi^1, \xi^2, \xi^3) \cdot \vec{g}_j(\xi^1, \xi^2, \xi^3) = \delta_i^j, \quad i, j = 1, 2, 3, \quad \forall (\xi^1, \xi^2, \xi^3) \in \Omega.$$

The components of the 3D metric tensor $\vec{\vec{g}}$ in covariant form are then given by

$$g_{ij} = \vec{g}_i \cdot \vec{g}_j.$$

Remark 2 *Introducing the* 3×3 *matrix* $G = (g_{ij})_{i,j=1,2,3}$*, we have* $G^{-1} = (g^{ij})_{i,j=1,2,3} = (\vec{g}^i \cdot \vec{g}^j)_{i,j=1,2,3}$.

For a differentiable scalar field defined over the surface the 2D gradient is given by

$$\underline{\nabla} u = \partial_1 u \, \vec{a}^1 + \partial_2 u \, \vec{a}^2,$$

whereas for a volume-defined quantity we have the 3D gradient

$$\vec{\nabla} u = \partial_1 u \, \vec{g}^1 + \partial_2 u \, \vec{g}^2 + \partial_3 u \, \vec{g}^3.$$

We need a number of preliminary results that relate 3D quantities to surface quantities. For the proofs of the three following propositions, see e.g. Ref. [4].

Proposition 1 *Introducing the second and third fundamental forms respectively defined for all* $\alpha, \beta = 1, 2$ *by*

$$b_{\alpha\beta} = \vec{a}_3 \cdot \partial_\beta \vec{a}_\alpha, \quad c_{\alpha\beta} = \sum_{\lambda,\mu=1,2} a^{\lambda\mu} b_{\lambda\alpha} b_{\mu\beta},$$

we have

$$g_{\alpha\beta} = a_{\alpha\beta} - 2\xi^3 b_{\alpha\beta} + (\xi^3)^2 c_{\alpha\beta}, \tag{3}$$

and, moreover,

$$g_{\alpha 3} = 0, \quad g_{33} = 1. \tag{4}$$

Proposition 2 *For the volume measure given by* $dV = \sqrt{g}\, d\xi^1 d\xi^2 d\xi^3$*, with* $g = (\det(\vec{g}_1, \vec{g}_2, \vec{g}_3))^2 = \det(G)$*, we have*

$$g = a\big(1 - 2H\xi^3 + K(\xi^3)^2\big)^2,$$

denoting by H *and* K *the mean and Gaussian curvatures of the surface* $\mathcal{S}$*, respectively. Assuming that* $1 - 2H\xi^3 + K(\xi^3)^2$ *is strictly positive over* $\bar{\mathcal{B}}$*, we then infer* $\sqrt{g} = \sqrt{a}\big(1 - 2H\xi^3 + K(\xi^3)^2\big)$ *and* $\exists \gamma > 0, \sqrt{g} > \gamma\sqrt{a}$.

Proposition 3 *There exist two strictly positive constants* c_1*,* c_2 *such that,* $\forall (\xi^1, \xi^2, \xi^3) \in \Omega$*,* $\forall (x_1, x_2) \in \mathbb{R}^2$*,*

$$c_1 \sum_{\alpha,\beta=1}^{2} a^{\alpha\beta}(\xi^1, \xi^2) x_\alpha x_\beta \le \sum_{\alpha,\beta=1}^{2} g^{\alpha\beta}(\xi^1, \xi^2, \xi^3) x_\alpha x_\beta \le c_2 \sum_{\alpha,\beta=1}^{2} a^{\alpha\beta}(\xi^1, \xi^2) x_\alpha x_\beta.$$

We can then show the following identity.

Lemma 1 $\forall \alpha, \beta = 1, 2$*, there exist* $d^{\alpha\beta}$ *and* $e^{\alpha\beta}$*, two functions of* (ξ^1, ξ^2) *in* $L^\infty(\omega)$*, such that*

$$g^{\alpha\beta} = \frac{1}{\big(1 - 2H\xi^3 + K(\xi^3)^2\big)^2} \big(a^{\alpha\beta} - 2\xi^3 d^{\alpha\beta} + (\xi^3)^2 e^{\alpha\beta}\big).$$

Proof According to Remarks 1 and 2, we have

$$A^{-1} = \begin{pmatrix} a^{11} & a^{12} \\ a^{12} & a^{22} \end{pmatrix} = \frac{1}{a} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{12} & a_{11} \end{pmatrix}, \quad G^{-1} = \begin{pmatrix} g^{11} & g^{12} & 0 \\ g^{12} & g^{22} & 0 \\ 0 & 0 & 1 \end{pmatrix} = \frac{1}{g} \begin{pmatrix} g_{22} & -g_{12} & 0 \\ -g_{12} & g_{11} & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Using Propositions 1 and 2, we can then conclude. □

Finally, this directly implies the following result.

Proposition 4 *$\forall \alpha, \beta = 1, 2$, $g^{\alpha\beta} = a^{\alpha\beta} + \xi^3 \bar{g}^{\alpha\beta}$, where $\bar{g}^{\alpha\beta}$ is a function bounded over $\mathcal{B}$.*

4 Limit model derivation by asymptotic analysis

4.1 Anisotropic diffusion model

We introduce the space $H^1(\mathcal{B})$ defined by

$$H^1(\mathcal{B}) = \left\{ u : \Omega \to \mathbb{R} \text{ measurable}, \int_\Omega u^2 \, dV, \int_\Omega \vec{\nabla} u \cdot \vec{\nabla} u \, dV < +\infty \right\},$$

with the natural norm

$$\|u\|_{H^1(\mathcal{B})} = \left(\int_\Omega u^2 \, dV + \int_\Omega \vec{\nabla} u \cdot \vec{\nabla} u \, dV \right)^{\frac{1}{2}}.$$

We need to introduce boundary conditions associated with the diffusion term that we want to study. We suppose $\partial\omega = \gamma_1 \cup \gamma_2$ with γ_1 of non-zero measure, and we define $\Gamma_i = \gamma_i \times \left]-\frac{d}{2}, +\frac{d}{2}\right[$, for $i = 1, 2$. We denote $\Gamma_3 = \partial\Omega \setminus \left(\Gamma_1 \cup \Gamma_2\right)$, namely, the top and bottom surfaces. We consider the following problem

$$\left\{ \begin{array}{rcll} -\operatorname{div}\left(\vec{\vec{\sigma}} \cdot \vec{\nabla} u\right) & = & f, & \text{in } \Omega, \\ u & = & 0, & \text{in } \Gamma_1, \\ \left(\vec{\vec{\sigma}} \cdot \vec{\nabla} u\right) \cdot \vec{n} & = & 0, & \text{in } \Gamma_2, \\ \left(\vec{\vec{\sigma}} \cdot \vec{\nabla} u\right) \cdot \vec{n} & = & 0, & \text{in } \Gamma_3, \end{array} \right. \tag{5}$$

where $f \in L^2(\mathcal{B})$.

Defining $\mathcal{V}^{3D} = H^1(\mathcal{B}) \cap (\mathcal{BC})$, where $(\mathcal{BC})$ corresponds to the Dirichlet boundary condition on Γ_1, the problem (5) can be rewritten in weak form as seeking $u \in \mathcal{V}^{3D}$ such that

$$A^{3D}(u, v) = F^{3D}(v), \quad \forall v \in \mathcal{V}^{3D}, \tag{6}$$

where

$$A^{3D}(u, v) = \int_\Omega \left(\vec{\vec{\sigma}} \cdot \vec{\nabla} u\right) \cdot \vec{\nabla} v \, dV, \quad F^{3D}(v) = \int_\Omega f v \, dV, \quad \forall u, v \in \mathcal{V}^{3D}.$$

Remark 3 *The variational problem is here written in an intrinsic form, namely, independently of any specific coordinate system, but of course it is very straightforward to obtain the corresponding expressions using the tensor components, which we will consider soon with the coordinates introduced in the previous section.*

Assumption 1 *We suppose that the diffusion tensor $\vec{\vec{\sigma}}(\xi)$ is:*

- *L^∞ on Ω;*
- *symmetric positive definite for almost all ξ in Ω;*
- *such that for almost all $\xi \in \Omega$, for any $\vec{\rho} \in \mathcal{E}$,*

$$c_1 \|\vec{\rho}\|^2 \leq \left(\vec{\vec{\sigma}}(\xi) \cdot \vec{\rho}\right) \cdot \vec{\rho},$$

for some strictly positive constant c_1.

Remark 4 *Due to the L^∞ character of $\vec{\vec{\sigma}}(\xi)$ we also have*

$$\left(\vec{\vec{\sigma}}(\xi)\cdot\vec{\rho}\right)\cdot\vec{\rho}\leq c_2\|\vec{\rho}\|^2,$$

for some strictly positive c_2, for any $\vec{\rho}\in\mathcal{E}$ and for almost all $\xi\in\Omega$.

A simple application of the Lax-Milgram theorem with the Poincaré inequality then provides the following result.

Theorem 1 *If Assumption 1 holds, there exists a unique $u\in\mathcal{V}^{3D}$ solution of (6).*

4.2 Asymptotic problem formulation

Denoting $dS=\sqrt{a(\xi^1,\xi^2)}\,d\xi^1d\xi^2$ the surface measure, we introduce the spaces $L^2(\mathcal{S})$ and $H^1(\mathcal{S})$ by

$$L^2(\mathcal{S})=\left\{u:\omega\to\mathbb{R}\text{ measurable},\int_\omega u^2\,dS<+\infty\right\},$$

$$H^1(\mathcal{S})=\left\{u:\omega\to\mathbb{R}\text{ measurable},\int_\omega u^2\,dS,\int_\omega \underline{\nabla}u\cdot\underline{\nabla}u\,dS<+\infty\right\},$$

with the natural norms

$$\|u\|_{L^2(\mathcal{S})}=\left(\int_\omega u^2\,dS\right)^{\frac{1}{2}},\quad \|u\|_{H^1(\mathcal{S})}=\left(\int_\omega u^2\,dS+\int_\omega \underline{\nabla}u\cdot\underline{\nabla}u\,dS\right)^{\frac{1}{2}}.$$

We assume that the source term f is smooth enough to provide

$$f(\xi^1,\xi^2,\xi^3)=f_0(\xi^1,\xi^2)+\xi^3\bar{f}(\xi^1,\xi^2,\xi^3),$$

where $f_0\in L^2(\mathcal{S})$ and $\bar{f}\in L^\infty(\mathcal{B})$. Regarding the modeling of the anisotropic conductivity, we make the following assumption on the diffusion tensor $\vec{\vec{\sigma}}$.

Assumption 2 *We suppose that $\vec{\vec{\sigma}}=\sigma^t\,\vec{\vec{g}}+(\sigma^l-\sigma^t)\,\vec{\tau}\otimes\vec{\tau}$, where σ^t and σ^l are two strictly positive constants such that $\sigma^t\leq\sigma^l$. The vector $\vec{\tau}$ is given by*

$$\vec{\tau}(\xi^1,\xi^2,\xi^3)=\underline{\tau}_0(\xi^1,\xi^2)\cos\left(\frac{2\theta(\xi^1,\xi^2)\xi^3}{d}\right)+\underline{\tau}_0^\perp(\xi^1,\xi^2)\sin\left(\frac{2\theta(\xi^1,\xi^2)\xi^3}{d}\right),\tag{7}$$

where $\underline{\tau}_0$ and $\underline{\tau}_0^\perp$ are unit vectors and

- *$\underline{\tau}_0$ belongs to the tangential plane to the midsurface $\mathcal{S}$, at each point in ω,*
- *$\underline{\tau}_0^\perp=\vec{a}_3\times\underline{\tau}_0$ i.e $\underline{\tau}_0^\perp$ is orthogonal to $\underline{\tau}_0$ and belongs to the tangential plane.*

Remark 5 *The assumptions on $\vec{\vec{\sigma}}$ translate the anisotropic behavior and heterogeneous distribution across the thickness of the body, see Figure 2. Note that*

- *$\sigma^l=\sigma^t$ gives the homogeneous case, namely, without any privileged direction;*
- *$\sigma^t<\sigma^l$ implies*
  - *if $\theta=0$, then $\vec{\tau}$ is independent of ξ^3, i.e. the fiber orientation is constant across the thickness, defined by the vector $\underline{\tau}_0$;*
  - *if $\theta\neq0$, the fibers rotate across the thickness by a total angle 2θ.*

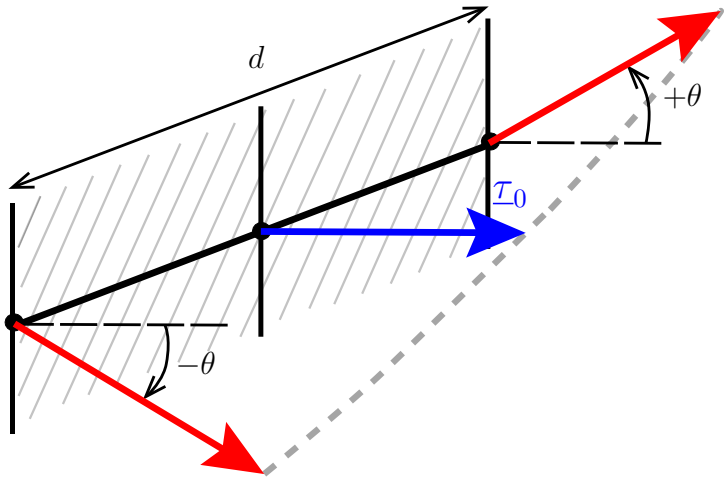

Figure 2: Fibers rotation across the thickness

Remark 6 *We commit a slight abuse of notation in* (7)*, since we have a 3D vector in the left-hand side and surface-attached 2D vectors in the right-hand side. We could have instead used a 3D notation for* $\underline{\tau}_0$ *and* $\underline{\tau}_0^{\perp}$*, but we made this choice to emphasize that they lie in the tangential plane.*

Remark 7 *Assumption 2 implies Assumption 1.*

An asymptotic analysis then consists in studying the behavior of the above diffusion model when varying the thickness parameter ε while keeping the midsurface $\mathcal{S}$ and the parameter field θ fixed. The complete asymptotic analysis of this 3D anisotropic diffusion problem is out of the scope of our preliminary analysis, which primarily aims at guiding the formulation of a relevant surface-based model. Hence, following similar strategies used in shell mechanical modeling, we will instead consider a Galerkin reduction of the problem (6) posed in a subspace of $\mathcal{V}^{3D}$ given by polynomial variations of the quantity of interest in the thickness variable – in our case quadratic polynomials – which can be seen as truncated Taylor expansions. We thus define $\mathcal{V} = (H^1(\mathcal{S}) \cap (\mathcal{BC}'))^3$, where $(\mathcal{BC}')$ corresponds to homogeneous Dirichlet boundary conditions prescribed on γ_1, and we seek $(u_0^\varepsilon, u_1^\varepsilon, u_2^\varepsilon) \in \mathcal{V}$ such that

$$A^{3D}\big(u_0^\varepsilon + \xi^3 u_1^\varepsilon + (\xi^3)^2 u_2^\varepsilon\,,\, v_0 + \xi^3 v_1 + (\xi^3)^2 v_2\big) = F^{3D}\big(v_0 + \xi^3 v_1 + (\xi^3)^2 v_2\big), \quad \forall (v_0, v_1, v_2) \in \mathcal{V}, \quad (8)$$

where we point out that the definitions of A^{3D} and F^{3D} include a dependence on the varying thickness parameter ε – via the dependence of the reference domain Ω itself – even though this is not reflected in the notation for compactness purposes.

Remark 8 *A 3D asymptotic analysis – namely, of a sequence of solutions of* (6) *associated with decreasing thickness parameter for the geometric domain – would require different tools, starting with a scaling of the domain in the transverse direction in order to work on a fixed domain, see e.g. Ref. [6]. Here, we adopt a somewhat simpler strategy with a Galerkin projection of the 3D model onto a subspace of functions that feature quadratic variations in the transverse direction. Note that this is also a classical asymptotic approach in structural mechanics and associated mathematical analysis, where it is known to provide important insight into the limit asymptotic behavior [8, 4]. As a matter of fact, in shell models such Galerkin projections*

have been established to be asymptotically consistent with 3D elasticity, namely, the respective asymptotic analyses provide the same limit models [5]. In addition, limit models obtained with polynomial assumptions are also interesting in relation to Galerkin discretizations of the variational problem, e.g. with finite elements used with a single element – of possibly high polynomial order – across the thickness.

If Assumption 2 holds, we can use Theorem 1 to show the existence and uniqueness of a solution of (8) because $\{u_0 + \xi^3 u_1 + (\xi^3)^2 u_2, (u_0, u_1, u_2) \in \mathcal{V}\}$ is a closed subspace of $\mathcal{V}^{3D}$.

Theorem 2 *For any $\varepsilon = \frac{d}{L} > 0$, there exists a unique $(u_0^\varepsilon, u_1^\varepsilon, u_2^\varepsilon) \in \mathcal{V}$ solution of (8).*

We will now denote $u^\varepsilon = u_0^\varepsilon + \xi^3 u_1^\varepsilon + (\xi^3)^2 u_2^\varepsilon \in \mathcal{V}^{3D}$ and study the behavior of this sequence when the parameter ε tends to zero.

4.3 Asymptotic analysis

In this section, we want to:

- identify a limit problem;
- show the existence and uniqueness of a solution to the limit problem;
- prove a preliminary weak convergence result;
- establish the strong convergence.

4.3.1 Limit problem

We start by considering the left-hand side of (8), i.e.

$$\begin{aligned}\int_\Omega \big(\vec{\vec{\sigma}} \cdot \vec{\nabla} u\big) \cdot \vec{\nabla} v \; dV &= \sigma^t \int_\Omega \big(\vec{\vec{g}} \cdot \vec{\nabla} u\big) \cdot \vec{\nabla} v \, dV + (\sigma^l - \sigma^t) \int_\Omega \big(\vec{\tau} \otimes \vec{\tau} \cdot \vec{\nabla} u\big) \cdot \vec{\nabla} v \, dV, \\ &= \sigma^t \int_\Omega \vec{\nabla} u \cdot \vec{\nabla} v \, dV + (\sigma^l - \sigma^t) \int_\Omega \big(\vec{\tau} \cdot \vec{\nabla} u\big)\big(\vec{\tau} \cdot \vec{\nabla} v\big) \, dV, \end{aligned} \tag{9}$$

where we denote $u = u_0 + \xi^3 u_1 + (\xi^3)^2 u_2$ and $v = v_0 + \xi^3 v_1 + (\xi^3)^2 v_2$, and we will analyze the two integrals separately.

According to Proposition 4, we have

$$\sigma^t \vec{\nabla} u \cdot \vec{\nabla} v = \sigma^t \left[\sum_{\alpha,\beta=1}^{2} \partial_\alpha u (a^{\alpha\beta} + \xi^3 \bar{g}^{\alpha\beta}) \partial_\beta v + u_1 v_1 + 2\xi^3 \big(u_1 v_2 + u_2 v_1\big) + 4\big(\xi^3\big)^2 u_2 v_2 \right].$$

We can integrate all the polynomial terms in ξ^3 between $-\frac{d}{2}$ and $+\frac{d}{2}$, and after integration the first term in the asymptotic expansion is given by

$$d\, \sigma^t \int_\omega \big(\underline{\nabla} u_0 \cdot \underline{\nabla} v_0 + u_1 v_1\big) \, dS.$$

We can then denote by A_l the bilinear form

$$A_l\big((u_0, u_1), (v_0, v_1)\big) = \sigma^t \int_\omega \big(\underline{\nabla} u_0 \cdot \underline{\nabla} v_0 + u_1 v_1\big) \, dS, \qquad \forall (u_0, u_1), (v_0, v_1) \in H^1(\mathcal{S}) \times L^2(\mathcal{S}).$$

Concerning next the second term in (9) we have

$$\begin{aligned}\big(\vec{\tau}\cdot\vec{\nabla}u\big)\big(\vec{\tau}\cdot\vec{\nabla}v\big) &= \sum_{i,j=1}^{3}(\vec{\tau}\cdot\vec{g}^{\,i})(\vec{\tau}\cdot\vec{g}^{\,j})\partial_i u\partial_j v\\ &= \sum_{\alpha,\beta=1}^{2}(\vec{\tau}\cdot\vec{g}^{\,\alpha})(\vec{\tau}\cdot\vec{g}^{\,\beta})\partial_\alpha u\partial_\beta v,\end{aligned}$$

according to Assumption 2. Using the geometric definitions, we have that $\vec{g}^{\,\alpha}(\xi^1,\xi^2,\xi^3) = \vec{a}^{\,\alpha}(\xi^1,\xi^2) + \xi^3\vec{H}^{\alpha}(\xi^1,\xi^2,\xi^3)$ with $\vec{H}^{\alpha}\in L^\infty(\mathcal{B})$, so that we can decompose $(\vec{\tau}\cdot\vec{g}^{\,\alpha})(\vec{\tau}\cdot\vec{g}^{\,\beta})$ into

$$\begin{aligned}(\vec{\tau}\cdot\vec{g}^{\,\alpha})(\vec{\tau}\cdot\vec{g}^{\,\beta}) &= \Big[(\underline{\tau}_0\cdot\vec{a}^{\,\alpha})(\underline{\tau}_0\cdot\vec{a}^{\,\beta}) + \xi^3 J_{\parallel}(\xi^1,\xi^2,\xi^3)\Big]\cos^2\!\left(\frac{2\theta\xi^3}{d}\right)\\ &\quad+\Big[(\underline{\tau}_0^{\perp}\cdot\vec{a}^{\,\alpha})(\underline{\tau}_0^{\perp}\cdot\vec{a}^{\,\beta}) + \xi^3 J_{\perp}(\xi^1,\xi^2,\xi^3)\Big]\sin^2\!\left(\frac{2\theta\xi^3}{d}\right)\\ &\quad+\Big[(\underline{\tau}_0\cdot\vec{a}^{\,\alpha})(\underline{\tau}_0^{\perp}\cdot\vec{a}^{\,\beta}) + (\underline{\tau}_0^{\perp}\cdot\vec{a}^{\,\alpha})(\underline{\tau}_0\cdot\vec{a}^{\,\beta})\\ &\qquad\qquad+\xi^3 J_{\dashv}(\xi^1,\xi^2,\xi^3)\Big]\sin\!\left(\frac{2\theta\xi^3}{d}\right)\cos\!\left(\frac{2\theta\xi^3}{d}\right),\end{aligned}$$

where $J_{\parallel}, J_{\perp}$ et $J_{\dashv}$ are all $L^\infty(\mathcal{B})$. When integrating over ξ^3, the following integrals appear

$$I_n(\theta) = \frac{1}{d^{n+1}}\int_{-\frac{d}{2}}^{\frac{d}{2}}(\xi^3)^n\cos^2\!\left(\frac{2\theta\xi^3}{d}\right)d\xi^3 = \frac{1}{2^{n+1}}\int_{-1}^{1}t^n\cos^2(t\theta)dt,$$

$$J_n(\theta) = \frac{1}{d^{n+1}}\int_{-\frac{d}{2}}^{\frac{d}{2}}(\xi^3)^n\sin^2\!\left(\frac{2\theta\xi^3}{d}\right)d\xi^3 = \frac{1}{2^{n+1}}\int_{-1}^{1}t^n\sin^2(t\theta)dt,$$

$$L_n(\theta) = \frac{1}{d^{n+1}}\int_{-\frac{d}{2}}^{\frac{d}{2}}(\xi^3)^n\cos\!\left(\frac{2\theta\xi^3}{d}\right)\sin\!\left(\frac{2\theta\xi^3}{d}\right)d\xi^3 = \frac{1}{2^{n+1}}\int_{-1}^{1}t^n\cos(t\theta)\sin(t\theta)dt.$$

We can now as before identify the first term of the asymptotic expansion, which gives $d\,A_l^f$ with

$$A_l^f(u_0,v_0) = (\sigma^l-\sigma^t)\int_\omega\Big(\big(I_0(\theta)\,\underline{\tau}_0\otimes\underline{\tau}_0 + J_0(\theta)\,\underline{\tau}_0^{\perp}\otimes\underline{\tau}_0^{\perp}\big)\cdot\underline{\nabla}u_0\Big)\cdot\underline{\nabla}v_0\,dS,$$

where $I_0(\theta) = \dfrac{1}{2}+\dfrac{1}{4\theta}\sin(2\theta)$ and $J_0(\theta) = \dfrac{1}{2}-\dfrac{1}{4\theta}\sin(2\theta) = 1-I_0(\theta)$. Note that these functions can be extended by continuity in 0, with values 1 and 0, respectively.

Remark 9

- $I_0(\theta)>0, \quad J_0(\theta)\geq 0, \quad \forall\theta\in\mathbb{R}$,
- $J_0(\theta)=0$ *(and* $I_0(\theta)=1$*) if and only if* $\theta=0$. *In this case* $\vec{\tau}(\xi^1,\xi^2,\xi^3)=\underline{\tau}_0(\xi^1,\xi^2)$, *i.e. the fiber direction is independent of the thickness.*

Using similar manipulations, the first term in the asymptotic expansion of the right-hand side of (8) is given by dF_0, with

$$F_0(v_0) = \int_\omega f_0 v_0\,dS, \quad \forall v_0\in H^1(S).$$

Defining $\mathcal{V}_{lim} = \big(H^1(\mathcal{S})\cap(\mathcal{BC}')\big)\times L^2(\mathcal{S})$, our candidate limit problem finally consists in seeking $(u_0^l,u_1^l)\in\mathcal{V}_{lim}$ such that

$$A_l\big((u_0^l,u_1^l),(v_0,v_1)\big) + A_l^f(u_0^l,v_0) = F_0(v_0), \quad \forall(v_0,v_1)\in\mathcal{V}_{lim}. \tag{10}$$

Theorem 3 *There exists a unique* $(u_0^l, u_1^l) \in \mathcal{V}_{lim}$ *solution of* (10).

Proof It is straightforward to see that the bilinear form in the left-hand side of (10) is continuous and coercive on $\mathcal{V}_{lim}$, and symmetric, hence we can directly apply the Lax-Milgram theorem. □

Note that, choosing $v_0 = 0$ in (10), we can show that $u_1^l = 0$, and $u_0^l \in H^1(\mathcal{S}) \cap (\mathcal{BC}')$ satisfies, $\forall v_0 \in H^1(\mathcal{S}) \cap (\mathcal{BC}')$,

$$\sigma^t \int_\omega \underline{\nabla} u_0^l \cdot \underline{\nabla} v_0 \, dS + (\sigma^l - \sigma^t) \int_\omega \Big(\big(I_0(\theta)\, \underline{\tau}_0 \otimes \underline{\tau}_0 + J_0(\theta)\, \underline{\tau}_0^\perp \otimes \underline{\tau}_0^\perp \big) \cdot \underline{\nabla} u_0^l \Big) \cdot \underline{\nabla} v_0 \, dS = \int_\omega f_0 v_0 \, dS. \quad (11)$$

4.3.2 Weak convergence

Theorem 4 *We have the following convergences:*

- $(u_0^\varepsilon, u_1^\varepsilon)$ *converges weakly to* $(u_0^l, 0)$ *in* $\mathcal{V}_{lim}$ *when* ε *tends to* 0,
- $\varepsilon u_1^\varepsilon$ *and* $\varepsilon^2 u_2^\varepsilon$ *converge weakly to* 0 *in* $H^1(\mathcal{S})$ *when* ε *tends to* 0.

The proof of Theorem 4 is divided into two steps:

(i) The first step consists in proving that a subsequence of $(u_0^\varepsilon, u_1^\varepsilon)_\varepsilon$ converges weakly to a limit denoted by $(u_0^m, u_1^m) \in \mathcal{V}_{lim}$ when ε tends to 0.

(ii) In the second step, we show that $u_0^m = u_0^l, u_1^m = 0$ and we can infer the theorem.

(i) Weak convergence of a subsequence

Lemma 2 *There exists one strictly positive constant* C *independent of* ε *such that* $\forall (v_0, v_1, v_2) \in \mathcal{V}$,

$$A^{3D}(v, v) \geq C \Big[\varepsilon \big(\|\underline{\nabla} v_0\|_{L^2(\mathcal{S})} + \|v_1\|_{L^2(\mathcal{S})} \big)^2 + \varepsilon^3 \big(\|\underline{\nabla} v_1\|_{L^2(\mathcal{S})} + \|v_2\|_{L^2(\mathcal{S})} \big)^2 + \varepsilon^5 \|\underline{\nabla} v_2\|^2_{L^2(\mathcal{S})} \Big].$$

Lemma 3 *There exists one strictly positive constant* C' *independent of* ε *such that* $\forall (v_0, v_1, v_2) \in \mathcal{V}$,

$$F^{3D}(v) \leq C' \Big[\varepsilon \big(\|\underline{\nabla} v_0\|_{L^2(\mathcal{S})} + \|v_1\|_{L^2(\mathcal{S})} \big) + \varepsilon^2 \big(\|\underline{\nabla} v_1\|_{L^2(\mathcal{S})} + \|v_2\|_{L^2(\mathcal{S})} \big) + \varepsilon^3 \|\underline{\nabla} v_2\|_{L^2(\mathcal{S})} \Big].$$

We directly have Lemma 3. We must prove Lemma 2.

Proof Using Propositions 2 and 3, we have $\forall (v_0, v_1, v_2) \in \mathcal{V}$,

$$\begin{aligned}
A^{3D}(v,v) &= \sigma^t \int_\Omega \vec{\nabla} v \cdot \vec{\nabla} v \, dV + (\sigma^l - \sigma^t) \int_\Omega \big(\vec{\tau} \otimes \vec{\tau} \cdot \vec{\nabla} v\big) \cdot \vec{\nabla} v \, dV \\
&\geq \sigma^t \int_\Omega \vec{\nabla} v \cdot \vec{\nabla} v \, dV \\
&\geq \gamma \, \sigma^t \int_\Omega \vec{\nabla} v \cdot \vec{\nabla} v \sqrt{a} \; d\xi \\
&\geq \gamma \, \sigma^t \int_\Omega \Big(\sum_{\alpha\beta=1,2} g^{\alpha\beta} \partial_\alpha v \partial_\beta v + v_1^2 + 4\xi^3 v_1 v_2 + 4\big(\xi^3\big)^2 v_2^2 \Big) \sqrt{a} \; d\xi \\
&\geq c \int_\Omega \Big(\sum_{\alpha\beta=1,2} a^{\alpha\beta} \partial_\alpha v \partial_\beta v + v_1^2 + 4\xi^3 v_1 v_2 + 4\big(\xi^3\big)^2 v_2^2 \Big) \sqrt{a} \; d\xi \\
&\geq c \int_\Omega \big(\underline{\nabla} v \cdot \underline{\nabla} v + v_1^2 + 4\xi^3 v_1 v_2 + 4\big(\xi^3\big)^2 v_2^2 \big) \sqrt{a} \; d\xi,
\end{aligned}$$

with Ω dependent on ε. We use the decomposition of v and we integrate over ξ^3

$$A^{3D}(v,v) \geq c \int_\omega \Big(\varepsilon (\underline{\nabla} v_0 \cdot \underline{\nabla} v_0 + v_1^2) + \frac{\varepsilon^3}{12} (\underline{\nabla} v_1 \cdot \underline{\nabla} v_1 + 4 v_2^2) + \frac{\varepsilon^3}{6} \underline{\nabla} v_0 \cdot \underline{\nabla} v_2 + \frac{\varepsilon^5}{80} \underline{\nabla} v_2 \cdot \underline{\nabla} v_2 \Big) \, dS.$$

Using the Young inequality on the cross term, we have for almost all $(\xi^1, \xi^2) \in \omega$, and for all $r > 0$,

$$2 \underline{\nabla} v_0(\xi^1, \xi^2) \cdot \underline{\nabla} v_2(\xi^1, \xi^2) \geq -\frac{1}{r} \underline{\nabla} v_0(\xi^1, \xi^2) \cdot \underline{\nabla} v_0(\xi^1, \xi^2) - r \underline{\nabla} v_2(\xi^1, \xi^2) \cdot \underline{\nabla} v_2(\xi^1, \xi^2).$$

With $r = \frac{1}{10} \varepsilon^2$, we obtain $\forall (v_0, v_1, v_2) \in \mathcal{V}$,

$$\begin{aligned}
A^{3D}(v,v) &\geq c \int_\omega \Big(\frac{\varepsilon}{6} \underline{\nabla} v_0 \cdot \underline{\nabla} v_0 + \varepsilon v_1^2 + \frac{\varepsilon^3}{12} (\underline{\nabla} v_1 \cdot \underline{\nabla} v_1 + 4 v_2^2) + \frac{\varepsilon^5}{240} \underline{\nabla} v_2 \cdot \underline{\nabla} v_2 \Big) \, dS \\
&\geq C \Big[\varepsilon \big(\|\underline{\nabla} v_0\|_{L^2(\mathcal{S})} + \|v_1\|_{L^2(\mathcal{S})} \big)^2 + \varepsilon^3 \big(\|\underline{\nabla} v_1\|_{L^2(\mathcal{S})} + \|v_2\|_{L^2(\mathcal{S})} \big)^2 \\
&\qquad + \varepsilon^5 \|\underline{\nabla} v_2\|^2_{L^2(\mathcal{S})} \Big].
\end{aligned}$$

□

These two lemmas directly allow to show the following result.

Lemma 4 *There exists a strictly positive constant C independent of ε such that*

$$\big(\|\underline{\nabla} u_0^\varepsilon\|_{L^2(\mathcal{S})} + \|u_1^\varepsilon\|_{L^2(\mathcal{S})} \big) + \varepsilon \big(\|\underline{\nabla} u_1^\varepsilon\|_{L^2(\mathcal{S})} + \|u_2^\varepsilon\|_{L^2(\mathcal{S})} \big) + \varepsilon^2 \|\underline{\nabla} u_2^\varepsilon\|_{L^2(\mathcal{S})} \leq C.$$

Using Lemma 4, we can quickly finish the first step. According to the Poincaré inequality (Γ_1 of non-zero measure) and Lemma 4, $u_0^\varepsilon, \varepsilon u_1^\varepsilon$ et $\varepsilon^2 u_2^\varepsilon$ are uniformly bounded in the $H^1(\mathcal{S})$-norm and u_1^ε and $\varepsilon u_2^\varepsilon$ are uniformly bounded in the $L^2(\mathcal{S})$-norm. We can then extract a subsequence (also denoted by ε) such that there exists $(u_0^m, u_1^m, \tilde{u}_1^m, u_2^m, \tilde{u}_2^m) \in H^1(\mathcal{S}) \times L^2(\mathcal{S}) \times H^1(\mathcal{S}) \times L^2(\mathcal{S}) \times H^1(\mathcal{S})$ such that $(u_0^\varepsilon, u_1^\varepsilon, \varepsilon u_1^\varepsilon, \varepsilon u_2^\varepsilon, \varepsilon^2 u_2^\varepsilon)$ converges weakly to $(u_0^m, u_1^m, \tilde{u}_1^m, u_2^m, \tilde{u}_2^m)$ in $H^1(\mathcal{S}) \times L^2(\mathcal{S}) \times H^1(\mathcal{S}) \times L^2(\mathcal{S}) \times H^1(\mathcal{S})$.

Remark 10 *We can show that $\tilde{u}_1^m = 0$, since by compact injection $\varepsilon u_1^\varepsilon \overset{L^2(\mathcal{S})}{\to} \tilde{u}_1^m$. Likewise, $\tilde{u}_2^m = 0$.*

(ii) Convergence to the solution of the limit problem

Lemma 5 $\forall (v_0, v_1, v_2) \in \mathcal{V}$,

- $\dfrac{1}{d} A^{3D}(u^\varepsilon, v) = A_l\big((u_0^\varepsilon, u_1^\varepsilon), (v_0, v_1)\big) + A_l^f(u_0^\varepsilon, v_0) + O(\varepsilon)$,
- $\dfrac{1}{d} F^{3D}(v) = F_0(v_0) + O(\varepsilon)$.

With Lemma 5 we can quickly conclude. According to Theorem 2, indeed, we have $\forall v \in \mathcal{V}$, $A^{3D}(u^\varepsilon, v) = F^{3D}(v)$ and Lemma 5 gives

$$A_l\big((u_0^\varepsilon, u_1^\varepsilon), (v_0, v_1)\big) + A_l^f(u_0^\varepsilon, v_0) = F_0(v_0) + O(\varepsilon), \quad \forall (v_0, v_1) \in \mathcal{V}_{lim}. \tag{12}$$

Making ε tend to 0 in (12) and according to the first step, we have

$$A_l\big((u_0^m, u_1^m), (v_0, v_1)\big) + A_l^f(u_0^m, v_0) = F_0(v_0), \quad \forall (v_0, v_1) \in \mathcal{V}_{lim}.$$

Using Theorem 3, we obtain that $(u_0^m, u_1^m) = (u_0^l, 0)$ where u_0^l verifies Problem (11). The limits $(u_0^m, u_1^m, \tilde{u}_1^m, \tilde{u}_2^m) = (u_0^l, 0, 0, 0)$ are independent of the subsequence so that we can infer the theorem. To complete the proof of Theorem 4, we must now show Lemma 5.

Proof

$$A^{3D}(u^\varepsilon, v) = \sigma^t \int_\Omega \vec{\nabla} u^\varepsilon \cdot \vec{\nabla} v \, dV + (\sigma^l - \sigma^t) \int_\Omega \big(\vec{\tau} \otimes \vec{\tau} \cdot \vec{\nabla} u^\varepsilon\big) \cdot \vec{\nabla} v \, dV, \quad \forall (v_0, v_1, v_2) \in \mathcal{V},$$

with Ω dependent on ε. We start by studying the first term of $A^{3D}(u^\varepsilon, v)$, i.e.

$$\sigma^t \int_\Omega \vec{\nabla} u^\varepsilon \cdot \vec{\nabla} v \, dV = \sigma^t \int_\Omega \vec{\nabla} u^\varepsilon \cdot \vec{\nabla} v \sqrt{a} \, d\xi + \sigma^t \int_\Omega \xi^3 \vec{\nabla} u^\varepsilon \cdot \vec{\nabla} v \sqrt{a} (-2H + K\xi^3) d\xi.$$

If we can show that $\sigma^t \displaystyle\int_\Omega \vec{\nabla} u^\varepsilon \cdot \vec{\nabla} v \sqrt{a} \, d\xi$ converges when ε tends to 0, then $\sigma^t \displaystyle\int_\Omega \xi^3 \vec{\nabla} u^\varepsilon \cdot \vec{\nabla} v \sqrt{a} (-2H + K\xi^3) d\xi$ converges to 0 when ε tends to 0, so that we just need to study the first term.

$$\begin{aligned}
&\sigma^t \int_\Omega \vec{\nabla} u^\varepsilon \cdot \vec{\nabla} v \sqrt{a} \, d\xi \\
&= \sigma^t \int_\Omega \Big(\sum_{\alpha\beta=1,2} g^{\alpha\beta} \partial_\alpha u^\varepsilon \partial_\beta v + u_1^\varepsilon v_1 + 2\xi^3 \big(u^\varepsilon v_2 + u_2^\varepsilon v_1\big) + 4\big(\xi^3\big)^2 u_2^\varepsilon v_2 \Big) \sqrt{a} \, d\xi \\
&= \sigma^t \int_\Omega \Big(\sum_{\alpha\beta=1,2} a^{\alpha\beta} \partial_\alpha u^\varepsilon \partial_\beta v + u_1^\varepsilon v_1 + 2\xi^3 \big(u_1^\varepsilon v_2 + u_2^\varepsilon v_1\big) + 4\big(\xi^3\big)^2 u_2^\varepsilon v_2 \Big) \sqrt{a} \, d\xi \\
&\qquad + \sigma^t \int_\Omega \sum_{\alpha\beta=1,2} \xi^3 \bar{g}^{\alpha\beta} \partial_\alpha u^\varepsilon \partial_\beta v \sqrt{a} \, d\xi.
\end{aligned}$$

For the same reason, we just need to show that $\sigma^t \displaystyle\int_\Omega \Big(\underline{\nabla} u^\varepsilon \cdot \underline{\nabla} v + u_1^\varepsilon v_1 + 2\xi^3 \big(u_1^\varepsilon v_2 + u_2^\varepsilon v_1\big) + 4\big(\xi^3\big)^2 u_2^\varepsilon v_2 \Big) \sqrt{a} \, d\xi$ converges when ε tends to 0.

We can apply the same method with the second term of $A^{3D}(u^\varepsilon, v)$ and we infer that we only need to show that $(\sigma^l - \sigma^t)\int_\Omega \big(\vec{\tau}\otimes\vec{\tau}\cdot\underline{\nabla}u^\varepsilon\big)\cdot\underline{\nabla}v\sqrt{a}\,d\xi$ converges when ε tends to 0. We define A_0^{3D} by $\forall (u_0,u_1,u_2),(v_0,v_1,v_2)\in\mathcal{V}$,

$$A_0^{3D}(u,v) = \sigma^t\int_\Omega \big(\underline{\nabla}u\cdot\underline{\nabla}v + u_1v_1 + 2\xi^3\big(u_1v_2+u_2v_1\big) + 4\big(\xi^3\big)^2u_2v_2\big)\sqrt{a}\,d\xi \\ + (\sigma^l-\sigma^t)\int_\Omega\big(\vec{\tau}\otimes\vec{\tau}\cdot\underline{\nabla}u\big)\cdot\underline{\nabla}v\sqrt{a}\,d\xi^3,$$

and we want to show that

$$\frac{1}{d}A_0^{3D}(u^\varepsilon,v) = A_l((u_0^\varepsilon,u_1^\varepsilon),(v_0,v_1)) + A_l^f(u_0^\varepsilon,v_0) + O(\varepsilon),\quad \forall (v_0,v_1,v_2)\in\mathcal{V}.$$

We decompose $A_0^{3D}(u^\varepsilon,v)$ as $A_0^{3D}(u^\varepsilon,v) = \sum_{n=0}^4 I_n(u^\varepsilon,v)$ with

$$\frac{1}{\varepsilon}I_0(u^\varepsilon,v) = L\Big(A_l\big((u_0^\varepsilon,u_1^\varepsilon),(v_0,v_1)\big) + A_l^f(u_0^\varepsilon,v_0)\Big),$$

$$\frac{1}{\varepsilon}I_1(u^\varepsilon,v) = \varepsilon L^2(\sigma^l-\sigma^t)\int_\omega L_1(\theta)\Big(\big(\underline{\tau}_0^\perp\otimes\underline{\tau}_0\cdot\underline{\nabla}u_0^\varepsilon\big)\cdot\underline{\nabla}v_1 \\ + \big(\underline{\tau}_0\otimes\underline{\tau}_0^\perp\cdot\underline{\nabla}u_0^\varepsilon\big)\cdot\underline{\nabla}v_1\Big)\,dS \\ + \varepsilon L^2(\sigma^l-\sigma^t)\int_\omega L_1(\theta)\Big(\big(\underline{\tau}_0^\perp\otimes\underline{\tau}_0\cdot\underline{\nabla}u_1^\varepsilon\big)\cdot\underline{\nabla}v_0 + \big(\underline{\tau}_0\otimes\underline{\tau}_0^\perp\cdot\underline{\nabla}u_1^\varepsilon\big)\cdot\underline{\nabla}v_0\Big)\,dS,$$

$$\frac{1}{\varepsilon}I_2(u^\varepsilon,v) = \frac{1}{12}\varepsilon^2L^3\sigma^t\int_\omega \underline{\nabla}u_0^\varepsilon\cdot\underline{\nabla}v_2 + \underline{\nabla}u_2^\varepsilon\cdot\underline{\nabla}v_0 + \underline{\nabla}u_1^\varepsilon\cdot\underline{\nabla}v_1 + 4u_2^\varepsilon v_2\,dS \\ + \varepsilon^2L^3(\sigma^l-\sigma^t)\int_\omega I_2(\theta)\big(\underline{\tau}_0\otimes\underline{\tau}_0\cdot\underline{\nabla}u_0^\varepsilon\big)\cdot\underline{\nabla}v_2 + J_2(\theta)\big(\underline{\tau}_0^\perp\otimes\underline{\tau}_0^\perp\cdot\underline{\nabla}u_0^\varepsilon\big)\cdot\underline{\nabla}v_2\,dS \\ + \varepsilon^2L^3(\sigma^l-\sigma^t)\int_\omega I_2(\theta)\big(\underline{\tau}_0\otimes\underline{\tau}_0\cdot\underline{\nabla}u_2^\varepsilon\big)\cdot\underline{\nabla}v_0 + J_2(\theta)\big(\underline{\tau}_0^\perp\otimes\underline{\tau}_0^\perp\cdot\underline{\nabla}u_2^\varepsilon\big)\cdot\underline{\nabla}v_0\,dS \\ + \varepsilon^2L^3(\sigma^l-\sigma^t)\int_\omega I_2(\theta)\big(\underline{\tau}_0\otimes\underline{\tau}_0\cdot\underline{\nabla}u_1^\varepsilon\big)\cdot\underline{\nabla}v_1 + J_2(\theta)\big(\underline{\tau}_0^\perp\otimes\underline{\tau}_0^\perp\cdot\underline{\nabla}u_1^\varepsilon\big)\cdot\underline{\nabla}v_1\,dS,$$

$$\frac{1}{\varepsilon}I_3(u^\varepsilon,v) = \varepsilon^3L^4(\sigma^l-\sigma^t)\int_\omega L_3(\theta)\Big(\big(\underline{\tau}_0^\perp\otimes\underline{\tau}_0\cdot\underline{\nabla}u_1^\varepsilon\big)\cdot\underline{\nabla}v_2 \\ + \big(\underline{\tau}_0\otimes\underline{\tau}_0^\perp\cdot\underline{\nabla}u_1^\varepsilon\big)\cdot\underline{\nabla}v_2\Big)\,dS \\ + \varepsilon^3L^4(\sigma^l-\sigma^t)\int_\omega L_3(\theta)\Big(\big(\underline{\tau}_0^\perp\otimes\underline{\tau}_0\cdot\underline{\nabla}u_2^\varepsilon\big)\cdot\underline{\nabla}v_1 + \big(\underline{\tau}_0\otimes\underline{\tau}_0^\perp\cdot\underline{\nabla}u_2^\varepsilon\big)\cdot\underline{\nabla}v_1\Big)\,dS,$$

$$\frac{1}{\varepsilon}I_4(u^\varepsilon,v) = \frac{1}{80}\varepsilon^4L^5(\sigma^l-\sigma^t)\int_\omega \underline{\nabla}u_2^\varepsilon\cdot\underline{\nabla}v_2\,dS \\ + \varepsilon^4L^5(\sigma^l-\sigma^t)\int_\omega I_4(\theta)\big(\underline{\tau}_0\otimes\underline{\tau}_0\cdot\underline{\nabla}u_2^\varepsilon\big)\cdot\underline{\nabla}v_2 + J_4(\theta)\big(\underline{\tau}_0^\perp\otimes\underline{\tau}_0^\perp\cdot\underline{\nabla}u_2^\varepsilon\big)\cdot\underline{\nabla}v_2\,dS.$$

We have that

- $\varepsilon \int_\omega L_1(\theta)\Big(\big(\underline{\tau}_0^\perp \otimes \underline{\tau}_0 \cdot \underline{\nabla} u_0^\varepsilon\big) \cdot \underline{\nabla} v_1 + \big(\underline{\tau}_0 \otimes \underline{\tau}_0^\perp \cdot \underline{\nabla} u_0^\varepsilon\big) \cdot \underline{\nabla} v_1 \Big)\, dS \underset{\varepsilon\to 0}{\to} 0$, because $\|u_0^\varepsilon\|_{H^1(\mathcal{S})}$ is bounded,

- $\varepsilon \int_\omega L_1(\theta)\Big(\big(\underline{\tau}_0^\perp \otimes \underline{\tau}_0 \cdot \underline{\nabla} u_1^\varepsilon\big) \cdot \underline{\nabla} v_0 + \big(\underline{\tau}_0 \otimes \underline{\tau}_0^\perp \cdot \underline{\nabla} u_1^\varepsilon\big) \cdot \underline{\nabla} v_0 \Big)\, dS \underset{\varepsilon\to 0}{\to} 0$, because $\tilde{u}_1^m = 0$.

We infer that $\frac{1}{\varepsilon} I_1(u^\varepsilon, v) \underset{\varepsilon\to 0}{\to} 0$, $\forall (v_0, v_1, v_2) \in \mathcal{V}$. With similar arguments, we can show that $\frac{1}{\varepsilon} I_2(u^\varepsilon, v)$, $\frac{1}{\varepsilon} I_3(u^\varepsilon, v)$ and $\frac{1}{\varepsilon} I_4(u^\varepsilon, v)$ all converge to 0 when ε tends to 0, for all $(v_0, v_1, v_2) \in \mathcal{V}$. We clearly have $\frac{1}{\varepsilon} F^{3D}(v) \underset{\varepsilon\to 0}{\to} F_0(v_0), \forall (v_0, v_1, v_2) \in \mathcal{V}$, and this concludes the proof of the theorem. □

4.3.3 Strong convergence

We will complete our asymptotic analysis by proving that the convergences established in the previous section also hold – in fact – in the strong sense. Defining the two bilinear forms

$$\tilde{A}_l^f(\underline{\rho}, \underline{\mu}) = (\sigma^l - \sigma^t) \int_\omega I_0(\theta) \left(\underline{\tau}_0 \otimes \underline{\tau}_0 \cdot \underline{\rho}\right) \cdot \underline{\mu}\, dS,$$

$$\bar{A}_l^f(\underline{\rho}, \underline{\mu}) = (\sigma^l - \sigma^t) \int_\omega J_0(\theta) \left(\underline{\tau}_0^\perp \otimes \underline{\tau}_0^\perp \cdot \underline{\rho}\right) \cdot \underline{\mu}\, dS,$$

and the rotation operator

$$\alpha : \left\{ \begin{array}{ccc} \mathrm{Vect}(\vec{a}_1, \vec{a}_2) & \longrightarrow & \mathrm{Vect}(\vec{a}_1, \vec{a}_2) \\ \underline{\tau} & \longmapsto & \vec{a}_3 \times \underline{\tau} \end{array} \right.$$

we will show that the following sum of perfect squares

$$\begin{aligned} D_\varepsilon = A_l\big(u_0^\varepsilon + \frac{1}{12} d^2 u_2^\varepsilon - u_0^l, u_1^\varepsilon\big)^2 + \tilde{A}_l^f\Big(\underline{\nabla}\big(u_0^\varepsilon + \frac{I_2}{I_0} d^2 u_2^\varepsilon\big) - \frac{L_1}{I_0} d\, \alpha\big(\underline{\nabla} u_1^\varepsilon\big) - \underline{\nabla} u_0^l\Big)^2 \\ + \bar{A}_l^f\Big(\underline{\nabla}\big(u_0^\varepsilon + \frac{I_2}{J_0} d^2 u_2^\varepsilon\big) + \frac{L_1}{J_0} d\, \alpha\big(\underline{\nabla} u_1^\varepsilon\big) - \underline{\nabla} u_0^l\Big)^2, \end{aligned}$$

tends to zero with ε, where we denote $A_l(u, v)^2 = A_l\big((u, v), (u, v)\big)$, and similarly for $\tilde{A}_l^f$ and $\bar{A}_l^f$. In fact, D_ε is constructed so that the expression

$$\begin{aligned} D'_\varepsilon = A_l\big(u_0^\varepsilon + \frac{1}{12} d^2 u_2^\varepsilon, u_1^\varepsilon\big)^2 + \tilde{A}_l^f\Big(\underline{\nabla}\big(u_0^\varepsilon + \frac{I_2}{I_0} d^2 u_2^\varepsilon\big) - \frac{L_1}{I_0} d\, \alpha\big(\underline{\nabla} u_1^\varepsilon\big)\Big)^2 \\ + \bar{A}_l^f\Big(\underline{\nabla}\big(u_0^\varepsilon + \frac{I_2}{J_0} d^2 u_2^\varepsilon\big) + \frac{L_1}{J_0} d\, \alpha\big(\underline{\nabla} u_1^\varepsilon\big)\Big)^2, \end{aligned}$$

appearing when developing D_ε, gathers the main lower-order terms in the expansion $A^{3D}(u^\varepsilon, u^\varepsilon)$ as expressed in the following lemma, shown like in the proof of Lemma 5.

Lemma 6

$$
\begin{aligned}
A^{3D}(u^\varepsilon,u^\varepsilon) = d\sigma^t\Bigg[&\int_\omega \underline{\nabla} u_0^\varepsilon\cdot\underline{\nabla} u_0^\varepsilon + (u_1^\varepsilon)^2\,dS\\
&+\frac{d^2}{12}\int_\omega 2\underline{\nabla} u_0^\varepsilon\cdot\underline{\nabla} u_2^\varepsilon + \underline{\nabla} u_1^\varepsilon\cdot\underline{\nabla} u_1^\varepsilon + (u_2^\varepsilon)^2\,dS + \frac{d^4}{80}\int_\omega \underline{\nabla} u_2^\varepsilon\cdot\underline{\nabla} u_2^\varepsilon\,dS\Bigg]\\
&+d(\sigma^l-\sigma^t)\int_\omega I_0(\theta)\big(\underline{\tau}_0\otimes\underline{\tau}_0\cdot\underline{\nabla} u_0^\varepsilon\big)\cdot\underline{\nabla} u_0^\varepsilon + J_0(\theta)\big(\underline{\tau}_0^\perp\otimes\underline{\tau}_0^\perp\cdot\underline{\nabla} u_0^\varepsilon\big)\cdot\underline{\nabla} u_0^\varepsilon\,dS\\
&+d(\sigma^l-\sigma^t)\Bigg[d\int_\omega 2L_1(\theta)\big(\underline{\tau}_0\otimes\underline{\tau}_0^\perp\cdot\underline{\nabla} u_0^\varepsilon\big)\cdot\underline{\nabla} u_1^\varepsilon + 2L_1(\theta)\big(\underline{\tau}_0^\perp\otimes\underline{\tau}_0\cdot\underline{\nabla} u_0^\varepsilon\big)\cdot\underline{\nabla} u_1^\varepsilon\,dS\Bigg]\\
&+d(\sigma^l-\sigma^t)\Bigg[d^2\int_\omega 2I_2(\theta)\big(\underline{\tau}_0\otimes\underline{\tau}_0\cdot\underline{\nabla} u_0^\varepsilon\big)\cdot\underline{\nabla} u_2^\varepsilon + 2J_2(\theta)\big(\underline{\tau}_0^\perp\otimes\underline{\tau}_0^\perp\cdot\underline{\nabla} u_0^\varepsilon\big)\cdot\underline{\nabla} u_2^\varepsilon\,dS\Bigg]\\
&+d(\sigma^l-\sigma^t)\Bigg[d^2\int_\omega I_2(\theta)\big(\underline{\tau}_0\otimes\underline{\tau}_0\cdot\underline{\nabla} u_1^\varepsilon\big)\cdot\underline{\nabla} u_1^\varepsilon + J_2(\theta)\big(\underline{\tau}_0^\perp\otimes\underline{\tau}_0^\perp\cdot\underline{\nabla} u_1^\varepsilon\big)\cdot\underline{\nabla} u_1^\varepsilon\,dS\Bigg]\\
&+d(\sigma^l-\sigma^t)L\Bigg[d^3\int_\omega 2L_3(\theta)\big(\underline{\tau}_0\otimes\underline{\tau}_0^\perp\cdot\underline{\nabla} u_1^\varepsilon\big)\cdot\underline{\nabla} u_2^\varepsilon + 2L_3(\theta)\big(\underline{\tau}_0^\perp\otimes\underline{\tau}_0\cdot\underline{\nabla} u_1^\varepsilon\big)\cdot\underline{\nabla} u_2^\varepsilon\,dS\Bigg]\\
&+d(\sigma^l-\sigma^t)L\Bigg[d^4\int_\omega I_4(\theta)\big(\underline{\tau}_0\otimes\underline{\tau}_0\cdot\underline{\nabla} u_2^\varepsilon\big)\cdot\underline{\nabla} u_2^\varepsilon + J_4(\theta)\big(\underline{\tau}_0^\perp\otimes\underline{\tau}_0^\perp\cdot\underline{\nabla} u_2^\varepsilon\big)\cdot\underline{\nabla} u_2^\varepsilon\,dS\Bigg]\\
&+R(u^\varepsilon),
\end{aligned}
$$

where the degree of ξ^3 *in the terms of* $\dfrac{1}{d}R(u^\varepsilon)=\dfrac{1}{L\varepsilon}R(u^\varepsilon)$ *is high enough to have the convergence to* 0 *when* ε *tends to* 0.

We can now state and prove our final convergence result.

Theorem 5 *We have the following convergences:*

- $(u_0^\varepsilon,u_1^\varepsilon)$ *converges strongly to* $(u_0^l,0)$ *in* $\mathcal{V}_{lim}$ *when* ε *tends to* 0,
- $\varepsilon u_1^\varepsilon$ *and* $\varepsilon^2 u_2^\varepsilon$ *converge strongly to* 0 *in* $H^1(\mathcal{S})$ *when* ε *tends to* 0.

In addition, $\varepsilon u_2^\varepsilon$ *converges strongly to* 0 *in* $L^2(\mathcal{S})$ *when* ε *tends to* 0.

Proof Defining

$$
\begin{aligned}
\tilde{D}_\varepsilon = \frac{1}{12}d^2\sigma^t\int_\omega \underline{\nabla} u_1^\varepsilon\cdot\underline{\nabla} u_1^\varepsilon + (u_2^\varepsilon)^2\,dS + \frac{1}{180}d^4\sigma^t\int_\omega \underline{\nabla} u_2^\varepsilon\cdot\underline{\nabla} u_2^\varepsilon\,dS\\
+d^2(\sigma^l-\sigma^t)\Bigg[\int_\omega\Big(I_2-\frac{L_1^2}{J_0}\Big)\big(\underline{\tau}_0\otimes\underline{\tau}_0\cdot\underline{\nabla} u_1^\varepsilon\big)\cdot\underline{\nabla} u_1^\varepsilon\,dS\\
+\int_\omega\Big(J_2-\frac{L_1^2}{I_0}\Big)\big(\underline{\tau}_0^\perp\otimes\underline{\tau}_0^\perp\cdot\underline{\nabla} u_1^\varepsilon\big)\cdot\underline{\nabla} u_1^\varepsilon\,dS\Bigg]\\
+2d^3(\sigma^l-\sigma^t)\Bigg[\int_\omega\Big(L_3-L_1\frac{I_2}{I_0}\Big)\big(\underline{\tau}_0\otimes\underline{\tau}_0^\perp\cdot\underline{\nabla} u_1^\varepsilon\big)\cdot\underline{\nabla} u_2^\varepsilon\,dS\\
+\int_\omega\Big(L_3-L_1\frac{J_2}{J_0}\Big)\big(\underline{\tau}_0^\perp\otimes\underline{\tau}_0\cdot\underline{\nabla} u_1^\varepsilon\big)\cdot\underline{\nabla} u_2^\varepsilon\,dS\Bigg]\\
+d^4(\sigma^l-\sigma^t)\Bigg[\int_\omega\Big(I_4-\frac{I_2^2}{I_0}\Big)\big(\underline{\tau}_0\otimes\underline{\tau}_0\cdot\underline{\nabla} u_2^\varepsilon\big)\cdot\underline{\nabla} u_2^\varepsilon\,dS\\
+\int_\omega\Big(J_4-\frac{J_2^2}{J_0}\Big)\big(\underline{\tau}_0^\perp\otimes\underline{\tau}_0^\perp\cdot\underline{\nabla} u_2^\varepsilon\big)\cdot\underline{\nabla} u_2^\varepsilon\,dS\Bigg],
\end{aligned}
$$

we will decompose the proof into two steps:

- in the first step, we show that $D_\varepsilon + \tilde{D}_\varepsilon \underset{\varepsilon\to 0}{\to} 0$,
- in the second step, we show that $\tilde{D}_\varepsilon \geq \eta\big(\|\varepsilon u_1^\varepsilon\|^2_{H^1(\mathcal{S})} + \|\varepsilon u_2^\varepsilon\|^2_{L^2(\mathcal{S})} + \|\varepsilon^2 u_2^\varepsilon\|^2_{H^1(\mathcal{S})}\big)$, with $\eta > 0$.

The conclusion is then at hand, because it implies that $D_\varepsilon \underset{\varepsilon\to 0}{\to} 0$ and $\tilde{D}_\varepsilon \underset{\varepsilon\to 0}{\to} 0$, and in particular we have the following convergences

$$A_l\big(u_0^\varepsilon + \frac{1}{3}d^2 u_2^\varepsilon - u_0^l, u_1^\varepsilon\big)^2 \underset{\varepsilon\to 0}{\to} 0,$$

$$\|\varepsilon u_1^\varepsilon\|_{H^1(\mathcal{S})} \underset{\varepsilon\to 0}{\to} 0, \quad \|\varepsilon u_2^\varepsilon\|_{L^2(\mathcal{S})} \underset{\varepsilon\to 0}{\to} 0, \quad \|\varepsilon^2 u_2^\varepsilon\|_{H^1(\mathcal{S})} \underset{\varepsilon\to 0}{\to} 0.$$

We infer

$$u_0^\varepsilon \overset{H^1(\mathcal{S})}{\underset{\varepsilon\to 0}{\to}} u_0^l, \quad u_1^\varepsilon \overset{L^2(\mathcal{S})}{\underset{\varepsilon\to 0}{\to}} 0, \quad \varepsilon u_1^\varepsilon \overset{H^1(\mathcal{S})}{\underset{\varepsilon\to 0}{\to}} 0, \quad \varepsilon u_2^\varepsilon \overset{L^2(\mathcal{S})}{\underset{\varepsilon\to 0}{\to}} 0, \quad \varepsilon^2 u_2^\varepsilon \overset{H^1(\mathcal{S})}{\underset{\varepsilon\to 0}{\to}} 0.$$

First step We can develop D_ε as

$$\begin{aligned} D_\varepsilon = D'_\varepsilon - 2A_l\Big(\big(u_0^\varepsilon + \frac{1}{12}d^2 u_2^\varepsilon, 0\big), (u_0^l, 0)\Big) \\ - 2\tilde{A}_l^f\Big(\underline{\nabla}\big(u_0^\varepsilon + \frac{I_2}{I_0}d^2 u_2^\varepsilon\big) - \frac{L_1}{I_0}d\,\alpha\big(\underline{\nabla}u_1^\varepsilon\big), \underline{\nabla}u_0^l\Big) \\ - 2\bar{A}_l^f\Big(\underline{\nabla}\big(u_0^\varepsilon + \frac{I_2}{J_0}d^2 u_2^\varepsilon\big) + \frac{L_1}{J_0}d\,\alpha\big(\underline{\nabla}u_1^\varepsilon\big), \underline{\nabla}u_0^l\Big) \\ + A_l\big((u_0^l, 0), (u_0^l, 0)\big) + \tilde{A}_l^f\big(\underline{\nabla}u_0^l, \underline{\nabla}u_0^l\big) + \bar{A}_l^f\big(\underline{\nabla}u_0^l, \underline{\nabla}u_0^l\big), \end{aligned}$$

We remark that $\tilde{A}_l^f\big(\underline{\nabla}u_0^l, \underline{\nabla}u_0^l\big) + \bar{A}_l^f\big(\underline{\nabla}u_0^l, \underline{\nabla}u_0^l\big) = A_l^f\big(u_0^l, u_0^l\big)$, and according to Theorem 3, we have $A_l\big((u_0^l, 0), (u_0^l, 0)\big) + A_l^f\big(u_0^l, u_0^l\big) = F_0(u_0^l)$. Furthermore, according to Theorem 4, we have

$$\begin{aligned} A_l\Big(\big(u_0^\varepsilon + \frac{1}{3}d^2 u_2^\varepsilon, 0\big), (u_0^l, 0)\Big) + \tilde{A}_l^f\Big(\underline{\nabla}\big(u_0^\varepsilon + \frac{I_2}{I_0}d^2 u_2^\varepsilon\big) - \frac{L_1}{I_0}d\,\alpha\big(\underline{\nabla}u_1^\varepsilon\big), \underline{\nabla}u_0^l\Big) \\ + \bar{A}_l^f\Big(\underline{\nabla}\big(u_0^\varepsilon + \frac{I_2}{J_0}d^2 u_2^\varepsilon\big) + \frac{L_1}{J_0}d\,\alpha\big(\underline{\nabla}u_1^\varepsilon\big), \underline{\nabla}u_0^l\Big) \\ \underset{\varepsilon\to 0}{\to} A_l\big(u_0^l, u_0^l\big) + A_l^f\big(u_0^l, u_0^l\big) = F_0(u_0^l). \end{aligned}$$

We infer

$$D_\varepsilon = D'_\varepsilon - F_0(u_0^l) + O(\varepsilon). \tag{13}$$

Developing D'_ε and using Lemma 6, we have

$$\begin{aligned}
D'_\varepsilon &= \frac{1}{d}A^{3D}(u^\varepsilon,u^\varepsilon)-\frac{1}{d}R(u^\varepsilon)\\
&\quad -\frac{1}{12}d^2\sigma^t\int_\omega \underline{\nabla}u_1^\varepsilon\cdot\underline{\nabla}u_1^\varepsilon+(u_2^\varepsilon)^2\,dS-\frac{1}{180}d^4\sigma^t\int_\omega \underline{\nabla}u_2^\varepsilon\cdot\underline{\nabla}u_2^\varepsilon\,dS\\
&\quad -d^2(\sigma^l-\sigma^t)\Big[\int_\omega \big(I_2-\frac{L_1^2}{J_0}\big)\underline{\tau}_0\otimes\underline{\tau}_0\cdot\underline{\nabla}u_1^\varepsilon\cdot\underline{\nabla}u_1^\varepsilon\,dS\\
&\qquad\qquad +\int_\omega \big(J_2-\frac{L_1^2}{I_0}\big)\underline{\tau}_0^\perp\otimes\underline{\tau}_0^\perp\cdot\underline{\nabla}u_1^\varepsilon\cdot\underline{\nabla}u_1^\varepsilon\,dS\Big]\\
&\quad -2d^3(\sigma^l-\sigma^t)\Big[\int_\omega \big(L_3-L_1\frac{I_2}{I_0}\big)\underline{\tau}_0\otimes\underline{\tau}_0^\perp\cdot\underline{\nabla}u_1^\varepsilon\cdot\underline{\nabla}u_2^\varepsilon\,dS\\
&\qquad\qquad +\int_\omega \big(L_3-L_1\frac{J_2}{J_0}\big)\underline{\tau}_0^\perp\otimes\underline{\tau}_0\cdot\underline{\nabla}u_1^\varepsilon\cdot\underline{\nabla}u_2^\varepsilon\,dS\Big]\\
&\quad -d^4(\sigma^l-\sigma^t)\Big[\int_\omega \big(I_4-\frac{I_2^2}{I_0}\big)\underline{\tau}_0\otimes\underline{\tau}_0\cdot\underline{\nabla}u_2^\varepsilon\cdot\underline{\nabla}u_2^\varepsilon\,dS\\
&\qquad\qquad +\int_\omega \big(J_4-\frac{J_2^2}{J_0}\big)\underline{\tau}_0^\perp\otimes\underline{\tau}_0^\perp\cdot\underline{\nabla}u_2^\varepsilon\cdot\underline{\nabla}u_2^\varepsilon\,dS\Big]\\
&= \frac{1}{d}F^{3D}(u^\varepsilon)-\frac{1}{d}R(u^\varepsilon)-\tilde{D}_\varepsilon\\
&= F_0(u_0^l)-\tilde{D}_\varepsilon+O(\varepsilon).
\end{aligned}$$

The equation (13) then becomes $D_\varepsilon+\tilde{D}_\varepsilon=O(\varepsilon)$, which concludes the first step.

Second step We want to show that $\tilde{D}_\varepsilon\geq\eta\big(\|\varepsilon u_1^\varepsilon\|^2_{H^1(\mathcal{S})}+\|\varepsilon u_2^\varepsilon\|^2_{L^2(\mathcal{S})}+\|\varepsilon^2 u_2^\varepsilon\|^2_{H^1(\mathcal{S})}\big)$.

$$\begin{aligned}
\tilde{D}_\varepsilon &= \frac{1}{12}d^2\sigma^t\int_\omega\big(\underline{\nabla}u_1^\varepsilon\cdot\underline{\nabla}u_1^\varepsilon+(u_2^\varepsilon)^2\big)\,dS+\frac{1}{180}d^4\sigma^t\int_\omega\underline{\nabla}u_2^\varepsilon\cdot\underline{\nabla}u_2^\varepsilon\,dS\\
&\quad +d^2(\sigma^l-\sigma^t)\Big[\int_\omega\big(I_2-\frac{L_1^2}{J_0}\big)\underline{\tau}_0\otimes\underline{\tau}_0\cdot\underline{\nabla}u_1^\varepsilon\cdot\underline{\nabla}u_1^\varepsilon\,dS\\
&\qquad\qquad +\int_\omega\big(J_2-\frac{L_1^2}{I_0}\big)\underline{\tau}_0^\perp\otimes\underline{\tau}_0^\perp\cdot\underline{\nabla}u_1^\varepsilon\cdot\underline{\nabla}u_1^\varepsilon\,dS\Big]\\
&\quad +2d^3(\sigma^l-\sigma^t)\Big[\int_\omega\big(L_3-L_1\frac{I_2}{I_0}\big)\underline{\tau}_0\otimes\underline{\tau}_0^\perp\cdot\underline{\nabla}u_1^\varepsilon\cdot\underline{\nabla}u_2^\varepsilon\,dS\\
&\qquad\qquad +\int_\omega\big(L_3-L_1\frac{J_2}{J_0}\big)\underline{\tau}_0^\perp\otimes\underline{\tau}_0\cdot\underline{\nabla}u_1^\varepsilon\cdot\underline{\nabla}u_2^\varepsilon\,dS\Big]\\
&\quad +d^4(\sigma^l-\sigma^t)\Big[\int_\omega\big(I_4-\frac{I_2^2}{I_0}\big)\underline{\tau}_0\otimes\underline{\tau}_0\cdot\underline{\nabla}u_2^\varepsilon\cdot\underline{\nabla}u_2^\varepsilon\,dS\\
&\qquad\qquad +\int_\omega\big(J_4-\frac{J_2^2}{J_0}\big)\underline{\tau}_0^\perp\otimes\underline{\tau}_0^\perp\cdot\underline{\nabla}u_2^\varepsilon\cdot\underline{\nabla}u_2^\varepsilon\,dS\Big].
\end{aligned}$$

We have

$$\begin{aligned}
&\frac{1}{3}d^2\sigma^t\int_\omega\big(\underline{\nabla}u_1^\varepsilon\cdot\underline{\nabla}u_1^\varepsilon+(u_2^\varepsilon)^2\big)\,dS+\frac{4}{45}d^4\sigma^t\int_\omega\underline{\nabla}u_2^\varepsilon\cdot\underline{\nabla}u_2^\varepsilon\,dS\\
&\qquad\qquad\geq\eta\big(\|\varepsilon u_1^\varepsilon\|^2_{H^1(\mathcal{S})}+\|\varepsilon u_2^\varepsilon\|^2_{L^2(\mathcal{S})}+\|\varepsilon^2 u_2^\varepsilon\|^2_{H^1(\mathcal{S})}\big),
\end{aligned}$$

so that the result holds if we can show the positiveness of $\tilde{\tilde{D}}_\varepsilon$ defined by

$$\begin{aligned}
\tilde{\tilde{D}}_\varepsilon &= (\sigma^l-\sigma^t)\Big[d^2\int_\omega \big(I_2-\frac{L_1^2}{J_0}\big)\underline{\tau}_0\otimes\underline{\tau}_0\cdot\underline{\nabla}u_1^\varepsilon\cdot\underline{\nabla}u_1^\varepsilon\,dS \\
&\qquad +2d^3\int_\omega\big(L_3-L_1\frac{J_2}{J_0}\big)\underline{\tau}_0^\perp\otimes\underline{\tau}_0\cdot\underline{\nabla}u_1^\varepsilon\cdot\underline{\nabla}u_2^\varepsilon\,dS\Big] \\
&\quad +(\sigma^l-\sigma^t)d^4\int_\omega\big(J_4-\frac{J_2^2}{J_0}\big)\underline{\tau}_0^\perp\otimes\underline{\tau}_0^\perp\cdot\underline{\nabla}u_2^\varepsilon\cdot\underline{\nabla}u_2^\varepsilon\,dS \\
&\quad +(\sigma^l-\sigma^t)\Big[d^2\int_\omega\big(J_2-\frac{L_1^2}{I_0}\big)\underline{\tau}_0^\perp\otimes\underline{\tau}_0^\perp\cdot\underline{\nabla}u_1^\varepsilon\cdot\underline{\nabla}u_1^\varepsilon\,dS \\
&\qquad +2d^3\int_\omega\big(L_3-L_1\frac{I_2}{I_0}\big)\underline{\tau}_0\otimes\underline{\tau}_0^\perp\cdot\underline{\nabla}u_1^\varepsilon\cdot\underline{\nabla}u_2^\varepsilon\,dS\Big] \\
&\quad +(\sigma^l-\sigma^t)d^4\int_\omega\big(I_4-\frac{I_2^2}{I_0}\big)\underline{\tau}_0\otimes\underline{\tau}_0\cdot\underline{\nabla}u_2^\varepsilon\cdot\underline{\nabla}u_2^\varepsilon\,dS.
\end{aligned}$$

We can decompose $\tilde{\tilde{D}}_\varepsilon$ into two sums of the form $au^2+2cuv+bv^2$. We then use the property that $au^2+2cuv+bv^2$ is positive for all (u,v) if and only if $a,b\geq 0$ and $c^2\leq ab$. Therefore, $\tilde{\tilde{D}}_\varepsilon$ is positive if and only if

$$\begin{aligned}
&I_2-\frac{L_1^2}{J_0}\geq 0,\quad J_4-\frac{J_2^2}{J_0}\geq 0,\quad \big(L_3-L_1\frac{J_2}{J_0}\big)^2\leq\big(I_2-\frac{L_1^2}{J_0}\big).\big(J_4-\frac{J_2^2}{J_0}\big),\\
&J_2-\frac{L_1^2}{I_0}\geq 0,\quad I_4-\frac{I_2^2}{I_0}\geq 0,\quad \big(L_3-L_1\frac{I_2}{I_0}\big)^2\leq\big(J_2-\frac{L_1^2}{I_0}\big).\big(I_4-\frac{I_2^2}{I_0}\big).
\end{aligned}$$

The inequalities in the left and center columns are simple consequences of the Cauchy-Schwarz inequality, while those in the right column can easily be checked analytically using, e.g., symbolic computation software. □

Remark 11 *Various possible extensions can be considered for the asymptotic setting assumptions:*

- *There is no particular difficulty in considering non-homogeneous Neumann boundary conditions on the lateral surface Γ_2 in the asymptotic analysis, with similar assumptions as for f regarding ξ^3-regularity.*
- *We conjecture that polynomial assumptions of higher degree could be handled in a similar manner and would provide similar convergence results to the same limits.*
- *We could also consider more general forms of the angular variations of the fibers across the thickness of the type*

$$\vec{\tau}(\xi^1,\xi^2,\xi^3)=\underline{\tau}_0(\xi^1,\xi^2)\cos\theta\big(\xi^1,\xi^2,2\xi^3/d\big)+\underline{\tau}_0^\perp(\xi^1,\xi^2)\sin\theta\big(\xi^1,\xi^2,2\xi^3/d\big),$$

which would lead to extended definitions of the geometric coefficients I_i, J_i and L_i. The convergence proofs could then be performed verbatim up to the final chains of inequalities to be satisfied by these coefficients, which would have to be checked on a case-by-case basis, except for those directly following from the Cauchy-Schwarz inequality. Other types of scaling of the transverse coordinate could also be investigated in the definition of the fiber direction, e.g. with d^α substituted for d, but our objective here was to obtain a finite variation of the angle across the thickness in view of the applications considered.

5 Surface-based bidomain model and numerical assessments

Following our above detailed asymptotic analysis, we directly adapt the result to propose the surface-based bidomain model, for all $t > 0$, find $V_m(\cdot,t) \in H^1(\mathcal{S})$, $u_e(\cdot,t) \in H^1(\mathcal{S})$ and $w(\cdot,t) \in L^\infty(\mathcal{S})$ with $\int_\omega u_e \, dS = 0$, such that

$$\begin{cases} A_m \displaystyle\int_\omega \Big(C_m \frac{\partial V_m}{\partial t} + I_{ion}(V_m, w) \Big) \phi \, dS + \int_\omega \Big(\underline{\underline{\sigma}}_i \cdot \big(\underline{\nabla} V_m + \underline{\nabla} u_e \big) \Big) \cdot \underline{\nabla} \phi \, dS \\ \qquad\qquad = A_m \displaystyle\int_\omega I_{app}(V_m, w) \phi \, dS, \\ \displaystyle\int_\omega \Big((\underline{\underline{\sigma}}_i + \underline{\underline{\sigma}}_e) \cdot \underline{\nabla} u_e \Big) \cdot \underline{\nabla} \psi \, dS + \int_\omega \Big(\underline{\underline{\sigma}}_i \cdot \underline{\nabla} V_m \Big) \cdot \underline{\nabla} \psi \, dS = 0, \\ \partial_t w + g(V_m, w) = 0, \ \text{in } \omega \times (0,T). \end{cases} \tag{14}$$

for all $\phi, \psi \in H^1(\mathcal{S})$ such that $\int_\omega \psi \, dS = 0$ and with $\underline{\underline{\sigma}}_{i,e} = \sigma^t_{i,e} \, \underline{\underline{a}} + (\sigma^l_{i,e} - \sigma^t_{i,e})(I_0(\theta) \underline{\tau}_0 \otimes \underline{\tau}_0 + J_0(\theta) \underline{\tau}_0^\perp \otimes \underline{\tau}_0^\perp)$, where $\underline{\underline{a}}$ denotes the surface-based metric tensor – namely, the restriction of the 3D metric tensor to vectors lying in the tangential plane – and recalling that $\underline{\tau}_0$ is a unit vector parallel to the local fiber direction on the midsurface, 2θ is the angle between the fiber directions on the lower and upper boundary surfaces – referred to as the endocardium and epicardium, respectively, in the heart – and $I_0(\theta) = \frac{1}{2} + \frac{1}{4\theta}\sin(2\theta)$, $J_0(\theta) = \frac{1}{2} - \frac{1}{4\theta}\sin(2\theta)$.

We will now perform a numerical assessment of our proposed model by comparing the results given by the 3D model (1) and the surface model (14) for two different geometries. For these comparisons, we limit ourselves to the Mitchell-Schaeffer model to avoid undue technicalities in the ionic model calibration and solution, although of course other ionic models could be considered, e.g. more adapted to the atria cells. The values of the parameters used in the simulations are given in Table 1. The tests are performed with the finite element library *FELiScE*, developed at Inria. We use $\mathbb{P}_1$-Lagrange finite elements, and as time discretization scheme a standard Backward Differentiation Formula (BDF) of order two, see e.g. Ref. [1]. Although we primarily aim at a *numerical* assessment of the reduced model, we will use realistic values characteristic of atrial electrophysiology for all dimensions and parameters.

Table 1: Cell membrane parameters

A_m (cm^{-1})	C_m (mF.cm^{-2})	τ_{in} ($\text{cm}^2\text{.mA}^{-1}$)	τ_{out} ($\text{cm}^2\text{.mA}^{-1}$)	τ_{open} (ms)	τ_{close} (ms)	V_{gate} (mV)	V_{min} (mV)	V_{max} (mV)
200.0	10^{-3}	4.0	90.0	100.0	100.0	-67.0	-80.0	20.0

5.1 Planar test case

In this first test case, the 3D domain is a rectangular cuboid with dimensions $10 \times 10 \times 0.2$ (all dimensions given in cm), hence the midsurface is a 10×10 square, which we discretize into 200 elements in each direction, namely, 40,401 vertices and 80,000 triangular elements. The 3D mesh is obtained by extrusion of the 2D mesh using 4 elements across the thickness, which gives 202,005 vertices and 960,000 tetrahedral elements. The conductivities are $\sigma_i^t = 4.0\,10^{-4}$, $\sigma_e^t = 2.2\,10^{-3}$, $\sigma_i^l = 4.0\,10^{-3}$, $\sigma_e^l = 4.0\,10^{-3}$, (all in S.cm^{-1}). The fiber directions in the 3D mesh vary across the thickness only, and by an angle $\frac{\pi}{4}$ between

$$\begin{pmatrix} \cos\left(\frac{\pi}{8}\right) \\ \sin\left(\frac{\pi}{8}\right) \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \cos\left(\frac{3\pi}{8}\right) \\ \sin\left(\frac{3\pi}{8}\right) \\ 0 \end{pmatrix}.$$

Therefore, on the midsurface the fiber direction is aligned with the vector

$$\underline{\tau}_0 = \frac{\sqrt{2}}{2}\begin{pmatrix}1\\1\\0\end{pmatrix},$$

and we have $\theta = \frac{\pi}{8}$.

The model is initialized at $t = 0$ with $V_m = -80\,\text{mV}$, $u_e = 0\,\text{mV}$, and we trigger the wave by applying a current $I_{app} = 0.5\,\text{mA.cm}^{-2}$ in a circle (or cylindrical region in 3D) centered at $(0\,,\,5)$ and of radius $r = 0.55$ during 1 ms.

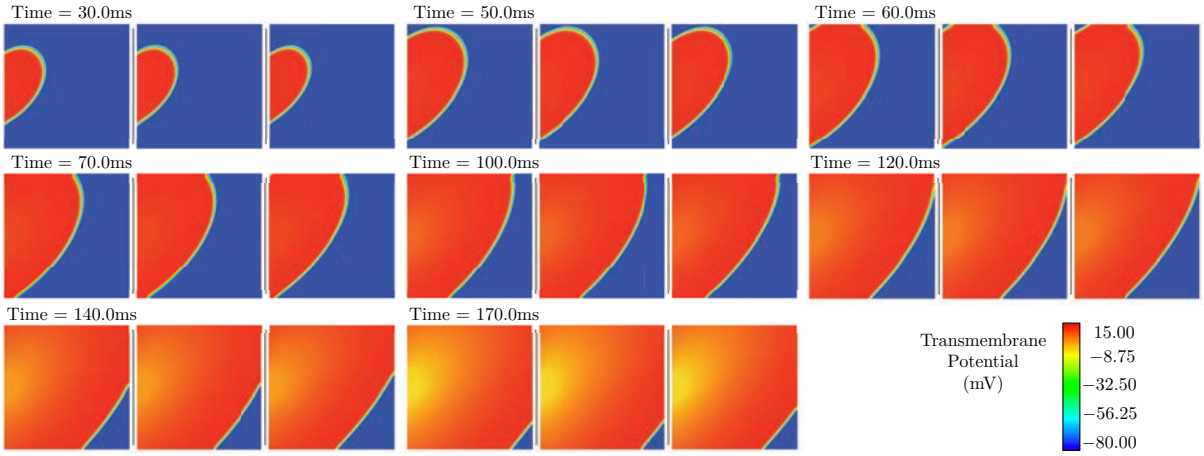

Figure 3: Planar test case – Comparison of asymptotic surface model (left) 3D model (center) and 2D naive model (right) on the midsurface at 8 consecutive times

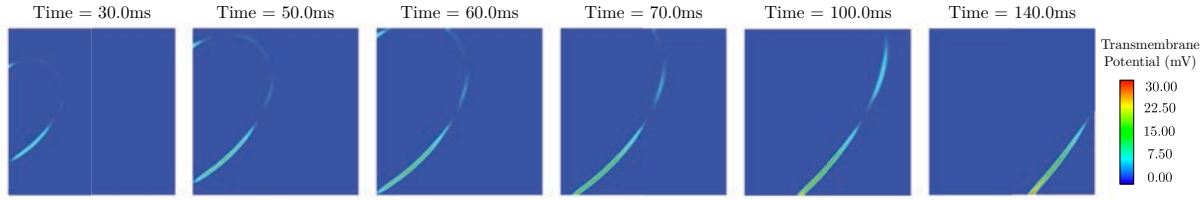

Figure 4: Planar test case – Point-wise difference between the 3D model and the asymptotic surface model on the midsurface (to be compared with the typical variation range of the transmembrane potential, i.e. about 100 mV)

In Figure 3, we compare the transmembrane potential given by the asymptotic surface model, the 3D model and a naive 2D model obtained by taking the value of the 3D conductivity tensor on the midsurface, namely, with $\underline{\underline{\sigma}}_{i,e} = \sigma^t_{i,e}\,\underline{\underline{a}} + (\sigma^l_{i,e} - \sigma^t_{i,e})\underline{\tau}_0 \otimes \underline{\tau}_0$. We observe an excellent agreement between the asymptotic surface model and the 3D model, whereas the naive model exhibits a slower front wave in the direction perpendicular to the midsurface fiber. In addition, we plot in Figure 4 the difference $|u^{ref,3D} - u^{2D}|$ on the midsurface. We notice that the differences between these two models are very limited and are narrowly concentrated near the wave front.

We also report in Table 2 the normalized l^∞ and l^2 differences between the asymptotic surface and the 3D models, *i.e.* $\frac{sup_{nodes}|u^{ref,3D} - u^{2D}|}{V_{max} - V_{min}}$ and $\frac{1}{V_{max} - V_{min}}\left(\frac{1}{\sharp Nodes}\sum_{nodes}|u^{ref,3D} - u^{2D}|^2\right)^{\frac{1}{2}}$, respectively. Note that $V_{max} - V_{min}$ is a typical normalizing constant, very similar

Table 2: Planar test case – Error norms between the 3D and asymptotic surface models, on the midsurface

T(ms)	30	50	60	70	100	140
$\lvert\cdot\rvert_\infty$	$1.081\,10^{-1}$	$1.582\,10^{-1}$	$1.792\,10^{-1}$	$2.364\,10^{-1}$	$3.172\,10^{-1}$	$3.687\,10^{-1}$
$\lvert\cdot\rvert_{l^2}$	$6.764\,10^{-3}$	$1.179\,10^{-2}$	$1.442\,10^{-2}$	$1.624\,10^{-2}$	$1.933\,10^{-2}$	$1.934\,10^{-2}$

Table 3: Planar test case – Error norms between the 3D and naive 2D models, on the midsurface

T(ms)	30	50	60	70	100	140
$\lvert\cdot\rvert_\infty$	$4.772\,10^{-1}$	$7.447\,10^{-1}$	$8.339\,10^{-1}$	$9.281\,10^{-1}$	1.000	1.000
$\lvert\cdot\rvert_{l^2}$	$4.020\,10^{-2}$	$8.218\,10^{-2}$	$1.035\,10^{-1}$	$1.151\,10^{-1}$	$1.530\,10^{-1}$	$1.757\,10^{-1}$

to the L^∞-norm of $u^{ref,3D}$ and also meaningful from a modeling standpoint since it appears in the ionic model. The relative l^∞ errors appear to be quite high, but when closely studying Figures 3 and 4 we see that this corresponds to very small shifts in the front location, which due to the dramatic steepness of the front induces significant – albeit very localized – errors. This is confirmed by the l^2 error values that are much smaller, indeed. We also point out that these errors are quite stable over time. Table 3 shows the same quantities when the asymptotic surface model is substituted by the naive 2D model. These results quantitatively confirm the superiority of the asymptotic model with respect to the naive model, and the excellent accuracy of the asymptotic model shows the relevance of the asymptotic analysis carried out in the previous sections when the fiber orientation rapidly varies across the thickness.

5.2 Spiral wave re-entry on a cylinder

This test case is motivated by so-called spiral waves, a fascinating phenomenon in cardiac electrophysiology, also very important in that it is often argued to be responsible for atrial or ventricular fibrillation, see e.g. Ref. [11]. This complex behavior provides a challenging test case for our proposed model, and we will also consider a non-planar geometry, indeed.

The domain is a 3D half-cylinder, with internal radius $r_{min} = 4.9$, external radius $r_{max} = 5.1$ and height $h = 10$ (all dimensions again given in cm). Thus, the midsurface is a half-cylinder of radius 5.0 and height 10. The 3D mesh contains 105,080 vertices and 499,080 tetrahedral elements with 4 elements across the thickness, and the 2D mesh has 21,016 vertices and 41,590 triangular elements. The fiber directions in the 3D geometry vary across the thickness by an angle $\frac{\pi}{2}$ around the midsurface direction given by the vector

$$\underline{\tau}_0 = \frac{1}{5}\begin{pmatrix} -y \\ x \\ 0 \end{pmatrix},$$

so that the fiber directions in the 2D mesh are characterized by this $\underline{\tau}_0$ and $\theta = \frac{\pi}{4}$. For this experiment, the conductivities are $\sigma_i^t = 4.0\,10^{-4}$, $\sigma_e^t = 2.2\,10^{-3}$, $\sigma_i^l = 4.0\,10^{-3}$, $\sigma_e^l = 4.0\,10^{-3}$ (in S.cm^{-1}).

In order to initiate a spiral wave, we follow the procedure used in Ref. [9]. The model is initialized at $t = 0$ with $V_m = -80\,\text{mV}$, $u_e = 0\,\text{mV}$, and we first generate a wave in the horizontal direction. Once the wave front has formed, it starts traveling along the cylinder (see Figure 5 at $t = 80.0\,\text{ms}$). Just after the end of the repolarization (see Figure 5 at $t = 300.0\,\text{ms}$), we depolarize a second region at time $t = 375.0\,\text{ms}$. The new depolarization front revolves around the region that has just started repolarizing, which triggers a spiral wave.

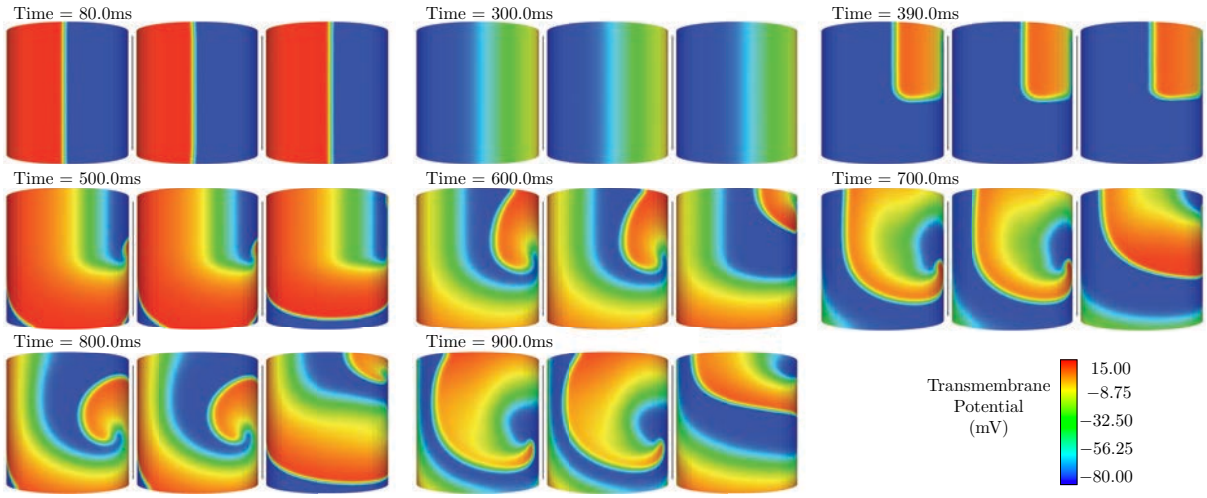

Figure 5: Spiral wave on cylinder – Comparison of asymptotic surface model (left), 3D model (center) and naive 2D model (right) on the midsurface at 8 consecutive times

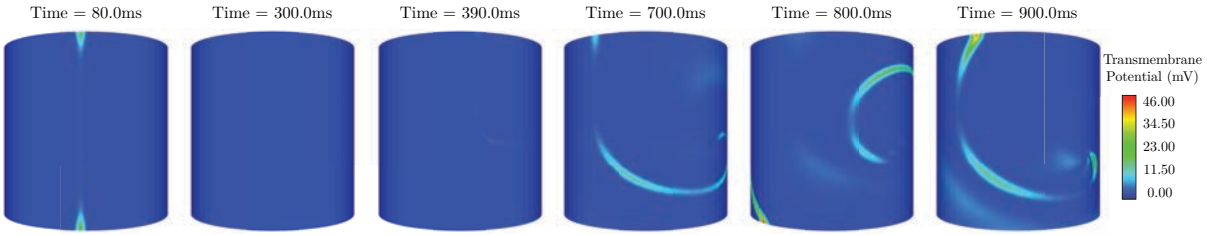

Figure 6: Spiral wave on cylinder – Point-wise difference between the 3D model and the asymptotic surface model on the midsurface (to be compared with the typical variation range of the transmembrane potential, i.e. about 100 mV)

Table 4: Spiral wave on cylinder – Error norms between the 3D and asymptotic surface models, on the midsurface

T(ms)	80	300	390	700	800	900
$\lvert\cdot\rvert_\infty$	$2.351\,10^{-1}$	$1.467\,10^{-2}$	$3.747\,10^{-2}$	$1.525\,10^{-1}$	$4.913\,10^{-1}$	$4.229\,10^{-1}$
$\lvert\cdot\rvert_{l^2}$	$8.375\,10^{-3}$	$2.315\,10^{-3}$	$2.797\,10^{-3}$	$2.019\,10^{-2}$	$3.744\,10^{-2}$	$3.228\,10^{-2}$

Table 5: Spiral wave on cylinder – Error norms between the 3D and naive 2D models, on the midsurface

T(ms)	80	300	390	700	800	900
$\lvert\cdot\rvert_\infty$	$9.110\,10^{-1}$	$1.026\,10^{-1}$	$2.338\,10^{-1}$	$9.548\,10^{-1}$	$9.525\,10^{-1}$	$8.788\,10^{-1}$
$\lvert\cdot\rvert_{l^2}$	$1.419\,10^{-1}$	$4.645\,10^{-2}$	$4.026\,10^{-2}$	$4.169\,10^{-1}$	$4.542\,10^{-1}$	$4.318\,10^{-1}$

We display in Figure 5 the results obtained with the 3D and surface models at successive times, and in Figure 6 the point-wise difference $|u^{ref,3D} - u^{2D}|$ is plotted on the midsurface. As in the planar test case experiment, the differences between the 3D and asymptotic surface models are quite limited and very narrowly concentrated along the wave front. By contrast, the naive 2D model appears to be of poor accuracy, and in fact altogether destroys the spiral wave after 40 ms. This is fully confirmed by Tables 4 and 5 showing the normalized differences

in norms l^∞ and l^2 (less than 4% in l^2-norm, here, for the asymptotic model). These results substantiate the validity of our proposed surface-based electrophysiology model to accurately represent such complex behaviors, and including over long time spans.

References

[1] U.M. Ascher and L.R. Petzold. *Level Set Methods and Dynamic Implicit Surfaces.* SIAM, 1998.

[2] Y. Bourgault, Y. Coudière, and C. Pierre. Existence and uniqueness of the solution for the bidomain model used in cardiac electrophysiology. *Nonlinear Analysis-Real World Applications*, 10(1):458–482, 2009.

[3] A. Carusi, K. Burrage, and B. Rodríguez. Bridging experiments, models and simulations: an integrative approach to validation in computational cardiac electrophysiology. *American Journal of Physiology - Heart and Circulatory Physiology*, 303(2):H144–H155, 2012.

[4] D. Chapelle and K.J. Bathe. *The Finite Element Analysis of Shells - Fundamentals.* Springer, second edition, 2011.

[5] D. Chapelle, A. Ferent, and K.J. Bathe. 3D-shell elements and their underlying mathematical model. *M3AS*, 14(1):105–142, 2004.

[6] P.G. Ciarlet. *Mathematical Elasticity - Volume III: Theory of Shells.* North-Holland, 2000.

[7] P. Colli Franzone, L.F. Pavarino, and B. Taccardi. Simulating patterns of excitation, repolarization and action potential duration with cardiac Bidomain and Monodomain models. *Mathematical Biosciences*, 197(1):35–66, 2005.

[8] M.C. Delfour. Intrinsic $P(2,1)$ thin shell model and Naghdi's models without a priori assumption on the stress tensor. In K.H. Hoffmann, G. Leugering, and F. Tröltzsch, editors, *Optimal Control of Partial Differential Equations*, pages 99–113, Basel, 1999. Birkhäuser.

[9] S. Göktepe and E. Kuhl. Computational modeling of cardiac electrophysiology: A novel finite element approach. *International Journal for Numerical Methods in Engineering*, 79(2):156–178, 2009.

[10] S.Y. Ho and D. Sanchez-Quintana. The Importance of Atrial Structure and Fibers. *Clinical Anatomy*, 22(1):52–63, 2009.

[11] A. Karma. Electrical alternans and spiral wave breakup in cardiac tissue. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 4(3):461–472, 1994.

[12] M. Krueger, V. Schmidt, C. Tobón, F. Weber, C. Lorenz, D. Keller, H. Barschdorf, M. Burdumy, P. Neher, G. Plank, K. Rhode, G. Seemann, D. Sanchez-Quintana, J. Saiz, R. Razavi, and O. Dössel. Modeling atrial fiber orientation in patient-specific geometries: a semi-automatic rule-based approach. *Functional Imaging and Modeling of the Heart*, pages 223–232, 2011.

[13] C.C. Mitchell and D.G. Schaeffer. A two-current model for the dynamics of cardiac membrane. *Bulletin Math. Bio.*, 65:767–793, 2003.

[14] C. Poignard. About the transmembrane voltage potential of a biological cell in time-harmonic regime. *ESAIM: Proceedings*, 26:162–179, 2009.

[15] A.J. Pullan, M.L. Buist, and L.K. Cheng. *Mathematically Modeling the Electrical Activity of the Heart*. World Scientific, 2005.

[16] F.B. Sachse. *Computational Cardiology: Modeling of Anatomy, Electrophysiology and Mechanics*. Springer-Verlag, 2004.

[17] M. Sermesant, R. Chabiniok, P. Chinchapatnam, T. Mansi, F. Billet, P. Moireau, J.-M. Peyrat, K. Wong, J. Relan, K. Rhode, M. Ginks, P. Lambiase, H. Delingette, M. Sorine, C.A. Rinaldi, D. Chapelle, R. Razavi, and N. Ayache. Patient-specific electromechanical models of the heart for the prediction of pacing acute effects in CRT: A preliminary clinical validation. *Medical Image Analysis*, 16(1):201–215, 2012.

[18] N. Smith, A. de Vecchi, M. McCormick, D. Nordsletten, O. Camara, A.F. Frangi, H. Delingette, M. Sermesant, J. Relan, N. Ayache, M.W. Krueger, W.H.W. Schulze, R. Hose, I. Valverde, P. Beerbaum, C. Staicu, M. Siebes, J. Spaan, P. Hunter, J. Weese, H. Lehmann, D. Chapelle, and R. Rezavi. euHeart: personalized and integrated cardiac care using patient-specific cardiovascular modelling. *Interface Focus*, 1(3):349–364, 2011.

[19] J. Sundnes, G.T. Lines, X. Cai, B.F. Nielsen, K.A. Mardal, and A. Tveito. *Computing the Electrical Activity in the Heart*, volume 1 of *Monographs in Computational Science and Engineering*. Springer-Verlag, 2006.

CHAPTER 4

Strong convergence results for the asymptotic behavior of the 3D-shell model

Abstract We revisit the asymptotic convergence properties – with respect to the thickness parameter – of the earlier-proposed 3D-shell model. This shell model is very attractive for engineering applications, in particular due to the possibility of directly using a general 3D constitutive law in the corresponding finite element formulations. We establish strong convergence results for the 3D-shell model in the two main types of asymptotic regimes, namely, bending- and membrane-dominated behavior. This is an important achievement, as it completely substantiates the asymptotic consistency of the 3D-shell model with 3D linearized isotropic elasticity.

Strong convergence results for the asymptotic behavior of the 3D-shell model

Dominique Chapelle, Annabelle Collin,
Inria Saclay Ile-de-France, Palaiseau, France

Journal of Elasticity, DOI:10.1007/s10659-013-9452-3

Abstract

We revisit the asymptotic convergence properties – with respect to the thickness parameter – of the earlier-proposed 3D-shell model. This shell model is very attractive for engineering applications, in particular due to the possibility of directly using a general 3D constitutive law in the corresponding finite element formulations. We establish strong convergence results for the 3D-shell model in the two main types of asymptotic regimes, namely, bending- and membrane-dominated behavior. This is an important achievement, as it completely substantiates the asymptotic consistency of the 3D-shell model with 3D linearized isotropic elasticity.

Keywords: Shells; 3D-shell model; Asymptotic analysis
Mathematics Subject Classification (2010): 74K25; 35B40

1 Introduction

The mechanical behavior of thin shells has been described and discussed since the early ages of structural mechanics, albeit only much more recently has the problem been mathematically addressed with the point of view of asymptotic analysis, see [11, 21, 8] and references therein. The major benefits of such asymptotic analyses – in which the convergence of sequences of solutions associated with decreasing thickness of the structure is investigated – have been to justify earlier-introduced limit models – in particular the so-called membrane and pure bending models – and to shed light on what type of convergence can be expected, with subtle accompanying features such as boundary layers, see, e.g., [12, 19].

Shell models are in essence surface-based, namely, they are described by equations posed on a surface, i.e., the midsurface of the structure, in general, see in particular [16] for a classical example of shell model. Via an adequate geometric description of this surface based on a chart that maps a two-dimensional (2D) reference domain to the actual surface immersed in the Euclidean space, the model equations can be expressed as partial differential equations (PDEs) posed in the 2D reference domain. Consequently, appropriate discretization procedures can be defined, in particular with finite elements [4]. However, such models and discretizations thereof are seldom used in engineering practice nowadays. A first type of reason for this lies in the difficulties associated with the geometric definition of the surface, not well-suited to geometries typically produced by computer-aided design (CAD) systems. A second limitation – probably even more important – of classical shell models in the engineering context pertains to the need of considering wide varieties of material properties generally characterized in three-dimensional (3D) continuum mechanics, whereas these shell models have "built-in" constitutive assumptions allowing little latitude for material variations.

By contrast, so-called "general shell elements" [3] or "degenerated solid shell elements" [1] are finite elements designed for shell structures – albeit directly based on 3D variational formulations, hence they can be accordingly used with 3D constitutive equations, without any reference to a given continuous shell model. In fact, it was later shown in [7] – see also [8] – that there exists an implicit shell model underlying these finite element methods. Furthermore, this underlying model was proven to be asymptotically consistent with 3D linearized isotropic elasticity, meaning that asymptotic analyses for this shell model and for 3D elasticity reveal the same limit models, whether the asymptotic behavior be of membrane- or bending-dominated type [6]. Nevertheless, general shell elements still feature a limitation regarding the material law, as they require the use of a "plane stress" assumption, which is not straightforward to implement in the case of non-linear stress-strain laws.

The 3D-shell model and corresponding finite elements as proposed in [9] were introduced to completely overcome any restriction in the use of 3D constitutive equations, namely, they are posed as a 3D variational formulation in which 3D material laws can be used without any prior transformation. In addition, these 3D-shell models and elements employ quadratic kinematical assumptions across the thickness – instead of essentially linear expressions of displacements for classical shell models and elements – which allows for a more accurate description of complex deformation fields. This model was also shown to be asymptotically consistent with 3D elasticity, albeit only based on weak convergence results [8]. We may also refer to [18, 20, 2] for earlier results in similar approaches applied to plate theory. In the present paper we revisit the convergence results of the continuous 3D-shell model, and we show that convergence also holds in the strong sense, again for the two main types of asymptotic behavior. We point out that our new results are obtained under the same assumptions as in [9, 8], albeit the difference lies in the improved techniques of proof.

In the next section, we recall the 3D-shell model formulation and properties, before proceeding to introduce the asymptotic setting and analyze the asymptotic convergence behavior in Section 3.

2 Model description

2.1 Geometry

We consider a thin three-dimensional (3D) volume assumed to be defined by a midsurface $\mathcal{S}$, and a thickness parameter t that may vary over the extent of $\mathcal{S}$. The surface $\mathcal{S}$ is described by a mapping $\vec{\phi}$ defined over $\overline{\omega}$, with ω a domain of $\mathbb{R}^2$, and with values in the three-dimensional Euclidean space $\mathcal{E}$. Namely, we have $\mathcal{S} = \vec{\phi}(\overline{\omega})$, and we assume that $\vec{\phi}$ is as regular as needed, so that we can introduce and employ the usual concepts of differential geometry, for which we will only give a brief summary, and refer to [8] – see also [14, 11] – for a more detailed exposition. We start with the so-called surface covariant basis

$$\vec{a}_\alpha = \vec{\phi}_{,\alpha} = \frac{\partial \vec{\phi}}{\partial \xi^\alpha}(\xi^1, \xi^2), \quad (\xi^1, \xi^2) \in \overline{\omega}. \tag{1}$$

Note that we employ the standard convention that Greek indices vary in $\{1, 2\}$, whereas Latin indices used later will vary in $\{1, 2, 3\}$. We then assume that the vectors $(\vec{a}_1, \vec{a}_2)$ provide a well-defined basis of the plane tangential to the midsurface, with the condition

$$\|\vec{a}_1 \wedge \vec{a}_2\| > 0, \quad \forall (\xi^1, \xi^2) \in \overline{\omega},$$

and we can thus introduce the unit normal vector

$$\vec{a}_3 = \frac{\vec{a}_1 \wedge \vec{a}_2}{\|\vec{a}_1 \wedge \vec{a}_2\|}. \tag{2}$$

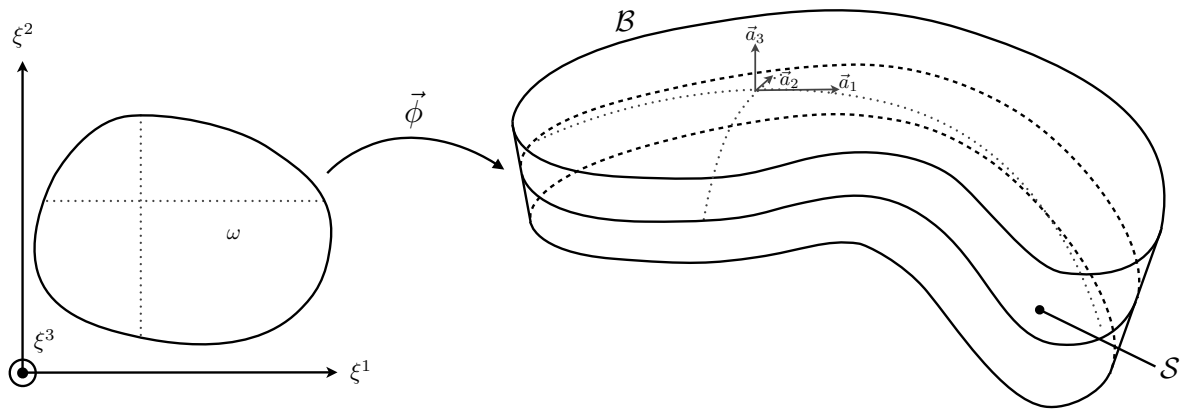

Figure 1: Geometrical description of thin 3D body

Alternatively, the contravariant basis – denoted with superscripts – is defined by

$$\vec{a}^{\alpha} \cdot \vec{a}_{\beta} = \delta^{\alpha}_{\beta},$$

where δ denotes the Kronecker symbol. The surface metric tensor – also called first fundamental form of the surface – is then associated with the components

$$a_{\alpha\beta} = \vec{a}_{\alpha} \cdot \vec{a}_{\beta}, \quad a^{\alpha\beta} = \vec{a}^{\alpha} \cdot \vec{a}^{\beta},$$

in covariant and contravariant forms, respectively. Intrinsic surface measures can be expressed as

$$dS = \sqrt{a}\, d\xi^1 d\xi^2, \quad \text{with } a = \|\vec{a}_1 \wedge \vec{a}_2\|^2 = a_{11}a_{22} - (a_{12})^2.$$

We next recall that the covariant components of the curvature tensor – also called second fundamental form – are given by

$$b_{\alpha\beta} = \vec{a}_3 \cdot \vec{a}_{\alpha,\beta}, \tag{3}$$

while the third fundamental form is defined by

$$c_{\alpha\beta} = b^{\lambda}_{\alpha} b_{\lambda\beta}, \quad \text{with } b^{\lambda}_{\alpha} = a^{\lambda\mu} b_{\mu\alpha},$$

where we use the Einstein convention of implicit summation for repeated indices – namely, for both λ and μ here. Finally, we will need the concept of surface-based covariant differentiation of vectors, defined by

$$u_{\alpha|\beta} = u_{\alpha,\beta} - \Gamma^{\lambda}_{\alpha\beta} u_{\lambda},$$

where the comma denotes standard derivatives, and with the so-called surface Christoffel symbols given by

$$\Gamma^{\lambda}_{\alpha\beta} = \vec{a}^{\lambda} \cdot \vec{a}_{\alpha,\beta}.$$

Defining now the 3D domain

$$\Omega = \left\{ (\xi^1, \xi^2, \xi^3) \in I\!R^3 \,|(\xi^1, \xi^2) \in \omega,\ \xi^3 \in \left]-\frac{t(\xi^1,\xi^2)}{2}, +\frac{t(\xi^1,\xi^2)}{2}\right[\right\}, \tag{4}$$

the 3D geometry is described by the mapping $\vec{\Phi}$

$$\vec{\Phi}(\xi^1, \xi^2, \xi^3) = \vec{\phi}(\xi^1, \xi^2) + \xi^3 \vec{a}_3(\xi^1, \xi^2), \quad (\xi^1, \xi^2, \xi^3) \in \overline{\Omega}, \tag{5}$$

and the region of the Euclidean space occupied by the thin body is given by $\mathcal{B} = \vec{\Phi}\big(\overline{\Omega}\big)$, see Figure 1. We will assume that the thickness does not degenerate anywhere, namely,

$$t(\xi^1, \xi^2) > 0, \quad \forall (\xi^1, \xi^2) \in \overline{\omega}. \tag{6}$$

Accordingly, we can now define the 3D covariant basis

$$\vec{g}_i = \vec{\Phi}_{,i} = \frac{\partial \vec{\Phi}}{\partial \xi^i}(\xi^1, \xi^2, \xi^3), \quad (\xi^1, \xi^2, \xi^3) \in \overline{\Omega}.$$

We infer [8]

$$\vec{g}_\alpha = (\delta_\alpha^\lambda - \xi^3 b_\alpha^\lambda)\vec{a}_\lambda, \quad \vec{g}_3 = \vec{a}_3, \tag{7}$$

and the following covariant components for the 3D metric tensor $g_{ij} = \vec{g}_i \cdot \vec{g}_j$

$$g_{\alpha\beta} = a_{\alpha\beta} - 2\xi^3 b_{\alpha\beta} + (\xi^3)^2 c_{\alpha\beta}, \quad g_{\alpha 3} = 0, \quad g_{33} = 1. \tag{8}$$

We will also use the contravariant 3D basis vectors such that

$$\vec{g}^{\,i} \cdot \vec{g}_j = \delta_j^i,$$

and the contravariant components of the 3D metric tensor given by $g^{ij} = \vec{g}^{\,i} \cdot \vec{g}^{\,j}$. Finally, the volume measure is expressed as

$$dV = \sqrt{g}\, d\xi^1 d\xi^2 d\xi^3,$$

with

$$g = [(\vec{g}_1 \wedge \vec{g}_2) \cdot \vec{g}_3]^2 = a\big(1 - 2H\xi^3 + K(\xi^3)^2\big)^2,$$

where H and K respectively denote the mean and Gaussian curvatures of the surface. Note in passing that this shows that the condition $g > 0$ necessary for the mapping $\vec{\Phi}$ to be well-defined is equivalent to $1 - 2H\xi^3 + K(\xi^3)^2 > 0$, namely also to

$$t(\xi^1, \xi^2) < 2|R_{\min}(\xi^1, \xi^2)|, \quad \forall (\xi^1, \xi^2) \in \overline{\omega},$$

where $R_{\min}$ denotes the radius of curvature of smallest absolute value at the specific point considered, condition that we will henceforth assume to be satisfied.

2.2 3D-shell model

We start by recalling the formulation of linearized isotropic elasticity expressed for $\mathcal{B}$ using the above-introduced curvilinear coordinates. The variational formulation reads

$$A^{3D}(\vec{U}, \vec{V}) = F^{3D}(\vec{V}), \tag{9}$$

for any $\vec{V}$ in a space of suitable test functions satisfying adequate boundary conditions, with

$$A^{3D}(\vec{U}, \vec{V}) = \int_\Omega H^{ijkl} e_{ij}(\vec{U}) e_{kl}(\vec{V})\, dV, \quad F^{3D}(\vec{V}) = \int_\Omega \vec{F} \cdot \vec{V}\, dV.$$

In these expressions, $\vec{F}$ represents the external loading – assumed to be volume-distributed here – while the linearized strain tensor appearing in A^{3D} is defined by

$$e_{ij}(\vec{V}) = \frac{1}{2}\big(\vec{V}_{,i} \cdot \vec{g}_j + \vec{V}_{,j} \cdot \vec{g}_i\big),$$

and the fourth-order constitutive tensor by

$$H^{ijkl} = L_1\, g^{ij} g^{kl} + L_2(g^{ik} g^{jl} + g^{il} g^{jk}), \tag{10}$$

with L_1 and L_2 the Lamé constants, i.e.,

$$L_1 = E\frac{\nu}{(1+\nu)(1-2\nu)}, \quad L_2 = \frac{E}{2(1+\nu)} \tag{11}$$

if we classically denote Young's modulus and Poisson's ratio by E and ν, respectively. Note that this entails the simplifications

$$\begin{cases} H^{\alpha\beta\gamma 3}(= H^{\alpha\beta 3\gamma} = H^{\gamma 3\alpha\beta} = H^{3\gamma\alpha\beta}) = 0, \quad \forall \alpha,\beta,\gamma = 1,2 \\ H^{\alpha 333}(= H^{3\alpha 33} = H^{33\alpha 3} = H^{333\alpha}) = 0, \quad \forall \alpha = 1,2 \end{cases} \tag{12}$$

The 3D-shell model is defined as the variational projection of the 3D model onto the space of displacement functions having quadratic expressions in the ξ^3-variable. Namely, defining

$$\begin{cases} A^{3Ds}(\vec{u},\vec{\theta},\vec{\varrho};\vec{v},\vec{\eta},\vec{\varsigma}) = A^{3D}\big(\vec{u}+\xi^3\vec{\theta}+(\xi^3)^2\vec{\varrho},\vec{v}+\xi^3\vec{\eta}+(\xi^3)^2\vec{\varsigma}\big) \\ F^{3Ds}(\vec{v},\vec{\eta},\vec{\varsigma}) = F^{3D}\big(\vec{v}+\xi^3\vec{\eta}+(\xi^3)^2\vec{\varsigma}\big) \end{cases}$$

we consider the variational equation

$$A^{3Ds}(\vec{u},\vec{\theta},\vec{\varrho};\vec{v},\vec{\eta},\vec{\varsigma}) = F^{3Ds}(\vec{v},\vec{\eta},\vec{\varsigma}), \tag{13}$$

see also [13, 5] for discussions on this type of kinematical assumption. With these specific displacement functions, we have the following identities for the linearized strains [9, 8]

$$\begin{cases} e_{\alpha\beta}(\vec{U}) = \gamma_{\alpha\beta}(\vec{u}) + \xi^3\,\hat{\chi}_{\alpha\beta}(\vec{u},\vec{\theta}) + (\xi^3)^2\,k_{\alpha\beta}(\vec{\theta},\vec{\varrho}) + (\xi^3)^3 l_{\alpha\beta}(\vec{\varrho}) \\ e_{\alpha 3}(\vec{U}) = \zeta_\alpha(\vec{u},\vec{\theta}) + \xi^3\,m_\alpha(\vec{\theta},\vec{\varrho}) + (\xi^3)^2\,n_\alpha(\vec{\varrho}) \\ e_{33}(\vec{U}) = \delta(\vec{\theta}) + \xi^3\,p(\vec{\varrho}) \end{cases} \tag{14}$$

where

$$\begin{cases} \gamma_{\alpha\beta}(\vec{u}) = \frac{1}{2}(u_{\alpha|\beta} + u_{\beta|\alpha}) - b_{\alpha\beta}u_3 \\ \hat{\chi}_{\alpha\beta}(\vec{u},\vec{\theta}) = \frac{1}{2}(\theta_{\alpha|\beta} + \theta_{\beta|\alpha} - b^\lambda_\alpha u_{\lambda|\beta} - b^\lambda_\beta u_{\lambda|\alpha}) - b_{\alpha\beta}\theta_3 + c_{\alpha\beta}u_3 \\ k_{\alpha\beta}(\vec{\theta},\vec{\varrho}) = \frac{1}{2}(\varrho_{\alpha|\beta} + \varrho_{\beta|\alpha} - b^\lambda_\alpha\theta_{\lambda|\beta} - b^\lambda_\beta\theta_{\lambda|\alpha}) - b_{\alpha\beta}\varrho_3 + c_{\alpha\beta}\theta_3 \\ l_{\alpha\beta}(\vec{\varrho}) = -\frac{1}{2}(b^\lambda_\alpha\varrho_{\lambda|\beta} + b^\lambda_\beta\varrho_{\lambda|\alpha}) + c_{\alpha\beta}\varrho_3 \\ \zeta_\alpha(\vec{u},\vec{\theta}) = \frac{1}{2}(\theta_\alpha + b^\lambda_\alpha u_\lambda + u_{3,\alpha}) \\ m_\alpha(\vec{\theta},\vec{\varrho}) = \frac{1}{2}(2\varrho_\alpha + \theta_{3,\alpha}) \\ n_\alpha(\vec{\varrho}) = \frac{1}{2}(-b^\lambda_\alpha\varrho_\lambda + \varrho_{3,\alpha}) \\ \delta(\vec{\theta}) = \theta_3 \\ p(\vec{\varrho}) = 2\varrho_3 \end{cases} \tag{15}$$

In our analyses we will use intrinsic Sobolev norms on $\mathcal{S}$ and $\mathcal{B}$, namely, integrating with surface and volume measures, respectively, and employing covariant derivatives whenever needed [8]. These norms are equivalent to the usual Sobolev norms defined on ω and Ω, respectively, albeit they are intrinsic, i.e., they do not depend on the particular choice of coordinate system used to describe the actual geometry. In the sequel, we will also use the subscripts 0 and 1 to denote the intrinsic norms of $L^2(\mathcal{S})$ and $H^1(\mathcal{S})$, respectively.

Then, defining the space $\mathcal{V}^{3Ds} = \big[H^1(\mathcal{S}) \cap \mathcal{BC}\big]^3$ where $\mathcal{BC}$ symbolically denotes functions having vanishing trace on a part of non-zero measure of the boundary $\partial\omega$ – namely, Dirichlet (i.e. essential) boundary conditions – we recall the following result [9, 8], where we use the symbol C to denote a generic strictly positive constant.

Proposition 1 *Assuming that* $\vec{F} \in L^2(\mathcal{B})$*, there is a unique* $(\vec{u}, \vec{\theta}, \vec{\varrho})$ *in* $\mathcal{V}^{3Ds}$ *that satisfies* (13) *for any* $(\vec{v}, \vec{\eta}, \vec{\varsigma})$ *in* $\mathcal{V}^{3Ds}$*, and we have*

$$\|\vec{u}, \vec{\theta}, \vec{\varrho}\|_{H^1(\mathcal{S})} \leq C \|\vec{F}\|_{L^2(\mathcal{B})}. \tag{16}$$

3 Asymptotic analysis

3.1 Asymptotic setting

We now proceed to recast the above problem in an asymptotic framework. Denoting by L an overall characteristic dimension of the surface, – e.g., the diameter – and

$$t_{\min} = \inf_{(\xi^1, \xi^2) \in \omega} t(\xi^1, \xi^2) > 0,$$

we define the dimensionless thickness parameter

$$\varepsilon = \frac{t_{\min}}{L},$$

and

$$l(\xi^1, \xi^2) = \frac{t(\xi^1, \xi^2)}{\varepsilon} = \frac{t(\xi^1, \xi^2)}{t_{\min}} L,$$

hence $\inf_{(\xi^1, \xi^2) \in \omega} l(\xi^1, \xi^2) = L$, and we call l the thickness profile. We then consider a sequence of shell geometries obtained by fixing the midsurface $\mathcal{S}$ and thickness profile l, and varying the ε parameter. We subsequently seek a scaling of the loading in the form

$$\vec{F} = \varepsilon^{\rho - 1} \vec{G}, \tag{17}$$

with $\vec{G}$ *independent* of ε, and the parameter ρ chosen so that the associated sequence of solutions $(\vec{u}^\varepsilon, \vec{\theta}^\varepsilon, \vec{\varrho}^\varepsilon) \in \mathcal{V}^{3Ds}$ satisfying

$$A^{3Ds}(\vec{u}^\varepsilon, \vec{\theta}^\varepsilon, \vec{\varrho}^\varepsilon; \vec{v}, \vec{\eta}, \vec{\varsigma}) = F^{3Ds}(\vec{v}, \vec{\eta}, \vec{\varsigma}), \quad \forall (\vec{v}, \vec{\eta}, \vec{\varsigma}) \in \mathcal{V}^{3Ds}, \tag{18}$$

has a non-zero finite limit energy-wise, namely, so that $\varepsilon^{-\rho} F^{3Ds}(\vec{u}^\varepsilon, \vec{\theta}^\varepsilon, \vec{\varrho}^\varepsilon)$ has a non-zero finite limit. Note that the expressions of A^{3Ds} and F^{3Ds} include a dependence on ε even though this is not reflected in their notations for the sake of compactness. It can be shown that there is at most one value of this scaling parameter for which this convergence holds, and that such a value necessarily satisfies [8]

$$1 \leq \rho \leq 3. \tag{19}$$

Several cases may then arise depending on the contents of the particular subspace

$$\begin{aligned} \mathcal{V}_0^{3Ds} = \{ & (\vec{v}, \vec{\eta}, \vec{\varsigma}) \in \mathcal{V}^{3Ds}, \text{ such that} \\ & \gamma_{\alpha\beta}(\vec{v}) = 0, \ \zeta_\alpha(\vec{v}, \vec{\eta}) = 0, \ \delta(\vec{\eta}) = 0, \ \forall \alpha, \beta = 1, 2 \}, \end{aligned} \tag{20}$$

called the subspace of *pure bending displacements*. More specifically, we will distinguish two situations according to whether we have

$$\mathcal{V}_0^{3Ds} \cap \{(\vec{v}, \vec{\eta}, \vec{0}) \in \mathcal{V}^{3Ds}\} = \{(\vec{0}, \vec{0}, \vec{0})\},$$

in which case we say that pure bending is inhibited. By contrast, the situation when $\mathcal{V}_0^{3Ds}$ has non-trivial contents is referred to as non-inhibited pure bending. Both cases are frequently

encountered in actual shells, depending on the surface geometry and on the boundary conditions considered, see [8] in particular.

Let us introduce some definitions and notation that will be needed in our asymptotic analysis. We will assume that the scaled loading $\vec{G}$ in (17) is of the form

$$\vec{G}(\xi^1,\xi^2,\xi^3) = \vec{G}_0(\xi^1,\xi^2) + \xi^3\vec{B}(\xi^1,\xi^2,\xi^3), \tag{21}$$

where $\vec{G}_0$ is in $L^2(\mathcal{S})$ and $\vec{B}$ is a bounded function over $\mathcal{B}$, uniformly in ε. We then define

$$G^{3Ds}(\vec{v}\,) = \int_\omega l\vec{G}_0\cdot\vec{v}\,dS. \tag{22}$$

We now define the bilinear forms

$$A_{\rm m}^{3Ds}(\vec{u},\vec{\theta};\vec{v},\vec{\eta}\,) = \int_\omega l\Big[{}^0H^{\alpha\beta\lambda\mu}\gamma_{\alpha\beta}(\vec{u}\,)\gamma_{\lambda\mu}(\vec{v}\,) + {}^0H^{\alpha\beta33}\big(\gamma_{\alpha\beta}(\vec{u}\,)\delta(\vec{\eta}\,) + \gamma_{\alpha\beta}(\vec{v}\,)\delta(\vec{\theta}\,)\big) + 4\,{}^0H^{\alpha3\beta3}\zeta_\alpha(\vec{u},\vec{\theta}\,)\zeta_\beta(\vec{v},\vec{\eta}\,) + {}^0H^{3333}\delta(\vec{\theta}\,)\delta(\vec{\eta}\,)\Big]\,dS, \tag{23}$$

$$A_{\rm b}^{3Ds}(\vec{u},\vec{\theta},\vec{\varrho};\vec{v},\vec{\eta},\vec{\varsigma}) = \int_\omega \frac{l^3}{12}\Big[{}^0H^{\alpha\beta\lambda\mu}\hat{\chi}_{\alpha\beta}(\vec{u},\vec{\theta}\,)\hat{\chi}_{\lambda\mu}(\vec{v},\vec{\eta}\,) + {}^0H^{\alpha\beta33}\big(\hat{\chi}_{\alpha\beta}(\vec{u},\vec{\theta}\,)p(\vec{\varsigma}\,) + \hat{\chi}_{\alpha\beta}(\vec{v},\vec{\eta}\,)p(\vec{\varrho}\,)\big) + 4\,{}^0H^{\alpha3\beta3}m_\alpha(\vec{\theta},\vec{\varrho}\,)m_\beta(\vec{\eta},\vec{\varsigma}\,) + {}^0H^{3333}p(\vec{\varrho}\,)p(\vec{\varsigma}\,)\Big]\,dS, \tag{24}$$

where

$${}^0H^{ijkl} = H^{ijkl}|_{\xi^3=0}. \tag{25}$$

We will also use the following expansion of this constitutive tensor

$$H^{ijkl}(\xi^1,\xi^2,\xi^3) = {}^0H^{ijkl}(\xi^1,\xi^2) + \xi^3\bar{H}^{ijkl}(\xi^1,\xi^2,\xi^3), \tag{26}$$

with $\bar{H}^{ijkl}(\xi^1,\xi^2,\xi^3)$ bounded over $\mathcal{B}$ since the geometry is smooth.

Let us introduce the norm

$$\|\vec{v},\vec{\eta},\vec{\varsigma}\|_{\rm b} = \Big(\|\vec{v}\|_1^2 + \|\underline{\eta}\|_1^2 + \|\eta_3\|_0^2 + \|\varsigma_3\|_0^2 + \|\underline{\varsigma} + \frac{1}{2}\underline{\nabla}\eta_3\|_0^2\Big)^{\frac{1}{2}}, \tag{27}$$

and the semi-norm

$$\|\vec{v},\vec{\eta}\,\|_{\rm m} = A_{\rm m}^{3Ds}(\vec{v},\vec{\eta};\vec{v},\vec{\eta}\,)^{\frac{1}{2}}. \tag{28}$$

Note that we classically use the underbar notation to denote tensors attached to the tangential plane of the surface $\mathcal{S}$, namely with component indices ranging in $\{1,2\}$ – e.g. first-order tensors correspond to vectors lying in the tangential plane, second-order tensors to linear mappings within this vector space, and so on – see [8] for more detail. As we will see in the following lemma, when pure bending is inhibited $\|\cdot\|_{\rm m}$ is a norm. In this lemma we recall some equivalence properties for the above norms or semi-norms – established in [8] in a straightforward manner – in which we dispense with indicating the (obvious) operands of strain tensors to alleviate the notation.

Lemma 1 *We have an equivalence of norms – or semi-norms, when applicable – between:*

1. $\|\vec{v},\vec{\eta}\,\|_{\rm m}$ *and* $\Big(\|\underline{\underline{\gamma}}\|_0^2 + \|\underline{\zeta}\|_0^2 + \|\delta\|_0^2\Big)^{1/2}$;

2. $\|\vec{v},\vec{\eta},\vec{\varsigma}\|_{\rm b}$ *and* $\Big(\|\underline{\underline{\gamma}}\|_0^2 + \|\underline{\zeta}\|_0^2 + \|\delta\|_0^2 + \|\underline{\underline{\hat{\chi}}}\|_0^2 + \|\underline{m}\|_0^2 + \|p\|_0^2\Big)^{\frac{1}{2}}$;

3. $\|\vec{v},\vec{\eta},\vec{\varsigma}\|_1$ *and* $\Big(\|\underline{\gamma}\|_0^2+\|\underline{\zeta}\|_0^2+\|\delta\|_0+\|\underline{\hat{\chi}}\|_0^2+\|\underline{m}\|_0^2+\|p\|_0^2+\|\underline{k}\|_0^2+\|\underline{n}\|_0^2\Big)^{\frac{1}{2}}$.

Note that none of the norm – or semi-norm – expressions used in this lemma depends on ε, hence all these equivalence properties hold independently of ε.

In our below asymptotic analyses we will use a decomposition of A^{3Ds} obtained by using (12), (14), and the change of variables $\xi^3=\varepsilon\xi$, viz.

$$A^{3Ds}(\vec{u},\vec{\theta},\vec{\varrho};\vec{v},\vec{\eta},\vec{\varsigma})=I_1+I_2+I_3+I_4+I_5, \tag{29}$$

with

$$I_1(\vec{U},\vec{V})=\varepsilon\int_\omega\int_{-l/2}^{l/2}H^{\alpha\beta\lambda\mu}\big[\gamma_{\alpha\beta}(\vec{u})+\varepsilon\xi\hat{\chi}_{\alpha\beta}(\vec{u},\vec{\theta})+\varepsilon^2(\xi)^2k_{\alpha\beta}(\vec{\theta},\vec{\varrho})+\varepsilon^3(\xi)^3l_{\alpha\beta}(\vec{\varrho})\big]$$
$$\times\big[\gamma_{\lambda\mu}(\vec{v})+\varepsilon\xi\hat{\chi}_{\lambda\mu}(\vec{v},\vec{\eta})+\varepsilon^2(\xi)^2k_{\lambda\mu}(\vec{\eta},\vec{\varsigma})+\varepsilon^3(\xi)^3l_{\lambda\mu}(\vec{\varsigma})\big]\sqrt{g}\,d\xi d\xi^1d\xi^2, \tag{30}$$

$$I_2(\vec{U},\vec{V})=\varepsilon\int_\omega\int_{-l/2}^{l/2}4H^{\alpha3\beta3}\big[\zeta_\alpha(\vec{u},\vec{\theta})+\varepsilon\xi m_\alpha(\vec{\theta},\vec{\varrho})+\varepsilon^2(\xi)^2n_\alpha(\vec{\varrho})\big]$$
$$\times\big[\zeta_\beta(\vec{v},\vec{\eta})+\varepsilon\xi m_\beta(\vec{\eta},\vec{\varsigma})+\varepsilon^2(\xi)^2n_\beta(\vec{\varsigma})\big]\sqrt{g}\,d\xi d\xi^1d\xi^2, \tag{31}$$

$$I_3(\vec{U},\vec{V})=\varepsilon\int_\omega\int_{-l/2}^{l/2}H^{3333}\big[\delta(\vec{\theta})+\varepsilon\xi p(\vec{\varrho})\big]\times\big[\delta(\vec{\eta})+\varepsilon\xi p(\vec{\varsigma})\big]\sqrt{g}\,d\xi d\xi^1d\xi^2, \tag{32}$$

$$I_4(\vec{U},\vec{V})=\varepsilon\int_\omega\int_{-l/2}^{l/2}H^{\alpha\beta33}\big[\gamma_{\alpha\beta}(\vec{u})+\varepsilon\xi\hat{\chi}_{\alpha\beta}(\vec{u},\vec{\theta})+\varepsilon^2(\xi)^2k_{\alpha\beta}(\vec{\theta},\vec{\varrho})+\varepsilon^3(\xi)^3l_{\alpha\beta}(\vec{\varrho})\big]$$
$$\times\big[\delta(\vec{\eta})+\varepsilon\xi p(\vec{\varsigma})\big]\sqrt{g}\,d\xi d\xi^1d\xi^2, \tag{33}$$

$$I_5(\vec{U},\vec{V})=\varepsilon\int_\omega\int_{-l/2}^{l/2}H^{\alpha\beta33}\big[\gamma_{\alpha\beta}(\vec{v})+\varepsilon\xi\hat{\chi}_{\alpha\beta}(\vec{v},\vec{\eta})+\varepsilon^2(\xi)^2k_{\alpha\beta}(\vec{\eta},\vec{\varsigma})+\varepsilon^3(\xi)^3l_{\alpha\beta}(\vec{\varsigma})\big]$$
$$\times\big[\delta(\vec{\theta})+\varepsilon\xi p(\vec{\varrho})\big]\sqrt{g}\,d\xi d\xi^1d\xi^2, \tag{34}$$

using the compact notation

$$\vec{U}=\vec{u}+\xi^3\vec{\theta}+(\xi^3)^2\vec{\varrho},\quad \vec{V}=\vec{v}+\xi^3\vec{\eta}+(\xi^3)^2\vec{\varsigma}.$$

Similarly, the linear form F^{3Ds} gives

$$F^{3Ds}(\vec{v},\vec{\eta},\vec{\varsigma})=\varepsilon\int_\omega\int_{-l/2}^{l/2}\vec{F}\cdot\big[\vec{v}+\varepsilon\xi\vec{\eta}+\varepsilon^2(\xi)^2\vec{\varsigma}\big]\sqrt{g}\,d\xi d\xi^1d\xi^2. \tag{35}$$

Finally, we will henceforth denote by C generic constants *independent* of ε, and by R_ε scalar quantities that tend to zero with ε.

3.2 Non-inhibited pure bending

We start with the case of non-inhibited pure bending, namely, when

$$\mathcal{V}_0^{3Ds} \cap \{(\vec{v}, \vec{\eta}, \vec{0}\,) \in \mathcal{V}^{3Ds}\} \neq \{(\vec{0}, \vec{0}, \vec{0})\}.$$

In this case, as in more standard shell models the expected asymptotic behavior is that of a penalization mechanism in which the solution "tends to belong" to the constrained subspace $\mathcal{V}_0^{3Ds}$, associated with a scaling parameter $\rho = 3$, see [8] and references therein.

We then define $\mathcal{V}_{\rm b}^{3Ds}$ as the completion of $\mathcal{V}^{3Ds}$ for the norm $\|\cdot\|_{\rm b}$. We will also use the space $\mathcal{V}_{0\rm b}^{3Ds}$, defined as the completion of $\mathcal{V}_0^{3Ds}$ for $\|\cdot\|_{\rm b}$, which is identified as

$$\begin{aligned}\mathcal{V}_{0\rm b}^{3Ds} = \{(\vec{v}, \vec{\eta}, \vec{\varsigma}) \in \mathcal{V}_{\rm b}^{3Ds} \text{ such that } \gamma_{\alpha\beta}(\vec{v}) = 0,\\ \zeta_\alpha(\vec{v}, \vec{\eta}\,) = 0, \delta(\vec{\eta}\,) = 0, \forall \alpha, \beta = 1, 2\}.\end{aligned}$$

Using the proposed scaling $\rho = 3$, the tentative limit problem reads:
Find $(\vec{u}^0, \vec{\theta}^0, \vec{\varrho}^{\,0}) \in \mathcal{V}_{0\rm b}^{3Ds}$ *such that*

$$A_{\rm b}^{3Ds}(\vec{u}^0, \vec{\theta}^0, \vec{\varrho}^{\,0}; \vec{v}, \vec{\eta}, \vec{\varsigma}) = G^{3Ds}(\vec{v}), \quad \forall (\vec{v}, \vec{\eta}, \vec{\varsigma}) \in \mathcal{V}_{0\rm b}^{3Ds}. \tag{36}$$

Note that the right-hand side of this variational formulation defines a linear form in $\mathcal{V}_{0\rm b}^{3Ds}$ (although this completed space may contain elements that do not belong to $\mathcal{V}^{3Ds}$) since

$$\begin{aligned}|\int_\omega l\vec{G}_0 \cdot \vec{v}\, dS| &\leq C\|\vec{G}_0\|_0\|\vec{v}\|_0\\ &\leq C\|\vec{G}_0\|_0\|\vec{v}, \vec{\eta}, \vec{\varsigma}\|_{\rm b}, \ \forall (\vec{v}, \vec{\eta}, \vec{\varsigma}) \in \mathcal{V}_{\rm b}^{3Ds},\end{aligned} \tag{37}$$

hence this variational problem is well-posed by construction. In this framework we have the following weak convergence [9, 8].

Lemma 2 *The solution* $(\vec{u}^\varepsilon, \vec{\theta}^\varepsilon, \vec{\varrho}^{\,\varepsilon})$ *of Problem (18) converges weakly in* $\mathcal{V}_{\rm b}^{3Ds}$*, as* ε *goes to* 0*, to* $(\vec{u}^0, \vec{\theta}^0, \vec{\varrho}^{\,0}) \in \mathcal{V}_{0\rm b}^{3Ds}$ *solution of* (36)*.*

We proceed to establish the corresponding strong convergence, for which we will use the following uniform bound obtained in the proof of Lemma 2

$$\frac{1}{\varepsilon}\|\vec{u}^\varepsilon, \vec{\theta}^\varepsilon\|_{\rm m} + \|\vec{u}^\varepsilon, \vec{\theta}^\varepsilon, \vec{\varrho}^{\,\varepsilon}\|_{\rm b} + \varepsilon\|\vec{u}^\varepsilon, \vec{\theta}^\varepsilon, \vec{\varrho}^{\,\varepsilon}\|_1 \leq C. \tag{38}$$

Proposition 2 *The solution* $(\vec{u}^\varepsilon, \vec{\theta}^\varepsilon, \vec{\varrho}^{\,\varepsilon})$ *of Problem* (18) *converges strongly in* $\mathcal{V}_{\rm b}^{3Ds}$ *to* $(\vec{u}^0, \vec{\theta}^0, \vec{\varrho}^{\,0}) \in \mathcal{V}_{0\rm b}^{3Ds}$ *solution of* (36)*, as* ε *goes to* 0*. Moreover, we have that* $\frac{1}{\varepsilon}\|\vec{u}^\varepsilon, \vec{\theta}^\varepsilon\|_{\rm m}$ *and* $\varepsilon\|\vec{u}^\varepsilon, \vec{\theta}^\varepsilon, \vec{\varrho}^{\,\varepsilon}\|_1$ *both tend to zero.*

Proof Considering the decomposition (29) we first have

$$\begin{aligned}I_1(\vec{U}^\varepsilon, \vec{U}^\varepsilon) &= \varepsilon\int_\omega l\, {}^0H^{\alpha\beta\lambda\mu}\Big(\gamma_{\alpha\beta}(\vec{u}^\varepsilon)\gamma_{\lambda\mu}(\vec{u}^\varepsilon) + \frac{t^2}{6}\gamma_{\alpha\beta}(\vec{u}^\varepsilon)k_{\lambda\mu}(\vec{\theta}^\varepsilon, \vec{\varrho}^{\,\varepsilon})\\ &\quad + \frac{t^2}{12}\hat{\chi}_{\alpha\beta}(\vec{u}^\varepsilon, \vec{\theta}^\varepsilon)\hat{\chi}_{\lambda\mu}(\vec{u}^\varepsilon, \vec{\theta}^\varepsilon) + \frac{t^4}{80}k_{\alpha\beta}(\vec{\theta}^\varepsilon, \vec{\varrho}^{\,\varepsilon})k_{\lambda\mu}(\vec{\theta}^\varepsilon, \vec{\varrho}^{\,\varepsilon})\Big)\sqrt{a}\, d\xi^1 d\xi^2 + \bar{I}_1(\vec{U}^\varepsilon, \vec{U}^\varepsilon)\\ &= {}^0I_1(\vec{U}^\varepsilon, \vec{U}^\varepsilon) + \bar{I}_1(\vec{U}^\varepsilon, \vec{U}^\varepsilon),\end{aligned}$$

in which we can show that $\frac{1}{\varepsilon^3}\bar{I}_1(\vec{U}^\varepsilon, \vec{U}^\varepsilon) = R_\varepsilon$ using (38) and Lemma 1. Defining next

$$D_{1,\varepsilon} = \int_\omega \frac{l^3}{12}\, {}^0H^{\alpha\beta\lambda\mu}\hat{\chi}_{\alpha\beta}(\vec{u}^\varepsilon - \vec{u}^0, \vec{\theta}^\varepsilon - \vec{\theta}^0)\hat{\chi}_{\lambda\mu}(\vec{u}^\varepsilon - \vec{u}^0, \vec{\theta}^\varepsilon - \vec{\theta}^0)\, dS,$$

and

$$D'_{1,\varepsilon} = \int_\omega \frac{l^3}{12}\, {}^0H^{\alpha\beta\lambda\mu}\hat{\chi}_{\alpha\beta}(\vec{u}^\varepsilon,\vec{\theta}^\varepsilon)\hat{\chi}_{\lambda\mu}(\vec{u}^\varepsilon,\vec{\theta}^\varepsilon)\, dS,$$

we develop $D_{1,\varepsilon}$, and by using Lemma 2 we obtain

$$\begin{aligned} D_{1,\varepsilon} &= D'_{1,\varepsilon} + \int_\omega \frac{l^3}{12}\, {}^0H^{\alpha\beta\lambda\mu}\Big(\hat{\chi}_{\alpha\beta}(\vec{u}^0,\vec{\theta}^0)\hat{\chi}_{\lambda\mu}(\vec{u}^0,\vec{\theta}^0) - 2\hat{\chi}_{\alpha\beta}(\vec{u}^\varepsilon,\vec{\theta}^\varepsilon)\hat{\chi}_{\lambda\mu}(\vec{u}^0,\vec{\theta}^0)\, dS\Big) \\ &= D'_{1,\varepsilon} - \int_\omega \frac{l^3}{12}\, {}^0H^{\alpha\beta\lambda\mu}\hat{\chi}_{\alpha\beta}(\vec{u}^0,\vec{\theta}^0)\hat{\chi}_{\lambda\mu}(\vec{u}^0,\vec{\theta}^0)\, dS + R_\varepsilon, \end{aligned} \tag{39}$$

with

$$\begin{aligned} D'_{1,\varepsilon} = \frac{1}{\varepsilon^3}{}^0I_1(\vec{U}^\varepsilon,\vec{U}^\varepsilon) - \frac{1}{\varepsilon^2}\int_\omega l\, {}^0H^{\alpha\beta\lambda\mu}\gamma_{\alpha\beta}(\vec{u}^\varepsilon)\gamma_{\lambda\mu}(\vec{u}^\varepsilon)\, dS \\ - \int_\omega \frac{l^3}{6}\, {}^0H^{\alpha\beta\lambda\mu}\gamma_{\alpha\beta}(\vec{u}^\varepsilon)k_{\lambda\mu}(\vec{\theta}^\varepsilon,\vec{\varrho}^\varepsilon)\, dS - \varepsilon^2\int_\omega \frac{l^5}{80}\, {}^0H^{\alpha\beta\lambda\mu}k_{\alpha\beta}(\vec{\theta}^\varepsilon,\vec{\varrho}^\varepsilon)k_{\lambda\mu}(\vec{\theta}^\varepsilon,\vec{\varrho}^\varepsilon)\, dS. \end{aligned} \tag{40}$$

We proceed likewise with the other terms in (29). For

$$\begin{aligned} I_2(\vec{U}^\varepsilon,\vec{U}^\varepsilon) &= \varepsilon\int_\omega 4l\, {}^0H^{\alpha3\beta3}\Big(\zeta_\alpha(\vec{u}^\varepsilon,\vec{\theta}^\varepsilon)\zeta_\beta(\vec{u}^\varepsilon,\vec{\theta}^\varepsilon) + \frac{t^2}{6}\zeta_\alpha(\vec{u}^\varepsilon,\vec{\theta}^\varepsilon)n_\beta(\vec{\varrho}^\varepsilon) \\ &\quad + \frac{t^2}{12}m_\alpha(\vec{\theta}^\varepsilon,\vec{\varrho}^\varepsilon)m_\beta(\vec{\theta}^\varepsilon,\vec{\varrho}^\varepsilon) + \frac{t^4}{80}n_\alpha(\vec{\varrho}^\varepsilon)n_\beta(\vec{\varrho}^\varepsilon)\Big)\sqrt{a}\, d\xi^1 d\xi^2 + \bar{I}_2(\vec{U}^\varepsilon,\vec{U}^\varepsilon) \\ &= {}^0I_2(\vec{U}^\varepsilon,\vec{U}^\varepsilon) + \bar{I}_2(\vec{U}^\varepsilon,\vec{U}^\varepsilon), \end{aligned}$$

we can show that $\frac{1}{\varepsilon^3}\bar{I}_2(\vec{U}^\varepsilon,\vec{U}^\varepsilon) = R_\varepsilon$, and defining

$$D_{2,\varepsilon} = \int_\omega \frac{l^3}{3}\, {}^0H^{\alpha3\lambda3}m_\alpha(\vec{\theta}^\varepsilon - \vec{\theta}^0, \vec{\varrho}^\varepsilon - \vec{\varrho}^0)m_\beta(\vec{\theta}^\varepsilon - \vec{\theta}^0, \vec{\varrho}^\varepsilon - \vec{\varrho}^0)\, dS,$$

and

$$D'_{2,\varepsilon} = \int_\omega \frac{l^3}{3}\, {}^0H^{\alpha3\lambda3}m_\alpha(\vec{\theta}^\varepsilon,\vec{\varrho}^\varepsilon)m_\beta(\vec{\theta}^\varepsilon,\vec{\varrho}^\varepsilon)\, dS,$$

we obtain

$$D_{2,\varepsilon} = D'_{2,\varepsilon} - \int_\omega \frac{l^3}{3}\, {}^0H^{\alpha3\lambda3}m_\alpha(\vec{\theta}^0,\vec{\varrho}^0)m_\beta(\vec{\theta}^0,\vec{\varrho}^0)\, dS + R_\varepsilon, \tag{41}$$

with

$$\begin{aligned} D'_{2,\varepsilon} = \frac{1}{\varepsilon^3}{}^0I_2(\vec{U}^\varepsilon,\vec{U}^\varepsilon) - \frac{1}{\varepsilon^2}\int_\omega 4l\, {}^0H^{\alpha3\lambda3}\zeta_\alpha(\vec{u}^\varepsilon,\vec{\theta}^\varepsilon)\zeta_\beta(\vec{u}^\varepsilon,\vec{\theta}^\varepsilon)\, dS \\ - \int_\omega \frac{2l^3}{3}\, {}^0H^{\alpha3\lambda3}\zeta_\alpha(\vec{u}^\varepsilon,\vec{\theta}^\varepsilon)n_\beta(\vec{\varrho}^\varepsilon)\, dS - \varepsilon^2\int_\omega \frac{l^5}{20}\, {}^0H^{\alpha3\lambda3}n_\alpha(\vec{\varrho}^\varepsilon)n_\beta(\vec{\varrho}^\varepsilon)\, dS. \end{aligned} \tag{42}$$

Next, for

$$\begin{aligned} I_3(\vec{U}^\varepsilon,\vec{U}^\varepsilon) &= \varepsilon\int_\omega l\, {}^0H^{3333}\Big(\big(\delta(\vec{\theta}^\varepsilon)\big)^2 + \frac{t^2}{12}\big(p(\vec{\varrho}^\varepsilon)\big)^2\Big)\sqrt{a}\, d\xi^1 d\xi^2 + \bar{I}_3(\vec{U}^\varepsilon,\vec{U}^\varepsilon) \\ &= {}^0I_3(\vec{U}^\varepsilon,\vec{U}^\varepsilon) + \bar{I}_3(\vec{U}^\varepsilon,\vec{U}^\varepsilon), \end{aligned}$$

we have $\frac{1}{\varepsilon^3}\bar{I}_3(\vec{U}^\varepsilon,\vec{U}^\varepsilon)=R_\varepsilon$, and defining

$$D_{3,\varepsilon}=\int_\omega \frac{l^3}{12}\,{}^0H^{3333}\big(p(\vec{\varrho}^{\,\varepsilon})-p(\vec{\varrho}^{\,0})\big)^2dS,$$

and

$$D'_{3,\varepsilon}=\int_\omega \frac{l^3}{12}\,{}^0H^{3333}\big(p(\vec{\varrho}^{\,\varepsilon})\big)^2dS,$$

we obtain

$$D_{3,\varepsilon}=D'_{3,\varepsilon}-\int_\omega \frac{l^3}{12}\,{}^0H^{3333}\big(p(\vec{\varrho}^{\,0})\big)^2dS+R_\varepsilon, \tag{43}$$

with

$$D'_{3,\varepsilon}=\frac{1}{\varepsilon^3}{}^0I_3(\vec{U}^\varepsilon,\vec{U}^\varepsilon)-\frac{1}{\varepsilon^2}\int_\omega l\,{}^0H^{3333}\big(\delta(\vec{\theta}^{\,\varepsilon})\big)^2dS. \tag{44}$$

Finally, for the last two terms $I_4(\vec{U}^\varepsilon,\vec{U}^\varepsilon)=I_5(\vec{U}^\varepsilon,\vec{U}^\varepsilon)$, denoting

$$\begin{aligned}I_4(\vec{U}^\varepsilon,\vec{U}^\varepsilon)=\varepsilon\int_\omega l\,{}^0H^{\alpha\beta33}\Big(&\gamma_{\alpha\beta}(\vec{u}^{\,\varepsilon})\delta(\vec{\theta}^{\,\varepsilon})+\frac{t^2}{12}\,\hat{\chi}_{\alpha\beta}(\vec{u}^{\,\varepsilon},\vec{\theta}^{\,\varepsilon})p(\vec{\varrho}^{\,\varepsilon})\\&+\frac{t^2}{12}k_{\alpha\beta}(\vec{\theta}^{\,\varepsilon},\vec{\varrho}^{\,\varepsilon})\delta(\vec{\theta}^{\,\varepsilon})\Big)\sqrt{a}\,d\xi^1d\xi^2+\bar{I}_4(\vec{U}^\varepsilon,\vec{U}^\varepsilon)\\&={}^0I_4(\vec{U}^\varepsilon,\vec{U}^\varepsilon)+\bar{I}_4(\vec{U}^\varepsilon,\vec{U}^\varepsilon),\end{aligned}$$

we can show that $\frac{1}{\varepsilon^3}\bar{I}_4(\vec{U}^\varepsilon,\vec{U}^\varepsilon)=R_\varepsilon$, and with

$$D_{4,\varepsilon}=\int_\omega \frac{l^3}{12}\,{}^0H^{\alpha\beta33}\hat{\chi}_{\alpha\beta}(\vec{u}^{\,\varepsilon}-\vec{u}^{\,0},\vec{\theta}^{\,\varepsilon}-\vec{\theta}^{\,0})\,p(\vec{\varrho}^{\,\varepsilon}-\vec{\varrho}^{\,0})\,dS,$$

and

$$D'_{4,\varepsilon}=\int_\omega \frac{l^3}{12}\,{}^0H^{\alpha\beta33}\hat{\chi}_{\alpha\beta}(\vec{u}^{\,\varepsilon},\vec{\theta}^{\,\varepsilon})p(\vec{\varrho}^{\,\varepsilon})\,dS,$$

we obtain

$$D_{4,\varepsilon}=D'_{4,\varepsilon}-\int_\omega \frac{l^3}{12}\,{}^0H^{\alpha\beta33}\hat{\chi}_{\alpha\beta}(\vec{u}^{\,0},\vec{\theta}^{\,0})p(\vec{\varrho}^{\,0})\,dS+R_\varepsilon, \tag{45}$$

with

$$\begin{aligned}D'_{4,\varepsilon}=\frac{1}{\varepsilon^3}{}^0I_4(\vec{U}^\varepsilon,\vec{U}^\varepsilon)-\frac{1}{\varepsilon^2}\int_\omega l\,{}^0H^{\alpha\beta33}\gamma_{\alpha\beta}(\vec{u}^{\,\varepsilon})\delta(\vec{\theta}^{\,\varepsilon})\,dS\\-\int_\omega \frac{l^3}{12}\,{}^0H^{\alpha\beta33}k_{\alpha\beta}(\vec{\theta}^{\,\varepsilon},\vec{\varrho}^{\,\varepsilon})\delta(\vec{\theta}^{\,\varepsilon})\,dS.\end{aligned} \tag{46}$$

Defining now $D_\varepsilon=D_{1,\varepsilon}+D_{2,\varepsilon}+D_{3,\varepsilon}+2D_{4,\varepsilon}$, $D'_\varepsilon=D'_{1,\varepsilon}+D'_{2,\varepsilon}+D'_{3,\varepsilon}+2D'_{4,\varepsilon}$ and $\bar{I}=\bar{I}_1+\bar{I}_2+\bar{I}_3+2\bar{I}_4$, and using (40), (42), (44) and (46), we obtain

$$D'_\varepsilon=\frac{1}{\varepsilon^3}A^{3Ds}(\vec{U}^\varepsilon,\vec{U}^{\,\varepsilon})-\frac{\bar{I}}{\varepsilon^3}-\tilde{D}_\varepsilon,$$

with

$$\begin{aligned}\tilde{D}_\varepsilon = &\frac{1}{\varepsilon^2}\int_\omega l\,{}^0H^{\alpha\beta\lambda\mu}\gamma_{\alpha\beta}(\vec{u}^\varepsilon)\gamma_{\lambda\mu}(\vec{u}^\varepsilon)\,dS + \int_\omega \frac{l^3}{6}\,{}^0H^{\alpha\beta\lambda\mu}\gamma_{\alpha\beta}(\vec{u}^\varepsilon)k_{\lambda\mu}(\vec{\theta}^\varepsilon,\vec{\varrho}^\varepsilon)\,dS\\ &+\varepsilon^2\int_\omega \frac{l^5}{80}\,{}^0H^{\alpha\beta\lambda\mu}k_{\alpha\beta}(\vec{\theta}^\varepsilon,\vec{\varrho}^\varepsilon)k_{\lambda\mu}(\vec{\theta}^\varepsilon,\vec{\varrho}^\varepsilon)\,dS + \frac{1}{\varepsilon^2}\int_\omega 4l\,{}^0H^{\alpha 3\lambda 3}\zeta_\alpha(\vec{u}^\varepsilon,\vec{\theta}^\varepsilon)\zeta_\beta(\vec{u}^\varepsilon,\vec{\theta}^\varepsilon)\,dS\\ &+\int_\omega \frac{2l^3}{3}{}^0H^{\alpha 3\lambda 3}n_\alpha(\vec{\varrho}^\varepsilon)\zeta_\beta(\vec{u}^\varepsilon,\vec{\theta}^\varepsilon)\,dS + \varepsilon^2\int_\omega \frac{l^5}{20}{}^0H^{\alpha 3\lambda 3}n_\alpha(\vec{\varrho}^\varepsilon)n_\beta(\vec{\varrho}^\varepsilon)\,dS\\ &+\frac{1}{\varepsilon^2}\int_\omega l\,{}^0H^{3333}\big(\delta(\vec{\theta}^\varepsilon)\big)^2\,dS + \frac{2}{\varepsilon^2}\int_\omega l\,{}^0H^{\alpha\beta 33}\gamma_{\alpha\beta}(\vec{u}^\varepsilon)\delta(\vec{\theta}^\varepsilon)\,dS\\ &+2\int_\omega \frac{l^3}{12}\,{}^0H^{\alpha\beta 33}k_{\alpha\beta}(\vec{\theta}^\varepsilon,\vec{\varrho}^\varepsilon)\delta(\vec{\theta}^\varepsilon)\,dS.\end{aligned}$$

Using the variational equation (18) and recalling $\varepsilon^{-3}\bar{I} = R_\varepsilon$ we get

$$D'_\varepsilon = \frac{1}{\varepsilon^3}F^{3Ds}(\vec{U}^\varepsilon) - \tilde{D}_\varepsilon + R_\varepsilon, \tag{47}$$

while we have by using (36) together with (39), (41), (43) and (45)

$$D_\varepsilon = D'_\varepsilon - A_{\rm b}^{3Ds}(\vec{u}^0,\vec{\theta}^0,\vec{\varrho}^0;\vec{u}^0,\vec{\theta}^0,\vec{\varrho}^0) + R_\varepsilon = D'_\varepsilon - G^{3Ds}(\vec{u}^0) + R_\varepsilon,$$

which yields, when substituting (47),

$$D_\varepsilon = \frac{1}{\varepsilon^3}F^{3Ds}(\vec{U}^\varepsilon) - G^{3Ds}(\vec{u}^0) - \tilde{D}_\varepsilon + R_\varepsilon = -\tilde{D}_\varepsilon + R_\varepsilon,$$

using the weak convergence. We have thus obtained

$$D_\varepsilon + \tilde{D}_\varepsilon = R_\varepsilon. \tag{48}$$

Clearly, $D_\varepsilon \geq 0$, and to complete the proof we will now show that

$$\tilde{D}_\varepsilon \geq C\Big(\frac{1}{\varepsilon^2}\|\underline{\underline{\gamma}}\|_0^2 + \frac{1}{\varepsilon^2}\|\underline{\zeta}\|_0^2 + \frac{1}{\varepsilon^2}\|\delta\|_0 + \varepsilon^2\|\underline{\underline{k}}\|_0^2 + \varepsilon^2\|\underline{n}\|_0^2\Big). \tag{49}$$

Defining the tensorial components

$$\begin{cases} f_{\alpha\beta} = \frac{1}{\varepsilon}\gamma_{\alpha\beta}(\vec{u}^\varepsilon) + \frac{l^2\varepsilon}{12}k_{\alpha\beta}(\vec{\theta}^\varepsilon,\vec{\varrho}^\varepsilon)\\ f_{\alpha 3} = \frac{1}{\varepsilon}\zeta_\alpha(\vec{u}^\varepsilon,\vec{\theta}^\varepsilon) + \frac{l^2\varepsilon}{12}n_\alpha(\vec{\varrho}^\varepsilon)\\ f_{33} = \frac{1}{\varepsilon}\delta(\vec{\theta}^\varepsilon)\end{cases}$$

we have

$$\begin{aligned}\int_\omega l\,{}^0H^{ijkl}f_{ij}f_{kl}\,dS = &\int_\omega l\,{}^0H^{\alpha\beta\lambda\mu}f_{\alpha\beta}f_{\lambda\mu}\,dS + 2\int_\omega l\,{}^0H^{\alpha\beta 33}f_{\alpha\beta}f_{33}\,dS\\ &+4\int_\omega l\,{}^0H^{\alpha 3\beta 3}f_{\alpha 3}f_{\beta 3}\,dS + \int_\omega l\,{}^0H^{3333}f_{33}f_{33}\,dS,\end{aligned}$$

hence,

$$\begin{aligned}\tilde{D}_\varepsilon &= \int_\omega l\,{}^0H^{ijkl}f_{ij}f_{kl}\,dS + \varepsilon^2\int_\omega l^5\Big(\frac{1}{80}-\frac{1}{144}\Big){}^0H^{\alpha\beta\lambda\mu}k_{\alpha\beta}(\vec{\theta}^\varepsilon,\vec{\varrho}^\varepsilon)k_{\lambda\mu}(\vec{\theta}^\varepsilon,\vec{\varrho}^\varepsilon)\,dS\\ &\quad+\varepsilon^2\int_\omega 4l^5\Big(\frac{1}{80}-\frac{1}{144}\Big){}^0H^{\alpha 3\lambda 3}n_\alpha(\vec{\varrho}^\varepsilon)n_\beta(\vec{\varrho}^\varepsilon)\,dS\\ &\geq C\int_\omega l\,{}^0H^{ijkl}f_{ij}f_{kl}\,dS,\end{aligned}$$

which implies (49). Therefore, due to (48) both D_ε and $\tilde{D}_\varepsilon$ tend to zero. Then, with $\tilde{D}_\varepsilon$ converging to zero and (38), we directly infer

$$\frac{1}{\varepsilon}\|\vec{u}^\varepsilon, \vec{\theta}^\varepsilon\|_{\rm m} \overset{\varepsilon\to 0}{\longrightarrow} 0, \quad \varepsilon\|\vec{u}^\varepsilon, \vec{\theta}^\varepsilon, \vec{\varrho}^\varepsilon\|_1 \overset{\varepsilon\to 0}{\longrightarrow} 0, \tag{50}$$

recalling Lemma 1. Hence, $\|\vec{u}^\varepsilon, \vec{\theta}^\varepsilon\|_{\rm m}$ tends to zero *a fortiori*, and combined with D_ε converging to zero this finally proves that $(\vec{u}^\varepsilon, \vec{\theta}^\varepsilon, \vec{\varrho}^\varepsilon)$ converges strongly in $\mathcal{V}_{\rm b}^{3Ds}$ to $(\vec{u}^0, \vec{\theta}^0, \vec{\varrho}^0)$, and all the claims of the Proposition are thus established. □

Remark 1 *The asymptotic behavior is essentially of a penalization type in this case – as for the bending-dominated behavior of classical shell models [8] – since the limit solution lies in the constrained subspace $\mathcal{V}_{\rm 0b}^{3Ds}$, albeit also with some secondary features of singular perturbation here, since the regularity of $\mathcal{V}_{\rm b}^{3Ds}$ is lower than that of the original space $\mathcal{V}^{3Ds}$, except for the first-order term $\vec{v}$ in the polynomial expression, recall* (27).

Remark 2 *The displacement of any point within the shell body being given by*

$$\vec{U}^\varepsilon = \vec{u}^\varepsilon(\xi^1,\xi^2) + \xi^3\vec{\theta}^\varepsilon(\xi^1,\xi^2) + (\xi^3)^2\vec{\varrho}^\varepsilon(\xi^1,\xi^2), \quad \xi^3 \in \varepsilon \times \Big]-\frac{l(\xi^1,\xi^2)}{2}, +\frac{l(\xi^1,\xi^2)}{2}\Big[, \tag{51}$$

the additional convergence results (50) *directly imply uniform convergence properties of $\vec{U}^\varepsilon$ to $\vec{u}^0$ in $H^1(\mathcal{S})$, e.g., when taking $\xi^3/(\varepsilon l)$ as a parameter, such as for the the top and bottom surfaces of the shell.*

3.3 Inhibited pure bending

We now assume

$$\mathcal{V}_0^{3Ds} \cap \{(\vec{v},\vec{\eta},\vec{0}\,) \in \mathcal{V}^{3Ds}\} = \{(\vec{0},\vec{0},\vec{0})\},$$

in which case $\|\vec{v},\vec{\eta}\,\|_{\rm m}$ gives a norm, and we accordingly introduce $\mathcal{V}_{\rm m}^{3Ds}$ as the completion of $\mathcal{V}^\natural = \{(\vec{v},\vec{\eta}\,)$ such that $(\vec{v},\vec{\eta},\vec{0}) \in \mathcal{V}^{3Ds}\}$ with respect to this norm. The convergence of the asymptotic behavior will then be obtained in this space for the scaling corresponding to $\rho = 1$, and the candidate limit problem reads:
Find $(\vec{u}^{\rm m}, \vec{\theta}^{\rm m}) \in \mathcal{V}_{\rm m}^{3Ds}$ such that

$$A_{\rm m}^{3Ds}(\vec{u}^{\rm m}, \vec{\theta}^{\rm m}; \vec{v}, \vec{\eta}) = G^{3Ds}(\vec{v}), \quad \forall (\vec{v},\vec{\eta}\,) \in \mathcal{V}_{\rm m}^{3Ds}. \tag{52}$$

We point out that, in order to obtain a well-posed limit problem, we need to enforce that $G^{3Ds} \in (\mathcal{V}_{\rm m}^{3Ds})'$, namely, that

$$\Big|\int_\omega l\vec{G}_0 \cdot \vec{v}\, dS\Big| \le C\|\vec{v},\vec{\eta}\,\|_{\rm m}, \quad \forall (\vec{v},\vec{\eta}\,) \in \mathcal{V}_{\rm m}^{3Ds}, \tag{53}$$

which we call the condition of admissible loading for membrane-dominated behavior, see [8, 21] and references therein for discussions on this condition. Provided this condition is satisfied, the variational problem (52) is clearly well-posed, by construction.

We start by proving weak convergence, in a more detailed form than in [9, 8].

Lemma 3 *Assuming that $G^{3Ds} \in (\mathcal{V}_{\rm m}^{3Ds})'$, we have that $\big(\vec{u}^\varepsilon, \vec{\theta}^\varepsilon\big)$ converges weakly in $\mathcal{V}_{\rm m}^{3Ds}$, as ε goes to 0, to $(\vec{u}^{\rm m}, \vec{\theta}^{\rm m})$ solution of* (52). *Moreover, $\varepsilon^2\vec{\varrho}^\varepsilon$ converges weakly in $H^1(\mathcal{S})$ to 0.*

Proof We divide the proof into two steps.
i) *Extraction of weakly-converging sequence and preliminary identifications.* We can establish the following uniform bound on the sequence of solutions [9, 8]

$$\|\vec{u}^\varepsilon, \vec{\theta}^\varepsilon\|_{\mathrm{m}} + \varepsilon\|\vec{u}^\varepsilon, \vec{\theta}^\varepsilon, \vec{\varrho}^\varepsilon\|_{\mathrm{b}} + \varepsilon^2\|\vec{u}^\varepsilon, \vec{\theta}^\varepsilon, \vec{\varrho}^\varepsilon\|_1 \leq C. \tag{54}$$

Hence, we can extract a subsequence – also denoted by $(\vec{u}^\varepsilon, \vec{\theta}^\varepsilon, \vec{\varrho}^\varepsilon)$ for convenience – such that we have the weak convergence of $(\vec{u}^\varepsilon, \vec{\theta}^\varepsilon)$ to some $(\vec{u}^w, \vec{\theta}^w)$ in $\mathcal{V}_{\mathrm{m}}^{3Ds}$, together with the weak convergence of $\varepsilon^2\vec{\varrho}^\varepsilon$ to some $\vec{\varrho}^w$ in $H^1(\mathcal{S})$. We will now show that

$$\vec{\varrho}^w = 0. \tag{55}$$

Recalling the second equivalence of norms in Lemma 1 and the expressions of m_α and p in (15), the boundedness of $\varepsilon\|\vec{u}^\varepsilon, \vec{\theta}^\varepsilon, \vec{\varrho}^\varepsilon\|_{\mathrm{b}}$ implies that $\varepsilon(2\varrho_\alpha^\varepsilon + \theta_{3,\alpha}^\varepsilon)$ and $\varepsilon\varrho_3^\varepsilon$ are both bounded in $L^2(\mathcal{S})$. Therefore, $\varepsilon^2(2\varrho_\alpha^\varepsilon + \theta_{3,\alpha}^\varepsilon)$ and $\varepsilon^2\varrho_3^\varepsilon$ tend to zero in $L^2(\mathcal{S})$, which already shows that $\rho_3^w = 0$. Moreover, since $\varepsilon^2\rho_\alpha$ tends to ρ_α^w in $L^2(\mathcal{S})$, we infer that $\varepsilon^2\theta_{3,\alpha}^\varepsilon$ tends to $-2\rho_\alpha^w$ in $L^2(\mathcal{S})$. On the other hand, the boundedness of $\|\vec{u}^\varepsilon, \vec{\theta}^\varepsilon\|_{\mathrm{m}}$ implies that θ_3^ε is bounded in $L^2(\mathcal{S})$. Hence, $\varepsilon^2\theta_3^\varepsilon$ tends to zero in $L^2(\mathcal{S})$, thus also in the distribution sense. As a consequence, $\varepsilon^2\theta_{3,\alpha}^\varepsilon$ also tends to zero in the distribution sense, which shows that $\rho_\alpha^w = 0$, hence (55) is proven.

ii) *Identification of weak limit* $(\vec{u}^w, \vec{\theta}^w)$. We will again use the decomposition (29), and similar notation as in the proof of Proposition 2 for

$$I_i(\vec{U}, \vec{V}) = {}^0I_i(\vec{U}, \vec{V}) + \bar{I}_i(\vec{U}, \vec{V}), \quad i = 1, \ldots, 5, \tag{56}$$

where of course all the mixed terms present in the above-defined expressions of ${}^0I_i(\vec{U}, \vec{U})$ are symmetrized in ${}^0I_i(\vec{U}, \vec{V})$. Starting with the first term, we have

$$\begin{aligned} {}^0I_1(\vec{U}^\varepsilon, \vec{V}) = \varepsilon \int_\omega l\, {}^0H^{\alpha\beta\lambda\mu} \Big(& \gamma_{\alpha\beta}(\vec{u}^\varepsilon)\gamma_{\lambda\mu}(\vec{v}) + \frac{t^2}{12}\gamma_{\alpha\beta}(\vec{u}^\varepsilon)k_{\lambda\mu}(\vec{\eta}, \vec{\varsigma}) \\ & + \frac{t^2}{12}\gamma_{\alpha\beta}(\vec{v})k_{\lambda\mu}(\vec{\theta}^\varepsilon, \vec{\varrho}^\varepsilon) + \frac{t^2}{12}\hat{\chi}_{\alpha\beta}(\vec{u}^\varepsilon, \vec{\theta}^\varepsilon)\hat{\chi}_{\lambda\mu}(\vec{v}, \vec{\eta}) \\ & + \frac{t^4}{80}k_{\alpha\beta}(\vec{\theta}^\varepsilon, \vec{\varrho}^\varepsilon)k_{\lambda\mu}(\vec{\eta}, \vec{\varsigma}) \Big)\, dS, \end{aligned}$$

and we can easily show for the remainder that $\varepsilon^{-1}\bar{I}_1$ tends to zero based on (54). Concerning $\varepsilon^{-1}\,{}^0I_1$, the weak convergence of $(\vec{u}^\varepsilon, \vec{\theta}^\varepsilon)$ implies that the first term under the integral tends to

$$\int_\omega l\, {}^0H^{\alpha\beta\lambda\mu}\, \gamma_{\alpha\beta}(\vec{u}^w)\gamma_{\lambda\mu}(\vec{v})\, dS,$$

and the uniform bound (54) directly implies that all other terms tend to zero, except for the third one which requires further analysis. We have

$$\varepsilon^2 k_{\lambda\mu}(\vec{\theta}^\varepsilon, \vec{\varrho}^\varepsilon) = \varepsilon^2\Big(\frac{1}{2}(\varrho_{\lambda|\mu}^\varepsilon + \varrho_{\mu|\lambda}^\varepsilon) - b_{\lambda\mu}\varrho_3^\varepsilon - \frac{1}{2}(b_\lambda^\alpha\theta_{\alpha|\mu}^\varepsilon + b_\mu^\alpha\theta_{\alpha|\lambda}^\varepsilon) + c_{\lambda\mu}\theta_3^\varepsilon\Big),$$

and the weak convergence of $\varepsilon^2\vec{\varrho}^\varepsilon$ to zero in $H^1(\mathcal{S})$ implies that the terms with $\vec{\varrho}^\varepsilon$ tend to zero weakly in $L^2(\mathcal{S})$. Then from the boundedness of $\varepsilon\|\vec{u}^\varepsilon, \vec{\theta}^\varepsilon, \vec{\varrho}^\varepsilon\|_{\mathrm{b}}$ in (54) and recalling (27) we infer that the terms with $\vec{\theta}^\varepsilon$ tend to zero strongly in $L^2(\mathcal{S})$. This shows that all terms but the first one vanish in $\varepsilon^{-1}\,{}^0I_1$ in the limit, hence

$$\varepsilon^{-1}I_1(\vec{U}^\varepsilon, \vec{V}) \overset{\varepsilon\to 0}{\longrightarrow} \int_\omega l\, {}^0H^{\alpha\beta\lambda\mu}\, \gamma_{\alpha\beta}(\vec{u}^w)\gamma_{\lambda\mu}(\vec{v})\, dS. \tag{57}$$

All the other terms in the decomposition (29) can be handled with very similar arguments, which leads to

$$\varepsilon^{-1} I_2(\vec{U}^\varepsilon, \vec{V}) \overset{\varepsilon\to 0}{\longrightarrow} \int_\omega 4l\, {}^0H^{\alpha 3\beta 3} \zeta_\alpha(\vec{u}^{\,w}, \vec{\theta}^{\,w}) \zeta_\beta(\vec{v}, \vec{\eta})\, dS, \tag{58}$$

$$\varepsilon^{-1} I_3(\vec{U}^\varepsilon, \vec{V}) \overset{\varepsilon\to 0}{\longrightarrow} \int_\omega l\, {}^0H^{3333} \delta(\vec{\theta}^{\,w}) \delta(\vec{\eta})\, dS, \tag{59}$$

$$\varepsilon^{-1} I_4(\vec{U}^\varepsilon, \vec{V}) \overset{\varepsilon\to 0}{\longrightarrow} \int_\omega l\, {}^0H^{\alpha\beta 33} \gamma_{\alpha\beta}(\vec{u}^{\,w}) \delta(\vec{\eta})\, dS, \tag{60}$$

$$\varepsilon^{-1} I_5(\vec{U}^\varepsilon, \vec{V}) \overset{\varepsilon\to 0}{\longrightarrow} \int_\omega l\, {}^0H^{\alpha\beta 33} \gamma_{\alpha\beta}(\vec{v}) \delta(\vec{\theta}^{\,w})\, dS. \tag{61}$$

Therefore, in the variational formulation (18) we have

$$\varepsilon^{-1} A^{3Ds}(\vec{u}^\varepsilon, \vec{\theta}^\varepsilon, \vec{\varrho}^\varepsilon; \vec{v}, \vec{\eta}, \vec{\varsigma}) \overset{\varepsilon\to 0}{\longrightarrow} A_{\rm m}^{3Ds}(\vec{u}^{\,w}, \vec{\theta}^{\,w}; \vec{v}, \vec{\eta}).$$

Moreover, it is straightforward to show that

$$\varepsilon^{-1} F^{3Ds}(\vec{v}, \vec{\eta}, \vec{\varsigma}) \overset{\varepsilon\to 0}{\longrightarrow} G^{3Ds}(\vec{v}).$$

Finally, taking the limit of (18) multiplied by ε^{-1}, we see that $(\vec{u}^{\,w}, \vec{\theta}^{\,w})$ satisfies

$$A_{\rm m}^{3Ds}(\vec{u}^{\,w}, \vec{\theta}^{\,w}; \vec{v}, \vec{\eta}) = G^{3Ds}(\vec{v}),$$

for any $(\vec{v}, \vec{\eta}) \in \mathcal{V}^\sharp$, hence also for any $(\vec{v}, \vec{\eta}) \in \mathcal{V}_{\rm m}^{3Ds}$ since $\mathcal{V}^\sharp$ is dense in $\mathcal{V}_{\rm m}^{3Ds}$ by construction of the latter space. This shows that $(\vec{u}^{\,w}, \vec{\theta}^{\,w}) = (\vec{u}^{\,\rm m}, \vec{\theta}^{\,\rm m})$ for the weak convergence of the extracted subsequence, hence this convergence holds for the whole original sequence since the limit is hereby uniquely characterized. □

We are now in a position to prove the final result of this paper, namely, that the convergence is in fact strong in this case also.

Proposition 3 *Assuming that $G^{3Ds} \in (\mathcal{V}_{\rm m}^{3Ds})'$, we have that $\big(\vec{u}^\varepsilon, \vec{\theta}^\varepsilon\big)$ converges strongly in $\mathcal{V}_{\rm m}^{3Ds}$, as ε goes to 0, to $(\vec{u}^{\,\rm m}, \vec{\theta}^{\,\rm m})$ solution of (52). Furthermore, $\varepsilon\|\vec{u}^\varepsilon, \vec{\theta}^\varepsilon, \vec{\varrho}^\varepsilon\|_{\rm b}$ and $\varepsilon^2\|\vec{u}^\varepsilon, \vec{\theta}^\varepsilon, \vec{\varrho}^\varepsilon\|_1$ both tend to zero.*

Proof We again use the decomposition (29) and the splitting (56), both defined as in the proof of Proposition 2. Starting with $I_1(\vec{U}^\varepsilon, \vec{U}^\varepsilon)$, based on (54) we can now easily see that $\varepsilon^{-1}\bar{I}_1(\vec{U}^\varepsilon, \vec{U}^\varepsilon) = R_\varepsilon$. We then define the following new quantities (different from those introduced in the proof of Proposition 2 with the same notation), where the key point of this proof lies in an appropriate combination of strain tensors at various orders, viz.

$$D_{1,\varepsilon} = \int_\omega l\, {}^0H^{\alpha\beta\lambda\mu} \Big(\gamma_{\alpha\beta}(\vec{u}^\varepsilon - \vec{u}^{\,\rm m}) + \frac{t^2}{12} k_{\alpha\beta}(\vec{\theta}^\varepsilon, \vec{\varrho}^\varepsilon)\Big)\Big(\gamma_{\lambda\mu}(\vec{u}^\varepsilon - \vec{u}^{\,\rm m}) + \frac{t^2}{12} k_{\lambda\mu}(\vec{\theta}^\varepsilon, \vec{\varrho}^\varepsilon)\Big) dS,$$

and

$$D'_{1,\varepsilon} = \int_\omega l\, {}^0H^{\alpha\beta\lambda\mu} \Big(\gamma_{\alpha\beta}(\vec{u}^\varepsilon) + \frac{t^2}{12} k_{\alpha\beta}(\vec{\theta}^\varepsilon, \vec{\varrho}^\varepsilon)\Big)\Big(\gamma_{\lambda\mu}(\vec{u}^\varepsilon) + \frac{t^2}{12} k_{\lambda\mu}(\vec{\theta}^\varepsilon, \vec{\varrho}^\varepsilon)\Big) dS,$$

and we have

$$D_{1,\varepsilon} = D'_{1,\varepsilon} + \int_\omega l\, {}^0H^{\alpha\beta\lambda\mu} \gamma_{\alpha\beta}(\vec{u}^{\,\rm m}) \gamma_{\lambda\mu}(\vec{u}^{\,\rm m}) dS - \int_\omega 2l\, {}^0H^{\alpha\beta\lambda\mu} \Big(\gamma_{\alpha\beta}(\vec{u}^\varepsilon) + \frac{t^2}{12} k_{\alpha\beta}(\vec{\theta}^\varepsilon, \vec{\varrho}^\varepsilon)\Big) \gamma_{\lambda\mu}(\vec{u}^{\,\rm m}) dS.$$

Using the weak convergence properties of Lemma 3 and (54), we infer that

$$\int_\omega 2l\,{}^0H^{\alpha\beta\lambda\mu}\Big(\gamma_{\alpha\beta}(\vec{u}^{\,\varepsilon})+\frac{t^2}{12}k_{\alpha\beta}(\vec{\theta}^{\,\varepsilon},\vec{\varrho}^{\,\varepsilon})\Big)\gamma_{\lambda\mu}(\vec{u}^{\,\mathrm{m}})dS\overset{\varepsilon\to 0}{\longrightarrow}\int_\omega 2l\,{}^0H^{\alpha\beta\lambda\mu}\gamma_{\alpha\beta}(\vec{u}^{\,\mathrm{m}})\gamma_{\lambda\mu}(\vec{u}^{\,\mathrm{m}})dS,$$

hence,

$$D_{1,\varepsilon}=D'_{1,\varepsilon}-\int_\omega l\,{}^0H^{\alpha\beta\lambda\mu}\gamma_{\alpha\beta}(\vec{u}^{\,\mathrm{m}})\gamma_{\lambda\mu}(\vec{u}^{\,\mathrm{m}})dS+R_\varepsilon. \tag{62}$$

and $D'_{1,\varepsilon}$ is related to ${}^0I_1(\vec{U}^\varepsilon,\vec{U}^\varepsilon)$ by

$$D'_{1,\varepsilon}=\frac{1}{\varepsilon}{}^0I_1(\vec{U}^\varepsilon,\vec{U}^\varepsilon)-\int_\omega l\,{}^0H^{\alpha\beta\lambda\mu}\Big(\frac{t^2}{12}\hat{\chi}_{\alpha\beta}(\vec{u}^{\,\varepsilon},\vec{\theta}^{\,\varepsilon})\hat{\chi}_{\lambda\mu}(\vec{u}^{\,\varepsilon},\vec{\theta}^{\,\varepsilon})+\frac{t^4}{180}k_{\alpha\beta}(\vec{\theta}^{\,\varepsilon},\vec{\varrho}^{\,\varepsilon})k_{\lambda\mu}(\vec{\theta}^{\,\varepsilon},\vec{\varrho}^{\,\varepsilon})\Big)dS. \tag{63}$$

Proceeding similarly for the other terms in the decomposition (29), we have

$$\varepsilon^{-1}\bar{I}_i(\vec{U}^\varepsilon,\vec{U}^\varepsilon)=R_\varepsilon,\quad i=1,\dots,5.$$

Also, defining

$$D_{2,\varepsilon}=\int_\omega 4l\,{}^0H^{\alpha3\beta3}\Big(\zeta_\alpha(\vec{u}^{\,\varepsilon}-\vec{u}^{\,\mathrm{m}},\vec{\theta}^{\,\varepsilon}-\vec{\theta}^{\,\mathrm{m}})+\frac{t^2}{12}n_\alpha(\vec{\varrho}^{\,\varepsilon})\Big)\Big(\zeta_\beta(\vec{u}^{\,\varepsilon}-\vec{u}^{\,\mathrm{m}},\vec{\theta}^{\,\varepsilon}-\vec{\theta}^{\,\mathrm{m}})+\frac{t^2}{12}n_\beta(\vec{\varrho}^{\,\varepsilon})\Big)dS,$$

$$D'_{2,\varepsilon}=\int_\omega 4l\,{}^0H^{\alpha3\beta3}\Big(\zeta_\alpha(\vec{u}^{\,\varepsilon},\vec{\theta}^{\,\varepsilon})+\frac{t^2}{12}n_\alpha(\vec{\varrho}^{\,\varepsilon})\Big)\Big(\zeta_\beta(\vec{u}^{\,\varepsilon},\vec{\theta}^{\,\varepsilon})+\frac{t^2}{12}n_\beta(\vec{\varrho}^{\,\varepsilon})\Big)dS,$$

we obtain

$$D_{2,\varepsilon}=D'_{2,\varepsilon}-\int_\omega 4l\,{}^0H^{\alpha3\beta3}\zeta_\alpha(\vec{u}^{\,\mathrm{m}},\vec{\theta}^{\,\mathrm{m}})\zeta_\beta(\vec{u}^{\,\mathrm{m}},\vec{\theta}^{\,\mathrm{m}})\,dS+R_\varepsilon, \tag{64}$$

$$D'_{2,\varepsilon}=\frac{1}{\varepsilon}{}^0I_2(\vec{U}^\varepsilon,\vec{U}^\varepsilon)-\int_\omega l\,{}^0H^{\alpha3\beta3}\Big(\frac{t^2}{12}m_\alpha(\vec{\theta}^{\,\varepsilon},\vec{\varrho}^{\,\varepsilon})m_\beta(\vec{\theta}^{\,\varepsilon},\vec{\varrho}^{\,\varepsilon})+\frac{t^4}{180}n_\alpha(\vec{\varrho}^{\,\varepsilon})n_\beta(\vec{\varrho}^{\,\varepsilon})\Big)dS. \tag{65}$$

Likewise, for

$$D_{3,\varepsilon}=\int_\omega l\,{}^0H^{3333}\big(\delta(\vec{\theta}^{\,\varepsilon}-\vec{\theta}^{\,\mathrm{m}})\big)^2dS,\quad D'_{3,\varepsilon}=\int_\omega l\,{}^0H^{3333}\big(\delta(\vec{\theta}^{\,\varepsilon})\big)^2dS,$$

we have

$$D_{3,\varepsilon}=D'_{3,\varepsilon}-\int_\omega l\,{}^0H^{3333}\big(\delta(\vec{\theta}^{\,\mathrm{m}})\big)^2dS+R_\varepsilon, \tag{66}$$

$$D'_{3,\varepsilon}=\frac{1}{\varepsilon}{}^0I_3(\vec{U}^\varepsilon,\vec{U}^\varepsilon)-\int_\omega l\,{}^0H^{3333}\frac{t^2}{12}\big(p(\vec{\varrho}^{\,\varepsilon})\big)^2dS. \tag{67}$$

Finally, for

$$D_{4,\varepsilon}=\int_\omega l\,{}^0H^{\alpha\beta33}\Big(\gamma_{\alpha\beta}(\vec{u}^{\,\varepsilon}-\vec{u}^{\,\mathrm{m}})+\frac{t^2}{12}\,k_{\alpha\beta}(\vec{\theta}^{\,\varepsilon},\vec{\varrho}^{\,\varepsilon})\Big)\delta(\vec{\theta}^{\,\varepsilon}-\vec{\theta}^{\,\mathrm{m}})\,dS,$$

$$D'_{4,\varepsilon}=\int_\omega l\,{}^0H^{\alpha\beta33}\Big(\gamma_{\alpha\beta}(\vec{u}^{\,\varepsilon})+\frac{t^2}{12}\,k_{\alpha\beta}(\vec{\theta}^{\,\varepsilon},\vec{\varrho}^{\,\varepsilon})\Big)\delta(\vec{\theta}^{\,\varepsilon})\,dS,$$

we infer

$$D_{4,\varepsilon}=D'_{4,\varepsilon}-\int_\omega l\,{}^0H^{\alpha\beta33}\gamma_{\alpha\beta}(\vec{u}^{\,\mathrm{m}})\delta(\vec{\theta}^{\,\mathrm{m}})\,dS+R_\varepsilon, \tag{68}$$

$$D'_{4,\varepsilon}=\frac{1}{\varepsilon}{}^0I_4(\vec{U}^\varepsilon,\vec{U}^\varepsilon)-\int_\omega l\,{}^0H^{\alpha\beta33}\frac{t^2}{12}\hat{\chi}_{\alpha\beta}(\vec{u}^{\,\varepsilon},\vec{\theta}^{\,\varepsilon})p(\vec{\varrho}^{\,\varepsilon})\,dS. \tag{69}$$

We again define $D_\varepsilon = D_{1,\varepsilon} + D_{2,\varepsilon} + D_{3,\varepsilon} + 2D_{4,\varepsilon}$, $D'_\varepsilon = D'_{1,\varepsilon} + D'_{2,\varepsilon} + D'_{3,\varepsilon} + 2D'_{4,\varepsilon}$ and $\bar{I} = \bar{I}_1 + \bar{I}_2 + \bar{I}_3 + 2\bar{I}_4$, and we recall that $\varepsilon^{-1}\bar{I} = R_\varepsilon$. Gathering now (63), (65), (67) and (69), we get

$$D'_\varepsilon = \frac{1}{\varepsilon}A^{3Ds}(\vec{U}^\varepsilon, \vec{U}^\varepsilon) - \frac{1}{\varepsilon}\bar{I}(\vec{U}^\varepsilon, \vec{U}^\varepsilon) - \tilde{D}_\varepsilon,$$

for

$$\begin{aligned}\tilde{D}_\varepsilon = &\int_\omega l\,{}^0H^{\alpha\beta\lambda\mu}\Big(\frac{t^2}{12}\hat{\chi}_{\alpha\beta}(\vec{u}^\varepsilon, \vec{\theta}^\varepsilon)\hat{\chi}_{\lambda\mu}(\vec{u}^\varepsilon, \vec{\theta}^\varepsilon) + \frac{t^4}{180}k_{\alpha\beta}(\vec{\theta}^\varepsilon, \vec{\varrho}^\varepsilon)k_{\lambda\mu}(\vec{\theta}^\varepsilon, \vec{\varrho}^\varepsilon)\Big)dS\\ &+\int_\omega l\,{}^0H^{\alpha3\beta3}\Big(\frac{t^2}{12}m_\alpha(\vec{\theta}^\varepsilon, \vec{\varrho}^\varepsilon)m_\beta(\vec{\theta}^\varepsilon, \vec{\varrho}^\varepsilon) + \frac{t^4}{180}n_\alpha(\vec{\varrho}^\varepsilon)n_\beta(\vec{\varrho}^\varepsilon)\Big)dS\\ &+\int_\omega l\,{}^0H^{3333}\frac{t^2}{12}\left(p(\vec{\varrho}^\varepsilon)\right)^2 dS + 2\int_\omega l\,{}^0H^{\alpha\beta33}\frac{t^2}{12}\hat{\chi}_{\alpha\beta}(\vec{u}^\varepsilon, \vec{\theta}^\varepsilon)p(\vec{\varrho}^\varepsilon)\,dS.\end{aligned}$$

hence,

$$D'_\varepsilon = \frac{1}{\varepsilon}F^{3Ds}(\vec{U}^\varepsilon) - \tilde{D}_\varepsilon + R_\varepsilon, \tag{70}$$

using the variational equation (18). On the other hand, using (52) together with (62), (64), (66) and (68) we have

$$D_\varepsilon = D'_\varepsilon - A^{3Ds}_{\rm m}(\vec{u}^{\rm m}, \vec{\theta}^{\rm m}; \vec{u}^{\rm m}, \vec{\theta}^{\rm m}) + R_\varepsilon = D'_\varepsilon - G^{3Ds}(\vec{u}^{\rm m}) + R_\varepsilon.$$

This gives when substituting (70)

$$D_\varepsilon = \frac{1}{\varepsilon}F^{3Ds}(\vec{U}^\varepsilon) - \tilde{D}_\varepsilon - G^{3Ds}(\vec{u}^{\rm m}) + R_\varepsilon = -\tilde{D}_\varepsilon + R_\varepsilon,$$

invoking the weak convergence. We have thus found

$$D_\varepsilon + \tilde{D}_\varepsilon = R_\varepsilon, \tag{71}$$

with $D_\varepsilon \geq 0$ by construction, and it is straightforward to see that

$$\tilde{D}_\varepsilon \geq C\Big(\varepsilon^2\|\underline{\underline{\hat{\chi}}}\|_0^2 + \varepsilon^2\|\underline{m}\|_0^2 + \varepsilon^2\|p\|_0^2 + \varepsilon^4\|\underline{\underline{k}}\|_0^2 + \varepsilon^4\|\underline{n}\|_0^2\Big). \tag{72}$$

Therefore, we infer that D_ε and $\tilde{D}_\varepsilon$ both tend to zero. Then, (72) implies that all the terms with coefficient t^2 appearing in the definitions of $D_{1,\varepsilon}$, $D_{2,\varepsilon}$ and $D_{4,\varepsilon}$ tend to zero, hence D_ε tending to zero leads to $\|\vec{u}^\varepsilon - \vec{u}^{\rm m}, \vec{\theta}^\varepsilon - \vec{\theta}^{\rm m}\|_{\rm m} \overset{\varepsilon\to0}{\longrightarrow} 0$. Finally, (72) also shows the additional convergence properties

$$\varepsilon\|\vec{u}^\varepsilon, \vec{\theta}^\varepsilon, \vec{\varrho}^\varepsilon\|_{\rm b} \overset{\varepsilon\to0}{\longrightarrow} 0, \quad \varepsilon^2\|\vec{u}^\varepsilon, \vec{\theta}^\varepsilon, \vec{\varrho}^\varepsilon\|_1 \overset{\varepsilon\to0}{\longrightarrow} 0, \tag{73}$$

recalling the equivalences of norms stated in Lemma 1. □

Remark 3 *The asymptotic behavior is entirely of singular perturbation type in this case [17], albeit with perturbation terms arising at several orders and with combinations thereof, which requires some rather subtle handling of these combinations in the convergence proofs.*

Remark 4 *In this case also, we have convergence of the displacements within the shell body – as expressed in* (51) *– to the limit solution* $\vec{u}^{\rm m}$*, uniformly with respect to the parameter given by* $\xi^3/(\varepsilon l)$*, albeit here in the norm provided by* $\|\underline{\underline{\gamma}}\|_0$*. This is a straightforward consequence of the additional convergence properties given in* (73)*, indeed.*

4 Concluding remarks

We have revisited the convergence properties of the 3D-shell model proposed in [9], and established strong convergence results for the two main types of asymptotic regimes, namely, bending- and membrane-dominated behavior. This is an important result, as it completely substantiates the asymptotic consistency of the 3D-shell model with 3D elasticity.

As mentioned in the introduction, the finite element discretization of the 3D-shell model is quite straightforward – even more so than for general shell elements due to the possibility of directly employing a general 3D constitutive law in the 3D-shell finite element formulation – and these elements are already available in some widely-used finite element software[1], indeed. Of course, special care must be exerted to avoid the very serious numerical pathologies associated with shell model discretizations – numerical locking phenomena, in particular – albeit the effective techniques already designed for general shell elements can be rather directly extended to 3D-shells [8, 15, 10].

References

[1] S. Ahmad, B.M. Irons, and O.C Zienkiewicz. Analysis of thick and thin shell structures by curved finite elements. *Internat. J. Numer. Methods Engrg*, 2:419–451, 1970.

[2] F. Auricchio, C. Lovadina, , and A.L. Madureira. An asymptotically optimal model for isotropic heterogeneous linearly elastic plates. *ESAIM: M2AN*, 38:877–897, 2004.

[3] K.J. Bathe. *Finite Element Procedures.* Prentice Hall, Englewood Cliffs, 1996.

[4] M. Bernadou. *Finite Element Methods for Thin Shell Problems.* John Wiley & Sons, New York, 1996.

[5] M. Bischoff and E. Ramm. On the physical significance of higher order kinematic and static variables in a three-dimensional shell formulation. *Internat. J. Solids Structures*, 37:6933–6960, 2000.

[6] D. Chapelle and K.J. Bathe. Fundamental considerations for the finite element analysis of shell structures. *Comput. & Structures*, 66:19–36, 1998.

[7] D. Chapelle and K.J. Bathe. The mathematical shell model underlying general shell elements. *Internat. J. Numer. Methods Engrg.*, 48(2):289–313, 2000.

[8] D. Chapelle and K.J. Bathe. *The Finite Element Analysis of Shells - Fundamentals.* Springer, second edition, 2011.

[9] D. Chapelle, A. Ferent, and K.J. Bathe. 3D-shell elements and their underlying mathematical model. *M3AS*, 14(1):105–142, 2004.

[10] D. Chapelle, A. Ferent, and P. Le Tallec. The treatment of "pinching locking" in 3D-shell elements. *ESAIM: M2AN*, 37:143–158, 2003.

[11] P.G. Ciarlet. *Mathematical Elasticity - Volume III: Theory of Shells.* North-Holland, 2000.

[12] M. Dauge and Z. Yosibash. Boundary layer realization in thin elastic 3D domains and 2D hierarchic plate models. *Internat. J. Solids Structures*, 37:2443–2471, 2000.

[1] See www.adina.com, in particular

[13] M.C. Delfour. Intrinsic $P(2,1)$ thin shell model and Naghdi's models without a priori assumption on the stress tensor. In K.H. Hoffmann, G. Leugering, and F. Tröltzsch, editors, *Optimal Control of Partial Differential Equations*, pages 99–113, Basel, 1999. Birkhäuser.

[14] A.E. Green and Z. Zerna. *Theoretical Elasticity*. Clarendon Press, Oxford, second edition, 1968.

[15] D.N. Kim and K.J. Bathe. A 4-node 3D-shell element to model shell surface tractions and incompressible behavior. *Comput. & Structures*, 86(21–22):2027–2041, 2008.

[16] W.T. Koiter. On the nonlinear theory of thin elastic shells. *Proc. Kon. Ned. Akad. Wetensch.*, B69:1–54, 1965.

[17] J.L Lions. *Perturbations Singulières dans les Problèmes aux Limites et en Contrôle Optimal*. pringer-Verlag, Berlin, New York, 1973.

[18] J.-C. Paumier and A. Raoult. Asymptotic consistency of the polynomial approximation in the linearized plate theory. *ESAIM: Proc.*, 2:203–213, 1997.

[19] J. Pitkäranta, A.M. Matache, and C. Schwab. Fourier mode analysis of layers in shallow shell deformations. *Comput. Methods Appl. Mech. Engrg.*, 190:2943–2975, 2001.

[20] A. Rössle, M. Bischoff, W. Wendland, and E. Ramm. On the mathematical foundation of the (1,1,2)-plate model. *Int. J. Solids Struct.*, 36(14):2143–2168, 1999.

[21] J. Sanchez-Hubert and E. Sanchez-Palencia. *Coques Elastiques Minces - Propriétés Asymptotiques*. Masson, Paris, 1997.

Part III

Applications of the Surface-based Bidomain Model

Introduction of Part III

Chapter 3 gives an innovative asymptotic electrophysiological model valid for thin cardiac structures as for example **atria**. We present in this part some medical applications of this asymptotic model.

In Chapter 5 – the first chapter of this part – we apply the **surface bidomain model** in a **real geometry** of the atria. We prescribe the positions at the fibers at the endocardium (inner surface) and at the epicardium (outer surface) of the atria using literature. Then the position of the fibers at the midsurface and also the value of the total angle variation – which are the two informations that we need for using the surface bidomain model – are deduced. The fast and the slow conductions bundles which are specific to the atria are determined. For the ionic current we use the Courtemanche-Ramirez-Nattel model [2] adapted to the atria. Simulations are presented in the following article [1]

Surface-based electrophysiology modeling and assessment of physiological simulations in atria

which is a proceeding of the FIMH conference in collaboration with J.-F. Gerbeau, M. Hocini, M. Haïssaguerre and D. Chapelle. Our surface simulation is compared and is in very good adequacy with 3D simulations founded in the literature [5, 4, 3], and we give these comparisons in the article.

Using a coupled bidomain model with the surface bidomain model for the atria and the classical 3D model for the ventricles, we obtain electrophysiological simulations of the **whole heart**. The simulations are used in order to produce full electrocardiograms which contain the PQRST waves. An electrocardiogram is a graphic representation of the electrical activity of the heart. The principe is to measure the differences between two different positions of the body surface. A classical electrocardiogram has twelve derivations. Using a weak coupling between the heart mesh and the body mesh – more precisely Robin boundary conditions – we consider a diffusion of the extra-cellular potential in the whole body. Then with the value of potential at 9 positions of the body surface, we obtain the 12 derivations of a **classical electrocardiogram**. This work takes the form of a pre-print article

Numerical simulations of full electrocardiogram cycles

in collaboration with J.-F. Gerbeau and E. Schenone given in Chapter 6.

As previously said, we take into account the fast and the slow conductions bundles of the atria in this work. As we will see, two different methods are used in this part. In the first chapter of this part, we modify the value of the conductivity parameters in these various bundles. This method is not very justified. That is why, we use a second method which consists in directly modifying a parameter of the Courtemanche-Ramirez-Nattel (CRN) model. In [6], we also modify some parameters of the 0D CRN model in order to adapt the model to 3D cases.

Bibliography

[1] A. Collin, J.-F. Gerbeau, M. Hocini, M. Haïssaguerre, and D. Chapelle. Surface-based electrophysiology modeling and assessment of physiological simulations in atria. *FIMH 2013*, 7945:352–359, 2013.

[2] M. Courtemanche, R.J. Ramirez, and S. Nattel. Ionic mechanisms underlying human atrial action potential properties: insights from a mathematical model. *American Journal of Physiology*, (275):H301–H321, 1998.

[3] D. Deng, Y. Gong, G. Shou, and P. Jiao. Simulation of biatrial conduction via different pathways during sinus rhythm with a detailed human atrial model. *Journal of Zheijiang University-SCIENCE B (Biomedicine and Biotechnology)*, pages 1862–1783, 2012.

[4] D.M. Harrild and S.H. Craig. A computer model of normal conduction in the human atria. *Circulation Research*, (87):e25–e36, 2000.

[5] M. Krueger, V. Schmidt, C. Tobón, F. Weber, C. Lorenz, D. Keller, H. Barschdorf, M. Burdumy, P. Neher, G. Plank, K. Rhode, G. Seemann, D. Sanchez-Quintana, J. Saiz, R. Razavi, and O. Dössel. Modeling atrial fiber orientation in patient-specific geometries: a semi-automatic rule-based approach. *Functional Imaging and Modeling of the Heart*, pages 223–232, 2011.

[6] G. Seemann, C. Höper, F. B. Sachse, O. Dössel, A. V. Holden, and H. Zhang. Heterogeneous three-dimensional anatomical and electrophysiological model of human atria. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 364(1843):1465–1481, 2006.

CHAPTER 5

Surface-based electrophysiology modeling and assessment of physiological simulations in atria

Abstract The objective of this paper is to assess a previously-proposed *surface-based* electrophysiology model with detailed atrial simulations. This model – derived and substantiated by mathematical arguments – is specifically designed to address thin structures such as atria, and to take into account strong anisotropy effects related to fiber directions with possibly rapid variations across the wall thickness. The simulation results are in excellent adequacy with previous studies, and confirm the importance of anisotropy effects and variations thereof. Furthermore, this surface-based model provides dramatic computational benefits over 3D models with preserved accuracy.

Surface-based electrophysiology modeling and assessment of physiological simulations in atria

Annabelle Collin,
Inria Saclay Ile-de-France, Palaiseau, France,
Jean-Frédéric Gerbeau,
Inria Paris-Rocquencourt, France,
Mélèze Hocini, Michel Haïssaguerre,
Hôpital Cardiologique du Haut-Lévêque and Institut LIRYC,
Bordeaux-Pessac, France
Dominique Chapelle,
Inria Saclay Ile-de-France, Palaiseau, France,

Proc. of FIMH 2013, DOI:10.1007/978-3-642-38899-6_42

Abstract

The objective of this paper is to assess a previously-proposed *surface-based* electrophysiology model with detailed atrial simulations. This model – derived and substantiated by mathematical arguments – is specifically designed to address thin structures such as atria, and to take into account strong anisotropy effects related to fiber directions with possibly rapid variations across the wall thickness. The simulation results are in excellent adequacy with previous studies, and confirm the importance of anisotropy effects and variations thereof. Furthermore, this surface-based model provides dramatic computational benefits over 3D models with preserved accuracy.

Keywords: Electrophysiology modeling; Asymptotic analysis; Cardiology

1 Introduction

There is a very important medical need for modeling the electrical activity of the heart in general, and in the atria in particular, e.g. with a view to therapy planning assistance in radiofrequency ablation for patients suffering from atrial fibrillation [10]. In addition to generic difficulties inherent to electrophysiology modeling, namely, modeling complexity and computational intensiveness, atria modeling features specific difficulties, in particular due to their very thin walls – mostly apparent as thick surfaces in medical imaging – which requires much refined meshes. Moreover, there is a major challenge in taking into account the anisotropy resulting from the preferred conduction direction along the muscle fibers, which is also known to vary extremely rapidly across the wall thickness [7].

The electrical wave propagating in the cardiac tissue can be represented by a nonlinear reaction-diffusion partial differential equation (PDE), coupled with ordinary differential equations (ODEs) representing cellular activity. In this study, we consider the 2D – namely, surface-based – model proposed and mathematically substantiated in [1], derived from the bidomain model (see for example [12, 11, 13]), and defined over the midsurface of the thin region. This surface-based model was specifically designed for thin cardiac structures – the atrial walls, in particular – and takes into account the strong anisotropy variations across the thickness. This

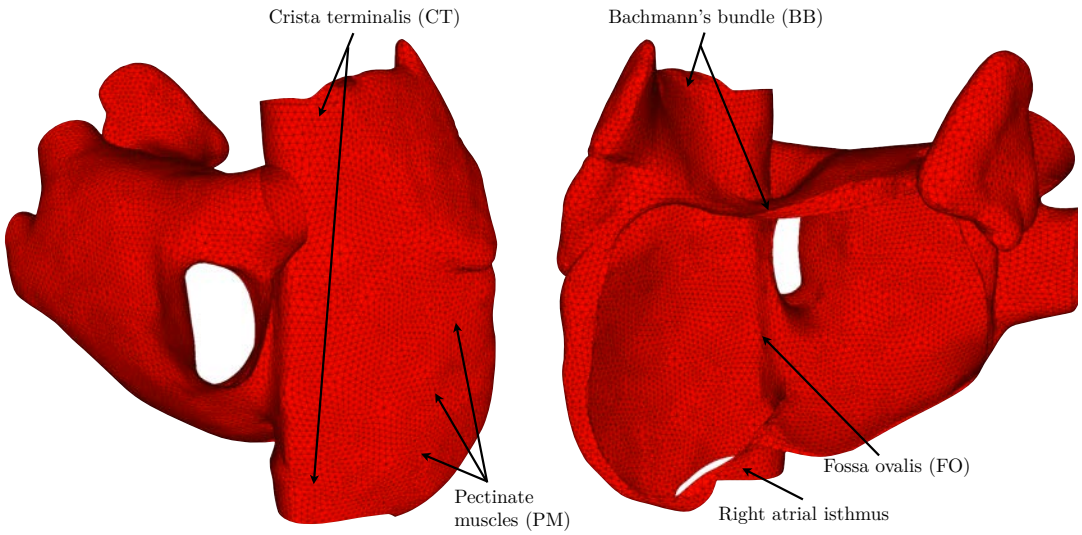

Figure 1: Atrial mesh posterior (left) and anterior (right) views, with main specific regions

LA indicators at ventricular end-systole	Normal (from [2, 8])	LA model dimensions
diameter (cm)	2.8–4.0	3.0
major axis (cm)	4.1–6.1	4.65
area (cm^2)	15.0–20.0	15.9
volume (ml)	41–75	47
diameters of pulmonary veins (cm)	1.1–1.5	1.14–1.45

Table 1: Assessment of left atrial dimension indicators

model was already successfully numerically assessed in [1] by comparing the resulting simulations with reference 3D simulations on thin domains of simple geometries, with dramatic benefits in computation times. Our objective here is to further assess this model with physiological simulations of the atrial electrophysiology.

2 Model

2.1 Atrial mesh

We produced a surface mesh representing the mid-surface of the two atria. Starting from the Zygote[1] heart model – a geometric model based on actual anatomical data – we used the `3-matic`[2] software to obtain a computationally-correct surface mesh, and the `Yams` [5] meshing software to further process and refine the surface mesh. Figure 1 shows the posterior and anterior views of the mesh, which contains about 26,000 triangles and 13,500 vertices. The anatomy corresponds to ventricular end-systole, namely, when the atrial chamber has its greatest internal volume. This anatomical surface was compared with literature reports of normal human atrial dimension indicators, see for example [2, 8], and found to be within standard ranges, see Table 1.

[1] `www.3dscience.com`
[2] `www.materialise.com`

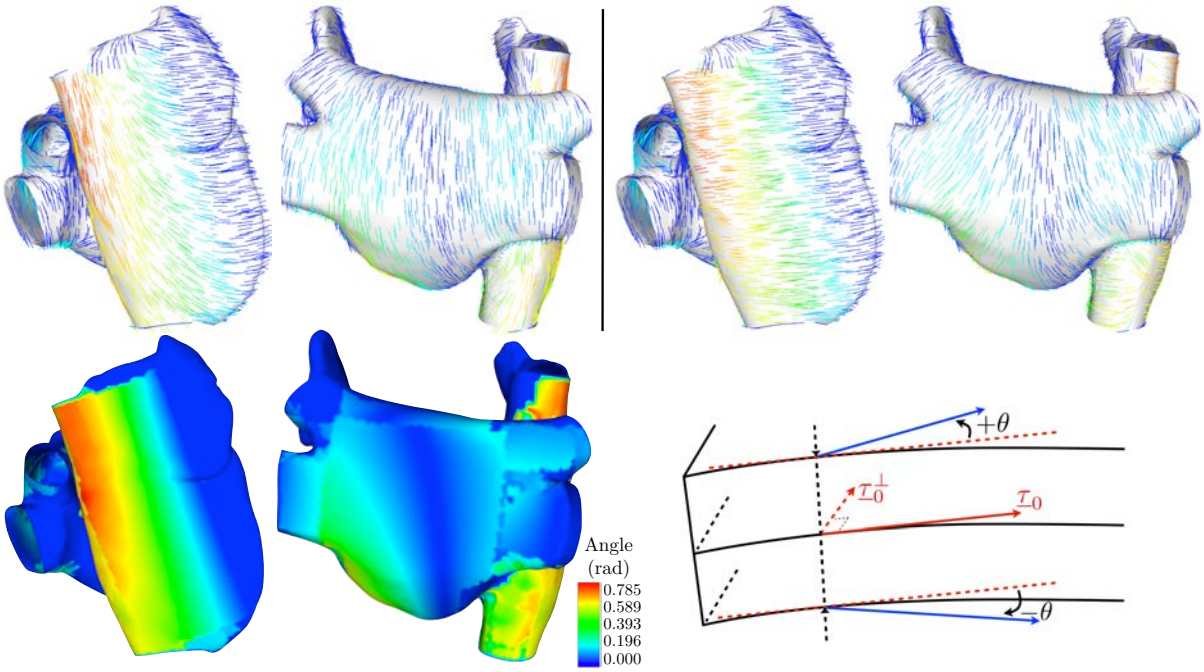

Figure 2: Fibers directions at the endocardium (top-left) and epicardium (top-right), angle θ (bottom-left), and angular variation visualization (bottom-right)

2.2 Fibers

Cardiac tissue has a fiber architecture. As the electrical conductivity is higher along than across the fiber direction, fiber orientation is very important in electric activation propagation. The specificities of the atria are that the walls are very thin, and that fibers orientations may vary extremely rapidly across the thickness. We use [7, 9] to identify and prescribe the fibers directions at the endocardium and epicardium, see Figure 2. This figure also displays the angle θ, defined as half of the angular variation between the endocardium and the epicardium.

2.3 Surface-based bidomain model

We denote by $\mathcal{S}$ the midsurface of the wall. The variational bidomain surface model that we propose can be written in terms of the extracellular potential u_e, the transmembrane potential $V_m = u_i - u_e$, with u_i the intracellular potential, as follows [1]. Find (V_m, u_e) with $\int_{\mathcal{S}} u_e \, dS = 0$, such that for all $t > 0$,

$$\begin{cases} A_m \displaystyle\int_{\mathcal{S}} \Big(C_m \frac{\partial V_m}{\partial t} + I_{ion}(V_m) \Big) \phi \, dS + \int_{\mathcal{S}} \Big(\underline{\underline{\sigma}}_i \cdot \big(\underline{\nabla} V_m + \underline{\nabla} u_e \big) \Big) \cdot \underline{\nabla} \phi \, dS \\ \qquad\qquad = A_m \displaystyle\int_{\mathcal{S}} I_{app} \phi \, dS, \quad \forall \phi \\ \displaystyle\int_{\mathcal{S}} \Big((\underline{\underline{\sigma}}_i + \underline{\underline{\sigma}}_e) \cdot \underline{\nabla} u_e \Big) \cdot \underline{\nabla} \psi \, dS + \int_{\mathcal{S}} \Big(\underline{\underline{\sigma}}_i \cdot \underline{\nabla} V_m \Big) \cdot \underline{\nabla} \psi \, dS = 0, \quad \forall \psi \end{cases} \tag{1}$$

also with $\int_{\mathcal{S}} \psi \, dS = 0$, and where A_m is a positive constant denoting the ratio of membrane area per unit volume, C_m the membrane capacitance per unit surface, $I_{ion}(V_m)$ a reaction term representing the ionic current across the membrane and also depending on local ionic variables satisfying additional ODEs, and I_{app} a given applied stimulus current.

	tissue	PM	CT	BB	IRA	FO	SN
σ_i^t	$2.2\,10^{-4}$	$3.3\,10^{-4}$	$6.6\,10^{-4}$	$1.045\,10^{-3}$	$8.250\,10^{-5}$	$2.2\,10^{-4}$	$2.2\,10^{-4}$
σ_i^l	$2.2\,10^{-3}$	$3.894\,10^{-3}$	$9.570\,10^{-3}$	$1.639\,10^{-2}$	$9.735\,10^{-4}$	$2.002\,10^{-3}$	$2.002\,10^{-3}$
σ_e^t	$9.0\,10^{-4}$	$1.350\,10^{-3}$	$2.7\,10^{-3}$	$4.275\,10^{-3}$	$3.375\,10^{-4}$	$9.0\,10^{-4}$	$9.0\,10^{-4}$
σ_e^l	$2.2\,10^{-3}$	$3.690\,10^{-3}$	$8.550\,10^{-3}$	$1.435\,10^{-2}$	$9.225\,10^{-4}$	$2.070\,10^{-3}$	$2.070\,10^{-3}$

Table 2: Conductivity parameters

We define the intra- and extra-cellular diffusion tensors $\underline{\underline{\sigma}}_i$ and $\underline{\underline{\sigma}}_e$ by

$$\underline{\underline{\sigma}}_{i,e} = \sigma_{i,e}^t \underline{\underline{I}} + (\sigma_{i,e}^l - \sigma_{i,e}^t)\big[I_0(\theta)\underline{\tau}_0 \otimes \underline{\tau}_0 + J_0(\theta)\underline{\tau}_0^\perp \otimes \underline{\tau}_0^\perp\big], \tag{2}$$

where $\underline{\underline{I}}$ denotes the identity tensor in the tangential plane – also sometimes called the surface metric tensor – $\underline{\tau}_0$ is a unit vector parallel to the local fiber direction on the atria midsurface, and $\underline{\tau}_0^\perp$ such that $(\underline{\tau}_0, \underline{\tau}_0^\perp)$ gives an orthonormal basis of the tangential plane. The effect of angular variations enters the model through the coefficients $I_0(\theta) = \frac{1}{2} + \frac{1}{4\theta}\sin(2\theta)$ and $J_0(\theta) = 1 - I_0(\theta)$. Note that $J_0(\theta) = 0$ (and $I_0(\theta) = 1$) if and only if $\theta = 0$ – namely, constant direction across the thickness – and then $\underline{\underline{\sigma}}_{i,e} = \sigma_{i,e}^t \underline{\underline{I}} + (\sigma_{i,e}^l - \sigma_{i,e}^t)\underline{\tau}_0 \otimes \underline{\tau}_0$ as expected for a single fiber direction. By contrast, angular variations make I_0 decrease and J_0 increase in (2), which renders diffusion "more isotropic". This model derived from a detailed asymptotic analysis thus allows to take into account the rapid variations of the fiber direction.

The current I_{ion} can be described by a physiological or a phenomenological model. In this study, the physiological model proposed by Courtemanche, Ramirez and Nattel in [3] – most widely accepted for atria modeling – is considered.

The values of the membrane parameters are $A_m = 200.0\,\mathrm{cm}^{-1}$ and $C_m = 10^{-3}$ mF.cm^{-2}. The values of the conductivity parameters vary substantially depending on the specific areas considered and are given in Table 2 (all in S.cm^{-1}). The Bachmann bundle (BB), the *Crista Terminalis* (CT) and the pectinate muscles (PM) are regions of established fast conduction. By contrast, the *Fossa Ovalis* (FO) and the *Isthmus* of the right atrial floor (IRA) are regions of known slow conduction. All these specific regions are depicted in Figure 1.

The two atria are connected via two regions only – in the mesh connectivity, hence also in terms of electrical conduction – namely, the Bachmann bundle and the *Fossa Ovalis*.

3 Simulations results

Figure 3 displays the simulation results obtained with the above-described model. Activation is initiated at the sinus node with a stimulus of 2 ms in duration and sufficient strength to cause the initiation of a propagating wavefront. We compare our results to several 3D modeling studies [9, 6, 4].

As observed in the figure, by 30 ms the wave quickly spreads along the *Crista Terminalis* as a consequence of the high conductivity in this part. Importantly, the depolarizing wave has now traversed the Bachmann bundle and the first activation of the left atrium has occurred at 29.6 ms. This compares very well with the findings of the modeling study [6] giving the left atrium activation at 29.7 ms. At $t = 40$ ms, we clearly see the effect of the anisotropy of the *Crista Terminalis*, and of the pectinate muscles already observed in [9], as the wavefront becomes nearly triangular – as also seen in [6]. The wavefront has encircled the superior *Vena Cava*. The *Fossa Ovalis*, which is the second atrial connection, becomes active (42.8 ms). By 50 ms, because of the rapid conduction in the Bachmann bundle, the wave spreads to the left

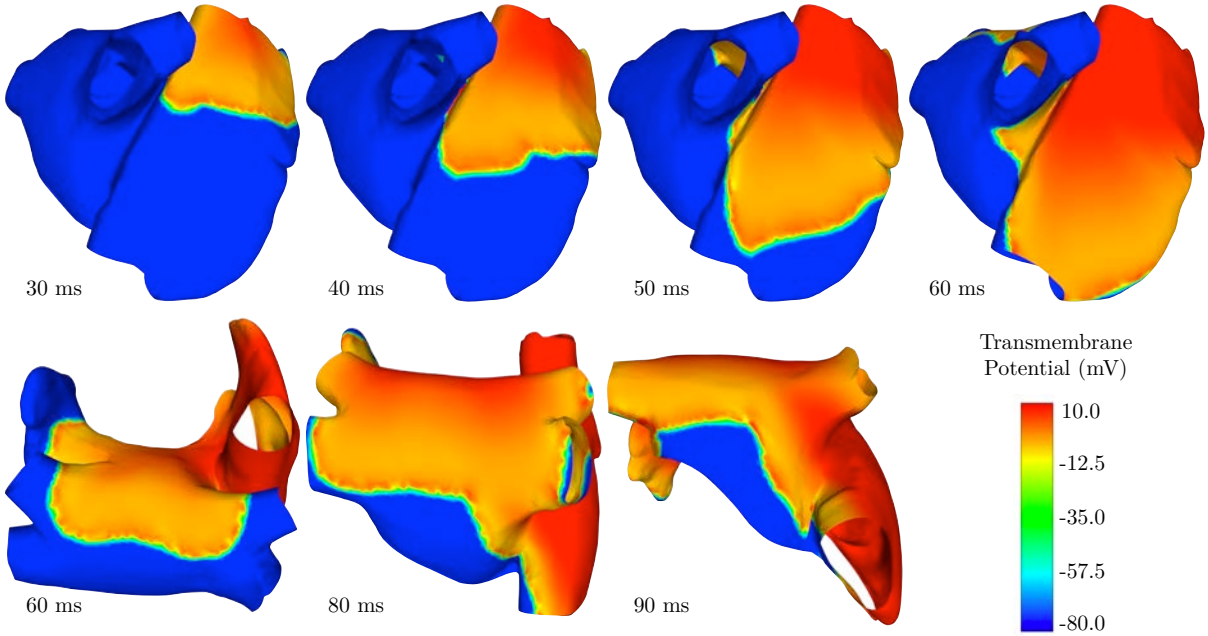

Figure 3: Simulation of atrial depolarization

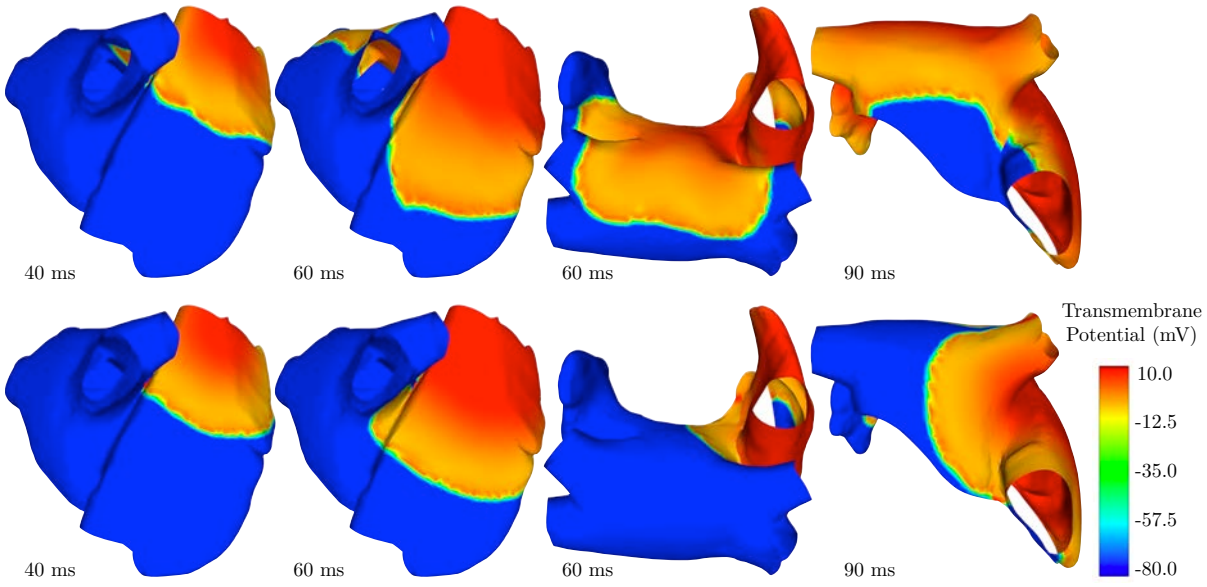

Figure 4: Simulations of atrial depolarization for homogeneous case i.e $\theta = 0$ (top) and for homogeneous and isotropic case (bottom)

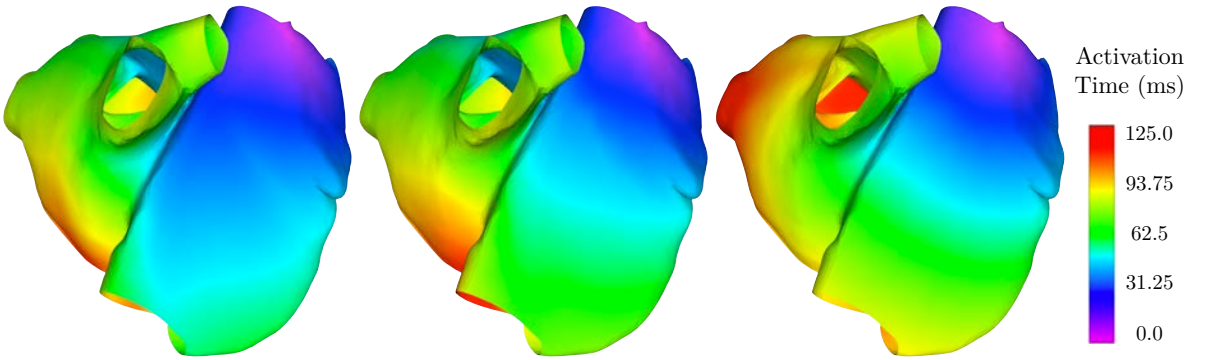

Figure 5: Activation time maps for the three cases: anisotropic-heterogeneous (left), anisotropic-homogeneous (middle) and isotropic-homogeneous (right)

atrial appendage and has activated a substantial part of the left atrial wall. The *Fossa Ovalis* is very active. The activation of the right atrial appendage is complete. At $t = 60$ ms, in the right atrium the floor is the last part that remains unexcited. The wave has encircled the mouth of the left atrial appendage at 72.8 ms. By $t = 80$ ms, the only part that remains unaffected by the wave in the right atrium is the isthmus, because of the reduced conductivity there. In the left atrium, at $t = 90$ ms we can see three separate wavefronts, also obtained in [6, 4]. The depolarization of the left atrial appendage is complete at 95.8 ms. At 100 ms, only a small part of the left atrium in the shape of a parallelepiped is still inactive. The depolarization of the right and left atrium are complete at 101.8 ms and 109.6 ms, respectively. This is in good adequacy with the timings found in [6], namely, 99.3 ms and 108.2 ms, respectively, while [9] gives 115.0 ms for the left atrium, and [4] 119.0 ms.

4 Discussion

For comparison purposes, we also display in Figure 4-top simulation results obtained when disregarding the fiber direction variations across the thickness, namely, when taking $\theta = 0$. We observe major differences with the above-discussed physiological simulation. In particular, the anisotropy effects characteristic of the *Crista Terminalis* are much less clearly seen, and the triangular shape is hardly observed, indeed. Overall, the activation timings are longer, in particular due to slower propagation in the *Crista Terminalis*, with complete depolarization occurring as late as 116.2 ms and 115.4 ms for the right and left atria, respectively. The last results displayed in Figure 4-bottom correspond to an homogeneous and isotropic tissue throughout. We increased the homogeneous conductivity with $\sigma_i = 8.0\,10^{-4}$ S.cm^{-1} and $\sigma_e = 3.1\,10^{-3}$ S.cm^{-1} to obtain roughly comparable depolarization timings. The resulting activation profile, however, is totally unrealistic – as already noted in [9] – see Figure 5 displaying the activation time maps for the three cases analyzed.

5 Concluding remarks

We have presented detailed simulation results of a surface-based electrophysiology bidomain model applied with an anatomical model of the atria. The results obtained are in excellent adequacy with previous studies, and confirm the importance of fiber-related anisotropy effects and of their strong variations across the wall thickness. Furthermore, this surface-based model provides dramatic computational benefits over 3D models with preserved accuracy [1]. Further

perspectives include applications to radio-frequency ablation planning assistance, for which computational effectiveness – less than 3 minutes for a complete simulation on a standard workstation – will be of great value.

References

[1] D. Chapelle, A. Collin, and J.-F. Gerbeau. A surface-based electrophysiology model relying on asymptotic analysis and motivated by cardiac atria modeling. *M3AS*, 23(14):2749–2776, 2013.

[2] G.I. Cohen, M. White, R.A. Sochowski, A.Z. Klein, P.D. Bridge, W.J. Steward, and K.L. Chang. Reference values for normal adult transesophageal echocardiographic measurements. *Journal of the American Society of Echocardiography*, (8):221–230, 1995.

[3] M. Courtemanche, R.J. Ramirez, and S. Nattel. Ionic mechanisms underlying human atrial action potential properties: insights from a mathematical model. *American Journal of Physiology*, (275):H301–H321, 1998.

[4] D. Deng, Y. Gong, G. Shou, and P. Jiao. Simulation of biatrial conduction via different pathways during sinus rhythm with a detailed human atrial model. *Journal of Zheijiang University-SCIENCE B (Biomedicine and Biotechnology)*, pages 1862–1783, 2012.

[5] P. Frey. Yams: A fully automatic adaptive isotropic surface remeshing procedure. Technical report 0252, Inria, Rocquencourt, France, Nov. 2001.

[6] D.M. Harrild and S.H. Craig. A computer model of normal conduction in the human atria. *Circulation Research*, (87):e25–e36, 2000.

[7] S.Y. Ho, R.H. Anderson, and D. Sánchez-Quintana. Atrial structure and fibres: morphologic bases of atrial conduction. *Cardiovascular Research*, (54):325–336, 2002.

[8] P. Jiamsripong, T. Honda, C.S. Reuss, R.T. Hurst, H.P. Chaliki, D.E. Grill, S.L. Schneck, R. Tyler, B.K. Khandheria, and S.J. Lester. Three methods for evaluation of left atrial volume. *European Journal of Echocardiography*, (9):351–355, 2008.

[9] M. Krueger, V. Schmidt, C. Tobón, F. Weber, C. Lorenz, D. Keller, H. Barschdorf, M. Burdumy, P. Neher, G. Plank, K. Rhode, G. Seemann, D. Sanchez-Quintana, J. Saiz, R. Razavi, and O. Dössel. Modeling atrial fiber orientation in patient-specific geometries: a semi-automatic rule-based approach. *Functional Imaging and Modeling of the Heart*, pages 223–232, 2011.

[10] S. Matsuo, N. Lellouche, M. Wright, M. Bevilacqua, S. Knecht, I. Nault, K.T. Lim, L. Arantes, M.D. O'Neill, P.G. Platonov, =. Carlson, J, F. Sacher, M. Hocini, P. Jaïs, and M. Haïssaguerre. Clinical predictors of termination and clinical outcome of catheter ablation for persistent atrial fibrillation. *Journal of the American College of Cardiology*, 54(9):788–795, 2009.

[11] A.J. Pullan, M.L. Buist, and L.K. Cheng. *Mathematically Modeling the Electrical Activity of the Heart*. World Scientific, 2005.

[12] F.B. Sachse. *Computational Cardiology: Modeling of Anatomy, Electrophysiology and Mechanics*. Springer-Verlag, 2004.

[13] J. Sundnes, G.T. Lines, X. Cai, B.F. Nielsen, K.A. Mardal, and A. Tveito. *Computing the Electrical Activity in the Heart*, volume 1 of *Monographs in Computational Science and Engineering*. Springer-Verlag, 2006.

CHAPTER 6
Numerical simulations of full electrocardiogram cycles

Abstract This work is dedicated to the simulation of the whole electrical activity and its corresponding body surface potential, including the standard electrocardiograms (ECGs). ECG is the most used exam providing an interpretation of the cardiac electrical activity, since it represents a non-invasive and inexpensive procedure. In order to reproduce a realistic ECG including P, QRS and T waves, we introduce a cardiac electrophysiology model which takes into account ventricles and atria. First, a realistic heart mesh is presented. The ventricular part is considered as a standard volume mesh, while the atria are modeled by a surface. The mathematical model is based on the bidomain equations, which represent one of the most used model for cardiac electrophysiology. As we deal with an hybrid mesh, the usual bidomain equations are used in the ventricle volume part of the domain, while an anisotropic atrial surface-based model is applied. The bidomain model is coupled with two ionic models: on the atrial surface we use the physiological model by Courtemanche, Ramirez and Nattel, while in the ventricular part we apply the *Minimal model for human Ventricular action potentials* (MV) by Bueno-Orovio, Cherry and Fenton. In order to obtain the ECGs, the heart is weakly coupled to the body by a Robin boundary condition based on a resistor-capacitor transmission condition. Then, a healthy ECG and some pathological cases - in particular a Wolff-Parkinson-White syndrome - are analyzed and we use our simulation tools to study more sophisticated signals as those provided by a vest of electrodes.

Numerical simulations of full electrocardiogram cycles

Elisa Schenone,
Inria Paris-Rocquencourt, France,
Annabelle Collin,
Inria Saclay Ile-de-France, Palaiseau, France,
Jean-Frédéric Gerbeau,
Inria Paris-Rocquencourt, France,

To be submitted

Abstract

This work is dedicated to the simulation of the whole electrical activity and its corresponding body surface potential, including the standard electrocardiograms (ECGs). ECG is the most used exam providing an interpretation of the cardiac electrical activity, since it represents a non-invasive and inexpensive procedure. In order to reproduce a realistic ECG including P, QRS and T waves, we introduce a cardiac electrophysiology model which takes into account ventricles and atria. First, a realistic heart mesh is presented. The ventricular part is considered as a standard volume mesh, while the atria are modeled by a surface. The mathematical model is based on the bidomain equations, which represent one of the most used model for cardiac electrophysiology. As we deal with an hybrid mesh, the usual bidomain equations are used in the ventricle volume part of the domain, while an anisotropic atrial surface-based model is applied. The bidomain model is coupled with two ionic models: on the atrial surface we use the physiological model by Courtemanche, Ramirez and Nattel, while in the ventricular part we apply the *Minimal model for human Ventricular action potentials* (MV) by Bueno-Orovio, Cherry and Fenton. In order to obtain the ECGs, the heart is weakly coupled to the body by a Robin boundary condition based on a resistor-capacitor transmission condition. Then, a healthy ECG and some pathological cases - in particular a Wolff-Parkinson-White syndrome - are analyzed and we use our simulation tools to study more sophisticated signals as those provided by a vest of electrodes.

Keywords: Electrocardiograms, bidomain model, atria, ventricles

Introduction

Cardiac electrophysiology is the study of the electrical activity of the heart [47, 50]. Over the past decades, huge progress has been made in the mathematical modeling of this field. Various models which represent the evolution of the potential in the membrane have been developed as for example [42, 34, 14] for the physiological models and [17, 40, 39, 6] for the phenomenological models. These models which correspond to one (or more) ordinary differential equation(s) can be coupled with diffusion terms in order to obtain 3D simulations as for example in [19] where a generalized FitzHugh–Nagumo model [1] is used and leads to realistic ventricular simulations. More complex reaction diffusion models – as the bidomain model initially introduced by [53] and its first approximation the monodomain model – are also available in the literature. In [12], a comparison between the bidomain model and the monodomain model is given in 3D test cases. For the atrial simulations we can cite [21, 32] which use a monodomain model. In this study, we want to introduce a detailed modeling of the human heart in order to obtain realistic simulations

of the whole heart (atria and ventricles). We consider a coupled bidomain model: a standard 3D bidomain model for the ventricles [53, 47] and an asymptotic surface-based bidomain model, presented in [8] and validated with simulations in [13], for the atria. Two different ionic models are used: the physiological model proposed by Courtemanche, Ramirez and Nattel in [14] on the atrial surface, and the *Minimal model for human Ventricular action potentials* (MV) model proposed by Bueno-Orovio, Cherry and Fenton in [6], in the ventricles. The coupled bidomain model allows us to obtain electrophysiological simulations of the whole heart. We want to assess these simulations by generating the full electrocardiograms which contain the PQRST waves.

The electrocardiogram (ECG) is a non-invasive recording of the cardiac electrical activity [35, 54]. The standard 12-lead ECG is constructed using 9 electrode measures of the body skin potential. This non-invasive and inexpensive procedure is the most used medical tool for the detection of pathologies of the electrical state of a patient. The modeling of realistic ECGs is an important step in the development of patient-specific models using medical ECG data. The mathematical modeling is known as the forward problem of electrocardiography (see for instance [33]) based on three main elements: a model for the heart, a model for the body diffusion and heart-body transmission conditions. Several simulations of ECGs are given in the literature. For example in [26, 44, 33, 31, 52], numerical simulations of ECGs using a bidomain or a monodomain model are given, and in particular the 12-lead ECG appear in [44, 52]. However, these simulations are not always realistic and only concern the ventricles. In [4], more realistic ventricular results based on detailed modeling assumptions, in particular for the torso transmission conditions, have been presented. More recently, we can cite [48] where full electrocardiogram cycles are obtained but for a very simple geometry and [28] where a study of the repolarization wave T is given. In our work, the coupled bidomain model presented previously and a signal diffusion in the whole body are considered. The coupling conditions between the heart and the body are based on the resistor-capacitor coupling condition presented in [4] and lead to a Robin boundary condition. The first contribution on this work concerns the detailed modeling assumptions with in particular a surface-volume bidomain model coupled with two adapted ionic models and also complex torso transmission conditions. The quality of our results is also a progress, in particular in the precordial leads V1-V6. Furthermore, using the classical criteria found in the literature, a detailed evaluation of our results is given and allows us to validate them. Pathological cases are also investigated in order to show the capability of our model to predict the features used by medical doctors to detect a disease.

Here is the outline of this article. In a first part, a whole surface (atria) and volume (ventricles) mesh adapted for finite elements is described. We compare this mesh with literature reports of normal human heart dimension indicators found for example in [10, 16, 11]. The second part concerns the modeling assumptions with the presentation of the coupled atria-ventricles model and we present the electrophysiological simulations of the whole heart. We conclude this part with the coupling conditions between the heart and the rest of the body. The third part concerns the simulations of the standard 12-lead ECG. A healthy case is given and validated by numerous criteria for a normal electrocardiogram. Some pathological cases are also studied. In particular we show simulations of the Wolff-Parkison-White syndrome, a pathology caused by the presence of an abnormal accessory electrical conduction pathway between the atria and the ventricles. In the last part, we give a principal components analysis of the potential measured by an electrode vest. In particular, the correlation between the signals of different electrodes is studied. We also give an analysis of the dependence of the electrode measures with respect to their positions on the body.

Left Ventricle	Volume (ml)	Mass (g)	Mitral (cm)	Aortic (cm)
Measures	53.7	111.3	2.9	2.3
Reference	46 ± 11	112 ± 27	2.5 ± 0.4	2.3 ± 0.2
Left Atrium	Major axis (cm)	Area (cm^2)	Volume (ml)	Pulmonary Veins (cm)
Measures	4.65	15.9	47	$1.14 - 1.45$
Reference	3.4 ± 0.6	17.5 ± 2.5	58 ± 34	1.3 ± 0.2

Table 1: Comparisons of the model dimensions with typical end-systolic values found in the literature [10, 25, 43, 16, 56, 11, 29]

1 Whole heart mesh

To obtain full cycle ECGs, the first step is to build a whole heart realistic mesh. The ventricles can be easily obtained from medical imaging and meshed in 3D. On the contrary, the atria have a very thin wall which makes them difficult to image in 3D. In addition, generating a 3D mesh on these very thin volumes would dramatically, and uselessly, increase the computational cost. For these reasons, we choose to model the geometry of the atria as a surface. We therefore obtain an hybrid mesh, made of tetrahedra in the ventricles and of triangles in the atria.

The heart model used in this work is obtained from an anatomical data set called *Zygote*[1]. The `3-matic` software is used to obtain a surface mesh satisfying the standard quality criteria of a finite element mesh, and the `Yams` software to refine the surface mesh [18]. Then, the volume of the two ventricles is meshed using `Gmsh` software. We can see in Figure 1 different views of the whole heart mesh, which contains about 230,600 tetrahedra, 73,500 triangles and 67,300 vertices. A simplified mesh of the body (Fig. 1), including the lungs and the ribs, is also built from the *Zygote* data set and the aforementioned software. The body mesh contains 408,171 tetrahedra, 89,222 triangles and 85,196 vertices.

The mechanical deformation of the heart is not taken into account in this work. The dimensions of the fixed domain correspond to the end of the systole (small ventricles, large atria). Table 1 shows a comparison of a few dimensions of the geometrical model with standard end-systolic values. The following quantities are compared: left ventricle volume and mass, mitral and aortic valves diameters, left atrium major axis, area, volume and four pulmonary veins diameters. We observe a good agreement with the values found in the literature [10, 25, 43]. We also have a good agreement for the diameters of mitral [16] and aortic [56] valves, the surface of the left atrium [11, 29].

Cardiac tissue has a fiber architecture. The electrical conductivity is higher along the fibers than in the transverse direction. This implies that the fiber orientation is very important in the study of the electrical activity of the heart. To identify and to prescribe the fibers at the endocardium and at the epicardium of the atria, we use [22, 23, 32]. As we can see in Figure 2 (top), the fibers orientation may vary extremely quickly across the thickness. The colors represent the angle θ defined as half of the angular difference between the endocardium and the epicardium. We use [41, 49] to prescribe the fibers in the ventricles, see Figure 2 (bottom).

Figure 3 represents a schematic view of the heart conduction system in a healthy heart: the sinus and atrioventricular nodes, the Bachmann's bundle and the Purkinje fibers. In this work, the atrio-ventricular node and the Purkinje fibers are not explicitly modeled (see below).

[1] `www.3dscience.com`

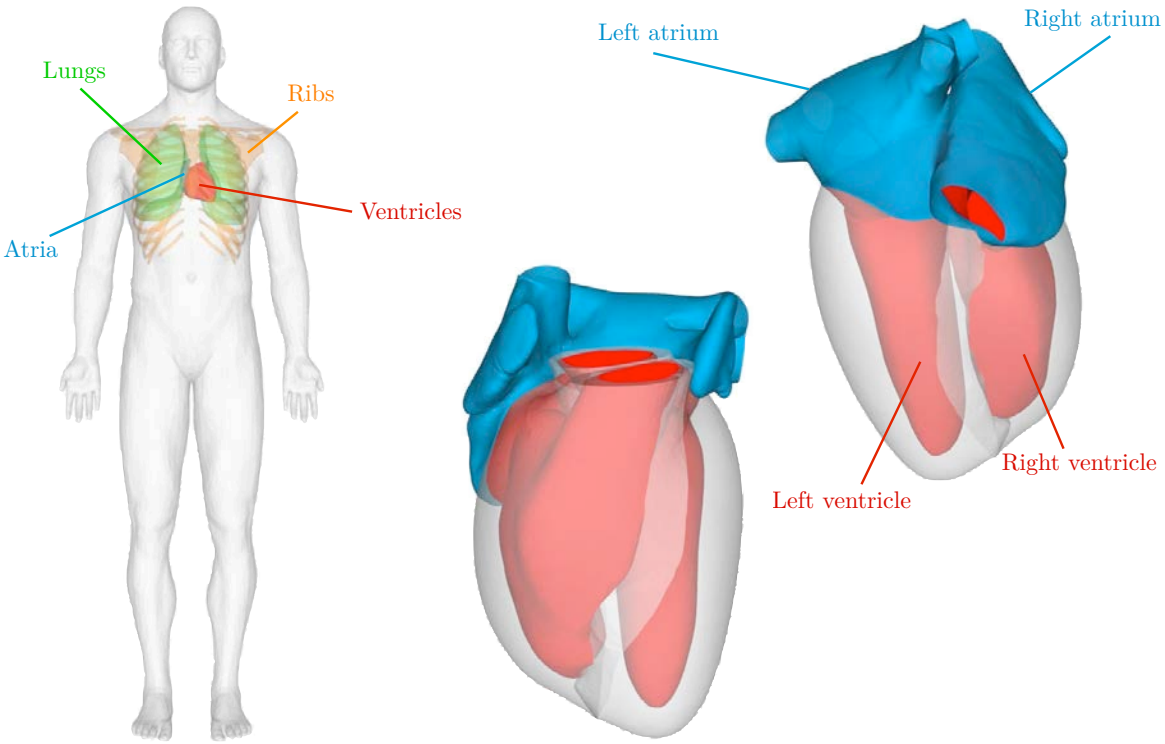

Figure 1: Whole heart mesh (right) and body mesh (left).

2 Modeling assumptions

In this section, we present the electrophysiology equations and the ionic models used in the ventricles and the atria. We also present the coupling conditions between the atria and the ventricles and between the heart and the body.

2.1 Bidomain model

In order to describe the electrical potential in the heart, we use the standard nonlinear reaction-diffusion equations known as the bidomain equations (see for instance [47, 50]). In terms of extracellular potential u_e and transmembrane potential $V_m = u_i - u_e$, with u_i the intracellular potential, the bidomain model reads

$$\left\{ \begin{array}{rcl} A_m\Big(C_m\dfrac{\partial V_m}{\partial t} + I_{ion}(V_m, w_1, \ldots, w_n)\Big) - \operatorname{div}\big(\vec{\vec{\sigma}}_i \cdot \vec{\nabla} V_m\big) & = & \operatorname{div}\big(\vec{\vec{\sigma}}_i \cdot \vec{\nabla} u_e\big) + A_m I_{app}, \\ \operatorname{div}\Big(\big(\vec{\vec{\sigma}}_i + \vec{\vec{\sigma}}_e\big) \cdot \vec{\nabla} u_e\Big) & = & -\operatorname{div}\big(\vec{\vec{\sigma}}_i \cdot \vec{\nabla} V_m\big), \end{array} \right. \tag{1}$$

in $\mathcal{B} \times (0, T)$, where $\mathcal{B}$ denotes the 3D domain of interest, A_m is a positive constant denoting the ratio of membrane area per unit volume, C_m the membrane capacitance per unit surface, I_{ion} the ionic current which depends on n ionic variables $w_1, \ldots, w_n$ and I_{app} a given applied stimulus current.

We assume that the heart is isolated from the body, so that the first boundary condition is

$$\big(\vec{\vec{\sigma}}_e \cdot \vec{\nabla} u_e\big) \cdot \vec{n} \;=\; 0, \quad \text{in } \partial\mathcal{B} \times (0, T), \tag{2}$$

while the second one comes from the fact that the intra-cellular current cannot propagate outside

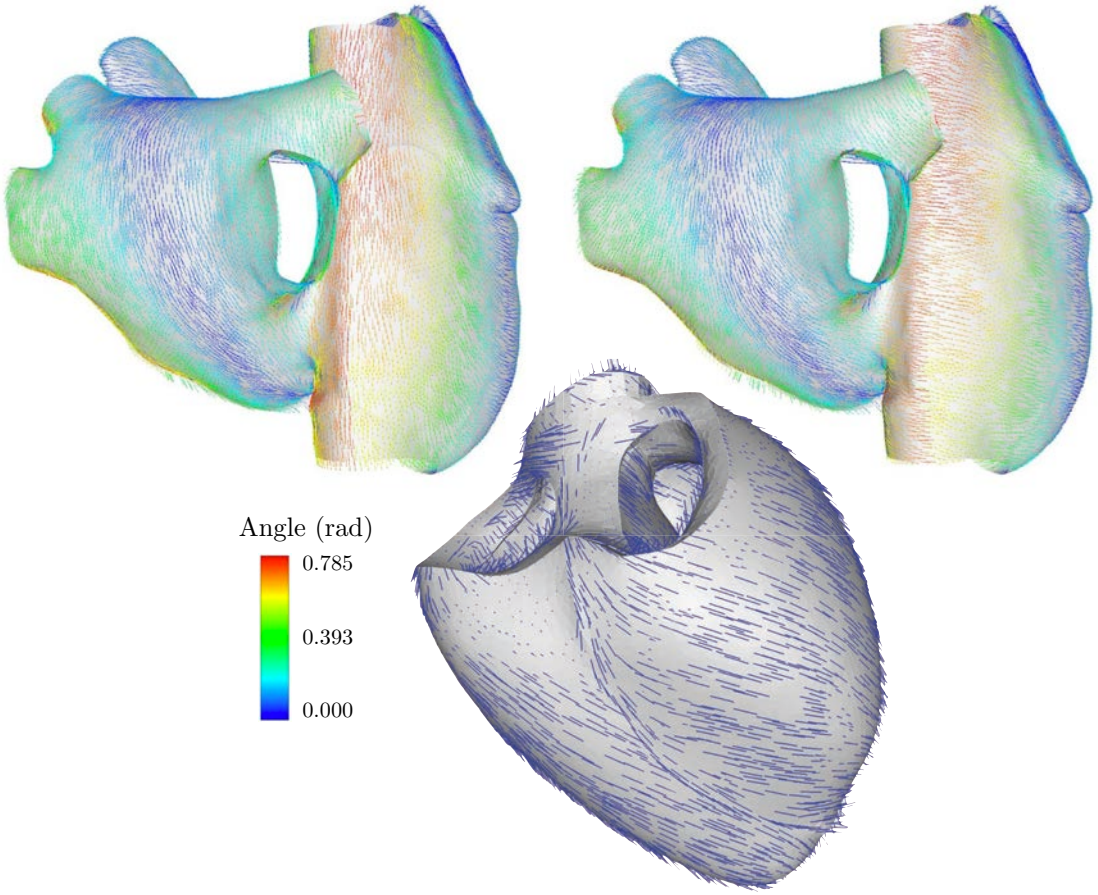

Figure 2: Fibers directions at the atrial endocardium (top-left) and atrial epicardium (top-right), and in ventricles (bottom).

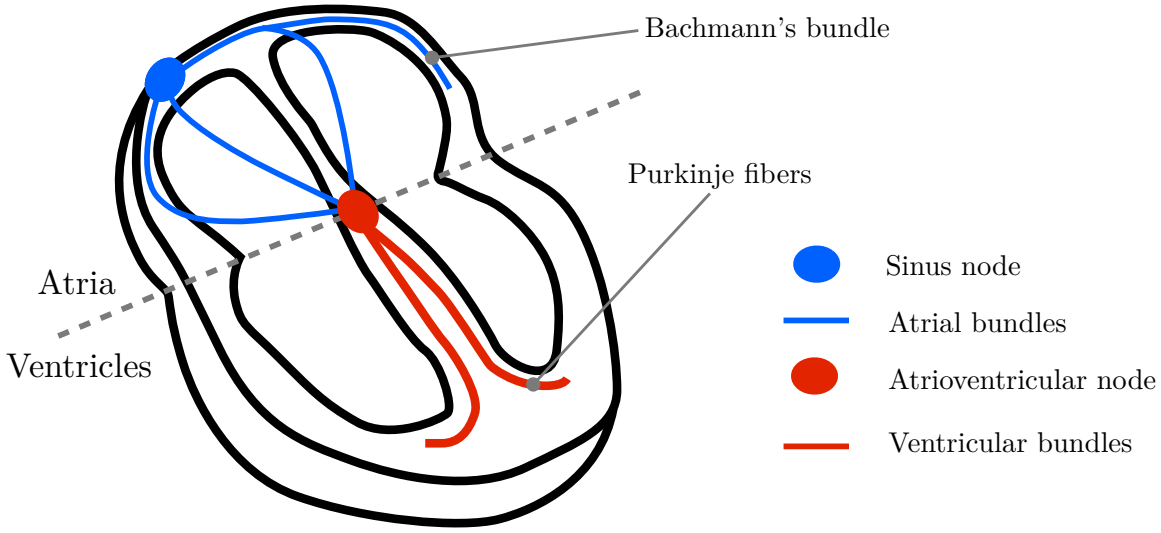

Figure 3: Heart conduction system

the heart tissue [53]

$$\left(\vec{\vec{\sigma}}_i \cdot \vec{\nabla} u_e\right) \cdot \vec{n} \;=\; -\left(\vec{\vec{\sigma}}_i \cdot \vec{\nabla} V_m\right) \cdot \vec{n}, \quad \text{in } \partial\mathcal{B} \times (0,T). \tag{3}$$

In order to define the I_{ion} term, equations (1) must be coupled with a ionic model, i.e. a system of nonlinear ordinary differential equations (ODEs). In this work, for the ventricular domain, we apply the *Minimal model for human Ventricular action potentials* (MV) introduced in [6]. MV is a phenomenological model associated with three ionic currents and three gate variables, and governed by 28 parameters.

As previously explained, the fiber architecture of the cardiac muscle is considered. In order to include the anisotropy between the orthogonal and the tangent direction of the fibers, the conductivity tensors $\vec{\vec{\sigma}}_i$ and $\vec{\vec{\sigma}}_e$ are defined by

$$\vec{\vec{\sigma}}_{i,e} = \sigma_{i,e}^{v,t}\, \vec{\vec{I}} + (\sigma_{i,e}^{v,l} - \sigma_{i,e}^{v,t})\, \vec{\tau} \otimes \vec{\tau},$$

where $\vec{\vec{I}}$ denotes the 3D identity matrix, the vector $\vec{\tau}$ is of unit length and parallel to the local fiber direction, and $\sigma_{i,e}^{v,l}$ and $\sigma_{i,e}^{v,t}$ are respectively the conductivity coefficients in the intra- and extra-cellular ventricular medium measured along and across the fiber direction.

The bidomain model can be rewritten in weak form as follows. For all $t > 0$, find $V_m(\cdot,t) \in H^1(\mathcal{B})$, $u_e(\cdot,t) \in H^1(\mathcal{B})$ and $w_1(\cdot,t), \ldots, w_n(\cdot,t) \in L^\infty(\mathcal{B})$ with $\int_{\mathcal{B}} u_e = 0$, such that

$$\begin{cases} A_m \displaystyle\int_{\mathcal{B}} \Big(C_m \frac{\partial V_m}{\partial t} + I_{ion}(V_m, w_1, \ldots, w_n) \Big) \phi + \int_{\mathcal{B}} \Big[\vec{\vec{\sigma}}_i \cdot \left(\vec{\nabla} V_m + \vec{\nabla} u_e \right) \Big] \cdot \vec{\nabla} \phi \\ \qquad\qquad = A_m \displaystyle\int_{\mathcal{B}} I_{app} \phi, \\ \displaystyle\int_{\mathcal{B}} \Big[(\vec{\vec{\sigma}}_i + \vec{\vec{\sigma}}_e) \cdot \vec{\nabla} u_e \Big] \cdot \vec{\nabla} \psi + \int_{\mathcal{B}} \Big[\vec{\vec{\sigma}}_i \cdot \vec{\nabla} V_m \Big] \cdot \vec{\nabla} \psi = 0, \end{cases} \tag{4}$$

for all $\phi, \psi \in H^1(\mathcal{B})$ such that $\int_{\mathcal{B}} \psi = 0$. Under some regularity assumptions, we have existence and uniqueness of a solution of the bidomain model, see e.g. [5]. The hypothesis $\int_{\mathcal{B}} u_e = 0$ is necessary in order to have uniqueness and we show in Section 2.3 how to adapt this condition when atria and ventricles are coupled.

2.2 Surface bidomain model

As explained in Section 1, we have a surface mesh for the atria. Then, we need a surface model. In this study, we consider the surface-based model presented and mathematically studied in [8], derived from the volume bidomain model, and defined over the midsurface of the thin region. This surface-based model was obtained from a rigorous asymptotic analysis and was specifically designed for thin cardiac structures, as the atria for example. It takes into account the strong anisotropy variations across the thickness and it is extremely attractive in term of computation time compared to its 3D counterpart. We denote by $\mathcal{S}$ the midsurface of the wall and we denote by $H^1(\mathcal{S})$ the associated space. The surface-based bidomain model can be rewritten in weak form as follows, for all $t > 0$, find $V_m(\cdot,t) \in H^1(\mathcal{S})$, $u_e(\cdot,t) \in H^1(\mathcal{S})$ and $w_1(\cdot,t), \ldots, w_n(\cdot,t) \in L^\infty(\mathcal{S})$ with $\int_{\mathcal{S}} u_e = 0$, such that

$$\begin{cases} A_m \displaystyle\int_{\mathcal{S}} \Big(C_m \frac{\partial V_m}{\partial t} + I_{ion}(V_m, w_1, \ldots, w_n) \Big) \phi + \int_{\mathcal{S}} \Big(\underline{\underline{\sigma}}_i \cdot \left(\underline{\nabla} V_m + \underline{\nabla} u_e \right) \Big) \cdot \underline{\nabla} \phi \\ \qquad\qquad = A_m \displaystyle\int_{\mathcal{S}} I_{app} \phi, \\ \displaystyle\int_{\mathcal{S}} \Big((\underline{\underline{\sigma}}_i + \underline{\underline{\sigma}}_e) \cdot \underline{\nabla} u_e \Big) \cdot \underline{\nabla} \psi + \int_{\mathcal{S}} \Big(\underline{\underline{\sigma}}_i \cdot \underline{\nabla} V_m \Big) \cdot \underline{\nabla} \psi = 0, \end{cases} \tag{5}$$

for all $\phi, \psi \in H^1(\mathcal{S})$ such that $\int_{\mathcal{S}} \psi = 0$. We denote by $\sigma_{i,e}^{a,l}$ and $\sigma_{i,e}^{a,t}$ the conductivity coefficients in the intra- and extra-cellular atrial medium measured along and across the fiber direction. We define the intra- and extra-cellular diffusion tensors $\underline{\underline{\sigma}}_i$ and $\underline{\underline{\sigma}}_e$ by

$$\underline{\underline{\sigma}}_{i,e} = \sigma_{i,e}^{a,t}\underline{\underline{I}} + (\sigma_{i,e}^{a,l} - \sigma_{i,e}^{a,t})\big[I_0(\theta)\underline{\tau}_0 \otimes \underline{\tau}_0 + J_0(\theta)\underline{\tau}_0^{\perp} \otimes \underline{\tau}_0^{\perp}\big], \tag{6}$$

where $\underline{\underline{I}}$ denotes the identity tensor in the tangential plane, $\underline{\tau}_0$ is a unit vector parallel to the local fiber direction on the atria midsurface, and $\underline{\tau}_0^{\perp}$ such that $(\underline{\tau}_0, \underline{\tau}_0^{\perp})$ gives an orthonormal basis of the tangential plane. We use the fibers direction at the endocardium and at the epicardium to define the fibers direction $\underline{\tau}_0$ on the atria midsurface and the angle variation θ between the endocardium and the epicardium. The effect of angular variations appears in the model with the coefficients $I_0(\theta) = \frac{1}{2} + \frac{1}{4\theta}\sin(2\theta)$ and $J_0(\theta) = 1 - I_0(\theta)$. Note that $J_0(\theta) = 0$ (and $I_0(\theta) = 1$) if and only if $\theta = 0$, which corresponds to a constant direction in the thickness and then $\underline{\underline{\sigma}}_{i,e} = \sigma_{i,e}^{a,t}\underline{\underline{I}} + (\sigma_{i,e}^{a,l} - \sigma_{i,e}^{a,t})\,\underline{\tau}_0 \otimes \underline{\tau}_0$. By contrast, important angular variations make I_0 decrease and J_0 increase in (6) and the diffusion becomes more isotropic. In [13] this model was compared to several 3D models proposed in the literature [15, 21, 37].

The physiological model introduced by Courtemanche, Ramirez and Nattel in [14] – a classic atrial model – is considered (12 ionic currents and 20 other variables). The two atria are connected only by two regions, the Bachmann bundle and the *Fossa Ovalis*. We refer to [13] for more details.

2.3 Simulations on the whole heart

Coupled model From a mathematical point of view, volume and surface models are incompatible. It would be erroneous to solve them separately because the uniqueness criterion for the first model is not consistent with the second one. As seen in Sections 2.1 and 2.2, the unique solution $u_e(\cdot, t) \in H^1(\mathcal{B})$ of (4) is s.t. $\int_{\mathcal{B}} u_e = 0$ and the unique solution $u_e(\cdot, t) \in H^1(\mathcal{S})$ of (5) is s.t. $\int_{\mathcal{S}} u_e = 0$. That is why we consider a whole domain $\mathcal{B} \cup \mathcal{S}$ and a new global criterion. The resulting coupled problem should be justified mathematically, but it is well-posed at discrete level. Let $\Omega_h = \mathcal{B}_h \cup \mathcal{S}_h$, where $\mathcal{B}_h$ is the mesh of the ventricles and $\mathcal{S}_h$ is the mesh of the atria, and let $\mathcal{L}_h$ be the line such that $\mathcal{B}_h \cap \mathcal{S}_h = \mathcal{L}_h$. We denote by $\gamma_{\tilde{\Omega}} u$ the restriction of a function u to a subdomain $\tilde{\Omega}$. The finite dimensional approximation space $\mathcal{V}_h$ is then defined by: $u_h \in \mathcal{V}_h$ if and only if u_h is continuous in Ω_h, $\gamma_{\mathcal{B}_h} u_h \in H^1(\mathcal{B}_h)$, $\gamma_{\mathcal{S}_h} u_h \in H^1(\mathcal{S}_h)$, and $\int_{\mathcal{B}_h} u_h = 0$. Using (4), (5), the full model reads, find $(u_{e,h},\, V_{m,h}) \in \mathcal{V}_h$ such that $\forall \phi, \psi \in \mathcal{V}_h$,

$$\left\{\begin{array}{l} A_m \displaystyle\int_{\mathcal{S}_h} \Big(C_m \frac{\partial V_{m,h}}{\partial t} + I_{ion}(V_{m,h}, \cdots)\Big)\phi + A_m \int_{\mathcal{B}_h} \Big(C_m \frac{\partial V_{m,h}}{\partial t} + I_{ion}(V_{m,h}, \cdots)\Big)\phi \\ + \displaystyle\int_{\mathcal{S}_h} \Big(\underline{\underline{\sigma}}_i \cdot \big(\underline{\nabla} V_{m,h} + \underline{\nabla} u_{e,h}\big)\Big) \cdot \underline{\nabla}\phi + \int_{\mathcal{B}_h} \Big[\vec{\vec{\sigma}}_i \cdot \big(\vec{\nabla} V_{m,h} + \vec{\nabla} u_{e,h}\big)\Big] \cdot \vec{\nabla}\phi \\ \qquad\qquad = A_m \displaystyle\int_{\mathcal{S}_h} I_{app}\phi + A_m \int_{\mathcal{B}_h} I_{app}\phi, \\ \displaystyle\int_{\mathcal{S}_h} \Big((\underline{\underline{\sigma}}_i + \underline{\underline{\sigma}}_e) \cdot \underline{\nabla} u_{e,h}\Big) \cdot \underline{\nabla}\psi + \int_{\mathcal{S}_h} \Big(\underline{\underline{\sigma}}_i \cdot \underline{\nabla} V_{m,h}\Big) \cdot \underline{\nabla}\psi \\ \qquad + \displaystyle\int_{\mathcal{B}_h} \Big[(\vec{\vec{\sigma}}_i + \vec{\vec{\sigma}}_e) \cdot \vec{\nabla} u_{e,h}\Big] \cdot \vec{\nabla}\psi + \int_{\mathcal{B}_h} \Big[\vec{\vec{\sigma}}_i \cdot \vec{\nabla} V_{m,h}\Big] \cdot \vec{\nabla}\psi = 0. \end{array}\right. \tag{7}$$

Connection surface As previously mentioned, the atrioventricular node is the only pathway for the electrical signal between the atria and the ventricles. From a physiological point of view, a fibrous skeleton separates atria boundaries from ventricles epicardium. This layer isolates the atrial cells from the ventricular ones [36, 38]. We propose to model this fibrous skeleton with

σ_T^{body}	σ_T^{bones}	σ_T^{lungs}
$3.0\,10^{-4}$	$1.2\,10^{-4}$	$2.0\,10^{-5}$

Table 2: Torso conductivity parameters (all in S.cm^{-1})

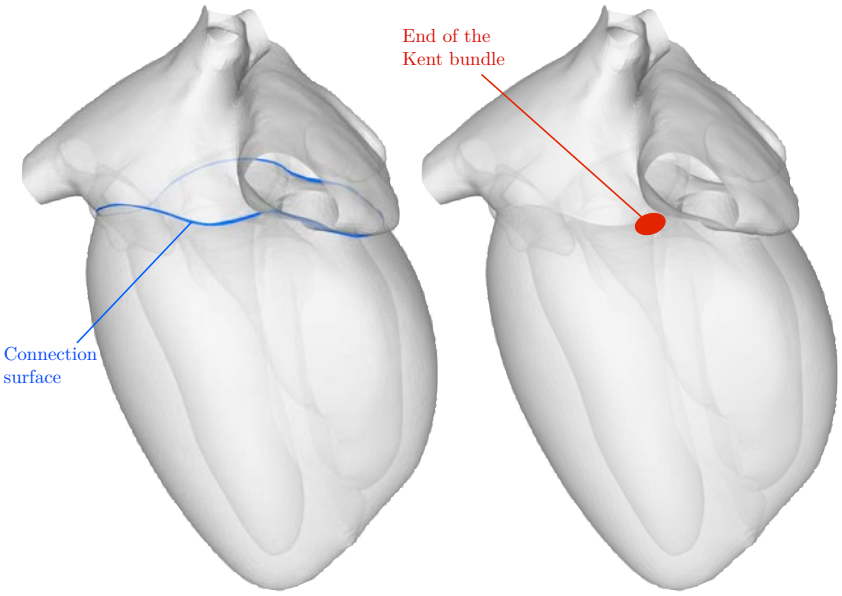

Figure 4: Left - Surface connection between the atria and the ventricles, Right - Kent bundle

a thin layer of the atrial surface, represented on the left of Figure 4. The idea is that in this area there is only a low conduction of the extracellular potential. In this region, denoted by $\mathcal{S}_c \subset \mathcal{S}$, the intracellular conductivity is set to zero and the extracellular conductivity denoted by $\sigma_e^{s,t}$ is very low (see Table 3). Finally, the surface and volume bidomain equations are solved simultaneously on this "hybrid" domain (7) with

$$\begin{aligned}
\vec{\vec{\sigma}}_{i,e} &= \sigma_{i,e}^{v,t}\,\vec{\vec{I}} + (\sigma_{i,e}^{v,l} - \sigma_{i,e}^{v,t})\,\vec{\tau}\otimes\vec{\tau}, \\
\underline{\underline{\sigma}}_{i,e} &= \sigma_{i,e}^{a,t}\underline{\underline{I}} + (\sigma_{i,e}^{a,l} - \sigma_{i,e}^{a,t})\big[I_0(\theta)\underline{\tau}_0\otimes\underline{\tau}_0 + J_0(\theta)\underline{\tau}_0^{\perp}\otimes\underline{\tau}_0^{\perp}\big], \text{ in } \mathcal{S}\setminus\overline{\mathcal{S}}_c, \\
\underline{\underline{\sigma}}_i &= \underline{\underline{0}}, \text{ and } \underline{\underline{\sigma}}_e = \sigma_e^{s,t}\underline{\underline{I}}, \text{ in } \mathcal{S}_c.
\end{aligned}$$

Parameters and applied currents The values of the membrane parameters are $A_m = 200.0\,\text{cm}^{-1}$ and $C_m = 10^{-3}$ mF.cm^{-2} for the whole heart. The conductivity takes different values depending on the region in the ventricles and atria (Table 3).

In the atria, the regions of fast conduction are the Bachmann bundle (BB), see Figure 3, the *Crista Terminalis* (CT) and the pectinate muscles (PM). By contrast, the *Fossa Ovalis* (FO) is a region of slow conduction. In order to model the different propagation velocities, we modify the values of g_{Na}, the maximal conductance of the Na^{2+} current I_{Na}. Table 4 gives the

$\sigma_e^{v,t}$	$\sigma_e^{v,l}$	$\sigma_i^{v,t}$	$\sigma_i^{v,l}$	$\sigma_e^{a,t}$	$\sigma_e^{a,l}$	$\sigma_i^{a,t}$	$\sigma_i^{a,l}$	$\sigma_e^{s,t}$
$6.0\,10^{-4}$	$2.0\,10^{-3}$	$2.0\,10^{-4}$	$2.0\,10^{-3}$	$9.0\,10^{-4}$	$2.5\,10^{-3}$	$2.5\,10^{-4}$	$2.5\,10^{-3}$	$7.5\,10^{-7}$

Table 3: Conductivity parameters (all in S.cm^{-1})

regular tissue	PM	CT	BB	FO
7.8	11.7	31.2	46.8	3.9

Table 4: Maximal conductance g_{Na} in the different atrial areas (all in nS.pF^{-1})

	EPI	ENDO	M	RV
[6]	30.0181	40.0	91.0	/
heart	25.0	40.0	61.0	26.0

Table 5: Changed ionic parameter τ_{so_1} of MV compared to [6]

parameters used for g_{Na}. Furthermore, the action potential duration is forced to be shorter, *i.e.* the parameter g_{k_s} is five times bigger than in [14].

In the ventricles, we modify the duration of the plateau too. In the MV model, we change the values of τ_{so_1} parameter in order to reduce the action potential duration for epicardial, endocardial and midmyocardial cells. This heterogeneity is considered in the left ventricle, for the positivity of T wave [55]. In the right ventricle, the cells are considered homogeneous and their parameters are taken as in the left ventricle epicardium, except for τ_{so_1} (Table 5).

Activation is initiated at the sinus node with a stimulus of 2ms which triggers a depolarization wavefront in the atria (Figure 3). For the sake of simplicity, the atrioventricular node, which is the only electrical connection between the atria and the ventricles, is not modeled with a sophisticated physiological model. Instead, the excitation is triggered in the ventricle after a parameterized delay (in healthy condition, we choose to start it at 190 ms). Similarly, the fast conduction in the Purkinje fibers (Figure 3) in modeled with a predefined stimulus pattern: a time-dependent thin subendocardial layer is activated by an external current on both right and left ventricles (see [4] for more details).

Simulation results The various simulations of this article are performed with the finite element library *FELiScE*[2], developed at Inria by the REO and MΞDISIM teams. The numerical methods used to solve problem (7) are presented in [4]. Figure 5 shows a full cardiac cycle. The corresponding first lead electrocardiogram is also represented. The electrical signal starts at the sinus node where the atrial depolarization (AD) begins. By 50 ms the wave quickly spreads along the *Crista Terminalis* as a consequence of the high conductivity in this part. Importantly, because of the rapid conduction in the Bachmann bundle, the wave spreads to the left atrial appendage and activates a substantial part of the left atrial wall. The depolarization of the right and left atria terminates at 100 ms and 110 ms, respectively. The ventricular depolarization begins at 190 ms. During this period, the atrial repolarization (AR) occurs. As we can see in the figure, at 200 ms the endocardium of the ventricles rapidly depolarizes. Then, the wave propagates across the ventricles. The repolarization ends at 430 ms in the right ventricle and at 470 ms in the left ventricle.

2.4 Coupling with the body

The last step in order to obtain an electrocardiogram is to couple the heart model with a diffusion problem in the rest of the body

$$-\operatorname{div}(\sigma_T \vec{\nabla} u_T) = 0, \text{ in } \Omega_T, \tag{8}$$

[2] `http://felisce.gforge.inria.fr`

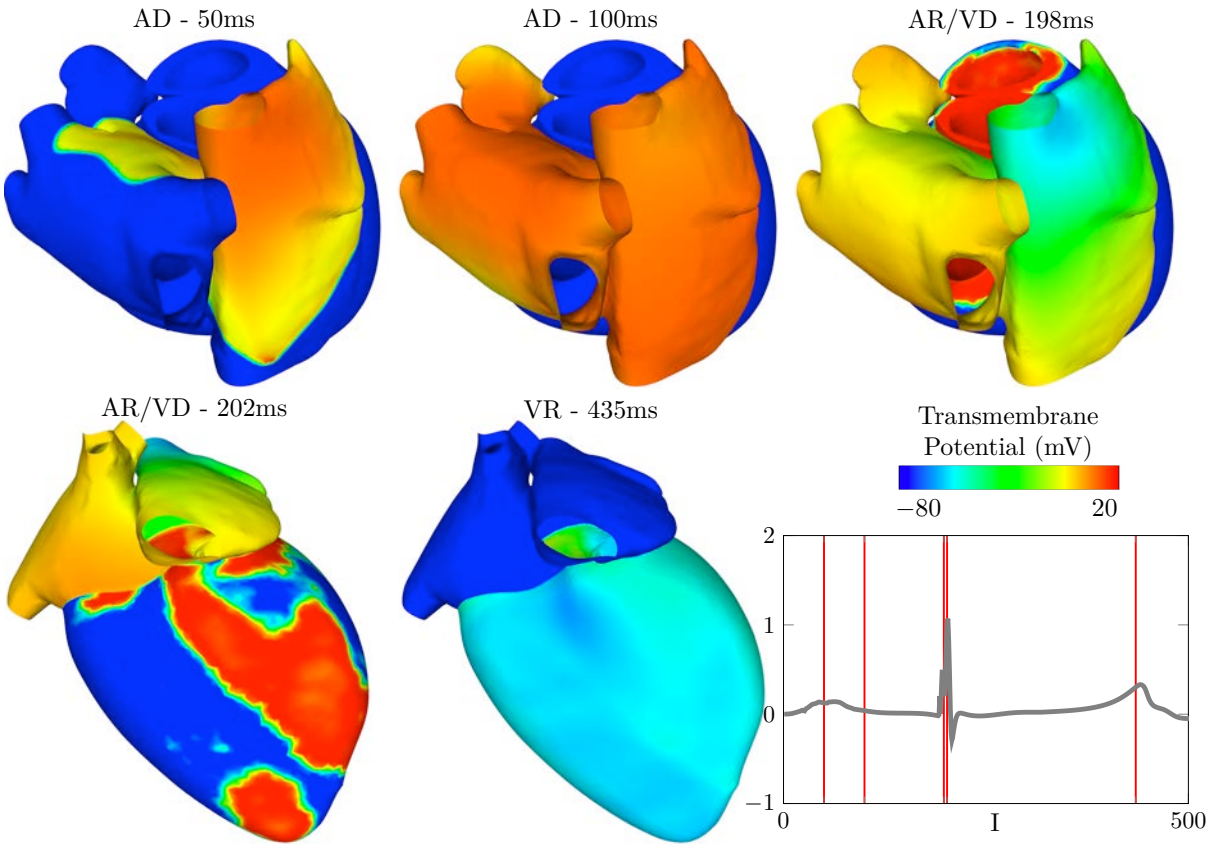

Figure 5: Simulations of heart depolarization in a healthy case with the corresponding electrocardiogram first lead

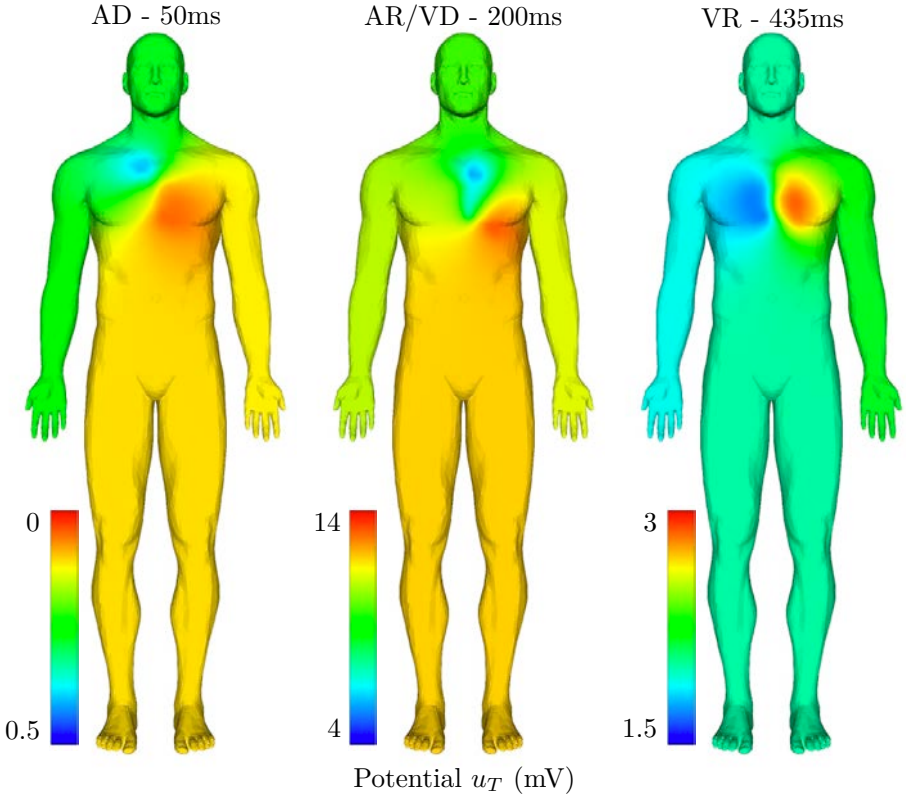

Figure 6: Body potential simulations of heart depolarization in a healthy case (Figure 5), with coupling condition (10)

where the electrical conductivity σ_T takes different scalar values in the ribs and the lungs (see [7] and Table 2).

On the body surface $\partial\Omega_T^{ext}$, an homogeneous Neumann boundary condition is imposed $\sigma_T\vec{\nabla}u_T \cdot \vec{n} = 0$. To define the transmission conditions at the heart-body interface $\partial\Omega_H$, we assume that the extracellular current does not flow through the pericardium (isolated heart assumption):

$$\vec{\bar{\sigma}}_e \cdot \vec{\nabla}u_e \cdot \vec{n} = 0 \tag{9}$$

and we consider the resistor-capacitor conditions presented in [4]:

$$R_p(\sigma_T\vec{\nabla}u_T) \cdot \vec{n} = R_pC_p\frac{\partial(u_e - u_T)}{\partial t} + (u_e - u_T) \tag{10}$$

where C_p and R_p stand for the capacitance and resistance of the pericardium, respectively.

Condition (9) is an approximation that has been shown not to affect too much the shape of the ECG in [4]. It allows us to solve the heart-body system as a one-way coupled problem, which dramatically reduced its computational cost.

Condition (10) allows us to model the fact that the transmission of potential through the pericardium is not perfect, and can be different for the ventricles and the atria. In this work, we take $R_p = 10^2$ $\Omega.\text{cm}^2$ on the surface in contact with the ventricles and $R_p = 10^5$ $\Omega.\text{cm}^2$ on the surface in contact with the atria. We neglect the capacitor effect by taking $C_p = 0$ mF.cm^2 in (10). The transmission between the heart and the body is therefore modeled as a Robin boundary condition:

$$R_p(\sigma_T\vec{\nabla}u_T) \cdot \vec{n} + u_T = u_e,\ \partial\Omega_H. \tag{11}$$

Figure 6 shows the body surface potential corresponding to the simulation shown in Figure 5.

2.5 Electrocardiogram computation

A standard electrocardiogram is based on the body surface potential recorded by 9 electrodes ($\Gamma_{ECG} = \{R, L, F, V_1, \ldots, V_6\}$, see Figure 7). These measures are combined to define 12 differences of potential, known as the 12 leads of the standard ECG:

$$\begin{array}{ll}
\text{I} = u_T(L) - u_T(R) & aVR = 1.5(u_T(R) - u_w) \\
\text{II} = u_T(F) - u_T(R) & aVL = 1.5(u_T(L) - u_w) \\
\text{III} = u_T(F) - u_T(L) & aVF = 1.5(u_T(F) - u_w) \\
V1 = u_T(V_1) - u_w & V4 = u_T(V_4) - u_w \\
V2 = u_T(V_2) - u_w & V5 = u_T(V_5) - u_w \\
V3 = u_T(V_3) - u_w & V6 = u_T(V_6) - u_w
\end{array}$$

where $u_w = \frac{1}{3}(u_T(L) + u_T(R) + u_T(F))$ is the Wilson potential [35].

3 Healthy and pathological numerical simulations of electrocardiograms

In this Section, we present the ECGs provided by the aforementioned model in healthy and pathological conditions. The healthy ECG is obtained by carefully choosing the parameters of the model in order to match most of the qualitative and quantitative features of a physiological ECG. To obtain the pathological ECGs, the approach is different: starting from the nominal values corresponding to a healthy ECGs, we modify the parameters in order to model the *physical characteristics* of the pathology. Then we observe the effects of these modifications on

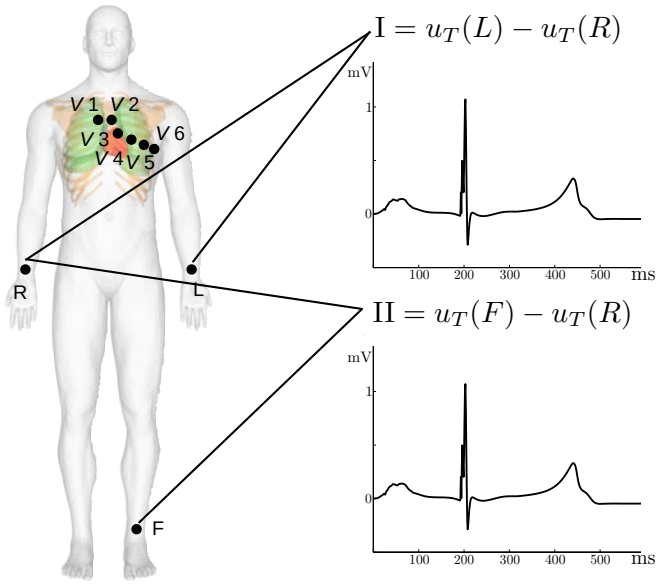

Figure 7: Standard 9 electrodes locations and first and second ECG leads.

the numerical ECG, and we compare its features with the ones described in the literature. It is important to emphasize that, for the pathological cases, the parameters are not intentionally fixed to match a given ECG. Thus, if the ECGs obtained after modeling the diseases look like those obtained on real patients, it gives confidence in the prediction capabilities of the model.

3.1 Healthy electrocardiograms

Figure 9 shows the simulated electrocardiogram in a healthy case, corresponding to the simulation of Figure 5. An electrocardiogram is typically described by distinguishing five events during the heartbeat, called P, Q, R, S and T "waves" (we will keep this standard terminology even though these events have nothing to do with waves). The P wave corresponds to the atrial depolarization, the QRS complex corresponds to the ventricular depolarization, the T wave corresponds to the ventricular repolarization. The typical durations of each wave, or each interval, are given in Table 6.

Table 6 also presents the durations of the simulated healthy ECG of Figure 9. These durations are obtained in the numerical ECG from the landmarks defined according to the following rules: the P wave (resp. QRS complex) starts if 1% of the atria (resp. ventricles) is activated (i.e. if the transmembrane potential V_m is greater than a threshold voltage V_{gate}); the P wave (resp. QRS complex) ends when 99% of the atria (resp. ventricles) are activated; the T wave starts when 20% of the ventricles are repolarized (*i.e.* $V_m \leq 0$); the T wave ends when 99% of the ventricles are fully repolarized ($V_m \leq 0$). If the minimal value of V_m is $V_{min} = -80$mV and its maximal value $V_{max} = 20$mV, we define $V_{gate} = -67$mV, which corresponds to a threshold voltage $\theta_w = 0.13$ in the MV model.

Table 7 gives the main features of each wave in a normal electrocardiogram. Interestingly, the simulated ECG verifies almost all the expected criteria. We only observe a discrepancy in the aVL lead, but this lead is not the most important one for the ECG interpretation.

To qualitatively assess the waves amplitude and orientation (but not their duration), Figure 10, which shows the normal variations of wave amplitude measured in adults [54], is extremely convenient. A visual comparison of Figures 9 and 10 shows that, for almost every lead,

	P wave	PR interval	Q wave	QR interval	S wave	QRS interval	QT interval
Typical Duration	< 0.12	0.12 to 0.21	< 0.04	< 0.03 $V1 - V2$ < 0.05 $V5 - V6$	< 0.04	< 0.10	0.35 to 0.45
Healthy ECG	0.08	0.19	0.015	0.015 $V1 - V2$ 0.02 $V5 - V6$	0.01	0.04	0.29

Table 6: Durations of the simulated healthy ECG of Figure 9 compared with typical durations [54] (all in s)

Wave/Interval	Description	Simulated ECG
P wave	≤ 0.25mV	✓ 0.2mV
	positive I, II, V3 to V6	✓
	negative aVR	✓
Q wave	limb leads $\leq 25\%$ of R	✓
	precordial leads $\leq 15\%$ of R	✓
	always negative	✓ except for aVL
R wave	limb leads ≤ 2mV	✓
	precordial leads ≤ 3mV	✓
	always positive, negative in aVR	✓
	R wave progression, see Figure 8	✓
S wave	always negative	✓
	small I, II, V5, V6	✓
	important V1 to V3	✓
ST interval	-0.05mV to 0.1mV	✓
	isoelectric	✓
	displacement of 0.02mV in V1, V3	✓
T wave	positive I, II, V3 to V6	✓
	negative aVR (follow the QRS)	✓

Table 7: Criteria for a typical electrocardiogram [54] compared with simulated ECG of Figure 9

each wave of our numerical ECG is in the range of the normal values. Note that in Figure 10, the length of each wave was arbitrarily chosen as its maximal normal duration. This is the reason why the full PQRST duration is so long in this schematic.

Here is another qualitative assessment. The R wave is known to have an important property in the precordial leads: it uniformly progresses from a RS complex in V1-V2 to a QRS complex in V5-V6 via a RS complex in V3-V4. The top of Figure 8, which represents this R variation, is extracted from [54]. The bottom of the same Figure shows the results of our simulation. Again, the qualitative agreement is very satisfactory.

A last qualitative comment is in order: we note that the P wave presents some oscillations in all the leads of the numerical ECG. The explanation of these oscillations is the brutal changes of the fibers' direction in the atria. It is also possible that the surface representation of the atria accentuate these oscillations.

3.2 Pathological electrocardiograms

In this Section, we modify the protocol of the simulation that provided the healthy ECG (Figure 9) in order to simulate different cardiac pathologies. Then we verify if the numerical

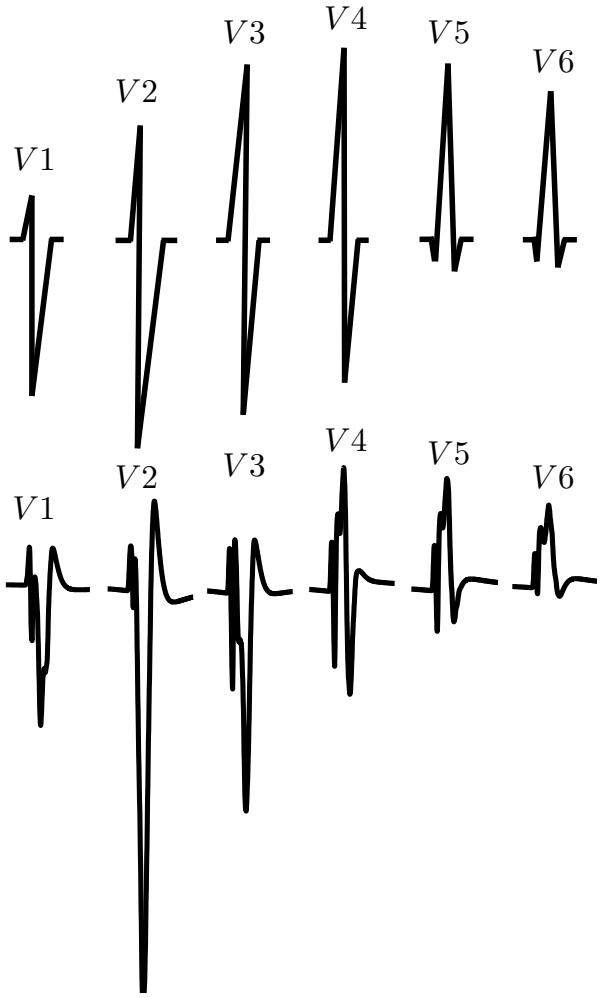

Figure 8: R wave progression in the precordial leads: schematic view from [54] in the top, and simulated ECG in the bottom

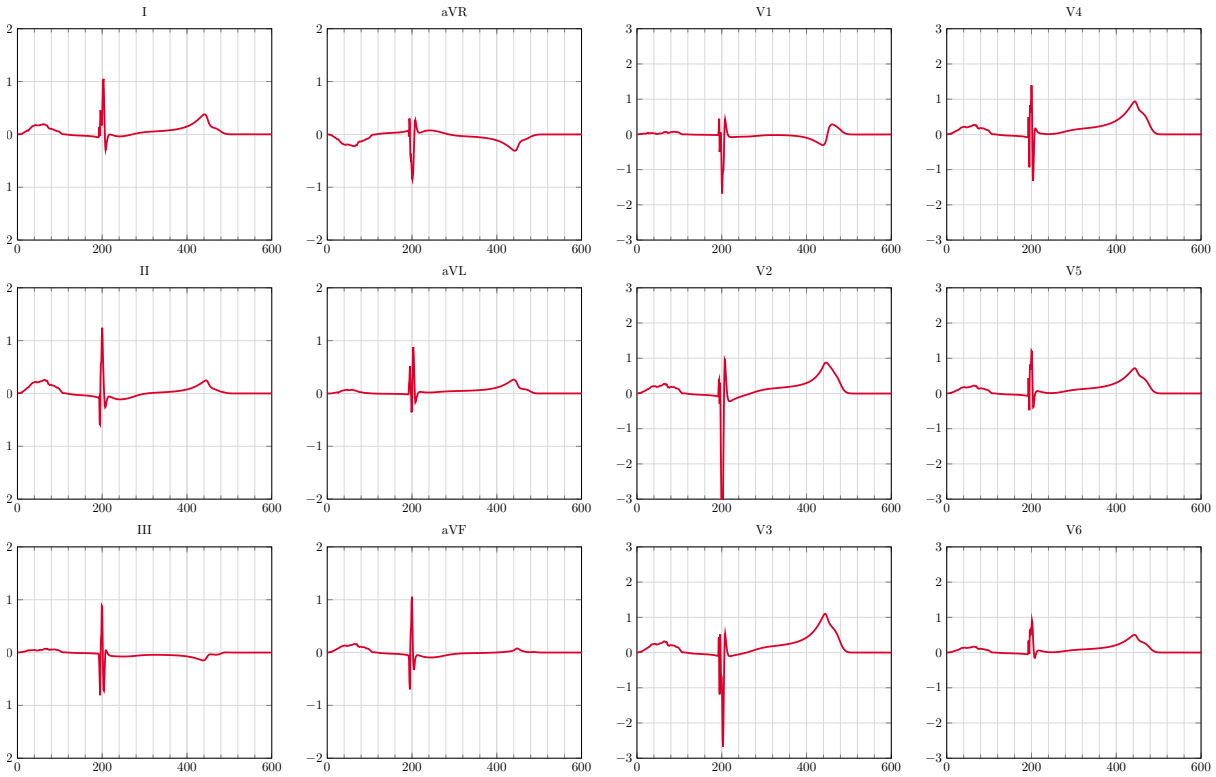

Figure 9: Healthy electrocardiogram corresponding to simulation of Figure 5 (voltages (mV) versus time (ms))

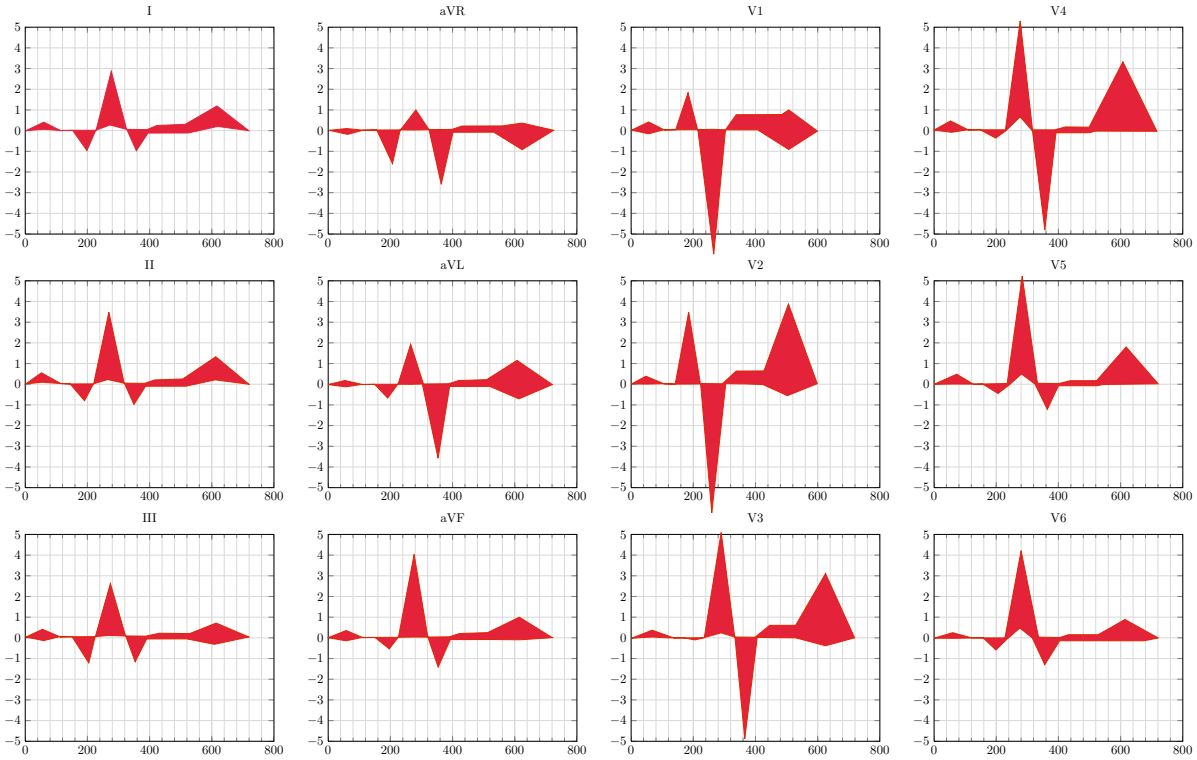

Figure 10: Normal variations of amplitude measurements in adults healthy ECG [54]

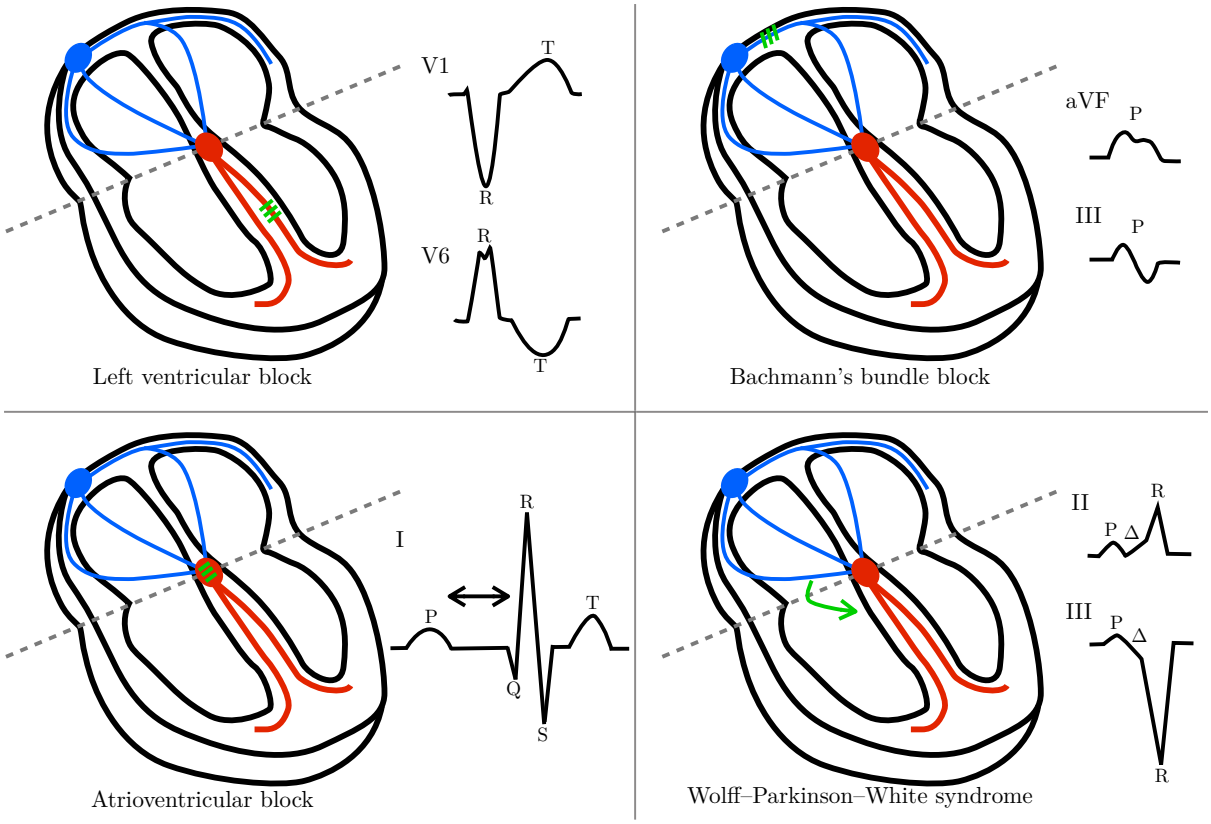

Figure 11: Different pathologies

ECGs present the main features that allow a medical doctor to detect the pathology. The different pathologies are schematically represented in Figure 11, along with the most important leads in each case.

3.2.1 Left and right ventricular block

We start with a left or a right bundle branch block. In a healthy case, the right and the left ventricles are activated simultaneously. Now, in order to simulate a left (or a right) bundle branch block, the initial activation is blocked in the left (resp. right) ventricle. In the top-left of Figure 11, we can see a left ventricular block. In order to obtain a left (resp. right) bundle branch block, the depolarization of the left (resp. right) Purkinje fibers is delayed as indicated in [4]. Results are reported in Figure 12 for the left and right bundle branch blocks. We recognize the main characteristics of a bundle branch block. In particular, we can see that leads V1 and V6 correspond to a left ventricular block described in Figure 11. We can see a larger QRS and a lead V1 without Q-wave [35]. The QRS-complex exceeds 0.12 seconds in both cases. Furthermore, it can be seen in Figure 12 that the duration between the beginning of the QRS complex and its last positive wave in V1 (resp. V6) exceeds 0.04 seconds which is a sign of right (resp. left) bundle branch block [35].

3.2.2 Bachmann's bundle block

In the heart conduction system, the Bachmann's bundle connects the left atrium with the right atrium and is the preferential path for electrical activation of the left atrium. A Bachmann's bundle block is represented at the top-right of Figure 11. It is characterized by the

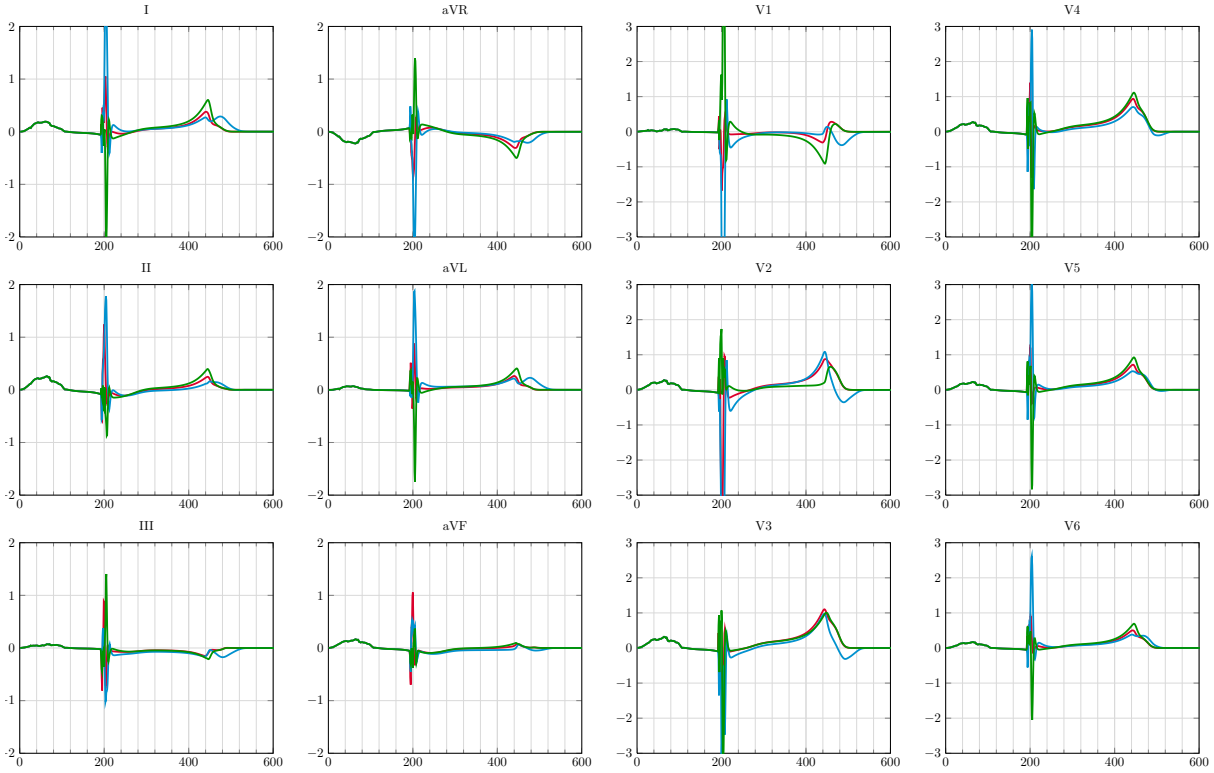

Figure 12: Left and Right Bundle Brunch Block – Healthy case in red, LBBB in blue and RBBB in green (voltages (mV) versus time (ms))

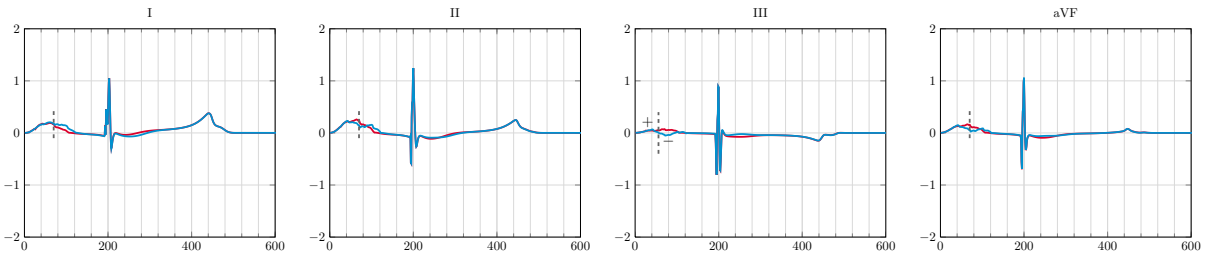

Figure 13: A Bachmann's Bundle Block – Healthy case in red, BBB in blue (voltages (mV) versus time (ms))

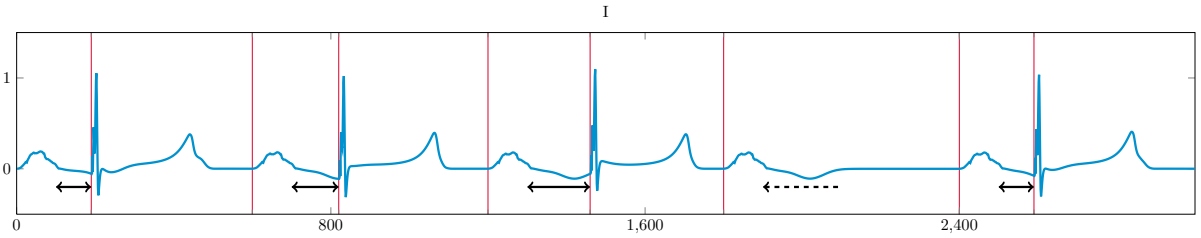

Figure 14: Wenckebach atrioventricular block (voltages (mV) versus time (ms))

presence of P-wave duration that equals or exceeds 0.12 seconds and presents usually a bimodal morphology, especially in leads I, II, aVF and the lead III becomes biphasic, as we can see in Figure 11. This is a very specific sign of left atrial enlargement [20, 35]. We simulate it by decreasing the maximal conductance $g_{Na} = 7.8$ in the Bachmann's bundle. The results are given in Figure 13. The more important the block, the more negative the P wave on lead III. A negative P wave in the third lead corresponds to the retrograde depolarization of the left atrium. The morphology of the simulated P wave is in very good agreement with the criteria given in [2, 3] for various degrees of Bachmann's bundle blocks.

3.2.3 Atrioventricular block

An atrioventricular block (or AV block) is a damage of the conduction between atria and ventricles. Under normal conductions, the atrioventricular node is activated by the atrial wave and it delays the activation by approximately 0.12 seconds. This delay is extremely important since it allows the atria to eject all their blood into the ventricles before the ventricular contraction. The bottom-left of Figure 11 represents an atrioventricular block. There are different degrees and types of atrioventricular blocks [35]. We consider here a Wenckebach atrioventricular block. To obtain it, we progressively increase the ventricular activation time: the delay between atrial and ventricular stimulus is 190 ms in the first beat, 220 ms in the second one, 260 ms in the third one, and at the fourth beat the ventricles are not stimulated at all. The Wenckebach block is characterized by a progressive prolongation of the PR interval on consecutive beats followed by a blocked P wave. After the missing QRS complex, the PR interval resets and the cycle repeats. Figure 14 correctly represents these features.

3.2.4 Wolff-Parkinson-White syndrome

The Wolff-Parkinson-White syndrome is one of the numerous pathologies of the conduction system of the heart. It corresponds to a pre-excitation syndrome and is caused by the presence

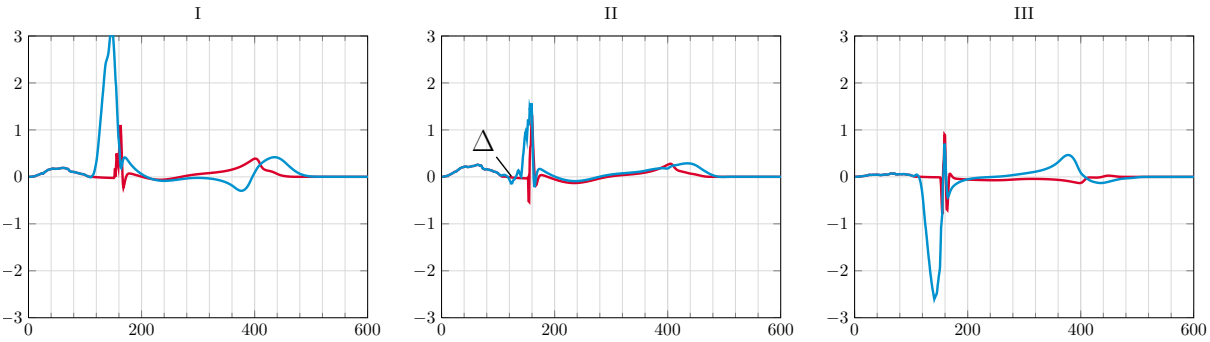

Figure 15: Wolff-Parkinson-White syndrome – Healthy case in red and WPW in blue (voltages (mV) versus time (ms))

of an abnormal electrical conduction pathway between the atria and the ventricles, named the Kent bundle. Electrical signals travel down this abnormal pathway and may stimulate the ventricles prematurely. In the bottom-right of Figure 11, we can see a schematic of the Wolff-Parkison-White syndrome. We model the abnormal pathway by stimulating a ventricle area near of the atria represented at the right of Figure 4. The Wolff-Parkinson-White syndrome is commonly diagnosed with the electrocardiogram [46]. It is characterized by a delta wave, a slurring of the initial segment of the QRS complex, due to the arrival of the impulses at ventricles via the abnormal route, which is associated with a short PR interval. Another feature is a QRS complex widening with a total duration greater than 0.12 seconds. We can indeed observe these characteristics, in particular the delta wave in Figure 15.

3.3 Comparison with the Mitchell-Schaeffer model

In this section we are interested in assessing the impact of the ionic model on the ECG simulation. In particular, we apply the technique described above to the Mitchell and Schaeffer model [39] which is a one-current phenomenological ionic model, appreciated for its simplicity.

The Mitchell and Schaeffer model is applied with the same conductivity parameters (excepted for some atrial areas, read below) and the same initial stimulus as described above in the case of more accurate models such as the Courtemanche and the MV models. Table 8 gives the value of the Mitchell-Schaeffer parameters. In order to correctly reproduce the T wave, we take into account three layers of cells in the left ventricles and an homogeneous tissue in the right ventricle as described in Section 2.3. The parameter that varies according to the type of cell is the τ_{close} parameter, whose values are reported in Table 8. On the atria surface, the repolarization "propagates" in the same direction as the depolarization. We therefore take a constant value for τ_{close}, equals to 100ms. As previously explained, we changed the values of the maximal conductance g_{Na} in the different atrial areas in the Courtemanche-Ramirez-Nattel model in order to take into account these bundles. The Mitchell-Schaeffer model does not allow the same flexibility then we decide to directly modify the value of conductivity parameters even if this is probably less physiological. The atrial conductivities are modified as reported in Table 9 in order to represent the different slow and fast bundles.

Figure 16 shows the ECGs obtained with the Mitchell and Schaffer (MS) model and the combined Courtemanche/MV model. We can see that the results are very close. With the MS model some oscillations in the P wave and the QRS complex are fixed, but the R wave progression in precordial leads is less precise and the T wave of V2 and V3 is not satisfactory. It is interesting to note that the results of the simulations are robust with respect to the choice

τ_{in}	τ_{out}	τ_{open}	$\tau_{close}^{\text{endo}}$	$\tau_{close}^{\text{Mcell}}$	$\tau_{close}^{\text{epi}}$	τ_{close}^{RV}	V_{min}	V_{max}	V_{gate}
($\text{cm}^2.\text{mA}^{-1}$)				(ms)				(mV)	
4.0	90.0	300.0	120.0	100.0	80.0	90.0	−80.0	20.	−67.0

Table 8: Mitchell and Schaeffer parameters and constants (different values of τ_{close} are given because of an heterogeneous tissue is considered, see [4])

	regular tissue	PM	CT	BB	FO
$\sigma_i^{a,t}$	$2.5\,10^{-4}$	$4.5\,10^{-4}$	$7.5\,10^{-4}$	$1.19\,10^{-3}$	$2.5\,10^{-4}$
$\sigma_e^{a,t}$	$9.0\,10^{-4}$	$1.35\,10^{-3}$	$2.7\,10^{-3}$	$4.3\,10^{-3}$	$9.0\,10^{-4}$
$\sigma_i^{a,l}$	$2.5\,10^{-3}$	$4.5\,10^{-3}$	$1.09\,10^{-2}$	$1.86\,10^{-2}$	$2.27\,10^{-3}$
$\sigma_e^{a,l}$	$2.5\,10^{-3}$	$4.5\,10^{-3}$	$1.09\,10^{-2}$	$1.86\,10^{-2}$	$2.27\,10^{-3}$

Table 9: Atrial conductivity parameters (all in S.cm^{-1}) for the Mitchell-Schaeffer model

of the ionic model: the Courtemanche/MV model in general gives better results, but it can be replaced by the MS model in order to reduce the computational costs without affecting too much the ECG. This remark is especially important if the ECG simulator has to be used for inverse problems: in that case, a model with a reduced number of parameters is probably more suitable.

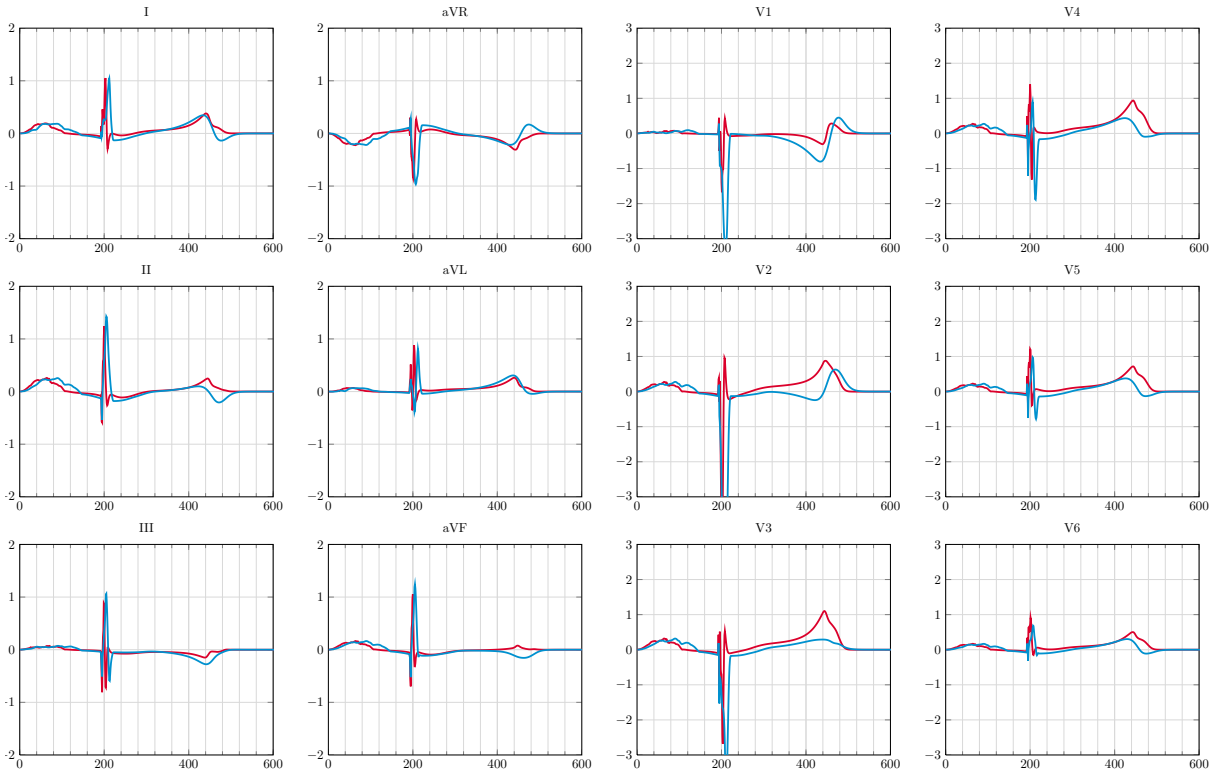

Figure 16: ECGs obtained with different ionic models – Courtemanche/MV ionic model in red and Mitchell and Schaeffer model in blue (voltages (mV) versus time (ms))

4 Electrodes vest

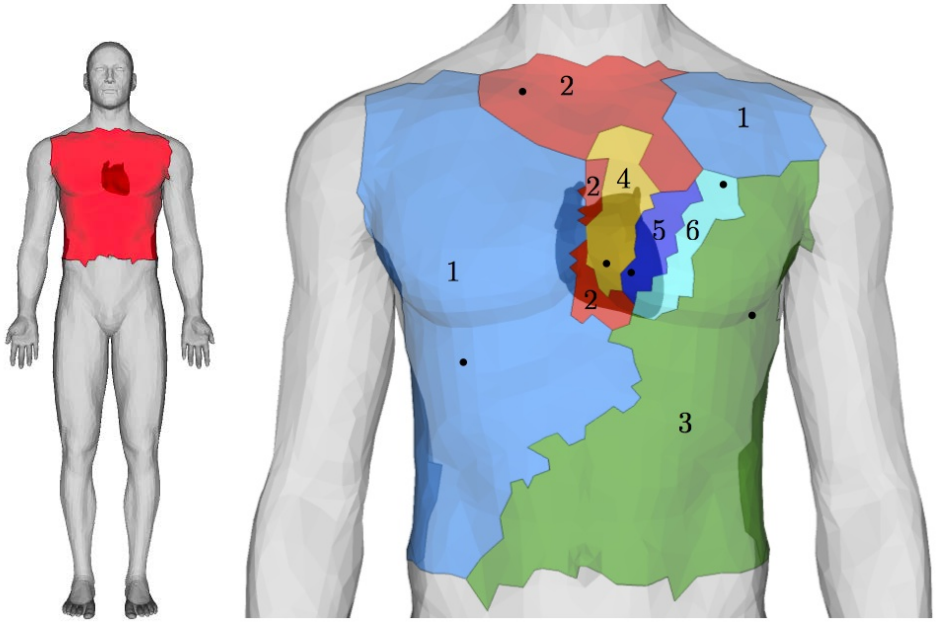

Figure 17: Human body mesh, electrodes vest (left) and clustering results (right).

In the previous sections, we focused on the 12-lead ECG because it is widely used in practice, and it can be easily assessed with the medical literature. But our simulator can of course provide more sophisticated measurements, like those obtained with electrodes vests.

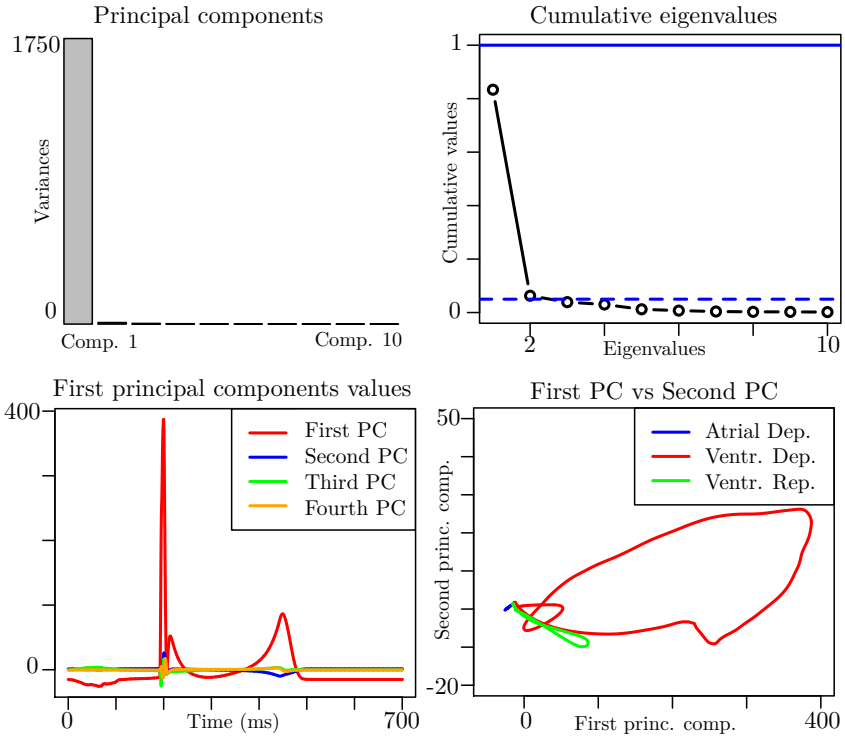

Figure 18: Principal components analysis on electrodes vest signals

Many studies on this topic have been carried out: on the forward problem and the analysis of useful number the electrodes to catch the signal (see for instance [27, 24]) and on the inverse problem of reconstruction of the potential on the heart surface, see *e.g.* [45, 51]. Our objective here is less ambitious: we just show an example of a statistical analysis that can be done with the measures provided by our ECG simulator.

To do so, we simulate a "virtual" electrodes vest which contains $N_{ECG} = 1,236$ electrodes. In Figure 17(left) the body part of the mesh is shown, the measures points are all the nodes of the mesh in the red part. The heart geometry used in this study contains $N_{\partial\Omega_H} = 28,510$ boundary vertices. We compute the body surface potential as described in Section 2.

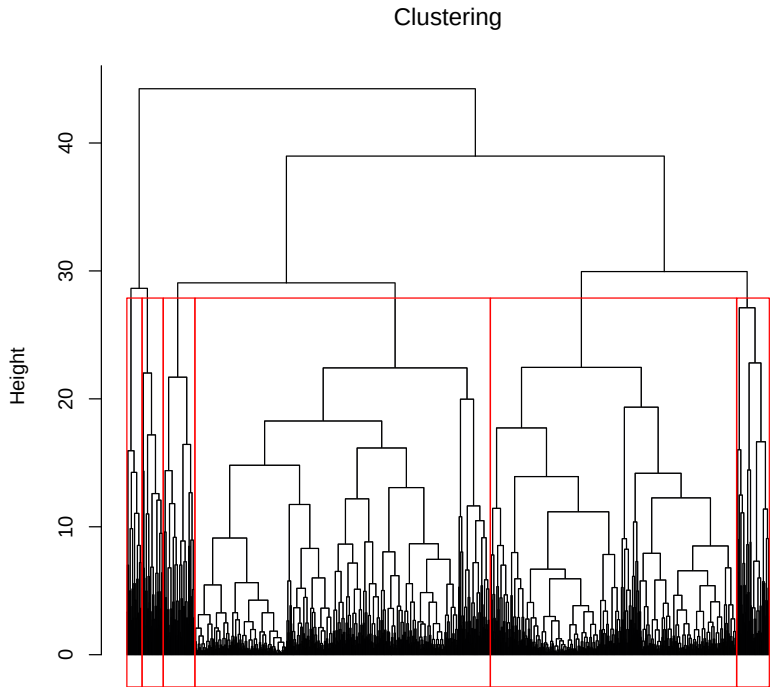

Figure 19: Clustering agglomeration of electrodes measures potential using euclidean distance and complete linkage. Red lines indicates the subdivision in six groups.

Finally, we are interested in analyzing the dependence on the electrodes measures compared to their positions on the body. A hierarchical clustering analysis is made, the $1,236$ measures are compared using an Euclidean distance between the curves and a complete linkage is applied (for further details on the clustering analysis see for instance [30]). In Figure 19, the tree produced by the clustering process is shown. We observe that a main subdivision into three wide groups is made while a second step of clustering suggests the subdivision into six smaller groups of signals. Then, let us divide the points measured into 6 groups, which is the number of precordial leads of a standard ECG too. In Figure 17 (right) the groups are shown. We observe that three bigger areas are defined and then a more precise definition of the cluster is given in the heart correspondence. In Figure 20 we plot the signals belonging to the six clusters. We can see a sort of P wave and QRS complex progressions between the first three clusters: the amplitude of P waves in the first cluster is higher than in cluster number 2, while in the third one is almost zero; the QRS complex changes in amplitude, it is higher in the first and third clusters and smaller in the second one. On the opposite, the clusters 4, 5, and 6 could be mismatched with the first ones.

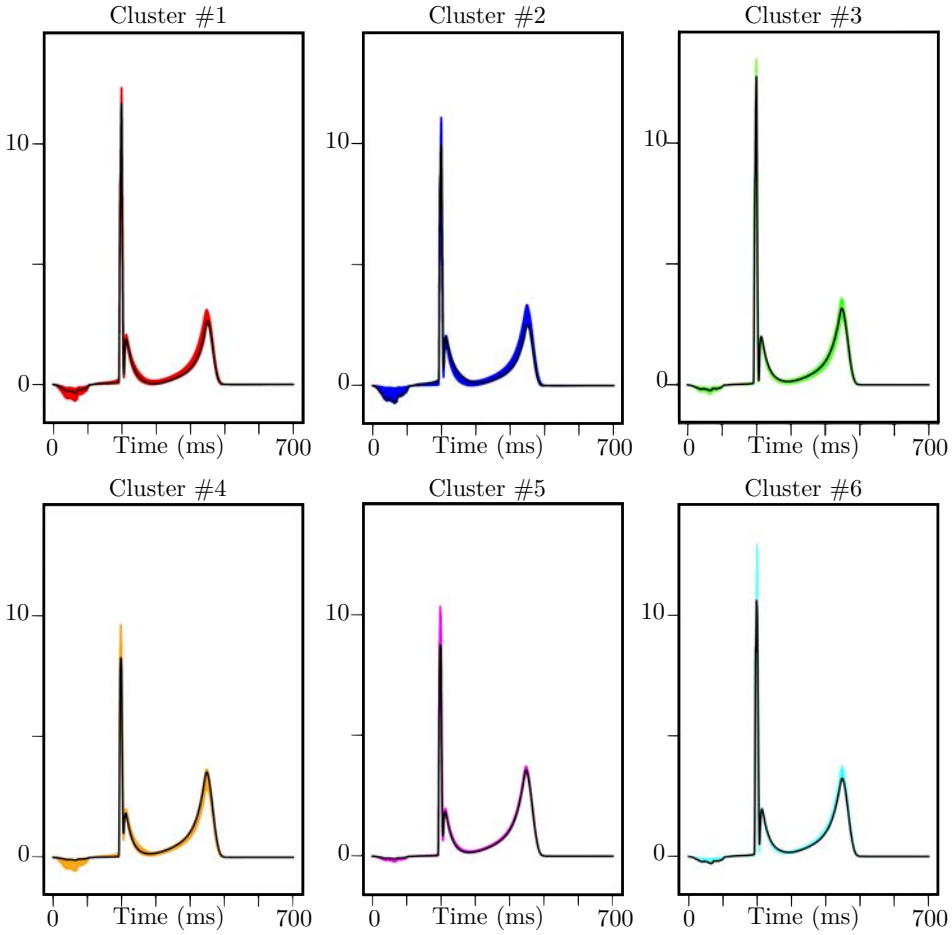

Figure 20: Signals in different clusters as shown in Figure 4. Black lines indicate the "mean" point (dots of Figure 4)

Then, we re-apply the principal components decomposition on the 6 centers of the cluster, *i.e.* the points indicated in Figure 17 (right). Comparing results of Figure 21 with Figure 18 we observe that the first principal component is again the most important one, and the same curves represents the first, second and third principal components.

We can conclude two different statistical approaches – the principal components decomposition and the hierarchical clustering analysis – suggest that it is not necessary to have a very high number of skin electrodes to describe the body surface potential. It seems that a limited number of correctly positioned electrodes would be enough to represent most of the features of the signal.

5 Conclusions

We have presented a comprehensive model for the simulation of full cycle ECGs. The main ingredients are: a volume bidomain system in the ventricles with the MV ionic model, a surface bidomain system in the atria with the Courtemanche ionic model, a one-way coupling between the heart and the torso through a resistor transmission condition.

This modeling effort leads to realistic ECGs in healthy and in four pathological cases: left and right bundle branch blocks, Bachmann's bundle block, AV block and Wolff-Parkinson-White

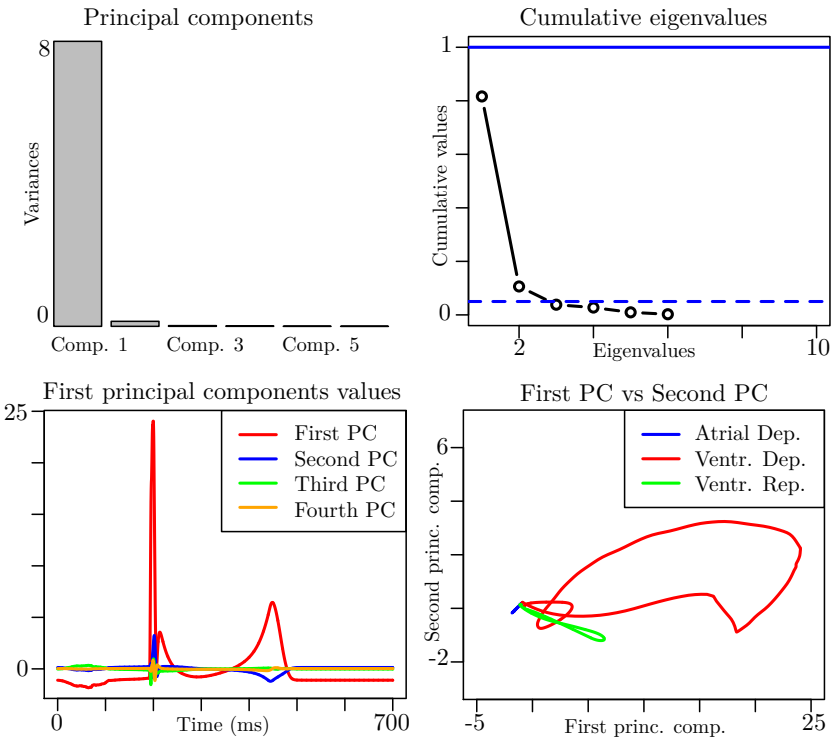

Figure 21: Principal components analysis on 6 electrodes signals.

syndrome. In the healthy case, we showed that similar results could been obtained with the one-current phenomenological ionic model proposed by Mitchell and Schaeffer.

This work could be improved in different directions. More physiological model of the atrio-ventricular node and the Purkinjie fibers could be included. A strong coupling with the torso could be done to assess the impact of the one-way coupling assumption. The electromechanical simulations presented in [9] could be extended to include the atria and performed with the present electrical model of the heart.

References

[1] R. R. Aliev and A. V. Panfilov. A simple two-variable model of cardiac excitation. *Chaos, Solitons & Fractals*, 7(3):293–301, 1996.

[2] A. Bayés de Luna, J. Guindo, X Viñolas, A. Martinez-Rubio, R. Oter, and Bayés-Genís A. Third-degree inter-atrial block and supraventricular tachyarrhythmias. *Eurospace*, 1:43–46, 1999.

[3] A. Bayés de Luna, P. Platonov, F. G. Cosio, I. Cygankiewicz, C. Pastore, R Baranowski, A. Bayés-Genis, J. Guindo, X Viñolas, J. Garcia-Niebla, R. Barbosa, S. Stern, and D. Spodick. Interatrial blocks. A separate entity from left atrial enlargement: a consensus report. *Journal of Electrocardiology*, 45(5):445–451, 2012.

[4] M. Boulakia, S. Cazeau, M.A. Fernández, J.-F. Gerbeau, and N. Zemzemi. Mathematical modeling of electrocardiograms: a numerical study. *Annals of Biomedical Engineering*, 38(3):1071–1097, 2010.

[5] Y. Bourgault, Y. Coudière, and C. Pierre. Existence and uniqueness of the solution for the bidomain model used in cardiac electrophysiology. *Nonlinear Analysis-Real World Applications*, 10(1):458–482, 2009.

[6] A. Bueno-Orovio, E. M. Cherry, and F. H. Fenton. Minimal model for human ventricular action potentials in tissue. *Journal of Theretical Biology*, 253:544–560, 2008.

[7] Martin L. Buist and Andrew J. Pullan. The effect of torso impedance on epicardial and body surface potentials: A modeling study. *EEE Trans. Biomed. Eng.*, 50(7):816 – 824, 2003.

[8] D. Chapelle, A. Collin, and J.-F. Gerbeau. A surface-based electrophysiology model relying on asymptotic analysis and motivated by cardiac atria modeling. *M3AS*, 23(14):2749–2776, 2013.

[9] D. Chapelle, M. A. Fernández, J-F. Gerbeau, P. Moireau, J. Sainte-Marie, and N. Zemzemi. Numerical simulation of the electromechanical activity of the heart. In *Functional Imaging and Modeling of the Heart*, number 5528 in Lecture Notes in Computer Science, pages 357–365. Springer-Verlag, 2009.

[10] K. Chatterjee and B. Massie. Systolic and diastolic heart failure: Differences and similarities. *Journal of Cardiac Failure*, 13(7):569–576, 2007.

[11] G.I. Cohen, M. White, R.A. Sochowski, A.Z. Klein, P.D. Bridge, W.J. Steward, and K.L. Chang. Reference values for normal adult transesophageal echocardiographic measurements. *Journal of the American Society of Echocardiography*, (8):221–230, 1995.

[12] P. Colli Franzone, L.F. Pavarino, and B. Taccardi. Simulating patterns of excitation, repolarization and action potential duration with cardiac Bidomain and Monodomain models. *Mathematical Biosciences*, 197(1):35–66, 2005.

[13] A. Collin, J.-F. Gerbeau, M. Hocini, M. Haïssaguerre, and D. Chapelle. Surface-based electrophysiology modeling and assessment of physiological simulations in atria. *FIMH 2013*, 7945:352–359, 2013.

[14] M. Courtemanche, R.J. Ramirez, and S. Nattel. Ionic mechanisms underlying human atrial action potential properties: insights from a mathematical model. *American Journal of Physiology*, (275):H301–H321, 1998.

[15] D. Deng, Y. Gong, G. Shou, and P. Jiao. Simulation of biatrial conduction via different pathways during sinus rhythm with a detailed human atrial model. *Journal of Zheijiang University-SCIENCE B (Biomedicine and Biotechnology)*, pages 1862–1783, 2012.

[16] D.B. Ennis, G.R. Rudd-Barnard, B. Li, C.G. Fonseca, A.A. Young, Cowan B.R., and R.A. Stewart. Changes in mitral annular geometry and dynamics with β- blockade in patients with degenerative mitral valve disease. *Circulation cardiovascular imaging*, 3(6):687–693, 2010.

[17] R. Fitzhugh. Impulses and physiological states in theoretical models of nerve membrane. *Biophysical Journal*, 1(6):445–466, 1961.

[18] P. Frey. Yams: A fully automatic adaptive isotropic surface remeshing procedure. Technical report 0252, Inria, Rocquencourt, France, Nov. 2001.

[19] S. Göktepe and E. Kuhl. Computational modeling of cardiac electrophysiology: A novel finite element approach. *International Journal for Numerical Methods in Engineering*, 79(2):156–178, 2009.

[20] Netter F. H. *The Heart - The Netter collection of medical illustrations*, volume 5. Saunders Elsevier, 1969.

[21] D.M. Harrild and S.H. Craig. A computer model of normal conduction in the human atria. *Circulation Research*, (87):e25–e36, 2000.

[22] S.Y. Ho, R.H. Anderson, and D. Sánchez-Quintana. Atrial structure and fibres: morphologic bases of atrial conduction. *Cardiovascular Research*, (54):325–336, 2002.

[23] S.Y. Ho and D. Sanchez-Quintana. The Importance of Atrial Structure and Fibers. *Clinical Anatomy*, 22(1):52–63, 2009.

[24] R. Hoekema, G. J. Uijen, and A. van Oosterom. On selecting a body surface mapping procedure. *Journal of Electrocardiology*, 32(2):93–101, 1999.

[25] L.E. Hudsmith, S.E. Petersen, J.M. Francis, M.D. Robson, and S. Neubauer. Normal human left and right ventricular and left atrial dimensions using steady state free precession magnetic resonance imaging. *Journal of cardiovascular magnetic resonance*, 7(5):775–782, 2005.

[26] G. Huiskamp. Simulation of depolarization in a membrane-equations-based model of the anisotropic ventricle. *Biomedical Engineering, IEEE Transactions on*, 45(7):847–855, 1998.

[27] Geertjan Huiskamp and Fred Greensite. A new method for myocardial activation imaging. *IEEE Trans. Biomed. Eng.*, 44(6):433–446, 1997.

[28] D. E. Hurtadoa and E. Kuhl. Computational modelling of electrocardiograms: repolarisation and T-wave polarity in the human heart. *Computer Methods in Biomechanics and Biomedical Engineering*, 17(9):2425–2435, 2014.

[29] P. Jiamsripong, T. Honda, C.S. Reuss, R.T. Hurst, H.P. Chaliki, D.E. Grill, S.L. Schneck, R. Tyler, B.K. Khandheria, and S.J. Lester. Three methods for evaluation of left atrial volume. *European Journal of Echocardiography*, (9):351–355, 2008.

[30] R.A. Johnson and D.W. Wichern. *Applied Multivariate Statistical Analysis*. Pearson Education, sixth edition, 2007.

[31] D. Keller, G. Seemann, D. Weiss, D. Farina, J. Zehelein, and O. Dössel. Computer based modeling of the congenital long-qt 2 syndrome in the visible man torso: From genes to ecg. In *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, pages 1410–1413, 2007.

[32] M. Krueger, V. Schmidt, C. Tobón, F. Weber, C. Lorenz, D. Keller, H. Barschdorf, M. Burdumy, P. Neher, G. Plank, K. Rhode, G. Seemann, D. Sanchez-Quintana, J. Saiz, R. Razavi, and O. Dössel. Modeling atrial fiber orientation in patient-specific geometries: a semi-automatic rule-based approach. *Functional Imaging and Modeling of the Heart*, pages 223–232, 2011.

[33] G. T. Lines, M. L. Buist, P. Grottum, A. J. Pullan, J. Sundnes, and A. Tveito. Mathematical models and numerical methods for the forward problem in cardiac electrophysiology. *Comput. Visual. Sci.*, 5(4):215–239, 2003.

[34] C.H. Luo and Y. Rudy. A dynamic model of the cardiac ventricular action potential. I. simulations of ionic currents and concentration changes. *Circulation Research*, 74:1071–1096, 1994.

[35] J. Malmivuo and R. Plonsey. *Bioelectromagnetism - Principles and Applications of Bioelectric and Biomagnetic Fields*. Oxford University Press, 1995.

[36] F.H. Martini, M.J. Timmons, and R.B. Tallitsch. *Human anatomy*. Pearson Benjamin Cummings, sixth edition, 2011.

[37] S. Matsuo, N. Lellouche, M. Wright, M. Bevilacqua, S. Knecht, I. Nault, K.T. Lim, L. Arantes, M.D. O'Neill, P.G. Platonov, =. Carlson, J, F. Sacher, M. Hocini, P. Jaïs, and M. Haïssaguerre. Clinical predictors of termination and clinical outcome of catheter ablation for persistent atrial fibrillation. *Journal of the American College of Cardiology*, 54(9):788–795, 2009.

[38] M.A. McGuire, J.M. de Bakker, J.T. Vermeulen, Moorman A.F., P. Loh, B. Thibault, J.L. Vermeulen, A.E. Becker, and M.J. Janse. Atrioventricular junctional tissue, discrepancy between histological and electrophysiological characteristics. *Circulation*, 94(3):571–577, 1996.

[39] C.C. Mitchell and D.G. Schaeffer. A two-current model for the dynamics of cardiac membrane. *Bulletin Math. Bio.*, 65:767–793, 2003.

[40] J.S. Nagumo, S. Arimoto, and S. Yoshizawa. An active pulse transmission line stimulating nerve axon. *Proceedings of the IEEE*, 50:2061–2071, 1962.

[41] M.P. Nash and P.J. Hunter. Computational mechanics of the heart, from tissue structure to ventricular function. *Journal of elasticity and the physical science of solids*, 61:113–141, 2000.

[42] D. Noble. A modification of the Hodgkin-Huxley equation applicable to Purkinje fiber action and pacemaker potentials. *Journal of Physiology*, 160:317–352, 1962.

[43] J.A. Ormiston, P.M. Shah, C. Tei, and M. Wong. Size and motion of the mitral valve annulus in man. II. Abnormalities in mitral valve prolapse. *Circulation*, 65:113–120, 1982.

[44] M. Potse, B. Dubé, and R. M. Gulrajani. Ecg simulations with realistic human membrane, heart, and torso models. In *Engineering in Medicine and Biology Society, 2003. Proceedings of the 25th Annual International Conference of the IEEE*, volume 1, pages 70–73, 2003.

[45] Andrew J. Pullan, Martin L. Buist, and Leo K. Cheng. *Mathematically modelling the electrical activity of the heart*. World Scientific, 2005.

[46] M. H. Rosner, W. J. Jr. Brady, M. P. Kefer, and M. L. Martin. Electrocardiography in the patient with the Wolff-Parkinson-White syndrome: Diagnostic and initial therapeutic issues. *The American Journal of Emergency Medicine*, 17(7):705–714, 1999.

[47] F.B. Sachse. *Computational Cardiology: Modeling of Anatomy, Electrophysiology and Mechanics*. Springer-Verlag, 2004.

[48] S. Sovilj, R. Magjarević, N.H. Lovell, and S. Dokos. A simplified 3D model of whole heart electrical activity and 12-lead ECG generation. *Computational and Mathematical Methods in Medicine*, 2013.

[49] D.D. Streeter. Gross morphology and fiber geometry of the heart. *Handbook Physiology, The cardiovascular system*, 1:61–112, 1979.

[50] J. Sundnes, G.T. Lines, X. Cai, B.F. Nielsen, K.A. Mardal, and A. Tveito. *Computing the Electrical Activity in the Heart*, volume 1 of *Monographs in Computational Science and Engineering*. Springer-Verlag, 2006.

[51] Joakim Sundnes, Glenn Terie Lines, Bjorn Fredrik Nielsen, Kent-Andre Mardal, and Aslak Tveito. *Computing the Electrical Activity in th Heart*. Springer, 2006.

[52] M.-C. Trudel, B. Dubé, M. Potse, R.M. Gulrajani, and L.J. Leon. Simulation of QRST integral maps with a membrane-based computer heart model employing parallel processing. *IEEE Trans. Biomed. Eng.*, 51(8):1319–1329, 2004.

[53] L. Tung. *A bi-domain model for describing ischemic myocardial d-c potentials*. PhD thesis, Massachusetts Institute of Technology. Dept. of Electrical Engineering and Computer Science, 1978.

[54] F. Wartak. *Electrocardiogram interpretation*. Medical Education Systems, 1975.

[55] G.-X. Yan and C. Antzelevitch. Cellular basis for the normal T wave and the electrocardiographic manifestations of the long-QT syndrome. *Circulation*, 98:1928–1936, 1998.

[56] T.R. Zwink, I.G. Burwash, C.Y. Miyake-Hull, and C.M. Otto. Changes in aortic annulus diameter during the cardiac cycle and its effect on predicting aortic valve prosthesis size. *Journal of diagnostic medical*, 10(5):262–267, 1994.

Part IV

Inverse Problems

Introduction of Part IV

As seen in Chapter 6, we are now able to obtain very realistic simulations of the electrophysiology of the whole heart. The objective of this last part is to adapt the bidomain models – 3D and asymptotic – to **each patient using available medical data**. A data assimilation procedure purports to take into account the various sources of information – **data and model** – in order to obtain an **estimation** of the real trajectory.

Chapter 7 *Introduction to data assimilation*, the first chapter of this part, is a data assimilation introduction. Variational and sequential methods are presented. Sequential methods are used in this work. In this context, a discrepancy compares the data with the solution of the observer model and the objective is then construct a gain operator. Joint state-parameter estimation allows to use two different gains, one for the state and one for the parameters of the model. The first strategy inspired from variational methods consists in constructing the gain following a particular criterion and leads to optimal observers. The second strategy – initially introduced by *Luenberger* – consists in choosing a correction as the simplest possible in order to circumvent the curse of dimensionality of the optimal observers. The presentation of a procedure of a joint state-parameter estimation with a *Luenberger*-like observer for the state space due to the large dimension of this space and a reduced-order Kalman-like filter for the parameters space is given.

Finally, the objective of Chapter 8 is to propose a **Luenberger-like state estimator** for the reaction-diffusion models adapted to available data – here the **depolarization maps** – coupled with a **parameter Kalman-like estimator**. This chapter takes the form of a pre-print article

A Luenberger observer for the reaction-diffusion model assimilating front position data

in collaboration with D. Chapelle and P. Moireau. Just for the sake of simplicity, we use in this chapter the monodomain model, see Remark 3, in Chapter 2 which is a simple reaction-diffusion model. As we want to assimilate the depolarization maps which give the position of the depolarization front at different times, the objective is to compare two propagation fronts. That is why we introduce the eikonal curvature equation which is an **asymptotic simplified model** of the monodomain model. Indeed, the curvature eikonal equation models the propagation of the action potential without the modeling of the action potential itself. This means that this model only focuses on the dynamics of the front propagation. Using methods developed in image processing in order to detect objects in an image, we present an **effective state estimator** for the eikonal curvature equation. Finally, using an **inverted asymptotic analysis**, the state observer is carried over to the initial monodomain model. Simulations validate this state estimator. We couple this state observer with a **Reduced-order Unscented Kalman Filter** (RoUKF) and we can find the model parameters for a specific patient. Theses problematics appear in other application fields where models (reaction-diffusion) and data (front position) are similar, as the fire propagation or the tumor growth. That is why the presentation of the Luenberger state observer is given in a more general case and only simulations are applied in the electrophysiological framework.

CHAPTER 7
Introduction to data assimilation

Abstract This chapter gives a quick overview of data assimilation methods in order to introduce the next chapter. The data assimilation framework began in the context of meteorology and oceanography and now appears in various application fields. Two actors are involved in data assimilation: the model and the data. The model is a set of coupled – non-linear – partial differential equations which is a good approximation of the target solution. However, there are uncertainties on the initial condition and/or on the model parameters, and the objective is to assimilate available data of the target solution. An introduction of the principles of data assimilation is given with a brief presentation of the variational and sequential methods. Optimal observers in the sense of *Bensoussan* – the *Kalman*-like filters – and their main drawback – the computation of a full matrix – are presented. *Luenberger* filters which offer an alternative of the optimal filters are also introduced. The presentation of a joint state and parameters strategy with a *Luenberger* filter for the state and an optimal filter – the Reduced-order Unscented Kalman Filter – for the parameters completes this chapter.

Introduction

Historically, the data assimilation framework began in the context of meteorology and oceanography, see [22, 33, 40] for details and environmental sciences continue to be a very active application field for data assimilation [5, 31]. In the context of life sciences, it has become very popular, in particular in cardiac mechanics, see [8, 36, 10] and also in electrophysiology, [11, 12]. In the field of fire propagation, the works [26, 35] can be directly compared with our work since similar models – a reaction diffusion model or an eikonal equation – are used.

In the data assimilation framework, there are two complementary actors: the model and the data. In most cases, the model is a set of coupled partial differential equations – as for example the bidomain model presented in Chapter 2 – which can be non-linear. The model is considered as a good potential approximation on the real trajectory but relies on uncertain parameters, initial condition and boundary conditions. Even the geometry can contain some errors. The data provide available information on the real trajectory but they are partial (in time and space) and often contain a non negligible part of error due to the measurement process. Then a data assimilation procedure aims at combining the various sources of information in order to construct an estimation of the real system.

The objective of Chapter 8 is to give an effective state observer – a model which contains an additional term which allows to incorporate the data – adapted to a reaction diffusion model. The available medical data that we consider are the depolarization maps, *a.k.a.* isochrones. In a more general case, these data correspond to the knowledge of the front position of a reaction-diffusion model at different times. We also want to couple this effective state observer with a parameter observer. The objective of this chapter is to introduce Chapter 8. It just aims at giving to the reader a brief overview of the numerous methods in data assimilation and refers to the literature for more details. We do not attempt to give a full or a mathematically-detailed presentation of data assimilation.

To that purpose, the first section of this chapter gives an introduction to the principles of data assimilation. In a second section, an overview of sequential and variational methods is presented. In Section 7.3 and 7.4, a presentation of optimal observers – the Kalman-like filters – is given. However the Kalman-like filters leads to the famous curse of dimensionality [2] which makes it numerically intractable for partial differential equations (PDE). An alternative method – initially introduced by *Luenberger* in finite dimension [25] – exists. In the PDE world, these alternative methods have been popularized with the nudging terminology [17, 39]. The literature shows that this type of method can be applied in numerous application fields, as for example conservation laws [23], waves [6, 7], elasticity [30] or fluid-structure interaction [4]. The determination of the nudging coefficients [42, 38, 41] or the restoring of the initial condition [1, 34, 16] have known new improvements. The last section gives a complete joint state and parameters estimation methodology [10]. We present the procedure where a nudging observer is defined on the state space – due to the large dimension of this space – and a reduced-order Kalman-like filter is used for the parameter space. Numerical prediction-correction time schemes are also presented.

7.1 Principles of data assimilation

To start with, let us introduce some standard notations used in data assimilation. Let x be the state of the system, *i.e* the solution of a system of partial differential equations. We denote by $\mathcal{X}$ the state space, and by A the dynamical operator of the system verified by x, *i.e* x satisfies the following model

$$\dot{\mathrm{x}}(t) = \mathrm{A}(\mathrm{x}, \theta, t). \tag{7.1}$$

The dynamical operator A can be non-linear. Here, θ represents the set of unknown model parameters. Contrary to the state space which could be infinite-dimensional or at least of a large dimension (for example the dimension of a finite element or finite difference space), the set of parameters is generally small. Indeed, even if we take into account distributed parameters, their variations can be considered as smooth and a discretization of the domain in large regions is sufficient. Prescribing $\mathrm{x}(0)$ and θ, we assume that the problem is well-posed. In practice, we just have an *a priori* on the initial condition $\mathrm{x}(0)$ and on the values of the parameters represented by θ. That is why we decompose them into two parts

$$\mathrm{x}(0) = \mathrm{x}_\bullet + \zeta^{\mathrm{x}} \text{ and } \theta = \theta_\bullet + \zeta^\theta,$$

where $\mathrm{x}_\bullet$ and $\theta_\bullet$ are the known parts (or *a priori*) and ζ^{x} and ζ^θ the uncertainties. The complete problem then reads

$$\left\{ \begin{array}{rcl} \dot{\mathrm{x}}(t) & = & \mathrm{A}(\mathrm{x}, \theta, t) \\ \mathrm{x}(0) & = & \mathrm{x}_\bullet + \zeta^{\mathrm{x}} \\ \theta(0) & = & \theta_\bullet + \zeta^\theta. \end{array} \right. \tag{7.2}$$

We can rewrite this problem considering together x and θ with

$$\left\{ \begin{array}{rcl} \dot{{}^\flat\mathrm{x}}(t) & = & {}^\flat\mathrm{A}({}^\flat\mathrm{x}, t) \\ {}^\flat\mathrm{x}(0) & = & \begin{pmatrix} \mathrm{x}_\bullet \\ \theta_\bullet \end{pmatrix} + \zeta, \end{array} \right. \tag{7.3}$$

where

$${}^\flat\mathrm{x} = \begin{pmatrix} \mathrm{x} \\ \theta \end{pmatrix}, \zeta = \begin{pmatrix} \zeta^{\mathrm{x}} \\ \zeta^\theta \end{pmatrix}$$

and

$${}^\flat\mathrm{A} = \begin{pmatrix} \mathrm{A} & 0 \\ 0 & 0 \end{pmatrix}$$

is the augmented model.

We assume that we have at our disposal some observations of a real trajectory denoted by $\mathrm{x}^\diamond$ and we define by z the observations of $\mathrm{x}^\diamond$. In many cases, an observation operator denoted by H represents the data generation procedure from $\mathrm{x}^\diamond$, *i.e.*

$$\mathrm{z}(t) = \mathrm{H}(\mathrm{x}^\diamond(t), t) + \chi(t), \tag{7.4}$$

where χ represents the error inherent to the measurement process. The parameters are generally not observed and we expect to retrieve them with the data assimilation procedure.

This observation operator allows to map the state space $\mathcal{X}$ to the observation space denoted by $\mathcal{Z}$. A simple case is when the observation corresponds to the restriction of a sub-domain of the state space $\mathcal{X}$. Let $\theta^\diamond$ the parameters of the real trajectory and we define

$$\mathrm{{}^\flat x}^\diamond = \begin{pmatrix} \mathrm{x}^\diamond \\ \theta^\diamond \end{pmatrix}$$

We denote by ${}^\flat\mathrm{H}$, the augmented operator defined such that

$$\mathrm{z}(t) = {}^\flat\mathrm{H}({}^\flat\mathrm{x}^\diamond(t), t) + \chi(t),$$

with a null part on the parameters as the parameters are not observed. In some cases, we just have an implicit relation [30] between the observations z and $\mathrm{x}^\diamond$ of the following form

$$\mathrm{D}(\mathrm{z}(t), \mathrm{x}^\diamond(t)) = 0,$$

where D is the discrepancy instead of (7.4). We denote by ${}^\flat\mathrm{D}$ the augmented discrepancy.

7.2 Overview of variational and sequential methods

7.2.1 From variational methods ...

We start with the presentation of variational methods. A good strategy in data assimilation consists in minimizing a criterion. This criterion is composed of two parts. Variational methods seek to minimize a functional that balances the confidence one has in the *a priori* values $\mathrm{x}_\bullet$ and $\theta_\bullet$ and the confidence in the data with the comparison of the observations z with the observations $\mathrm{H}(\hat{\mathrm{x}})$ derived from the solution $\hat{\mathrm{x}}$ of the observer. The minimization of the criterion, typically least-square based, reads

$$\mathcal{J}_T(\zeta) = \frac{1}{2}\|\zeta\|^2_{\mathcal{X}_{ic}} + \frac{1}{2}\int_0^T \|\mathrm{z}(t) - {}^\flat\mathrm{H}({}^\flat\mathrm{x}^\zeta_T(t), t)\|^2_{\mathcal{Z}}\mathrm{d}t, \tag{7.5}$$

where $\mathcal{X}_{ic}$ is the space of initial conditions. We denote by M a metric on the observation space $\mathcal{Z}$ and by ${}^\flat\mathrm{P}_0$, the inverse of a metric of $\mathcal{X}_{ic}$ which can be interpreted in a stochastic context as an initial uncertainty covariance. Both operators M and ${}^\flat\mathrm{P}_0$ are self-adjoint. Estimating ζ by minimizing $\mathcal{J}_T$ under the constraint of the dynamics (7.1) is known as *4D-var* strategy [5, 8]. Obviously, in more general cases, (7.5) reads

$$\mathcal{J}_T(\zeta) = \frac{1}{2}\|\zeta\|^2_{\mathcal{X}_{ic}} + \frac{1}{2}\int_0^T \|{}^\flat\mathrm{D}(\mathrm{z}(t), {}^\flat\mathrm{x}^\zeta_T(t))\|^2_{\mathcal{Z}}\mathrm{d}t. \tag{7.6}$$

Linear case In what follows, we assume for the sake of simplicity that the dynamical operator ${}^\flat\mathrm{A}$ and the observation operator ${}^\flat\mathrm{H}$ are linear. Furthermore we suppose, for simplicity, that the state space $\mathcal{X}$ and the observation space $\mathcal{Z}$ are finite dimensional spaces. The minimization of $\mathcal{J}_T$ is performed with respect to the unknown ζ. Using a minimizing sequence (existence) and the fact that $\mathcal{J}_T$ is convex (uniqueness), we can show that the minimization

of $\mathcal{J}_T$ admits a unique solution. We denote by ζ_T^{min} this optimal value. The dependence in the time parameter T is important. The condition which characterizes the optimal value is

$$\mathrm{d}\mathcal{J}_T(\zeta)|_{\zeta_T^{min}} = 0.$$

We differentiate the functional and we obtain, $\forall \bar{\zeta}$,

$$\mathrm{d}\mathcal{J}_T(\zeta)\cdot\bar{\zeta} = \zeta^{\intercal\,\flat}\mathrm{P}_0^{-1}\bar{\zeta} - \int_0^T \left(\mathrm{z} - {}^{\flat}\mathrm{H}({}^{\flat}\mathrm{x}_T^{\zeta})\right)^{\intercal}\mathrm{M}\,{}^{\flat}\mathrm{H}\,\mathrm{d}_{\zeta}{}^{\flat}\mathrm{x}_T^{\zeta}\cdot\bar{\zeta}.$$

Remarking that

$$\partial_t(\mathrm{d}_{\zeta}{}^{\flat}\mathrm{x}_T^{\zeta}) = {}^{\flat}\mathrm{A}\mathrm{d}_{\zeta}{}^{\flat}\mathrm{x}_T^{\zeta} \text{ and } \mathrm{d}_{\zeta}({}^{\flat}\mathrm{x}_T^{\zeta})_{|t=0} = \mathbb{1},$$

and defining the so-called adjoint variable ${}^{\flat}\mathrm{p}_T$ as the solution of the adjoint problem

$$\left\{\begin{array}{rcl} {}^{\flat}\dot{\mathrm{p}}_T(t) + {}^{\flat}\mathrm{A}^{\intercal}\,{}^{\flat}\mathrm{p}_T(t) &=& -{}^{\flat}\mathrm{H}^{\intercal}\mathrm{M}(\mathrm{z} - {}^{\flat}\mathrm{H}{}^{\flat}\mathrm{x}_T^{\zeta}), \\ {}^{\flat}\mathrm{p}_T(T) &=& 0, \end{array}\right.$$

the gradient of the functional $\mathcal{J}_T$ becomes

$$\mathrm{d}\mathcal{J}_T(\zeta)\cdot\bar{\zeta} = \zeta^{\intercal\,\flat}\mathrm{P}_0^{-1}\bar{\zeta} - {}^{\flat}\mathrm{p}_T(0)^{\intercal}\bar{\zeta}.$$

The optimal value ζ_T^{min} is then defined by

$$\zeta_T^{min} = {}^{\flat}\mathrm{P}_0\,{}^{\flat}\mathrm{p}_T(0).$$

To conclude, the variational method consists in solving the two ends problem, find ${}^{\flat}\mathrm{x}_T, {}^{\flat}\mathrm{p}_T$ such that

$$\left\{\begin{array}{rcl} \widehat{{}^{\flat}\dot{\mathrm{x}}_T(t)} &=& {}^{\flat}\mathrm{A}\,{}^{\flat}\mathrm{x}_T(t) \\ {}^{\flat}\mathrm{x}_T(0) &=& {}^{\flat}\mathrm{x}_{\bullet} + {}^{\flat}\mathrm{P}_0\,{}^{\flat}\mathrm{p}_T(0) \\ \widehat{{}^{\flat}\dot{\mathrm{p}}_T(t)} &=& -{}^{\flat}\mathrm{A}^{\intercal}\,{}^{\flat}\mathrm{p}_T(t) - {}^{\flat}\mathrm{H}^{\intercal}\mathrm{M}(\mathrm{z}(t) - {}^{\flat}\mathrm{H}\,{}^{\flat}\mathrm{x}_T(t)) \\ {}^{\flat}\mathrm{p}_T(T) &=& 0 \end{array}\right. \tag{7.7}$$

The variational method can be used for every model and that is why this method is very interesting. However, as seen previously, the computation of the adjoint problem is necessary in order to minimize of $\mathcal{J}_T$ under constraint. This may lead to prohibitive computational times. Indeed, the gradient based descent algorithm used for the minimization requires numerous iterations of the direct and adjoint models.

7.2.2 ... To sequential methods

Sequential methods consist in constructing an estimator $(\hat{\mathrm{x}}, \hat{\theta})$ by coupling the model and the observations z. As previously presented, if the observation space $\mathcal{Z}$ is equipped with standard operations, we can compare z with the value of $\mathrm{H}(\hat{\mathrm{x}}(t), t)$ by

$$\mathrm{z} - \mathrm{H}(\hat{\mathrm{x}}(t), t)$$

and in some cases, we just have an implicit relation defined by

$$\mathrm{D}(\mathrm{z}, \hat{\mathrm{x}}(t)) = 0,$$

where D is called the discrepancy. Then, in sequential methods, the dynamics of system is modified. This new system is named the observer and reads

$$\left\{ \begin{array}{rcl} {}^{\flat}\dot{\hat{\mathrm{x}}}(t) & = & {}^{\flat}\mathrm{A}({}^{\flat}\hat{\mathrm{x}}, \theta, t) + {}^{\flat}\mathrm{G}\Big({}^{\flat}\mathrm{D}(\mathrm{z}, {}^{\flat}\hat{\mathrm{x}})\Big) \\ {}^{\flat}\hat{\mathrm{x}}(0) & = & \begin{pmatrix} \hat{\mathrm{x}}_\bullet \\ \theta_\bullet \end{pmatrix}, \end{array} \right. \tag{7.8}$$

where ${}^{\flat}\mathrm{G}$ is the gain operator, also called the filter. The goal of a sequential method is thus to find a discrepancy and a gain ${}^{\flat}\mathrm{G}$ such that

$$\lim_{t \to +\infty} {}^{\flat}\hat{\mathrm{x}} = {}^{\flat}\mathrm{x},$$

where ${}^{\flat}\mathrm{x}$ is the solution of (7.3).

Remark 4. *We can use two different gains – one for the state space and one for the parameters space – using the joint state and parameter strategy introduced in [10].*

Two different strategies that we will present in Sections 7.3 and 7.4 can be used in order to find an effective gain ${}^{\flat}\mathrm{G}$. The first method, called the *optimal observer*, consists in constructing the gain from the optimal criterion by defining

$${}^{\flat}\hat{\mathrm{x}}(t) = {}^{\flat}\mathrm{x}^{\zeta_t^{min}}(t),$$

where

$$\zeta_t^{min} = \mathrm{Argmin}_\zeta \mathcal{J}_t(\zeta) = \mathrm{Argmin}_\zeta \left(\frac{1}{2} \|\zeta\|_{\mathcal{X}_{ic}} + \frac{1}{2} \int_0^t \|\mathrm{z} - {}^{\flat}\mathrm{H}({}^{\flat}\mathrm{x}_T^\zeta)\|_{\mathcal{Z}} \right)$$

and ${}^{\flat}\mathrm{x}^{\zeta_t^{min}}$ the solution of

$$\left\{ \begin{array}{rcl} \widehat{{}^{\flat}\mathrm{x}^{\zeta_t^{min}}}(t) & = & {}^{\flat}\mathrm{A}({}^{\flat}\mathrm{x}^{\zeta_t^{min}}, t) \\ {}^{\flat}\mathrm{x}^{\zeta_t^{min}}(0) & = & \begin{pmatrix} \mathrm{x}_\bullet \\ \theta_\bullet \end{pmatrix} + \zeta_t^{min}. \end{array} \right. \tag{7.9}$$

In linear cases, it leads to the well-known *Kalman-Bucy filter*, [20, 3] and in non-linear cases, it derives from a *Hamilton-Jacobi-Bellman solution* [15]. Due to the complexity of this equation, many approximated solutions exist as for example the *Extended Kalman filter* [37], the *Unscented Kalman filter*, [19] and the *Ensemble Kalman filter*, [14]. We can remark that the great advantage of the optimal filter is that it can be defined for every model and every observation operator. The biggest drawback is that without an uncertainty space reduction, the filter is very costly and often impossible to compute. Concerning the reduction of the uncertainty space, we will present later in this document a strategy called *reduced-order optimal filters* that allows to circumvent the computation of the full optimal gain. Besides, a second strategy exists in order to circumvent the curse of dimensionality. This method is not

relying on the optimal criterion minimization. Initially introduced by *Luenberger* [24, 25], the idea is to introduce a correction such that the error between the observed trajectory and the observer system tends to zero. This correction is chosen as the simplest possible to avoid prohibitive additional computational times with respect to the original computational times necessary for the resolution of the direct model.

A joint state and parameter estimation can be applied. We recall that

$$
{}^{\flat}\mathrm{x} = \begin{pmatrix} \mathrm{x} \\ \theta \end{pmatrix},
$$

where x is the state and θ are the parameters. We can use two different gains in (7.8), one for the state and one for the parameters following the strategy presented in [29] and the observer model becomes

$$
\left\{ \begin{array}{rcl} \dot{\hat{\mathrm{x}}}(t) & = & \mathrm{A}(\hat{\mathrm{x}}, \theta, t) + \mathrm{G}^{\mathrm{x}}\big(\mathrm{D}(\mathrm{z}, \hat{\mathrm{x}})\big) \\ \dot{\hat{\theta}}(t) & = & \mathrm{G}^{\theta}\big(\mathrm{D}(\mathrm{z}, \hat{\mathrm{x}})\big) \\ \hat{\mathrm{x}}(0) & = & \hat{\mathrm{x}}_{\bullet}, \\ \theta(0) & = & \theta_{\bullet}. \end{array} \right. \tag{7.10}
$$

The strategy consists in taking an optimal-based filter for the parameters space due to the relatively small dimension of the parameters space and taking a physically-based *Luenberger* observer on the large dimensional state. The coupled observer is proved to converge under assumptions in linear cases and can be extended to non-linear cases.

7.3 Optimal observers

We are interested in sequential methods. As previously explained, the first strategy – corresponding to optimal observers – consists in constructing the gain operator such that

$$
{}^{\flat}\hat{\mathrm{x}}(t) = {}^{\flat}\mathrm{x}^{\zeta_t^{min}}(t).
$$

7.3.1 Linear case: The Kalman-Bucy filter

In linear cases, it leads to the well-known *Kalman-Bucy filter* – also called the *Kalman filter* – initially introduced in [20]. We use the same context as for the variational methods, i.e the dynamical operator ${}^{\flat}\mathrm{A}$ and the observation operator ${}^{\flat}\mathrm{H}$ are linear and the state space $\mathcal{X}$ and the observation space $\mathcal{Z}$ are finite dimensional spaces. The Kalman-Bucy method consists in reformulating the minimization problem (7.7) into a sequential formulation. In [3], the following theorem is proved

Theorem 1. *There is a relation between both solutions ${}^{\flat}x_T$ and ${}^{\flat}p_T$ of the minimization problem (7.7) of the following form*

$$
{}^{\flat}\mathrm{x}_T(t) = {}^{\flat}\hat{\mathrm{x}}(t) + {}^{\flat}\mathrm{P}(t){}^{\flat}\mathrm{p}_T(t),
$$

with ${}^{\flat}\hat{x}$ and ${}^{\flat}P$ independent of T.

Problem (7.7) becomes

$$\left\{\begin{array}{rcl} {}^{\flat}\dot{\hat{\mathrm{x}}}(t) + {}^{\flat}\dot{\mathrm{P}}(t){}^{\flat}\mathrm{p}_T(t) + {}^{\flat}\mathrm{P}(t){}^{\flat}\dot{\mathrm{p}}_T(t) & = & {}^{\flat}\mathrm{A}{}^{\flat}\hat{\mathrm{x}}(t) + {}^{\flat}\mathrm{A}{}^{\flat}\mathrm{P}(t){}^{\flat}\mathrm{p}_T(t) \\ {}^{\flat}\dot{\mathrm{p}}_T(t) & = & -{}^{\flat}\mathrm{A}^{\intercal}{}^{\flat}\mathrm{p}_T(t) - {}^{\flat}\mathrm{H}^{\intercal}\mathrm{M}\mathrm{z}(t) + {}^{\flat}\mathrm{H}^{\intercal}\mathrm{M}{}^{\flat}\mathrm{H}{}^{\flat}\hat{\mathrm{x}}(t) + {}^{\flat}\mathrm{H}^{\intercal}\mathrm{M}{}^{\flat}\mathrm{H}{}^{\flat}\mathrm{P}(t){}^{\flat}\mathrm{p}_T(t), \end{array}\right.$$

and this implies

$$\begin{aligned} {}^{\flat}\dot{\hat{\mathrm{x}}}(t) &= -{}^{\flat}\mathrm{P}(t)\big(-{}^{\flat}\mathrm{A}^{\intercal}{}^{\flat}\mathrm{p}_T(t) - {}^{\flat}\mathrm{H}^{\intercal}\mathrm{M}\mathrm{z}(t) + {}^{\flat}\mathrm{H}^{\intercal}\mathrm{M}{}^{\flat}\mathrm{H}{}^{\flat}\hat{\mathrm{x}}(t) + {}^{\flat}\mathrm{H}^{\intercal}\mathrm{M}{}^{\flat}\mathrm{H}{}^{\flat}\mathrm{P}(t){}^{\flat}\mathrm{p}_T(t)\big) && (7.11)\\ &\quad + {}^{\flat}\mathrm{A}{}^{\flat}\hat{\mathrm{x}}(t) - {}^{\flat}\dot{\mathrm{P}}(t){}^{\flat}\mathrm{p}_T(t) + {}^{\flat}\mathrm{A}{}^{\flat}\mathrm{P}(t){}^{\flat}\mathrm{p}_T(t) && (7.12)\\ &= \big(-{}^{\flat}\dot{\mathrm{P}}(t) + {}^{\flat}\mathrm{A}{}^{\flat}\mathrm{P}(t) + {}^{\flat}\mathrm{P}(t){}^{\flat}\mathrm{A}^{\intercal} - {}^{\flat}\mathrm{P}(t){}^{\flat}\mathrm{H}^{\intercal}\mathrm{M}{}^{\flat}\mathrm{H}{}^{\flat}\mathrm{P}(t)\big){}^{\flat}\mathrm{p}_T(t) && (7.13)\\ &\quad + {}^{\flat}\mathrm{A}{}^{\flat}\hat{\mathrm{x}}(t) - {}^{\flat}\mathrm{P}(t){}^{\flat}\mathrm{H}^{\intercal}\mathrm{M}({}^{\flat}\mathrm{H}{}^{\flat}\hat{\mathrm{x}}(t) - \mathrm{z}(t)). && (7.14) \end{aligned}$$

If ${}^{\flat}\mathrm{P}$ verifies the following Riccati equation

$$\left\{\begin{array}{rcl} {}^{\flat}\dot{\mathrm{P}}(t) & = & {}^{\flat}\mathrm{A}{}^{\flat}\mathrm{P}(t) + {}^{\flat}\mathrm{P}(t){}^{\flat}\mathrm{A}^{\intercal} - {}^{\flat}\mathrm{P}{}^{\flat}\mathrm{H}^{\intercal}\mathrm{M}{}^{\flat}\mathrm{H}{}^{\flat}\mathrm{P}(t), \\ {}^{\flat}\mathrm{P}(0) & = & {}^{\flat}\mathrm{P}_0, \end{array}\right. \tag{7.15}$$

the coupled model (7.11) becomes

$$\left\{\begin{array}{rcl} {}^{\flat}\dot{\hat{\mathrm{x}}}(t) & = & {}^{\flat}\mathrm{A}{}^{\flat}\hat{\mathrm{x}}(t) + {}^{\flat}\mathrm{P}(t){}^{\flat}\mathrm{H}^{\intercal}\mathrm{M}(\mathrm{z}(t) - {}^{\flat}\mathrm{H}{}^{\flat}\hat{\mathrm{x}}(t)), \\ {}^{\flat}\hat{\mathrm{x}}(0) & = & {}^{\flat}\mathrm{x}_{\bullet}(= {}^{\flat}\mathrm{x}_T(0) - {}^{\flat}\mathrm{P}(0){}^{\flat}\mathrm{p}_T(0)). \end{array}\right.$$

We can deduce the expression of the *Kalman* gain ${}^{\flat}\mathrm{G} = {}^{\flat}\mathrm{P}{}^{\flat}\mathrm{H}^{\intercal}\mathrm{M}$. The matrix ${}^{\flat}\mathrm{P}$ following the differential Riccati equation can be interpreted, from a probabilistic point of view, as the covariance matrix of the estimation error. A prediction-correction scheme called the Kalman-Bucy algorithm [3, 37] is used to discretize in time the Kalman-Bucy filter.

7.3.2 Non linear case

7.3.2.1 Approximate Kalman filters

The original Kalman filter is limited to a linear assumption. However, more complex systems can be non-linear. Different approximations are developed for these non-linear cases.

Extended Kalman filter The most classical choice of gain in non-linear configuration is given by the *Extended Kalman filter* (EKF) [37],

$$ {}^{\flat}\mathrm{G} = {}^{\flat}\mathrm{P}({\rm d}{}^{\flat}\mathrm{H})^{\intercal}\mathrm{M}, $$

where ${}^{\flat}\mathrm{P}$ verifies the following augmented Riccati equation

$$ {}^{\flat}\dot{\mathrm{P}} - {\rm d}{}^{\flat}\mathrm{A}{}^{\flat}\mathrm{P} - {}^{\flat}\mathrm{P}({\rm d}{}^{\flat}\mathrm{A})^{\intercal} + {}^{\flat}\mathrm{P}({\rm d}{}^{\flat}\mathrm{H})^{\intercal}\mathrm{M}({\rm d}{}^{\flat}\mathrm{H}){}^{\flat}\mathrm{P} = 0. $$

We recall that the operator ${}^\flat\mathrm{P}$ solution of the previous equation is called *covariance*. As the reader can see, the idea is to replace the non-linear mapping by their tangent operators. The convergence of the EKF observer is verified for small errors by studying the linearized dynamics of the estimation error ${}^\flat\tilde{\mathrm{x}} = {}^\flat\mathrm{x} - {}^\flat\hat{\mathrm{x}}$, see for example [21]. Note that if we decompose the operator ${}^\flat\mathrm{P}$ on the state and on the parameter,

$$
{}^\flat\mathrm{P} = \begin{pmatrix} \mathrm{P}^{\mathrm{xx}} & \mathrm{P}^{\mathrm{x}\theta} \\ (\mathrm{P}^{\mathrm{x}\theta})^\intercal & \mathrm{P}^{\theta\theta} \end{pmatrix},
$$

we have

$$
\left\{
\begin{array}{rcl}
\dot{\mathrm{P}^{\mathrm{xx}}} & = & \mathrm{d_x A P^{xx}} + \mathrm{P^{xx}(d_x A)^\intercal} + \mathrm{d_\theta A (P^{x\theta})^\intercal} + \mathrm{P^{x\theta}(d_\theta A)^\intercal} - \mathrm{P^{xx}(d_x H)^\intercal M d_x H P^{xx}} \\
\dot{\mathrm{P}^{\theta\theta}} & = & -\mathrm{(P^{x\theta})^\intercal (d_x H)^\intercal M d_x H P^{x\theta}} \\
\dot{\mathrm{P}^{\mathrm{x}\theta}} & = & \mathrm{d_x A P^{x\theta}} + \mathrm{d_\theta A P^{\theta\theta}} - \mathrm{P^{xx}(d_x H)^\intercal M d_x H P^{x\theta}} \\
\mathrm{P^{xx}}(0) & = & \mathrm{P_x},\ \mathrm{P}^{\theta\theta}(0) = \mathrm{P}_\theta \text{ and } \mathrm{P}^{\mathrm{x}\theta}(0) = 0.
\end{array}
\right.
$$

where the two gains are given by

$$
\mathrm{G^x} = \mathrm{P^{xx}(d_x H)^\intercal M}
$$

and

$$
\mathrm{G}^\theta = \mathrm{(P^{x\theta})^\intercal (d_x H)^\intercal M}.
$$

The time-discretization of the optimal observers is based on the fact that the optimality should be preserved also at the discrete level. We consider a stable and consistent discretization of the original model

$$
\left\{
\begin{array}{rcl}
\dot{\mathrm{x}}(t) & = & \mathrm{A}(\mathrm{x}, \theta, t) \\
\mathrm{x}(0) & = & \mathrm{x}_\bullet + \zeta^{\mathrm{x}} \\
\theta(0) & = & \theta_\bullet + \zeta^\theta
\end{array}
\right. \tag{7.16}
$$

denoted by

$$
\left\{
\begin{array}{rcl}
\mathrm{x}_{n+1} & = & \mathrm{A}_{n+1|n}(\mathrm{x}_n, \theta_n, t) \\
\mathrm{x}_0 & = & \mathrm{x}_\bullet + \zeta^{\mathrm{x}} \\
\theta_0 & = & \theta_\bullet + \zeta^\theta.
\end{array}
\right. \tag{7.17}
$$

We define a discrete-time functional

$$
\mathcal{J}_T^N(\zeta) = \frac{1}{2}\|\zeta\|^2_{\mathcal{X}_{ic}} + \frac{1}{2}\sum_{k=0}^{N} \|\mathrm{z}_k(t) - \mathrm{H}(\mathrm{x}_k)\|^2_{M_k}, \tag{7.18}
$$

with $M_{n+1} = \Delta t \mathrm{M}$ a discretization of M since a fixed time-step Δt is considered. The discrete-time functional $\mathcal{J}_T^N$ is consistent with respect to the continuous-time functional $\mathcal{J}_T$ with this choice of discrete observation norm M_n. The tangent of $\mathrm{A}_{n+1|n}$ is defined by

$$
\mathrm{d}\,{}^\flat\mathrm{A}_{n+1|n} = \begin{pmatrix} \mathrm{d_x A}_{n+1|n} & \mathrm{d}_\theta \mathrm{A}_{n+1|n} \\ 0 & \mathbb{1} \end{pmatrix}.
$$

As said previously, the time-discretization of an observer ${}^\flat\hat{\mathrm{x}}$ is a prediction-correction scheme.

Prediction-Correction Scheme

1. Prediction

$$\left\{\begin{array}{rcl}
\hat{\mathrm{x}}_{n+1}^- &=& \mathrm{A}_{n+1|n}(\hat{\mathrm{x}}_n^+, \hat{\theta}_n^+) \\
\hat{\theta}_{n+1}^- &=& \hat{\theta}_n^+ \\
\mathrm{P}_{n+1}^{\mathrm{xx}-} &=& \mathrm{d}_{\mathrm{x}}\mathrm{A}_{n+1|n}\mathrm{P}_n^{\mathrm{xx}+}(\mathrm{d}_{\mathrm{x}}\mathrm{A}_{n+1|n})^\intercal + \mathrm{d}_{\mathrm{x}}\mathrm{A}_{n+1|n}\mathrm{P}_n^{\mathrm{x}\theta+}(\mathrm{d}_\theta\mathrm{A}_{n+1|n})^\intercal \\
&& \qquad + \mathrm{d}_\theta\mathrm{A}_{n+1|n}(\mathrm{P}_n^{\mathrm{x}\theta+})^\intercal(\mathrm{d}_{\mathrm{x}}\mathrm{A}_{n+1|n})^\intercal + \mathrm{d}_\theta\mathrm{A}_{n+1|n}\mathrm{P}_n^{\theta\theta+}(\mathrm{d}_\theta\mathrm{A}_{n+1|n})^\intercal \\
\mathrm{P}_{n+1}^{\theta\theta-} &=& \mathrm{P}_n^{\theta\theta+} \\
\mathrm{P}_{n+1}^{\mathrm{x}\theta-} &=& \mathrm{d}_{\mathrm{x}}\mathrm{A}_{n+1|n}\mathrm{P}_n^{\mathrm{x}\theta+} + \mathrm{d}_\theta\mathrm{A}_{n+1|n}\mathrm{P}_n^{\theta\theta+}
\end{array}\right.$$

2. Correction

$$\left\{\begin{array}{rcl}
{}^\flat\mathrm{P}_{n+1}^+ &=& {}^\flat\mathrm{P}_{n+1}^- - {}^\flat\mathrm{P}_{n+1}^+(\mathrm{d}{}^\flat\mathrm{H}_{n+1})^\intercal M_{n+1}\mathrm{d}{}^\flat\mathrm{H}_{n+1}{}^\flat\mathrm{P}_{n+1}^- \\
\multicolumn{3}{l}{\quad \text{Explicit scheme: } {}^\flat\mathrm{P}_{n+1}^+ = \left((\mathrm{d}{}^\flat\mathrm{H}_{n+1})^\intercal M_{n+1}\mathrm{d}{}^\flat\mathrm{H}_{n+1} + ({}^\flat\mathrm{P}_{n+1}^-)^{-1}\right)^{-1}} \\
\mathrm{G}_{n+1}^{\mathrm{x}} &=& \mathrm{P}_{n+1}^{\mathrm{xx}+}(\mathrm{d}_{\mathrm{x}}\mathrm{H}_{n+1}))^\intercal M_{n+1} \\
\mathrm{G}_{n+1}^{\hat{\theta}} &=& \mathrm{P}_{n+1}^{\mathrm{x}\theta+}(\mathrm{d}_{\mathrm{x}}\mathrm{H}_{n+1}))^\intercal M_{n+1} \\
\hat{\mathrm{x}}_{n+1}^+ &=& \hat{\mathrm{x}}_{n+1}^- + \mathrm{G}_{n+1}^{\mathrm{x}}(\mathrm{z}_{n+1} - \mathrm{H}_{n+1}\left(\hat{\mathrm{x}}_{n+1}^-\right)) \\
\hat{\theta}_{n+1}^+ &=& \hat{\theta}_{n+1}^- + \mathrm{G}_{n+1}^{\hat{\theta}}(\mathrm{z}_{n+1} - \mathrm{H}_{n+1}\left(\hat{\mathrm{x}}_{n+1}^-\right))
\end{array}\right.$$

Unscented Kalman filter The Unscented Kalman Filter (UKF) [19] is another alternative for state and parameters estimation in non-linear cases. This more recent method was initially introduced to circumvent the main computational disadvantage of the EKF method – namely the computation of the tangent operators of the model and observation operators. From a stochastic point of view, ${}^\flat\mathrm{P}_n^-$ and ${}^\flat\mathrm{P}_n^+$ can be seen as *a priori* and *a posteriori* covariances and ${}^\flat\mathrm{x}_n^-$ and ${}^\flat\mathrm{x}_n^+$ can be interpreted as *a priori* and *a posteriori* means. The UKF is a discrete-time estimator based on sampling particles – called *sigma points* – whose role is to replace the tangent computations. We introduce the so-called p *unitary sampling points* $\mathrm{sp}^{[i]}$ and the p weights α^i as

$$\sum_{i=1}^{p}\alpha^i\mathrm{sp}^{[i]} = 0 \text{ and } \sum_{i=1}^{p}\alpha^i\mathrm{sp}^{[i]}\cdot(\mathrm{sp}^{[i]})^\intercal = \mathbb{1}.$$

The main idea of UKF is in the choice of these sampling particles. Indeed, the particular choice of sampling points can be justified from a stochastic point of view and is explained for example in [18]. The sigma points are generated around the estimated values based on the covariance estimation. This choice can also be interpreted in a deterministic framework. The method consists in replacing the tangent operators by finite difference approximations.

We then need to define the points which allow to compute a correct finite difference at a given point. The choice can be made such that the distance between the points – that we need to determine – and the given point corresponds to the order of magnitude of the uncertainties [28]. We also have a sampling-prediction-correction scheme with a simple numerical implementation.

Sampling-Prediction-Correction Scheme

1. Sampling

$$\begin{pmatrix} \hat{\mathrm{x}}_n^{[i]+} \\ \hat{\theta}_n^{[i]+} \end{pmatrix} = \begin{pmatrix} \hat{\mathrm{x}}_n^{+} \\ \hat{\theta}_n^{+} \end{pmatrix} + \sqrt{{}^{\flat}\mathrm{P}_n^{+}}\,\mathrm{sp}^{[i]}$$

2. Prediction

$$\left\{ \begin{array}{rcl} \hat{\mathrm{x}}_{n+1}^{-} & = & \sum_{i=1}^{p} \alpha^i \mathrm{A}_{n|n+1}(\hat{\mathrm{x}}_n^{[i]+}, \hat{\theta}_n^{[i]+}) \text{ and } \hat{\mathrm{x}}_{n+1}^{[i]-} = \mathrm{A}_{n|n+1}(\hat{\mathrm{x}}_n^{[i]+}, \hat{\theta}_n^{[i]+}) \\ \hat{\theta}_{n+1}^{-} & = & \sum_{i=1}^{p} \alpha^i \hat{\theta}_n^{[i]+} = \hat{\theta}_n^{+} \text{ and } \hat{\theta}_{n+1}^{[i]-} = \hat{\theta}_n^{[i]+} \\ {}^{\flat}\mathrm{P}_{n+1}^{-} & = & \sum_{i=1}^{p} \alpha^i \begin{pmatrix} \hat{\mathrm{x}}_{n+1}^{[i]-} - \hat{\mathrm{x}}_{n+1}^{-} \\ \hat{\theta}_n^{[i]+} - \hat{\theta}_{n+1}^{-} \end{pmatrix} \begin{pmatrix} \hat{\mathrm{x}}_{n+1}^{[i]-} - \hat{\mathrm{x}}_{n+1}^{-} \\ \hat{\theta}_n^{[i]+} - \hat{\theta}_{n+1}^{-} \end{pmatrix}^{\intercal} \end{array} \right.$$

3. Observations (of the observer model)

$$\mathrm{z}_{n+1}^{[i]-} = \mathrm{H}_{n+1}(\hat{\mathrm{x}}_{n+1}^{[i]-}) \text{ and } \mathrm{z}_{n+1}^{-} = \sum_{i=1}^{p} \alpha^i \mathrm{z}_{n+1}^{[i]-}$$

4. Correction

$$\left\{ \begin{array}{rcl} \mathrm{P}^{\flat\mathbf{XZ}} & = & \sum_{i=1}^{p} \alpha^i \begin{pmatrix} \hat{\mathrm{x}}_{n+1}^{[i]-} - \hat{\mathrm{x}}_{n+1}^{-} \\ \hat{\theta}_{n+1}^{[i]-} - \hat{\theta}_{n+1}^{-} \end{pmatrix} (\mathrm{z}_{n+1}^{[i]-} - \mathrm{z}_{n+1}^{-})^{\intercal} = \begin{pmatrix} \mathrm{P}_{n+1}^{\mathbf{XZ}} & \times \\ \times & \mathrm{P}_{n+1}^{\theta\mathbf{Z}} \end{pmatrix} \\ \mathrm{P}^{\mathbf{ZZ}} & = & \sum_{i=1}^{p} \alpha^i (\mathrm{z}_{n+1}^{[i]-} - \mathrm{z}_{n+1}^{-})(\mathrm{z}_{n+1}^{[i]-} - \mathrm{z}_{n+1}^{-})^{\intercal} \\ {}^{\flat}\mathrm{G}_{n+1}^{\mathbf{X}} & = & \mathrm{P}_{n+1}^{\mathbf{XZ}} (\mathrm{P}_{n+1}^{\mathbf{ZZ}})^{-1} \\ {}^{\flat}\mathrm{G}_{n+1}^{\theta} & = & \mathrm{P}_{n+1}^{\theta\mathbf{Z}} (\mathrm{P}_{n+1}^{\mathbf{ZZ}})^{-1} \\ \hat{\mathrm{x}}_{n+1}^{+} & = & \hat{\mathrm{x}}_{n+1}^{-} + {}^{\flat}\mathrm{G}_{n+1}^{\mathbf{X}} (\mathrm{z}_{n+1} - \mathrm{z}_{n+1}^{-}) \\ \hat{\theta}_{n+1}^{+} & = & \hat{\theta}_{n+1}^{-} + {}^{\flat}\mathrm{G}_{n+1}^{\theta} (\mathrm{z}_{n+1} - \mathrm{z}_{n+1}^{-}) \\ {}^{\flat}\mathrm{P}_{n+1}^{+} & = & {}^{\flat}\mathrm{P}_{n+1}^{-} - \mathrm{P}^{\flat\mathbf{XZ}} (\mathrm{P}^{\mathbf{ZZ}})^{-1} (\mathrm{P}^{\flat\mathbf{XZ}})^{\intercal} \end{array} \right.$$

Other approximated Kalman filters have been developed as for example the *Ensemble Kalman filter* (EnKF) [14] which follows a similar approach. The main difference concerns the choice of the particles and their numbers since in practice the EnKF relies on a significantly increased number of particles.

7.3.2.2 Reduced-order optimal filters

As seen in the previous section, the optimal filter in the linear case and its various approximate filters in non-linear cases can be formulated. We recall that their great advantage

is that they can be applied with any dynamics. However, these optimal filters have an important drawback which is the prohibitive computation of the full covariance matrix ${}^\flat\mathrm{P}$. One first strategy is to use reduced-order for the covariance operators. A typical example of order reduction occurs when we restrict the uncertainty space to a parameter space. In other words, we consider only the problem of identification without considering other type of errors like some initial condition errors.

To begin with, let us consider the general reduced order formulation. We will then see its implications in an identification context. For the sake of simplicity we recall that the state space is considered as finite dimensional even if its dimension can be very large. The main idea behind the reduced order strategy is to consider a *SVD* decomposition of ${}^\flat\mathrm{P}$ of the form

$$\forall t,\ {}^\flat\mathrm{P}(t) = {}^\flat\mathrm{L}(t)\mathrm{U}^{-1}(t){}^\flat\mathrm{L}(t)^\intercal,$$

with U an invertible matrix of small size r and ${}^\flat\mathrm{L}$ an extension operator. For linear operators, we can prove, see [29], that this decomposition is stable over time and that the Riccati equation (7.15) leads to the two following systems with admissible computational times

$$\dot{{}^\flat\mathrm{L}} = {}^\flat\mathrm{A}{}^\flat\mathrm{L} \text{ and } \dot{\mathrm{U}} = {}^\flat\mathrm{L}^\intercal{}^\flat\mathrm{H}^\intercal\mathrm{M}_{\mathcal{Z}}{}^\flat\mathrm{H}{}^\flat\mathrm{L}. \tag{7.19}$$

In non-linear cases, extensions of these two systems have been developed. Note that a specific analysis of the error produced by these approximations is then necessary. We present here the Reduced-order EKF (RoEKF) and the Reduced-order UKF (RoUKF).

Reduced-order EKF The RoEKF replaces the non-linear operators by their tangent counterparts in (7.19) following the idea of the EKF

$$\dot{{}^\flat\mathrm{L}} = \mathrm{d}{}^\flat\mathrm{A}{}^\flat\mathrm{L} \text{ and } \dot{\mathrm{U}} = {}^\flat\mathrm{L}^\intercal(\mathrm{d}{}^\flat\mathrm{H})^\intercal\mathrm{M}_{\mathcal{Z}}(\mathrm{d}{}^\flat\mathrm{H}){}^\flat\mathrm{L}.$$

A time and space discretization is directly given by the following prediction-correction scheme.

Prediction-Correction Scheme

1. Prediction

$$\left\{\begin{array}{rcl} {}^\flat\hat{\mathrm{x}}^-_{n+1} & = & {}^\flat\mathrm{A}_{n+1|n}({}^\flat\hat{\mathrm{x}}^+_n) \\ {}^\flat\mathrm{L}_{n+1} & = & \mathrm{d}{}^\flat\mathrm{A}_{n+1|n}{}^\flat\mathrm{L}_n \end{array}\right.$$

2. Correction

$$\left\{\begin{array}{rcl} \mathrm{U}_{n+1} & = & \mathrm{U}_n + {}^\flat\mathrm{L}^\intercal_{n+1}(\mathrm{d}{}^\flat\mathrm{H}_{n+1})^\intercal\mathrm{M}_{n+1}(\mathrm{d}{}^\flat\mathrm{H}_{n+1}){}^\flat\mathrm{L}_{n+1} \\ {}^\flat\mathrm{G}_{n+1} & = & {}^\flat\mathrm{P}_{n+1}(\mathrm{d}{}^\flat\mathrm{H}_{n+1})^\intercal\mathrm{M}_{n+1} \\ {}^\flat\hat{\mathrm{x}}^+_{n+1} & = & {}^\flat\hat{\mathrm{x}}^-_{n+1} + {}^\flat\mathrm{G}_{n+1}(\mathrm{z}_{n+1} - {}^\flat\mathrm{H}_{n+1}({}^\flat\hat{\mathrm{x}}^-_{n+1})) \end{array}\right.$$

Reduced-order UKF Now the Reduced-order strategy for the UKF filter denoted by RoUKF (Reduced-order UKF) is presented [28, 27, 32]. We can directly use the Sampling-Prediction-Correction scheme of the UKF and the decomposition of the covariance operator ${}^{\flat}\mathrm{P}$ given by

$$\forall t,\ {}^{\flat}\mathrm{P}(t) = {}^{\flat}\mathrm{L}(t)\mathrm{U}^{-1}(t){}^{\flat}\mathrm{L}(t)^{\intercal}$$

in order to obtain the RoUKF. Let the p *unitary sampling points* $\mathrm{sp}^{[i]}$ and the p weights α^i as defined above

$$\sum_{i=1}^{p} \alpha^i \mathrm{sp}^{[i]} = 0 \text{ and } \sum_{i=1}^{p} \alpha^i \mathrm{sp}^{[i]} \cdot (\mathrm{sp}^{[i]})^{\intercal} = \mathbb{1}.$$

We denote by D_α, the diagonal matrix of the p weights α^i. Let $(v^{[i]})_{1\leq i\leq p}$ p values, we denote by $v^{[*]}$, the following vector

$$v^{[*]} = \begin{pmatrix} v^{[1]} & \cdots & v^{[p]} \end{pmatrix}^{\intercal}.$$

In [28, 27], a study of the estimation error is given and in [10], the link with the singular evolutive interpolated Kalman filter (SEIK) introduced in [32] is given. The Sampling-Prediction-Correction scheme of RoUKF is given. Since, in what follows, we apply the RoUKF only on the parameters space, the time scheme given here separates the parameters from the state.

Sampling-Prediction-Correction Scheme

1. Sampling

$$\left\{ \begin{array}{rcl} \mathrm{C}_n & = & \sqrt{\mathrm{U}_n^{-1}} \\ \begin{pmatrix} \hat{\mathrm{x}}_n^{[i]+} \\ \hat{\theta}_n^{[i]+} \end{pmatrix} & = & \begin{pmatrix} \hat{\mathrm{x}}_n^{+} \\ \hat{\theta}_n^{+} \end{pmatrix} + \begin{pmatrix} \mathrm{L}_n^{\mathrm{X}} \\ \mathrm{L}_n^{\theta} \end{pmatrix} \mathrm{C}_n^{\intercal} \mathrm{sp}^{[i]} \end{array} \right.$$

2. Prediction

$$\left\{ \begin{array}{rcl} \hat{\mathrm{x}}_{n+1}^{-} & = & \sum_{i=1}^{p} \alpha^i \mathrm{A}_{n|n+1}(\hat{\mathrm{x}}_n^{[i]+}, \hat{\theta}_n^{[i]+}) \text{ and } \hat{\mathrm{x}}_{n+1}^{[i]-} = \mathrm{A}_{n|n+1}(\hat{\mathrm{x}}_n^{[i]+}, \hat{\theta}_n^{[i]+}) \\ \hat{\theta}_{n+1}^{-} & = & \sum_{i=1}^{p} \alpha^i \hat{\theta}_n^{[i]+} = \hat{\theta}_n^{+} \text{ and } \hat{\theta}_{n+1}^{[i]-} = \hat{\theta}_n^{[i]+} \end{array} \right.$$

3. Observations (of the observer model)

$$\mathrm{z}_{n+1}^{[i]-} = \mathrm{H}_{n+1}(\hat{\mathrm{x}}_{n+1}^{[i]-}) \text{ and } \mathrm{z}_{n+1}^{-} = \sum_{i=1}^{p} \alpha^i \mathrm{z}_{n+1}^{[i]-}$$

4. Correction

$$\left\{ \begin{array}{rcl} \mathrm{L}_{n+1}^{\mathrm{X}} & = & \hat{\mathrm{x}}_{n+1}^{[*]-} D_\alpha (\mathrm{sp}^{[*]})^{\intercal} \\ \mathrm{L}_{n+1}^{\theta} & = & \hat{\theta}_{n+1}^{[*]-} D_\alpha (\mathrm{sp}^{[*]})^{\intercal} \\ \Gamma_{n+1} & = & \mathrm{z}_{n+1}^{[*]-} D_\alpha (\mathrm{sp}^{[*]})^{\intercal} \\ \mathrm{U}_{n+1} & = & 1 + \Gamma_{n+1}^{\intercal} \mathrm{M}_{n+1} \Gamma_{n+1} \\ \hat{\mathrm{x}}_{n+1}^{+} & = & \hat{\mathrm{x}}_{n+1}^{-} + \mathrm{L}_{n+1}^{\mathrm{X}} \mathrm{U}_{n+1}^{-1} \Gamma_{n+1}^{\intercal} \mathrm{M}_{n+1} (\mathrm{z}_{n+1} - \mathrm{z}_{n+1}^{-}) \\ \hat{\theta}_{n+1}^{+} & = & \hat{\theta}_{n+1}^{-} + \mathrm{L}_{n+1}^{\theta} \mathrm{U}_{n+1}^{-1} \Gamma_{n+1}^{\intercal} \mathrm{M}_{n+1} (\mathrm{z}_{n+1} - \mathrm{z}_{n+1}^{-}) \end{array} \right.$$

7.4 Luenberger observers

An alternative method to the reduced-order strategy consists in building a filter which is not based on an optimal criterion. This idea – originally introduced by *Luenberger* [24, 25] – relies on the definition of the simplest possible filter such that the error between the observed trajectory and the observer system tends to zero.

The so-called Luenberger filter in the finite dimension framework – also named nudging in the PDE community [17, 39] – consists in stabilizing the dynamics satisfied by the estimation error

$$\tilde{\mathrm{x}} = \mathrm{x} - \hat{\mathrm{x}}.$$

Most of the Luenberger filter are defined only for estimating the state – and we will see in the sequel the possible strategy when we also expect to identify some parameters. We thus consider a real trajectory $\{\mathrm{x}(t), t \in [0,T]\}$ resulting of the dynamics

$$\left\{ \begin{array}{rcl} \dot{\mathrm{x}}(t) & = & \mathrm{A}(\mathrm{x},t) \\ \mathrm{x}(0) & = & \mathrm{x}_\bullet + \zeta^{\mathrm{X}}. \end{array} \right. \tag{7.20}$$

The available observations of x are denoted by z and we assume that there exists an observation operator such that

$$\mathrm{z}(t) = \mathrm{H}(\mathrm{x}(t),t) + \chi(t),$$

with χ an additive noise induced by the measurement process. The objective is to find a state filter G^{X} for the observer model

$$\left\{ \begin{array}{rcl} \dot{\hat{\mathrm{x}}}(t) & = & \mathrm{A}(\hat{\mathrm{x}},t) + \mathrm{G}^{\mathrm{X}}\big(\mathrm{z}(t) - \mathrm{H}(\hat{\mathrm{x}}(t))\big), \\ \hat{\mathrm{x}}(0) & = & \hat{\mathrm{x}}_\bullet. \end{array} \right. \tag{7.21}$$

The estimation error $\tilde{\mathrm{x}}$ verifies

$$\left\{ \begin{array}{rcl} \dot{\tilde{\mathrm{x}}}(t) & = & \mathrm{A}(\mathrm{x},t) - \mathrm{A}(\hat{\mathrm{x}},t) - \mathrm{G}^{\mathrm{X}}\big(\mathrm{z}(t) - \mathrm{H}(\hat{\mathrm{x}}(t))\big), \\ \hat{\tilde{\mathrm{x}}}(0) & = & \zeta^{\mathrm{X}}. \end{array} \right. \tag{7.22}$$

First assume that A and H are linear operators, then the model of the estimation error becomes

$$\left\{ \begin{array}{rcl} \dot{\tilde{\mathrm{x}}}(t) & = & (\mathrm{A} - \mathrm{G}^{\mathrm{X}}\mathrm{H})\tilde{\mathrm{x}} - \mathrm{G}^{\mathrm{X}}\,\chi, \\ \hat{\tilde{\mathrm{x}}}(0) & = & \zeta^{\mathrm{X}}. \end{array} \right. \tag{7.23}$$

We want to find a state filter which allows to control the estimation error dynamics. This means that the objective is to suggest a gain operator such that $\mathrm{A} - \mathrm{G}^{\mathrm{X}}\mathrm{H}$ is a dissipative operator.

We want to show that with the gain operator $\mathrm{G}^{\mathrm{X}} = \lambda \mathrm{H}^*$ – where H^* is the adjoint operator of the observation operator (for the corresponding metrics in the state and observations spaces), and λ a strictly positive value – is a possible choice. The dynamical system verified

by the estimation error becomes

$$\left\{\begin{array}{rcl} \dot{\tilde{\mathrm{x}}}(t) & = & (\mathrm{A}-\lambda \mathrm{H}^*\mathrm{H})\tilde{\mathrm{x}}-\lambda \mathrm{H}^*\,\chi \\ \hat{\tilde{\mathrm{x}}}(0) & = & \zeta^{\mathrm{x}} \end{array}\right. \tag{7.24}$$

This system can be rewritten on the variational formulation as, find $\tilde{\mathrm{x}}$ such that $\forall \psi$,

$$\int_\Omega \dot{\tilde{\mathrm{x}}}\,\psi = \int_\Omega \mathrm{A}\tilde{\mathrm{x}}\,\psi - \lambda \int_\Omega \mathrm{H}^*\mathrm{H}\tilde{\mathrm{x}}\,\psi - \lambda \int_\Omega \mathrm{H}^*\,\chi\,\psi$$

and with $\psi = \tilde{\mathrm{x}}$, we have

$$\frac{1}{2}\int_\Omega \partial_t \tilde{\mathrm{x}}^2 = \int_\Omega \mathrm{A}\tilde{\mathrm{x}}\cdot\tilde{\mathrm{x}} - \lambda \int_\Omega \|\mathrm{H}\tilde{\mathrm{x}}\|^2_{\mathcal{Z}} - \lambda \int_\Omega \chi\,\mathrm{H}\tilde{\mathrm{x}},$$

and this automatically implies that the observer term $\lambda \mathrm{H}^*\mathrm{H}$ is a stabilizing term for the system of the estimation error assuming the noise χ is controlled.

In the non-linear case, a classical strategy is to replace the operators by their tangent counterparts. Indeed, when the error is small, we have the following approximation if we neglect the noise

$$\mathrm{z}-\mathrm{H}(\hat{\mathrm{x}}) \approx \mathrm{z}-\mathrm{H}(\mathrm{x})+\mathrm{dH}(\hat{\mathrm{x}})(\mathrm{x}-\hat{\mathrm{x}}) \approx \mathrm{dH}(\hat{\mathrm{x}})\tilde{\mathrm{x}}$$

and in this case the gain is defined by

$$\mathrm{G}^{\mathrm{x}} = \lambda \mathrm{dH}(\hat{\mathrm{x}})^*.$$

In some cases, we just have an implicit relation between the observations z and the estimator $\hat{\mathrm{x}}$ of the following form

$$\mathrm{D}(\mathrm{z}(t),\hat{\mathrm{x}}(t)) = 0,$$

where D is the discrepancy. This leads to the following definition of the gain

$$\mathrm{G}^{\mathrm{x}} = -\lambda(\mathrm{dD}(\mathrm{z},\hat{\mathrm{x}}))^*.$$

Heat model Let us illustrate this Luenberger observer in the case of the Heat equation. We consider the diffusion problem

$$\left\{\begin{array}{rcl} \dot{\mathrm{x}} & = & \Delta \mathrm{x} + f,\ \Omega\times[0,T] \\ \mathrm{x}(0) & = & \mathrm{x}_\bullet + \zeta^{\mathrm{x}},\ \Omega \\ \mathrm{x} & = & 0,\ \partial\Omega\times[0,T]. \end{array}\right. \tag{7.25}$$

We consider two real separable Hilbert spaces V and U and we suppose that V is dense in U so that by identifying U and its dual U', we have $V \subset U \subset V'$, where each space is dense in the following. We denote by $W([0,T],V)$ the space

$$W([0,T],V) = \{\mathrm{x},\ \mathrm{x}\in L^2([0,T],V), \dot{\mathrm{x}}\in L^2([0,T],V')\}.$$

Here we consider that $U = L^2(\Omega)$, $V = H_0^1(\Omega)$ and then $V' = H^{-1}(\Omega)$. Using Theorem 1 and Theorem 2 given in [13] on page 512, there exists a unique solution of the associated variational problem of (7.25) in $W([0,T], H_0^1(\Omega))$ if $f \in L^2([0,T], H^{-1}(\Omega))$ and $\mathrm{x}(0) \in L^2(\Omega)$. Using Theorem 1 on page 473 in [13], we have $\mathrm{x} \in \mathcal{C}([0,T], L^2(\Omega))$. This implies that for all $t \in [0,T]$, $\mathrm{x}(t) \in L^2(\Omega)$. We assume that we have at our disposal the observations on a sub-domain denoted by ω and the following observer is then well-defined

$$\mathrm{H} : \left\{ \begin{array}{ccc} L^2(\Omega) & \to & L^2(\omega) \\ \mathrm{x} & \mapsto & \mathrm{x}_{|\omega} \end{array} \right. . \tag{7.26}$$

We have

$$\mathrm{z} = \mathrm{x}_{|\omega} + \chi = \mathrm{H}\mathrm{x} + \chi.$$

We search the adjoint of this observer

$$\Big\langle \mathrm{H}^*\mathrm{z}, \phi \Big\rangle_{L^2(\Omega)} = \Big\langle \mathrm{z}, \mathrm{H}\phi \Big\rangle_{L^2(\omega)} = \int_\omega \mathrm{z}\phi_{|\omega} = \int_\Omega \mathrm{z}\phi_{|\omega}$$

and then the adjoint is given by

$$\forall \mathrm{z} \in L^2(\omega),\ \mathrm{H}^*\mathrm{z} = \mathbb{1}_\omega \mathrm{z} \in L^2(\Omega).$$

This implies that the observer model verifies

$$\left\{ \begin{array}{rcl} \dot{\hat{\mathrm{x}}} & = & \Delta\hat{\mathrm{x}} + \lambda \mathbb{1}_\omega(\mathrm{z} - \hat{\mathrm{x}}_{|\omega}) + f, \text{ in } \Omega \\ \hat{\mathrm{x}}(0) & = & \mathrm{x}_\bullet, \text{ in } \Omega \\ \hat{\mathrm{x}} & = & 0, \text{ in } \partial\Omega \end{array} \right. \tag{7.27}$$

By applying (7.4) at the previous observer, the variational problem of the error model is given by find $\tilde{\mathrm{x}} \in H_0^1(\Omega)$ such that $\forall \psi \in H_0^1(\Omega)$,

$$\int_\Omega \dot{\tilde{\mathrm{x}}}\psi = -\int_\Omega \nabla\tilde{\mathrm{x}} \cdot \nabla\psi - \lambda \int_\omega \tilde{\mathrm{x}}\psi - \lambda \int_\omega \chi\psi.$$

The correction term then adds dissipation. We denote by μ_n, the eigenvalues of the diffusion problem (7.25), *i.e.* the eigenvalues of the Laplacian operator Δ. We also define $\tilde{\mu}_n$ as the eigenvalues of the error problem, *i.e.* the eigenvalues of the operator $\Delta - \lambda \mathrm{H}^*\mathrm{H}$. We can prove the following result

$$\tilde{\mu}_n < \mu_n, \tag{7.28}$$

where $\tilde{\mu}_n$ decreases when λ increases. The proof is based on the Courant-Fish theorem in infinite dimension.

Theorem 2. *(Courant-Fisher) Let M be a coercive symmetric operator and $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_n \geq \cdots$ its eigenvalues. Then*

$$\mu_n = \inf_{\substack{W \subset H_0^1(\Omega) \\ dim(W) = k}} \sup_{y \in W \setminus \{0\}} \frac{(My, y)_{L^2}}{\|y\|_{L^2}}.$$

Applying this theorem to $\mathrm{M} = -\Delta$ and $\tilde{\mathrm{M}} = \mathrm{M} + \lambda \mathrm{H}^*\mathrm{H}$, we can prove that

$$-\tilde{\mu}_n > -\mu_n,$$

and this directly implies (7.28).

We can quickly discuss the choice of λ in this case. As the reader can see in the error model, the term containing the noise χ increases proportionally to λ. The value of λ is then balanced versus the value of χ. We also study the limit when $\lambda \to +\infty$ in order to see if we can find a second condition for the choice of λ. If we neglect the noise χ, the limit corresponds to $\mathrm{z} = \mathrm{H}\hat{\mathrm{x}}$. Then the best method to choose the value of λ is by controlling the value of the noise.

7.5 A complete joint state and parameters methodology

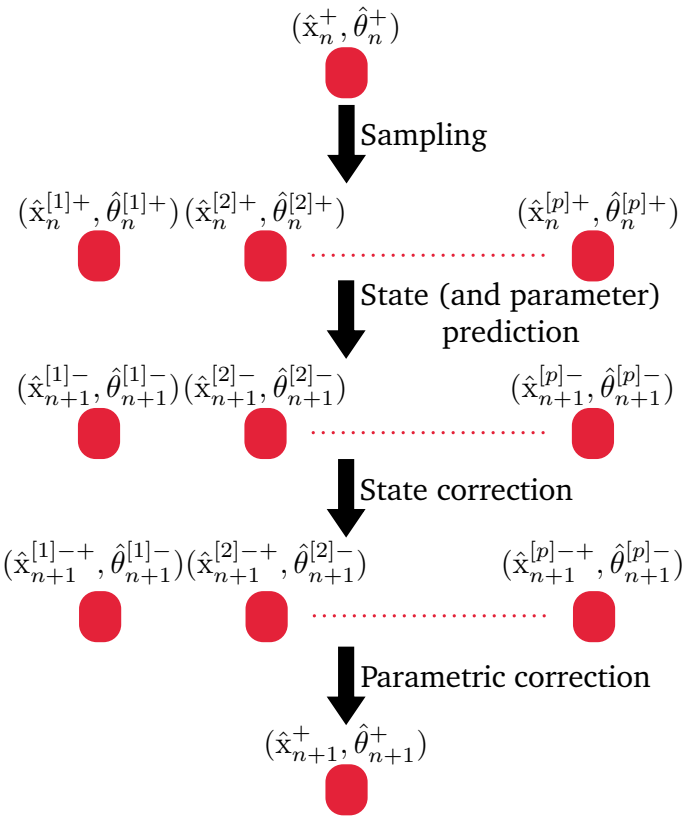

Figure 7.1: Sampling-Prediction-Correction with a nudging filter on the state space and a RoUKF for the parameters space

We will now combined the sequential strategies presented below in order to propose a joint state and parameters estimation of the form of (7.10). Hence, two different gains – one for the state and one for the parameters – have to be defined. On the state space, we build

a Luenberger filter and the following chapter will be devoted to the presentation of a very efficient observer for reaction diffusion models based on a nudging strategy. Then, the small dimension of the parameters space allows us to use a reduced-order optimal filter typically a RoUKF. The resulting joint sampling-prediction-correction scheme given below is illustrated in Figure 7.1.

Joint state (Luenberger) - parameters (RoUKF) scheme

1. Sampling
$$\left\{\begin{array}{rcl} \mathrm{C}_n & = & \sqrt{\mathrm{U}_n^{-1}} \\ \begin{pmatrix}\hat{\mathrm{x}}_n^{[i]+} \\ \hat{\theta}_n^{[i]+}\end{pmatrix} & = & \begin{pmatrix}\hat{\mathrm{x}}_n^{+} \\ \hat{\theta}_n^{+}\end{pmatrix} + \begin{pmatrix}\mathrm{L}_n^{\mathrm{x}} \\ \mathrm{L}_n^{\theta}\end{pmatrix} \mathrm{C}_n^{\intercal}\mathrm{sp}^{[i]} \end{array}\right.$$

2. Prediction
$$\left\{\begin{array}{rcl} \hat{\mathrm{x}}_{n+1}^{[i]-} & = & \mathrm{A}_{n|n+1}(\hat{\mathrm{x}}_n^{[i]+}, \hat{\theta}_n^{[i]+}) \\ \hat{\theta}_{n+1}^{-} & = & \sum_{i=1}^{p} \alpha^i \hat{\theta}_n^{[i]+} = \hat{\theta}_n^{+} \text{ and } \hat{\theta}_{n+1}^{[i]-} = \hat{\theta}_n^{[i]+} \end{array}\right.$$

3. Observations (of the observer model)
$$\mathrm{z}_{n+1}^{[i]-} = \mathrm{H}_{n+1}(\hat{\mathrm{x}}_{n+1}^{[i]-}) \text{ and } \mathrm{z}_{n+1}^{-} = \sum_{i=1}^{p} \alpha^i \mathrm{z}_{n+1}^{[i]-}$$

4. State correction
$$\left\{\begin{array}{rcl} \hat{\mathrm{x}}_{n+1}^{[i]+-} & = & \hat{\mathrm{x}}_{n+1}^{[i]-} + \lambda \mathrm{H}_{n+1}^{*}(\mathrm{z}_{n+1} - \mathrm{z}_{n+1}^{-}) \\ \hat{\mathrm{x}}_{n+1}^{+-} & = & \sum_{i=1}^{p} \alpha^i \hat{\mathrm{x}}_{n+1}^{[i]+-} \end{array}\right.$$

5. Correction
$$\left\{\begin{array}{rcl} \mathrm{L}_{n+1}^{\mathrm{x}} & = & \hat{\mathrm{x}}_{n+1}^{[*]+-} D_\alpha(\mathrm{sp}^{[*]})^{\intercal} \\ \mathrm{L}_{n+1}^{\theta} & = & \hat{\theta}_{n+1}^{[*]-} D_\alpha(\mathrm{sp}^{[*]})^{\intercal} \\ \Gamma_{n+1} & = & \mathrm{z}_{n+1}^{[*]-} D_\alpha(\mathrm{sp}^{[*]})^{\intercal} \\ \mathrm{U}_{n+1} & = & 1 + \Gamma_{n+1}^{\intercal} \mathrm{M}_{n+1} \Gamma_{n+1} \\ \hat{\mathrm{x}}_{n+1}^{+} & = & \hat{\mathrm{x}}_{n+1}^{+-} + \mathrm{L}_{n+1}^{\mathrm{x}} \mathrm{U}_{n+1}^{-1} \Gamma_{n+1}^{\intercal} \mathrm{M}_{n+1}(\mathrm{z}_{n+1} - \mathrm{z}_{n+1}^{-}) \\ \hat{\theta}_{n+1}^{+} & = & \hat{\theta}_{n+1}^{-} + \mathrm{L}_{n+1}^{\theta} \mathrm{U}_{n+1}^{-1} \Gamma_{n+1}^{\intercal} \mathrm{M}_{n+1}(\mathrm{z}_{n+1} - \mathrm{z}_{n+1}^{-}) \end{array}\right.$$

In practice, the discretization of the adjoint operator H_{n+1}^{*} is more complicated (with implicit terms see [9]) but the objective here is to give an idea of the discretization of a joint state and parameters estimator.

This strategy was initially introduced and analyzed by [29, 28]. In a few words, the idea is that the state observer stabilizes the state uncertainties. This allows to restrict the uncertainty to the parameter space, and then the reduced-order filtering strategy is valid to handle the identification.

Bibliography

[1] D. Auroux and J. Blum. Back and forth nudging algorithm for data assimilation problems. *Comptes Rendus Mathematique*, 340(12):873–878, 2005.

[2] R. Bellman. Dynamic programming and Lagrange multipliers. *Proceedings of the National Academy of Sciences of the United States of America*, 42(10):767, 1956.

[3] A. Bensoussan. *Filtrage optimal des systèmes linéaires*. Dunod, 1971.

[4] C. Bertoglio, D. Chapelle, M. A. Fernández, J.-F. Gerbeau, and P. Moireau. State observers of a vascular fluid–structure interaction model through measurements in the solid. *Computer Methods in Applied Mechanics and Engineering*, 256:149–168, 2013.

[5] J. Blum, F.-X. Le Dimet, and I. M. Navon. Data assimilation for geophysical fluids. In R. Temam and J. Tribbia, editors, *Handbook of Numerical Analysis: Computational Methods for the Atmosphere and the Oceans*. Elsevier, 2008.

[6] D. Chapelle, N. Cîndea, M. De Buhan, and P. Moireau. Exponential convergence of an observer based on partial field measurements for the wave equation. *Mathematical Problems in Engineering*, 2012, 2012.

[7] D. Chapelle, N. Cîndea, and P. Moireau. Improving convergence in numerical analysis using observers—the wave-like equation case. *Mathematical Models and Methods in Applied Sciences*, 22(12), 2012.

[8] D. Chapelle, M. Fragu, V. Mallet, and P. Moireau. Fundamental principles of data assimilation underlying the verdandi library: applications to biophysical model personalization within euheart. *Medical & Biological Eng & Computing*, pages 1–13, 2012.

[9] D. Chapelle, A. Imperiale, and P. Moireau. Cardiac estimation from tagged-MR images. In preparation.

[10] D. Chapelle, P. Moireau, and P. Le Tallec. Robust filtering for joint state-parameter estimation in distributed mechanical systems. *Discrete and Continuous Dynamical Systems*, (1–2):65–84, 2009.

[11] C. Corrado, J.-F. Gerbeau, and P. Moireau. Identification of an electrophysiological model from combined ECG and MRI measurements. *Conf. Bioengineering 2012, Oxford*, 2012.

[12] C. Corrado, J.-F. Gerbeau, and P. Moireau. Identification of weakly coupled multiphysics problems. application to the inverse problem of electrocardiology. *Journal of Computational Physics*, 2014. In review.

[13] R. Dautray and J.-L. Lions. *Mathematical Analysis and Numerical Methods for Science and Technology: Volume 3 Spectral Theory and Applications*, volume 3. Springer, 2000.

[14] G. Evensen. *Data Assimilation: The Ensemble Kalman Filter*. Springer, 2009.

[15] W. H. Fleming. Deterministic nonlinear filtering. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 25(3-4):435–454, 1997.

[16] G. Haine and K. Ramdani. Reconstructing initial data using observers: error analysis of the semi-discrete and fully discrete approximations. *Numerische Mathematik*, 120(2):307–343, 2012.

[17] J. E. Hoke and R. A. Anthes. The initialization of numerical models by a dynamic-initialization technique. *Monthly Weather Review*, 104(12):1551–1556, 1976.

[18] S. J. Julier and J. K. Uhlmann. Reduced sigma point filters for the propagation of means and covariances through nonlinear transformations. In *In Proceedings of the 2002 American Control Conference*, volume 2, pages 887–892, 2002.

[19] S.J. Julier and J.K. Uhlmann. A new extension of the Kalman filter to nonlinear systems. In *Proc. of AeroSense: The 11th Int. Symp. on Aerospace/Defence Sensing, Simulation and Controls*, 1997.

[20] R. Kalman and R. Bucy. New results in linear filtering and prediction theory. *Trans. ASME J. Basic. Eng.*, 83:95—108, 1961.

[21] A. J. Krener. The convergence of the Extended Kalman Filter. *In Lecture Notes in Control and Inform. Sci. Berlin: Springer*, 286:173–182, 2003.

[22] F.-X. Le Dimet and O. Talagrand. Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus A*, 38(2):97–110, 1986.

[23] K. Liu. Locally distributed control and damping for the conservative systems. *SIAM Journal on Control and Optimization*, 35(5):1574–1590, 1997.

[24] D.G. Luenberger. *Determining the State of a Linear with Observers of Low Dynamic Order*. PhD thesis, Stanford University, 1963.

[25] D.G. Luenberger. An introduction to observers. *IEEE Transactions on Automatic Control*, 16:596–602, 1971.

[26] J. Mandel, L.S. Bennethum, J.D. Beezley, J.L. Coen, C.C Douglas, L.P. Franca, M. Kim, and A. Vodacek. A wildland fire model with data assimilation. *Mathematics and Computers in Simulation*, 79(3):584–606, 2008.

[27] P. Moireau and D. Chapelle. Erratum of article "Reduced-order Unscented Kalman Filtering with application to parameter identification in large-dimensional systems". *ESAIM: Control, Optimisation and Calculus of Variations*, 17(2):406–409, 2011.

[28] P. Moireau and D. Chapelle. Reduced-order Unscented Kalman Filtering with application to parameter identification in large-dimensional systems. *ESAIM: Control, Optimisation and Calculus of Variations*, 17(2):380–405, 2011.

[29] P. Moireau, D. Chapelle, and P. Le Tallec. Joint state and parameter estimation for distributed mechanical systems. *Computer Methods in Applied Mechanics and Engineering*, 197:659–677, 2008.

[30] P. Moireau, D. Chapelle, and P. Le Tallec. Filtering for distributed mechanical systems using position measurements: Perspectives in medical imaging. *Inverse Problems*, 25(3):035010 (25pp), 2009.

[31] I. M. Navon. Data assimilation for numerical weather prediction: A review. In *Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications*. Springer, 2009.

[32] D. T. Pham. Stochastic methods for sequential data assimilation in strongly nonlinear systems. *Monthly Weather Review*, 129(5):1194–1207, 2001.

[33] Daley R. *Atmospheric Data Analysis*. Cambridge University Press, 1993.

[34] K. Ramdani, M. Tucsnak, and G. Weiss. Recovering the initial state of an infinite-dimensional system using observers. *Automatica*, 46(10):1616–1625, 2010.

[35] C.M. Rochoux, B. Cuenot, S. Ricci, A. Trouvé, B. Delmotte, Se. Massart, R Paoli, and R. Paugam. Data assimilation applied to combustion. *Comptes Rendus Mécanique*, 341:266–276, 2013.

[36] J. Sainte-Marie, D. Chapelle, R. Cimrman, and M. Sorine. Modeling and estimation of the cardiac electromechanical activity. *Computers and Structures*, 84:1743–1759, 2006.

[37] D. Simon. *Optimal State Estimation: Kalman, H^∞, and Nonlinear Approaches*. Wiley-Interscience, 2006.

[38] D. R. Stauffer and J.-W. Bao. Optimal determination of nudging coefficients using the adjoint equations. *Tellus A*, 45(5):358–369, 1993.

[39] D. R. Stauffer and N. L. Seaman. Use of four-dimensional data assimilation in a limited-area mesoscale model. Part I: Experiments with synoptic-scale data. *Monthly Weather Review*, 118(6):1250–1277, 1990.

[40] O. Talagrand. Assimilation of observations, an introduction. *Journal-meteorological society of japan series 2*, 75:81–99, 1997.

[41] P.A. Vidard, F.-X. Le Dimet, and A. Piacentini. Determination of optimal nudging coefficients. *Tellus A*, 55(1):1–15, 2003.

[42] X. Zou, I.M. Navon, and F.-X. Le Dimet. An optimal nudging data assimilation scheme using parameter estimation. *Quarterly Journal of the Royal Meteorological Society*, 118(508):1163–1186, 1992.

CHAPTER 8

A joint state and parameter observer for the reaction diffusion model. Application in cardiac electrophysiology.

Abstract Data assimilation is a very active subject with various applications. The objective of this work is to adapt complex reaction-diffusion models to some data in order to obtain very realistic simulations for each case considered. A strategy which allows to estimate the state and the parameters from available data corresponding to the front position at different times is given. Different data assimilation methods are combined for the inverse problem: a Luenberger observer for the state space and a Reduced-order Unscented Kalman Filter (RoUKF) for the parameters space. The origin of the efficient state observer is given. The data consist of front maps such levels sets which naturally appears in imaging techniques in the detection of objet (*i.e.* the segmentation) where an eikonal equation – representing the evolution of the object contour which pursues the one sought – is solved. Moreover the reaction-diffusion equation can be linked by an asymptotic analysis in the front direction to an eikonal equation. Using strategies developed in image processing in order to compare the contours, we present a state observer of the eikonal equation. This observer is carried over to the initial reaction-diffusion model by an inverted asymptotic analysis. Using shape derivatives theory, this state observer is mathematically justified. One of the various application fields of this work is cardiac electrophysiology where the bidomain or the monodomain models – both reaction-diffusion models – are used and depolarization maps (ECGI) are available. Simulations on realistic synthetic data demonstrate in the context of cardiac electrophysiology the potential of this joint state-parameter approach and bring us closer to the personalization of models to each patient.

A Luenberger observer for the reaction-diffusion model assimilating front position data

Dominique Chapelle, Annabelle Collin, Philippe Moireau
Inria Saclay Ile-de-France, Palaiseau, France

To be submitted

Abstract

Data assimilation is a very active subject with various applications. The objective of this work is to adapt complex reaction-diffusion models to some data in order to obtain very realistic simulations for each case considered. A strategy which allows to estimate the state and the parameters from available data corresponding to the front position at different times is given. Different data assimilation methods are combined for the inverse problem: a Luenberger observer for the state space and a Reduced-order Unscented Kalman Filter (RoUKF) for the parameters space. The origin of the efficient state observer is given. The data consist of front maps such levels sets which naturally appears in imaging techniques in the detection of objet (*i.e.* the segmentation) where an eikonal equation – representing the evolution of the object contour which pursues the one sought – is solved. Moreover the reaction-diffusion equation can be linked by an asymptotic analysis in the front direction to an eikonal equation. Using strategies developed in image processing in order to compare the contours, we present a state observer of the eikonal equation. This observer is carried over to the initial reaction-diffusion model by an inverted asymptotic analysis. Using shape derivatives theory, this state observer is mathematically justified. One of the various application fields of this work is cardiac electrophysiology where the bidomain or the monodomain models – both reaction-diffusion models – are used and depolarization maps (ECGI) are available. Simulations on realistic synthetic data demonstrate in the context of cardiac electrophysiology the potential of this joint state-parameter approach and bring us closer to the personalization of models to each patient.

Keywords: Reaction diffusion model; Data assimilation; Image processing; Front propagation; Cardiac electrophysiology
Mathematics Subject Classification (2010):

Introduction

Reaction-diffusion models and simplified asymptotic models as the curvature eikonal equation [35, 29] appear in various application fields, for example in cardiac electrophysiology with the bidomain model [56, 49] and its first approximation the monodomain model [11, 50]. In wildland fire propagation, reaction diffusion models are also used [2] and a third field is for example the growth of tumors, see [55, 20]. As for all natural and physical systems, a great difficulty consists in dealing with the many uncertain quantities which must be preserved for running model simulations. These quantities include initial conditions and physical parameters of the model – as for example the conductivity parameter – which are difficult to measure. Fortunately, other types of measurement exist and the objective is to use these available data in order to circumvent the uncertainties corresponding to the dynamical model. Here, we consider the propagation maps which give the position of the front at different times.

Data assimilation aims at reducing the uncertainties and also at estimating uncertain quantities as for example the model parameters. As we do not wish to precisely retrieve the initial condition error and parameters error but just have a correct approximation of the state and of the parameters, sequential – also called filtering – methods are adapted [32, 21, 37]. The objective of this article is to give a joint state and parameter estimation with admissible computing times where a Luenberger observer corrects the state and an optimal Kalman-like filter converges to the real parameter. The main difficulties concern the non-linearity of a reaction-diffusion model and the nature of available data. Indeed, the comparison of the solution of a reaction-diffusion model with a propagation front is difficult and a so-called *discrepancy* measure needs to be defined. Data assimilation for the reaction diffusion model is an active subject and represents a real challenge because of the various application fields. In cardiac electrophysiology in [15], a data assimilation combining electrical and mechanical measurements is given. The electrical data considered is the classical electrocardiogram [39] using a Luenberger observer for the mechanical state and an Unscented Kalman Filter for the electrical state and parameters. For example, in wildland fire propagation, some studies [40, 48] assimilate front propagation data with Kalman-like filters in order to predict the state. In [10] data assimilation methods are given in order to identify some model parameters in application field of tumor growth. In the same application field, we can also cite [27, 36] where optimization methods are developed in order to identify parameters or find the spatial location of the origin of the tumor (initial condition), whereas in [54] a statistical approach is considered in order to predict the patient survival.

The outline of this paper is as follows. In a first section, we present the reaction-diffusion model and the available data that we want to assimilate. The second section concerns the state estimator. We start with the presentation of an observer suitable for a reaction-diffusion model and front propagation data. Then we give the origin of the state estimator starting with the presentation of an asymptotic eikonal equation. Using image processing techniques [8, 58], we present an observer valid for the eikonal equation and we carry over the observer to the initial reaction-diffusion model by an inverted asymptotic analysis. The last but not least part of this section gives a mathematically-justified study of the observer model and we prove that the correction term stabilizes the model of the error. The third section concerns the parameter estimation and the Reduced-order Unscented Kalman Filtering (RoUKF) [43, 42] is used. In the last section, several numerical simulations in cardiac electrophysiology in 1D and 2D cases are presented. These simulations validate, in a first part, the very efficient state estimator. In a second part, using a joint state-parameter strategy, we identify the conductivity parameter. We finish this section with a state estimator simulation in a realistic case.

1 Position of the problem

1.1 Reaction-diffusion model

We consider the following reaction-diffusion model

$$\left\{\begin{array}{rcll} \partial_t u - \vec{\nabla}\cdot(\vec{\vec{\sigma}}\cdot\vec{\nabla}u) & = & kf(u), & \mathcal{B}\times(0,T), \\ (\vec{\vec{\sigma}}\cdot\vec{\nabla}u)\cdot\vec{n} & = & 0, & \partial\mathcal{B}\times(0,T), \\ u(\vec{x},0) & = & u_0(\vec{x}), & \mathcal{B}, \end{array}\right. \tag{1}$$

where u is the unknown, $\vec{\vec{\sigma}}$ the diffusion tensor. The reaction term f represents the variation of u between two states: burned/not burned in fire propagation, sick/healthy in tumor growth and depolarized/polarized in cardiac electrophysiology. For the sake of simplicity in what follows, we denote by *traveled-through region*, the part of the domain $\mathcal{B}$ which is already burned, sick or

depolarized, according to the application considered. A classical and simple term considered for the reaction term is defined by $f(u) = u(1-u)$ but more complicated cases can be considered. For example in cardiac electrophysiology, f depends on other variables and the reaction-diffusion model is then coupled with one or more ordinary differential equations which model the evolution of these other variables.

1.2 Available data

The objective of this work is to obtain simulations – adapted for example to an individual patient or to an occurrence of wildland fire – by correcting and adapting the reaction-diffusion model with available data. Here, the data that we want to assimilate are the isochrone maps of u. These maps give the front position at successive times. This means that data correspond to the boundary of the traveled-through region at successive times. In practice the observations are often partial in space and in time but using interpolations the whole observations can be reconstructed.

2 State observer for the monodomain model

We denote by d the dimension of the domain $\mathcal{B}$. Using these observations, we can define an evolving manifold of dimension $d-1$ (set of isolated points in 1D, curve in 2D and surface in 3D) as the boundary of the traveled-through domain. The objective is to find an observer able to assimilate these data in the reaction-diffusion model (1), *i.e.* an observer able to adapt to the evolving manifold defined by the observation. Object detection in image processing uses similar ideas [45, 8]. That is why we start with geometric preliminaries used in image processing.

2.1 Geometric preliminaries

The use of level sets to detect objects in an image is a classical method, see [45, 8]. We define Γ which corresponds to the boundary of a domain Ω^{in} and which evolves in time. We define ϕ, a level set function associated with Ω^{in} by

$$\phi > 0 \text{ in } \Omega^{\text{in}},\ \phi < 0 \text{ in } \mathcal{B} \setminus \overline{\Omega^{\text{in}}} \text{ and } \phi = 0 \text{ in } \Gamma.$$

We denote by z_{data} the object sought in image processing and by Γ_{data} the boundary of this object, with $\Omega^{\text{in}}_{data}$ the interior defined by Γ_{data}. In this context, z_{data} is constant in time. In fact in image processing, the objective is to detect a fixed object. All these notions are represented in Figure 1 in a 2D case. The definition of an object will be discussed.

As a preliminary, we start by the presentation of various notations. We define H the one dimensional Heaviside function by

$$H(\phi) = \begin{cases} 0 & \text{if } \phi < 0, \\ 1 & \text{if } \phi > 0, \\ \frac{1}{2} & \text{if } \phi = 0. \end{cases}$$

The directional derivative of the Heaviside function H in the normal direction $\vec{n}$ is defined by

$$\hat{\delta}(\Gamma, \vec{x}) = \vec{\nabla}\big(H \circ \phi\big)(\vec{x}) \cdot \vec{n}.$$

This function is the Dirac delta function of the multidimensional variable $\vec{x}$ and this distribution

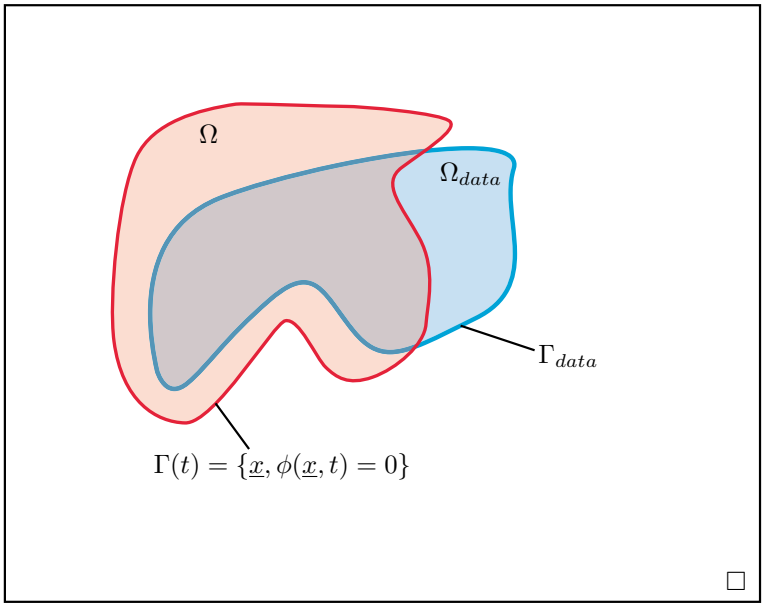

Figure 1: In blue, the object to be detected is represented and the curve $\Gamma(t)$ representing the domain in red which evolves to the object.

is non-zero only on the interface Γ, *i.e.* where $\phi = 0$. We have

$$\begin{aligned}\hat{\delta}(\Gamma, \vec{x}) &= \vec{\nabla}\big(H \circ \phi\big)(\vec{x}) \cdot \vec{n} \\ &= H'\big(\phi(\vec{x})\big)\vec{\nabla}\phi(\vec{x}) \cdot \frac{\vec{\nabla}\phi(\vec{x})}{|\vec{\nabla}\phi(\vec{x})|} \\ &= H'\big(\phi(\vec{x})\big)|\vec{\nabla}\phi(\vec{x})|.\end{aligned}$$

The Dirac delta function in one spatial dimension denoted by δ, is defined as the derivative of the Heaviside function H, *i.e.* $\delta(\phi) = H'(\phi)$. This allows to finally obtain

$$\hat{\delta}(\Gamma, \vec{x}) = \delta\big(\phi(\vec{x})\big)|\vec{\nabla}\phi(\vec{x})|. \tag{2}$$

This gives the following property

$$\int_{\mathcal{B}} \hat{\delta}(\Gamma, \vec{x})\, \psi \, d\mathcal{B} = \int_{\Gamma} \psi \, d\Gamma,$$

which shows that $\hat{\delta}$ is independent of ϕ.

2.2 State estimator the reaction-diffusion model

We suppose that the target solution u satisfies

$$\begin{cases} \partial_t u - \vec{\nabla} \cdot (\vec{\sigma} \cdot \vec{\nabla}\, u) &= kf(u), & \mathcal{B} \times (0, T), \\ (\vec{\sigma} \cdot \vec{\nabla}\, u) \cdot \vec{n} &= 0, & \partial\mathcal{B} \times (0, T), \\ u(\vec{x}, 0) &= u_0(\vec{x}), & \mathcal{B}. \end{cases} \tag{3}$$

In practice, we just have an *a priori* of the initial condition u_0 denoted by $\hat{u}_0$, and this means that u_0 is decomposed into two parts

$$u_0 = \hat{u}_0 + \zeta,$$

where ζ is the uncertainty. If we compute the solution of (3) with the initial condition $\hat{u}_0$, we make an error. However, we have observations of the target solution and we want to define an observer denoted by $\hat{u}$ which converges to the target solution u. We recall that these observations correspond to the front position over time. The objective is to propose a *Luenberger* observer [38] – also called nudging in PDE world [28, 52] – for the reaction-diffusion model. Using these observations, we can define $t \mapsto \Gamma_u(t)$ as the boundary of a region $\Omega_u^{\text{in}}(t)$ which corresponds to the traveled-through region. We suppose that $t \mapsto \Gamma_u(t)$ is a closed boundary. From this front corresponding to the observations, we define an object z_u associated with u. The choice of the exact definition of the object will be discussed. Of course in our case, the object z_u varies in time. In order to compare the observer solution $\hat{u}$ with the observations z_u, we give some definitions. Let c_{th} be the value above which the solution changes state. This allows to define the time-dependent region $\Omega_{\hat{u}}^{\text{in}}$ – corresponding to the traveled-through region – by

$$\Omega_{\hat{u}}^{\text{in}}(t) = \{\vec{x} \in \mathcal{B},\, \hat{u}(\vec{x},t) > c_{\text{th}}\}$$

with its boundary

$$\Gamma_{\hat{u}}(t) = \{\vec{x} \in \mathcal{B},\, \hat{u}(\vec{x},t) = c_{\text{th}}\}.$$

The averages of the z_u in the interior of $\Omega_{\hat{u}}^{\text{in}}$ and in the exterior are given by

$$C_1(\Omega_{\hat{u}}^{\text{in}}) = \frac{\int_{\Omega_{\hat{u}}^{\text{in}}} z_u d\vec{x}}{\int_{\Omega_{\hat{u}}^{\text{in}}} d\vec{x}} \text{ and } C_2(\Omega_{\hat{u}}^{\text{in}}) = \frac{\int_{\mathcal{B}\setminus\overline{\Omega_{\hat{u}}^{\text{in}}}} z_u d\vec{x}}{\int_{\mathcal{B}\setminus\overline{\Omega_{\hat{u}}^{\text{in}}}} d\vec{x}}.$$

Using the definitions of Section 2.1, we have $\hat{\delta}(\Gamma_{\hat{u}},\, \vec{x}) = \hat{\delta}_{\{\hat{u}=c_{\text{th}}\}}$. The model verified by $\hat{u}$ that we propose is

$$\begin{cases} \partial_t \hat{u} - \vec{\nabla}\cdot(\vec{\vec{\sigma}}\cdot\vec{\nabla}\,\hat{u}) &= kf(\hat{u}) & \\ +\lambda\hat{\delta}(\Gamma_{\hat{u}},\,\vec{x})\alpha(|\vec{\nabla}\,\hat{u}|)\Big(-\big(z_u - C_1(\Omega_{\hat{u}}^{\text{in}})\big)^2 + \big(z_u - C_2(\Omega_{\hat{u}}^{\text{in}})\big)^2\Big), & & \mathcal{B}\times(0,T), \\ (\vec{\vec{\sigma}}\cdot\vec{\nabla}\,\hat{u})\cdot\vec{n} &= 0, & \partial\mathcal{B}\times(0,T), \\ \hat{u}(\vec{x},0) &= \hat{u}_0(\vec{x}), & \mathcal{B}, \end{cases} \tag{4}$$

where λ is a positive constant and α corresponds to a strictly positive function to be chosen as discussed later.

The correction term in the observer model does not have the classical form of the *Luenberger* observer

$$\mathrm{G}_u\big(\mathrm{D}(\mathrm{z}_u, \hat{u})\big),$$

where G_u is called the gain and D is the discrepancy which compares the solution of the observer with the observations. However, a discrepancy D can be extracted formally and it is essential for a joint state and parameter estimation as we will see in Section 3. This discrepancy is then defined by

$$\mathrm{D}(z_u,\hat{u}) = \int_{\Omega_{\hat{u}}^{\text{in}}} \big(z_u - C_1(\Omega_{\hat{u}}^{\text{in}})\big)^2 d\vec{x} + \int_{\mathcal{B}\setminus\overline{\Omega_{\hat{u}}^{\text{in}}}} \big(z_u - C_2(\Omega_{\hat{u}}^{\text{in}})\big)^2 d\vec{x}. \tag{5}$$

As we will see, this discrepancy D naturally appears.

The following section gives the main ideas that justify the state estimator given in (4). In a first step, the curvature eikonal equation, a simplified model derived from the reaction-diffusion model is introduced. The unknown of this eikonal equation is a level set of the depolarization front of u. The eikonal equation gives the evolution of the depolarization front. Using methods developed in image processing in order to detect objects in an image, a state observer for the eikonal equation is presented. The third step consists in getting back to the reaction-diffusion model.

2.3 Origin of the state estimator

2.3.1 Reaction-diffusion model to the eikonal curvature equation

Simplified models of the reaction-diffusion model exist. The curvature eikonal equation model is one of them and this model focuses on the evolution of the front propagation. In [35, 34], a derivation of an eikonal curvature equation is given for a reaction-diffusion model. As its name suggests, the eikonal curvature equation is obtained using a change of variable which allows to only take into account the front change along the curvature. The change of variable presented in [35, 34] is

$$\vec{x} = \vec{X}(\xi, t) \text{ and } t = \tau,$$

with

- ξ_1 the normal coordinate to ϕ_u a level set of u, *i.e.* the wave front location,
- ξ_2, ξ_3 the coordinates of the moving level set surfaces.

For the sake of simplicity, we consider the diffusion tensor $\vec{\vec{\sigma}} = \sigma$, *i.e.* constant and isotropic in space for this section. Using the Einstein implicit summation convention, the reaction-diffusion model (3) becomes in this new coordinate system

$$\frac{\partial u}{\partial \tau} = \sigma a_{il} a_{ik} \frac{\partial^2 u}{\partial \xi_l \partial \xi_k} + \sigma \frac{\partial a_{il}}{\partial x_i} \frac{\partial u}{\partial \xi_l} + \frac{\partial X_j}{\partial \tau} a_{jk} \frac{\partial u}{\partial \xi_k} + kf,$$

where $(a_{ij})_{ij}$ is the inverse of the matrix with entries $\left(\frac{\partial X_j}{\partial \xi_i}\right)_{ij}$. In [35, 34], the spatial scale variation of ξ_1 is supposed to be much shorter than the spatial scale variations of ξ_2 and ξ_3, namely there exists a small parameter ε such that

$$a_{i1} = O(1) \text{ and } a_{ij} = O(\varepsilon), \forall i, \forall j \neq 1.$$

This hypothesis means that the normal of the front changes faster than the tangent space. This explains the name of the eikonal curvature equation. A second assumption is that u is independent of ξ_2, ξ_3 and τ in the leading order ε.

Remark 1 *These assumptions correspond to choosing the coordinate system such that u only depends on ξ_1. We can do an asymptotic analysis with $u = u^0(\xi_1) + u^1(\xi_1)P^1(\xi_2, \xi_3, \tau) + u^2(\xi_1)P^2(\xi_2, \xi_3, \tau)$ with P^i polynomial functions with all terms of order i.*

Denoting by $\vec{a} = (a_{i1})_i$ the reaction-diffusion model becomes

$$\sigma |\vec{a}|^2 \frac{\partial^2 u}{\partial \xi_1^2} + \left(\sigma \vec{\nabla}_{\vec{x}} \cdot \vec{a} + \frac{\partial \vec{X}}{\partial \tau} \cdot \vec{a}\right) \frac{\partial u}{\partial \xi_1} + kf(u) = O(\varepsilon), \tag{6}$$

under the asymptotic assumptions. We choose the new coordinate system such that $k = \sigma |\vec{a}|^2$ and we identify this equation to

$$ku'' + kc_0 u' + kf(u) = 0. \tag{7}$$

In [1] the authors show that under assumptions on f, there exists a constant c_0 depending on f such that there exists a strictly decreasing function u solution of (7). In the special case $f(u) = u(1-u)$, we have $c_0 = 2\sqrt{f'(0)} = 2$. We use this mathematical theory and there exists a constant c_0 depending on f such that

$$kc_0 = \sigma \vec{\nabla}_{\vec{x}} \cdot \vec{a} + \frac{\partial \vec{X}}{\partial \tau} \cdot \vec{a}. \tag{8}$$

We define from the observer solution u a level set associated with u, *i.e.*

$$\phi_u > 0 \text{ if } u > c_{\text{th}},\ \phi_u < 0 \text{ if } u < c_{\text{th}} \text{ and } \phi_u = 0 \text{ if } u = c_{\text{th}}.$$

With ϕ_u a level set of u, we have

$$\frac{\vec{a}}{|\vec{a}|} = -\frac{\vec{\nabla}_{\vec{x}}\,\phi_u}{|\vec{\nabla}_{\vec{x}}\,\phi_u|},$$

and it follows that

$$\vec{a} = -\frac{\sqrt{k}}{\sqrt{\sigma}}\frac{\vec{\nabla}_{\vec{x}}\,\phi_u}{|\vec{\nabla}_{\vec{x}}\,\phi_u|}.$$

Then (8) becomes

$$kc_0 = -\sqrt{\sigma k}\,\vec{\nabla}_{\vec{x}} \cdot \Big(\frac{\vec{\nabla}_{\vec{x}}\,\phi_u}{|\vec{\nabla}_{\vec{x}}\,\phi_u|}\Big) - \frac{\sqrt{k}}{\sqrt{\sigma}}\frac{\partial \vec{X}}{\partial \tau} \cdot \frac{\vec{\nabla}_{\vec{x}}\,\phi_u}{|\vec{\nabla}_{\vec{x}}\,\phi_u|}. \tag{9}$$

As ϕ_u is constant along the wave front we have

$$\vec{\nabla}_{\vec{x}}\phi_u \cdot \frac{\partial \vec{X}}{\partial \tau} + \frac{\partial \phi_u}{\partial \tau} = 0.$$

We then obtain

$$\frac{\sqrt{k}}{\sqrt{\sigma}}\frac{\partial_\tau \phi_u}{|\vec{\nabla}_{\vec{x}}\phi_u|} = \sqrt{\sigma k}\,\vec{\nabla}_{\vec{x}} \cdot \Big(\frac{\vec{\nabla}_{\vec{x}}\,\phi_u}{|\vec{\nabla}_{\vec{x}}\,\phi_u|}\Big) + kc_0 \tag{10}$$

The curvature eikonal equation is then defined as

$$\partial_t \phi_u = |\vec{\nabla}_{\vec{x}}\,\phi_u|\Big(\sigma \vec{\nabla}_{\vec{x}} \cdot \Big(\frac{\vec{\nabla}_{\vec{x}}\,\phi_u}{|\vec{\nabla}_{\vec{x}}\,\phi_u|}\Big) + \sqrt{\sigma k}\,c_0\Big),\ \mathcal{B} \times (0,T). \tag{11}$$

2.3.2 Energy minimization (link with the active contours without edges)

The method developed in image processing – for the detection of objects in an image – allows us to propose a state estimator of the eikonal equation presented in (11). The definitions of the geometrical notions used in this Section have been given in Section 2.1. The active contours without edges method, see [45, 8] consists in minimizing the following energy

$$E(\tilde{C}_1, \tilde{C}_2, \phi) = \mu\ \text{length}(\Gamma) + \nu\ \text{area}(\Omega^{\text{in}}) + \lambda_1 \int_{\Omega^{\text{in}}} \big(z(\vec{x}) - \tilde{C}_1\big)^2 d\vec{x} + \lambda_2 \int_{\mathcal{B}\setminus\overline{\Omega^{\text{in}}}} \big(z(\vec{x}) - \tilde{C}_2\big)^2 d\vec{x}. \tag{12}$$

The constants $\mu, \nu, \lambda_1, \lambda_2$ are positive in the active contours theory. Using (2), the length of Γ and the area of Ω^{in} are defined by

$$\text{length}(\Gamma) = \int \delta(\phi)|\vec{\nabla}\,\phi|d\vec{x} \text{ and } \text{area}(\Omega^{\text{in}}) = \int H(\phi)d\vec{x}.$$

The first two terms of this energy are regularization terms and the last two terms correspond to data discrepancy. Of course, in what follows we are interested in the last two terms. In a first stage, C_1 and C_2 are determined in order to have $(C_1, C_2) = \underset{\tilde{C}_1, \tilde{C}_2}{\text{Argmin}}\, E(\tilde{C}_1, \tilde{C}_2, \phi)$. This implies that

$$\frac{\partial E}{\partial \tilde{C}_1} = \lambda_1 \int_{\Omega^{\text{in}}} -2(z - \tilde{C}_1)d\vec{x} = 0 \text{ and } \frac{\partial E}{\partial \tilde{C}_2} = \lambda_2 \int_{\mathcal{B}\setminus\overline{\Omega^{\text{in}}}} -2(z - \tilde{C}_2)d\vec{x} = 0. \tag{13}$$

This allows to determine C_1 and C_2.

$$C_1 = \frac{\int_{\Omega^{\text{in}}} z d\vec{x}}{\int_{\Omega^{\text{in}}} d\vec{x}} = \frac{\int z H(\phi) d\vec{x}}{\int H(\phi) d\vec{x}} \text{ and } C_2 = \frac{\int_{\mathcal{B}\setminus\overline{\Omega^{\text{in}}}} z d\vec{x}}{\int_{\mathcal{B}\setminus\overline{\Omega^{\text{in}}}} d\vec{x}} = \frac{\int z(1-H(\phi)) d\vec{x}}{\int (1-H(\phi)) d\vec{x}}.$$

We remark that C_1 is the average of z inside Γ and C_2 is the average of z outside of Γ. In a second stage, we are interested in the minimum of $E_C(\phi) = E(C_1, C_2, \phi)$. In [58], a computation of the Fréchet derivative of E is done. Here, we just give the final result

$$\begin{aligned}\Big(\frac{\partial E_C}{\partial \phi}, \psi\Big) &= -\mu \int \delta(\phi) \vec{\nabla} \cdot \Big(\frac{\vec{\nabla}\,\phi}{|\vec{\nabla}\,\phi|}\Big)\psi d\vec{x} + \mu \int_{\partial\mathcal{B}} \frac{\delta(\phi)}{|\vec{\nabla}\,\phi|}\frac{\partial \phi}{\partial n}\psi d\vec{x}. + \nu \int \delta(\phi)\psi d\vec{x} \\ &\quad + \lambda_1 \int \big(z(\vec{x}) - C_1\big)^2 \delta(\phi)\psi d\vec{x} - \lambda_2 \int \big(z(\vec{x}) - C_2\big)^2 \delta(\phi)\psi d\vec{x}.\end{aligned}$$

However we use similar methods for the computation of others Fréchet derivatives presented in Appendix B. We remark that the two averages C_1 and C_2 verify (13) and that this implies that

$$\Big(\frac{\partial E_C}{\partial \phi}, \psi\Big) = \Big(\frac{\partial E}{\partial \phi}, \psi\Big). \tag{14}$$

Using a gradient projection method, see also [58], the minimization of the energy E_C (and also E using (14)) amounts to finding ϕ such that

$$\begin{cases} \dfrac{\partial \phi}{\partial t} &= \mu\,\delta(\phi)\vec{\nabla} \cdot \Big(\dfrac{\vec{\nabla}\,\phi}{|\vec{\nabla}\,\phi|}\Big) - \nu\delta(\phi) - \lambda_1 \delta(\phi)\big(z - C_1\big)^2 + \lambda_2 \delta(\phi)\big(z - C_2\big)^2, \text{ in } \mathcal{B} \\ \dfrac{\delta(\phi)}{|\vec{\nabla}\,\phi|}\dfrac{\partial \phi}{\partial n} &= 0, \ \textit{i.e.} \ \dfrac{\partial \phi}{\partial n} = 0 \text{ in } \partial\mathcal{B} \cap \Gamma. \end{cases} \tag{15}$$

We want to find a state estimator – denoted by $\phi_{\hat{u}}$ – of the curvature eikonal equation (11) using time activation maps. Using these observations, we define an object z_{ϕ_u} which varies in time. This object corresponds to the observations for the observer of the eikonal equation. As previously said, the choice of the exact definition of the object will be discussed. Using the previous analysis, we define

$$\Omega^{\text{in}}_{\phi_{\hat{u}}} = \{\vec{x} \in \mathcal{B},\, \phi_{\hat{u}}(\vec{x}) > 0\},$$

its boundary

$$\Gamma_{\phi_{\hat{u}}} = \{\vec{x} \in \mathcal{B},\, \phi_{\hat{u}}(\vec{x}) = 0\}.$$

and

$$C_1(\Omega^{\text{in}}_{\phi_{\hat{u}}}) = \frac{\int_{\Omega^{\text{in}}_{\phi_{\hat{u}}}} z_{\phi_u} d\vec{x}}{\int_{\Omega^{\text{in}}_{\phi_{\hat{u}}}} d\vec{x}} \text{ and } C_2(\Omega^{\text{in}}_{\phi_{\hat{u}}}) = \frac{\int_{\mathcal{B}\setminus\overline{\Omega^{\text{in}}}_{\phi_{\hat{u}}}} z_{\phi_u} d\vec{x}}{\int_{\mathcal{B}\setminus\overline{\Omega^{\text{in}}}_{\phi_{\hat{u}}}} d\vec{x}}$$

As we do not apprehend the influence of $|\vec{\nabla}\,\phi_{\hat{u}}|$ in

$$\delta(\phi_{\hat{u}}) = \frac{1}{|\vec{\nabla}\,\phi_{\hat{u}}|}\hat{\delta}(\Gamma_{\phi_{\hat{u}}}, \vec{x})$$

the state estimator for the eikonal equation is given by

$$\begin{aligned}\partial_t \phi_{\hat{u}} &= |\vec{\nabla}\,\phi_{\hat{u}}|\Big(\sigma\vec{\nabla} \cdot \Big(\frac{\vec{\nabla}\,\phi_{\hat{u}}}{|\vec{\nabla}\,\phi_{\hat{u}}|}\Big) + \sqrt{\sigma k}\, c_0\Big) \\ &\quad + \lambda\alpha(|\vec{\nabla}\,\phi_{\hat{u}}|)\hat{\delta}(\Gamma_{\phi_{\hat{u}}}, \vec{x})\Big(-\big(z_{\phi_u} - C_1(\Omega^{\text{in}}_{\phi_{\hat{u}}})\big)^2 + \big(z_{\phi_{\hat{u}}} - C_2(\Omega^{\text{in}}_{\phi_{\hat{u}}})\big)^2\Big),\ \mathcal{B} \times (0,T),\end{aligned} \tag{16}$$

where λ is a positive constant and α corresponds to a strictly positive function which will based on the mathematical analysis.

2.3.3 Observer for the reaction-diffusion model

We are interested in a state estimator for the reaction-diffusion model inspired from the state estimator of the curvature eikonal equation (11) defined in (16). We want to find the correction terms of the reaction-diffusion model which give the data terms in (16) during the asymptotic analysis. We want that the terms of the observer appear in the parenthesis before $\frac{\partial \hat{u}}{\partial \xi_1}$ in (6). We then assume that the state estimator of the reaction-diffusion has the following form

$$\begin{aligned}\partial_t \hat{u} = \vec{\nabla}_{\vec{x}} \cdot (\sigma \vec{\nabla}_{\vec{x}}\, \hat{u}) + k f(\hat{u}) \\ + \gamma \hat{\delta}(\Gamma_{\phi_{\hat{u}}}, \vec{x}) \frac{\alpha(|\vec{\nabla}_{\vec{x}}\, \phi_{\hat{u}}|)}{|\vec{\nabla}_{\vec{x}}\, \phi_{\hat{u}}|} \vec{\nabla}_{\vec{x}} g \cdot \vec{\nabla}_{\vec{x}}\, \hat{u} \big(\big(z_{\phi_{\hat{u}}} - C_1(\Omega^{\text{in}}_{\phi_{\hat{u}}})\big)^2 - \big(z_{\phi_{\hat{u}}} - C_2(\Omega^{\text{in}}_{\phi_{\hat{u}}})\big)^2 \big). \quad (17)\end{aligned}$$

and we use the asymptotic analysis presented in Section 2.3.1 in order to determine the function g. The same moving coordinate system gives

$$\begin{aligned}\frac{\partial \hat{u}}{\partial \tau} = \sigma a_{il} a_{ik} \frac{\partial^2 \hat{u}}{\partial \xi_l \partial \xi_k} + \sigma \frac{\partial a_{il}}{\partial x_i} \frac{\partial \hat{u}}{\partial \xi_l} + \frac{\partial X_j}{\partial \tau} a_{jk} \frac{\partial \hat{u}}{\partial \xi_k} + k f(\hat{u}) \\ + \gamma \hat{\delta}(\Gamma_{\phi_{\hat{u}}}, \vec{x}) \frac{\alpha(|\vec{\nabla}_{\vec{x}}\, \phi_{\hat{u}}|)}{|\vec{\nabla}_{\vec{x}}\, \phi_{\hat{u}}|} \big(\big(z_{\phi_{\hat{u}}} - C_1\big)^2 - \big(z_{\phi_{\hat{u}}} - C_2\big)^2 \big) a_{il} a_{ik} \frac{\partial g}{\partial \xi_k} \frac{\partial \hat{u}}{\partial \xi_l}. \quad (18)\end{aligned}$$

We suppose that $\hat{u}$ and g only depend on ξ_1 and the previous equation becomes

$$\begin{aligned}\sigma |\vec{a}|^2 \frac{\partial^2 \hat{u}}{\partial \xi_1^2} \\ + \left[\sigma \vec{\nabla}_{\vec{x}} \cdot \vec{a} + \frac{\partial \vec{X}}{\partial \tau} \cdot \vec{a} + \gamma \hat{\delta}(\Gamma_{\phi_{\hat{u}}}, \vec{x}) \frac{\alpha(|\vec{\nabla}_{\vec{x}}\, \phi_{\hat{u}}|)}{|\vec{\nabla}_{\vec{x}}\, \phi_{\hat{u}}|} \Big(\big(z_{\phi_{\hat{u}}} - C_1\big)^2 - \big(z_{\phi_{\hat{u}}} - C_2\big)^2 \Big) |\vec{a}|^2 \frac{\partial g}{\partial \xi_1} \right] \frac{\partial \hat{u}}{\partial \xi_1} \\ + k f(\hat{u}) = 0. \quad (19)\end{aligned}$$

As in Section 2.3.1, we choose the moving coordinate system in order to have $k = \sigma |\vec{a}|^2$ and using [1], there exists a constant c_0 depending on f such that

$$k c_0 = \sigma \vec{\nabla}_{\vec{x}} \cdot \vec{a} + \frac{\partial \vec{X}}{\partial \tau} \cdot \vec{a} + \gamma \hat{\delta}(\Gamma_{\phi_{\hat{u}}}, \vec{x}) \frac{\alpha(|\vec{\nabla}_{\vec{x}}\, \phi_{\hat{u}}|)}{|\vec{\nabla}_{\vec{x}}\, \phi_{\hat{u}}|} \Big(\big(z_{\phi_{\hat{u}}} - C_1\big)^2 - \big(z_{\phi_{\hat{u}}} - C_2\big)^2 \Big) \frac{k}{\sigma} \frac{\partial g}{\partial \xi_1}.$$

We suppose that $\partial_{\xi_1} g$ equals a positive constant $\tilde{c}$ and it follows that

$$k c_0 = \sigma \vec{\nabla}_{\vec{x}} \cdot \vec{a} + \frac{\partial \vec{X}}{\partial \tau} \cdot \vec{a} + \gamma \hat{\delta}(\Gamma_{\phi_{\hat{u}}}, \vec{x}) \frac{\alpha(|\vec{\nabla}_{\vec{x}}\, \phi_{\hat{u}}|)}{|\vec{\nabla}_{\vec{x}}\, \phi_{\hat{u}}|} \Big(\big(z_{\phi_{\hat{u}}} - C_1\big)^2 - \big(z_{\phi_{\hat{u}}} - C_2\big)^2 \Big) \frac{k}{\sigma} \tilde{c}. \quad (20)$$

We recall that

$$\vec{a} = -\frac{\sqrt{k}}{\sqrt{\sigma}} \frac{\vec{\nabla}_{\vec{x}}\, \phi_{\hat{u}}}{|\vec{\nabla}_{\vec{x}}\, \phi_{\hat{u}}|} \text{ and } \vec{\nabla}_{\vec{x}} \phi_{\hat{u}} \cdot \frac{\partial \vec{X}}{\partial \tau} + \frac{\partial \phi_{\hat{u}}}{\partial \tau} = 0.$$

We finally obtain for the eikonal equation derived from (17)

$$\begin{aligned}\partial_t \phi_{\hat{u}} = |\vec{\nabla}_{\vec{x}}\, \phi_{\hat{u}}| \Big(\sigma \vec{\nabla}_{\vec{x}} \cdot \Big(\frac{\vec{\nabla}_{\vec{x}}\, \phi_{\hat{u}}}{|\vec{\nabla}_{\vec{x}}\, \phi_{\hat{u}}|} \Big) + \sqrt{D k}\, c_0 \\ + \gamma \hat{\delta}(\Gamma_{\phi_{\hat{u}}}, \vec{x}) \frac{\alpha(|\vec{\nabla}_{\vec{x}}\, \phi_{\hat{u}}|)}{|\vec{\nabla}_{\vec{x}}\, \phi_{\hat{u}}|} \Big(-\big(z_{\phi_{\hat{u}}} - C_1(\Omega^{\text{in}}_{\phi_{\hat{u}}})\big)^2 + \big(z_{\phi_{\hat{u}}} - C_2(\Omega^{\text{in}}_{\phi_{\hat{u}}})\big)^2 \Big) \frac{\sqrt{k}}{\sqrt{D}} \tilde{c} \Big).\end{aligned}$$

and with $\lambda = \gamma \frac{\sqrt{k}}{\sqrt{D}} \tilde{c}$ for $i = 1, 2$, we find (16). Using

$$\partial_{\xi_1} g = \tilde{c} \geq 0 \text{ and } \partial_{\xi_2} g = \partial_{\xi_3} g = 0,$$

we obtain

$$\partial_{x_1} g = \tilde{c} a_{11},\ \partial_{x_2} g = \tilde{c} a_{21} \text{ and } \partial_{x_3} g = \tilde{c} a_{31}.$$

and we can determine

$$\vec{\nabla}_{\vec{x}} g = \tilde{c} \vec{a} = -\tilde{c} \frac{\sqrt{k}}{\sqrt{D}} \frac{\vec{\nabla}_{\vec{x}} \phi_{\hat{u}}}{|\vec{\nabla}_{\vec{x}} \phi_{\hat{u}}|}.$$

With this analysis, we obtain an observer for the reaction-diffusion model (with the more general case for the diffusion tensor)

$$\begin{aligned} \partial_t \hat{u} = {} & \vec{\nabla}_{\vec{x}} \cdot (\vec{\vec{\sigma}} \cdot \vec{\nabla}_{\vec{x}} \hat{u}) + k f(\hat{u}) \\ & + \lambda \hat{\delta}(\Gamma_{\phi_{\hat{u}}}, \vec{x}) \frac{\alpha(|\vec{\nabla}_{\vec{x}} \phi_{\hat{u}}|)}{|\vec{\nabla}_{\vec{x}} \phi_{\hat{u}}|} \frac{\vec{\nabla}_{\vec{x}} \phi_{\hat{u}}}{|\vec{\nabla}_{\vec{x}} \phi_{\hat{u}}|} \cdot \vec{\nabla}_{\vec{x}} \hat{u} \Big(-\big(z_{\phi_{\hat{u}}} - C_1(\Omega^{\text{in}}_{\phi_{\hat{u}}})\big)^2 + \big(z_{\phi_{\hat{u}}} - C_2(\Omega^{\text{in}}_{\phi_{\hat{u}}})\big)^2 \Big), \end{aligned} \quad (21)$$

with $\phi_{\hat{u}}$ a level set associated with $\hat{u}$, $z_{\phi_{\hat{u}}}$ an object defined from the data and λ a positive constant.

This equation can be simplified using $\hat{u} - c_{\text{th}}$ as a level set associated with $\hat{u}$. Furthermore, we directly have $\Omega^{\text{in}}_{\phi_{\hat{u}}} = \Omega^{\text{in}}_{\hat{u}}$, $\Gamma_{\phi_{\hat{u}}} = \Gamma_{\hat{u}}$ and $z_{\phi_{\hat{u}}} = z_{\hat{u}}$ using the definitions given in Section 2.2. Then (21) becomes (4) as expected.

2.4 Mathematical analysis of the state observer

2.4.1 Model of $\tilde{u} = u - \hat{u}$

We want to study the error defined by $\tilde{u} = u - \hat{u}$. The target model verifying (3) and the observer model (4), the error verifies the following model

$$\begin{cases} \partial_t \tilde{u} - \vec{\nabla} \cdot (\vec{\vec{\sigma}} \cdot \vec{\nabla} \tilde{u}) &= k(f(u) - f(\hat{u})) & \\ +\lambda \hat{\delta}(\Gamma_{\hat{u}}, \vec{x}) \alpha(|\vec{\nabla} \hat{u}|) \Big(\big(z_u - C_1(\Omega^{\text{in}}_{\hat{u}})\big)^2 - \big(z_u - C_2(\Omega^{\text{in}}_{\hat{u}})\big)^2 \Big), & & \mathcal{B} \times (0,T), \\ (\vec{\vec{\sigma}} \cdot \vec{\nabla} \tilde{u}) \cdot \vec{n} &= 0, & \partial\mathcal{B} \times (0,T), \\ \tilde{u}(\vec{x}, 0) &= u_0(\vec{x}) - \hat{u}_0(\vec{x}), & \mathcal{B}. \end{cases}$$

This problem can be rewritten in variational form and it becomes, find $\tilde{u}(\cdot, t) \in H^1(\mathcal{B})$ such that $\forall \chi \in H^1(\mathcal{B})$, $\forall t \in [0, T]$,

$$\begin{aligned} \int_{\mathcal{B}} \partial_t \tilde{u}\, \chi \, d\vec{x} + \int_{\mathcal{B}} \vec{\vec{\sigma}} \cdot \vec{\nabla} \tilde{u} \cdot \vec{\nabla} \chi \, d\vec{x} &= \int_{\mathcal{B}} k(f(u) - f(\hat{u}))\, \chi \, d\vec{x} \\ &+ \lambda \int_{\Gamma_{\hat{u}}} \alpha(|\vec{\nabla} \hat{u}|) \Big(\big(z_u - C_1(\Omega^{\text{in}}_{\hat{u}})\big)^2 - \big(z_u - C_2(\Omega^{\text{in}}_{\hat{u}})\big)^2 \Big) \chi \, d\Gamma_{\hat{u}}, \end{aligned} \quad (22)$$

where $u(\cdot, t)$, $\hat{u}(\cdot, t) \in H^1(\mathcal{B})$ respectively satisfy, $\forall \chi \in H^1(\mathcal{B})$, $\forall t \in [0, T]$,

$$\int_{\mathcal{B}} \partial_t u\, \chi \, d\vec{x} + \int_{\mathcal{B}} \vec{\vec{\sigma}} \cdot \vec{\nabla} u \cdot \vec{\nabla} \chi \, d\vec{x} = \int_{\mathcal{B}} k f(u)\, \chi \, d\vec{x}, \quad (23)$$

and

$$\begin{aligned} \int_{\mathcal{B}} \partial_t \hat{u}\, \chi \, d\vec{x} + \int_{\mathcal{B}} \vec{\vec{\sigma}} \cdot \vec{\nabla} \hat{u} \cdot \vec{\nabla} \chi \, d\vec{x} &= \int_{\mathcal{B}} k f(\hat{u})\, \chi \, d\vec{x} \\ &+ \lambda \int_{\Gamma_{\hat{u}}} \alpha(|\vec{\nabla} \hat{u}|) \Big(-\big(z_u - C_1(\Omega^{\text{in}}_{\hat{u}})\big)^2 + \big(z_u - C_2(\Omega^{\text{in}}_{\hat{u}})\big)^2 \Big) \chi \, \Gamma_{\hat{u}}. \end{aligned} \quad (24)$$

In order to have a consistent observer we expect

$$\hat{\delta}(\Gamma_u,\,\vec{x})\alpha(|\vec{\nabla}\,u|)\Big(\big(z_u-C_1(\Omega_u^{\rm in})\big)^2-\big(z_u-C_2(\Omega_u^{\rm in})\big)^2\Big)=0,$$

i.e. the correction term equals 0 for the target solution. In order to have this property, we need some assumptions on z_u.

Assumption 1 *If $\bar{C}_1$ is the average of z_u in the interior of the object defined by Ω_u^{in} and $\bar{C}_2$ the average of z_u in the exterior $\mathcal{B}\setminus\overline{\Omega_u^{in}}$, the value of z_u on the front equals*

$$\frac{\bar{C}_1+\bar{C}_2}{2}.$$

If Assumption 1 holds, the term

$$\hat{\delta}(\Gamma_u,\,\vec{x})\Big(\big(z_u-C_1(\Omega_u^{\rm in})\big)^2-\big(z_u-C_2(\Omega_u^{\rm in})\big)^2\Big),$$

then becomes

$$\left(\frac{\bar{C}_1+\bar{C}_2}{2}-\bar{C}_1\right)^2-\left(\frac{\bar{C}_1+\bar{C}_2}{2}-\bar{C}_2\right)^2=0,\ \text{on}\ \Gamma_u.$$

Now, we want to justify the observer model by showing that the correction term of the observer model adds dissipation in the model of the error defined by $\tilde{u}=u-\hat{u}$. Indeed, we want to prove that by adding the correction term, the error between $\hat{u}$ and u decreases. We compute the energy of the error model (22) applying $\chi=\tilde{u}$

$$\begin{aligned}\frac{1}{2}\int_{\mathcal{B}}\partial_t\tilde{u}^2d\vec{x} &= -\int_{\mathcal{B}}\vec{\vec{\sigma}}\cdot\vec{\nabla}\,\tilde{u}\cdot\vec{\nabla}\,\tilde{u}\,d\vec{x}+\int_{\mathcal{B}}k(f(u)-f(\hat{u}))\,\tilde{u}\,d\vec{x}\\ &+ \lambda\int_{\Gamma_{\hat{u}}}\alpha(|\vec{\nabla}\,\hat{u}|)\Big(\big(z_u-C_1(\Omega_{\hat{u}}^{\rm in})\big)^2-\big(z_u-C_2(\Omega_{\hat{u}}^{\rm in})\big)^2\Big)\,\tilde{u}\,d\Gamma_{\hat{u}}\end{aligned}\tag{25}$$

We denote by $\mathcal{Q}^{\Gamma_{\hat{u}}}(\chi)$ the following form

$$\mathcal{Q}^{\Gamma_{\hat{u}}}(\chi)=\int_{\Gamma_{\hat{u}}}\alpha(|\vec{\nabla}\,\hat{u}|)\Big(\big(z_u-C_1(\Omega_{\hat{u}}^{\rm in})\big)^2-\big(z_u-C_2(\Omega_{\hat{u}}^{\rm in})\big)^2\Big)\,\chi\,d\Gamma_{\hat{u}}.$$

If $\mathcal{Q}^{\Gamma_{\hat{u}}}(\tilde{u})$ is negative, the correction term adds dissipation in the error model and this provides a preliminary mathematical justification for the choice of this observer. However, this sign cannot be determined directly and we have to linearize $\mathcal{Q}^{\Gamma_{\hat{u}}}(\tilde{u})$.

Link between the shape derivative and the level set formalism This section gives a mathematically justified linearization of

$$\mathcal{Q}^{\Gamma_{\hat{u}}}(\chi)=\int_{\Gamma_{\hat{u}}}\alpha(|\vec{\nabla}\,\hat{u}|)\Big(\big(z_u-C_1(\Omega_{\hat{u}}^{\rm in})\big)^2-\big(z_u-C_2(\Omega_{\hat{u}}^{\rm in})\big)^2\Big)\,\chi\,d\Gamma_{\hat{u}}.\tag{26}$$

using the theory of the shape derivative. The main definitions and the main propositions of the shape derivative theory are given in Appendix A. In Appendix B we provide a comparison with a formal differentiation of $\mathcal{Q}^{\Gamma_{\hat{u}}}$ written as an integral over the whole domain, *i.e.*

$$\mathcal{Q}^{\Gamma_{\hat{u}}}(\chi)=\int_{\mathcal{B}}\delta(\hat{u}-c_{\rm th})|\vec{\nabla}\,\hat{u}|\alpha(|\vec{\nabla}\,\hat{u}|)\Big(\big(z_u-C_1(\hat{u}-c_{\rm th})\big)^2-\big(z_u-C_2(\hat{u}-c_{\rm th})\big)^2\Big)\,\chi\,d\vec{x},$$

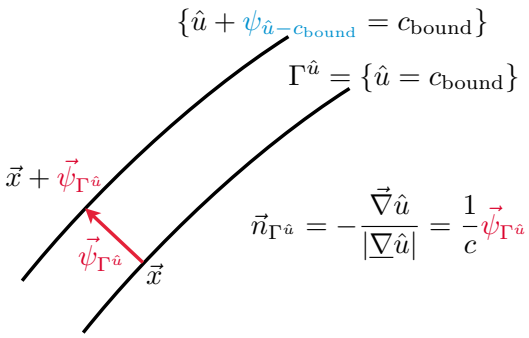

Figure 2: Link between the level set and shape derivatives

where C_1 and C_2 as seen dependent of $\hat{u}$ by

$$C_1(\hat{u} - c_{\text{th}}) = \frac{\int_{\mathcal{B}} z_{\phi_u} H(\hat{u} - c_{\text{th}}) d\vec{x}}{\int_{\mathcal{B}} H(\hat{u} - c_{\text{th}}) d\vec{x}} \text{ and } C_2(\hat{u} - c_{\text{th}}) = \frac{\int_{\mathcal{B}} z_{\phi_u} \big(1 - H(\hat{u} - c_{\text{th}})\big) d\vec{x}}{\int_{\mathcal{B}} \big(1 - H(\hat{u} - c_{\text{th}})\big) d\vec{x}}.$$

The two methods of differentiation are related and the link is important in order to compare

$$d_{\Gamma_{\hat{u}}} \mathcal{Q}^{\Gamma_{\hat{u}}} \text{ and } d_{\hat{u}} \mathcal{Q}^{\Gamma_{\hat{u}}}.$$

Indeed, the shape derivative theory allows to obtain a justified differentiation of the first term whereas the second term appears more directly when we linearize $\mathcal{Q}^{\Gamma_{\hat{u}}}$ with respect to $\hat{u}$. Using Figure 2, we define $\vec{\psi}_{\Gamma_{\hat{u}}}$ the vector flow of shape derivative, and $\psi_{\hat{u}-c_{\text{th}}}$ the corresponding level set variation. We want to find the relation between them and we have

$$\big((\hat{u} - c_{\text{th}}) + \psi_{\hat{u}-c_{\text{th}}}\big)(\vec{x} + \vec{\psi}_{\Gamma_{\hat{u}}}) = 0,$$

and this implies

$$\hat{u}(\vec{x}) - c_{\text{th}} + \vec{\nabla}\,\hat{u}(\vec{x}) \cdot \vec{\psi}_{\Gamma_{\hat{u}}} + \psi_{\hat{u}-c_{\text{th}}}(\vec{x}) + O(\|(\psi_{\hat{u}-c_{\text{th}}}, \vec{\psi}_{\Gamma_{\hat{u}}})\|^2) = 0,$$

and

$$\vec{\nabla}\,\hat{u}(\vec{x}) \cdot \vec{\psi}_{\Gamma_{\hat{u}}} + \psi_{\hat{u}-c_{\text{th}}}(\vec{x}) + O(\|(\psi_{\hat{u}-c_{\text{th}}}, \vec{\psi}_{\Gamma_{\hat{u}}})\|^2) = 0.$$

We see that we can look for a constant c such that

$$\vec{\psi}_{\Gamma_{\hat{u}}} = c\,\vec{n}_{\Gamma_{\hat{u}}} = -c\,\frac{\vec{\nabla}\hat{u}}{|\vec{\nabla}\hat{u}|}.$$

We obtain

$$c = \frac{\psi_{\hat{u}-c_{\text{th}}}}{|\vec{\nabla}\hat{u}|}$$

and we finally have

$$\vec{\psi}_{\Gamma_{\hat{u}}} = \frac{\psi_{\hat{u}-c_{\text{th}}}}{|\vec{\nabla}\hat{u}|}\vec{n}_{\Gamma_{\hat{u}}}. \tag{27}$$

The relation between the shape and the level set derivatives is then given by

$$\Big(d_{\hat{u}}\mathcal{Q}^{\Gamma_{\hat{u}}}(\chi);\, \psi_{\hat{u}-c_{\text{th}}}\Big) = \Big(d_{\Gamma_{\hat{u}}}\mathcal{Q}^{\Gamma_{\hat{u}}}(\chi);\, \frac{\psi_{\hat{u}-c_{\text{th}}}}{|\vec{\nabla}\hat{u}|}\vec{n}_{\Gamma_{\hat{u}}}\Big). \tag{28}$$

2.4.2 Linearization of $\mathcal{Q}^{\Gamma_{\hat{u}}}$

Shape derivative of $\mathcal{Q}^{\Gamma_{\hat{u}}}$ In Appendix A the shape derivative theory is presented. More precisely, the proof of the following proposition is given in Section A.2.

Proposition 1 *The shape derivative of $\mathcal{Q}^{\Gamma_{\hat{u}}}$ is*

$$
\begin{aligned}
\big(d_{\Gamma_{\hat{u}}}\mathcal{Q}^{\Gamma_{\hat{u}}}(\chi);\vec{\psi}_{\Gamma_{\hat{u}}}\big) &= \int_{\Gamma_{\hat{u}}} \kappa\,\alpha(|\vec{\nabla}\,\hat{u}|)\Big(\big(z_u - C_1(\Omega_{\hat{u}}^{in})\big)^2 - \big(z_u - C_2(\Omega_{\hat{u}}^{in})\big)^2\Big)\,\chi\;\vec{\psi}_{\Gamma_{\hat{u}}}\cdot\vec{n}_{\Gamma_{\hat{u}}}\,d\Gamma_{\hat{u}}\\
&- \int_{\Gamma_{\hat{u}}} \alpha'(|\vec{\nabla}\,\hat{u}|)\Big(\big(z_u - C_1(\Omega_{\hat{u}}^{in})\big)^2 - \big(z_u - C_2(\Omega_{\hat{u}}^{in})\big)^2\Big)\,\chi\;\partial_{\vec{n}_\Gamma}\big(\vec{\psi}_{\Gamma_{\hat{u}}}\cdot\vec{n}_\Gamma|\vec{\nabla}\,\hat{u}|\big)\,d\Gamma_{\hat{u}}\\
&- \int_{\Gamma_{\hat{u}}} \alpha'(|\vec{\nabla}\,\hat{u}|)\partial^2_{\vec{n}_{\Gamma_{\hat{u}}}}\hat{u}\Big(\big(z_u - C_1(\Omega_{\hat{u}}^{in})\big)^2 - \big(z_u - C_2(\Omega_{\hat{u}}^{in})\big)^2\Big)\,\chi\;\vec{\psi}_{\Gamma_{\hat{u}}}\cdot\vec{n}_{\Gamma_{\hat{u}}}\,d\Gamma_{\hat{u}}\\
&+2\int_{\Gamma_{\hat{u}}} \alpha(|\vec{\nabla}\,\hat{u}|)\vec{\nabla}z_u\cdot\vec{n}_{\Gamma_{\hat{u}}}\big(C_2(\Omega_{\hat{u}}^{in}) - C_1(\Omega_{\hat{u}}^{in})\big)\,\chi\;\vec{\psi}_{\Gamma_{\hat{u}}}\cdot\vec{n}_{\Gamma_{\hat{u}}}\,d\Gamma_{\hat{u}}\\
&-2\frac{1}{|\Omega_{\hat{u}}^{in}|}\int_{\Gamma_{\hat{u}}}\big(z_u - C_1(\Omega_{\hat{u}}^{in})\big)\;\vec{\psi}_{\Gamma_{\hat{u}}}\cdot\vec{n}_{\Gamma_{\hat{u}}}\,d\Gamma_{\hat{u}}\int_{\Gamma_{\hat{u}}}\alpha(|\vec{\nabla}\,\hat{u}|)\big(z_u - C_1(\Omega_{\hat{u}}^{in})\big)\chi\,d\Gamma_{\hat{u}}\\
&-2\frac{1}{|\mathcal{B}\setminus\overline{\Omega_{\hat{u}}^{in}}|}\int_{\Gamma_{\hat{u}}}\big(z_u - C_2(\Omega_{\hat{u}}^{in})\big)\;\vec{\psi}_{\Gamma_{\hat{u}}}\cdot\vec{n}_{\Gamma_{\hat{u}}}\,d\Gamma_{\hat{u}}\int_{\Gamma_{\hat{u}}}\alpha(|\vec{\nabla}\,\hat{u}|)\big(z_u - C_2(\Omega_{\hat{u}}^{in})\big)\chi\,d\Gamma_{\hat{u}}\\
&+\int_{\Gamma_{\hat{u}}}\alpha(|\vec{\nabla}\,\hat{u}|)\Big(\big(z_u - C_1(\Omega_{\hat{u}}^{in})\big)^2 - \big(z_u - C_2(\Omega_{\hat{u}}^{in})\big)^2\Big)\,\vec{\nabla}\chi\cdot\vec{\psi}_{\Gamma_{\hat{u}}}\,d\Gamma_{\hat{u}}.
\end{aligned}
$$

Linearization of $\mathcal{Q}^{\Gamma_{\hat{u}}}$ We recall that the objective is to linearize $\mathcal{Q}^{\Gamma_{\hat{u}}}(\tilde{u})$. We have

$$
\mathcal{Q}^{\Gamma_{\hat{u}}}(\tilde{u}) \;\approx\; \mathcal{Q}^{\Gamma_u}(u) + \Big(d_u\mathcal{Q}^{\Gamma_u}(\tilde{u});\,\hat{u}-u\Big).
$$

Using (28) and Assumption 1, we obtain

$$
\begin{aligned}
\mathcal{Q}^{\Gamma_{\hat{u}}}(\tilde{u}) &\approx \mathcal{Q}^{\Gamma_u}(u) + \Big(d_{\Gamma_u}\mathcal{Q}^{\Gamma_u}(\tilde{u});\,\frac{\hat{u}-u}{|\vec{\nabla}u|}\vec{n}_{\Gamma_u}\Big)\\
&\approx -\Big(d_{\Gamma_u}\mathcal{Q}^{\Gamma_u}(\tilde{u});\,\frac{\tilde{u}}{|\vec{\nabla}u|}\vec{n}_{\Gamma_u}\Big).
\end{aligned}
$$

As a consequence, we are interested in the term

$$
\big(d_{\Gamma_u}\mathcal{Q}^{\Gamma_u}(\tilde{u});\,\frac{\tilde{u}}{|\vec{\nabla}u|}\vec{n}_{\Gamma_u}\big).
$$

Applying Proposition 1 and using again Assumption 1, we have

$$
\begin{aligned}
\big(d\mathcal{Q}^{\Gamma_u}(\tilde{u});\,\frac{\tilde{u}}{|\vec{\nabla}u|}\vec{n}_{\Gamma_u}\big) &= 2\int_{\Gamma_u}\frac{\alpha(|\vec{\nabla}\,u|)}{|\vec{\nabla}u|}\vec{\nabla}z_u(\vec{x})\cdot\vec{n}_{\Gamma_u}\big(C_2(\Gamma_u) - C_1(\Gamma_u)\big)\,\tilde{u}^2\,d\Gamma_u\\
&-2\frac{1}{|\Omega_{\hat{u}}^{\mathrm{in}}|}\int_{\Gamma_u}\frac{1}{|\vec{\nabla}u|}\big(z_u(\vec{x}) - C_1(\Gamma_u)\big)\,\tilde{u}\,d\Gamma_u\int_{\Gamma_u}\alpha(|\vec{\nabla}\,u|)\big(z_u(\vec{x}) - C_1(\Gamma_u)\big)\tilde{u}\,d\Gamma_u\\
&-2\frac{1}{|\mathcal{B}\setminus\overline{\Omega_{\hat{u}}^{\mathrm{in}}}|}\int_{\Gamma_u}\frac{1}{|\vec{\nabla}u|}\big(z_u(\vec{x}) - C_2(\Gamma_u)\big)\,\tilde{u}\,d\Gamma_u\int_{\Gamma_u}\alpha(|\vec{\nabla}\,u|)\big(z_u(\vec{x}) - C_2(\Gamma_u)\big)\tilde{u}\,d\Gamma_u \qquad (29)\\
&=2\big(\bar{C}_2 - \bar{C}_1\big)\int_{\Gamma_u}\frac{\alpha(|\vec{\nabla}\,u|)}{|\vec{\nabla}u|}\vec{\nabla}z_u(\vec{x})\cdot\vec{n}_{\Gamma_u}\,\tilde{u}^2\,d\Gamma_u\\
&-2\Big(\frac{\bar{C}_1 - \bar{C}_2}{2}\Big)^2\Big(\frac{1}{|\Omega_{\hat{u}}^{\mathrm{in}}|} + \frac{1}{|\mathcal{B}\setminus\overline{\Omega_{\hat{u}}^{\mathrm{in}}}|}\Big)\int_{\Gamma_u}\frac{1}{|\vec{\nabla}u|}\,\tilde{u}\,d\Gamma_u\int_{\Gamma_u}\alpha(|\vec{\nabla}\,u|)\tilde{u}\,d\Gamma_u.
\end{aligned}
$$

In Appendix B, the same result is found by an alternative method using a direct differentiation of the expression in (26).

2.4.3 Stabilizing state estimator

We recall that we want to prove that $\mathcal{Q}^{\Gamma_{\hat{u}}}(\tilde{u})$ which appears in the error model (25) is negative in order to show that this correction term adds dissipation in this model. The sign cannot be directly determined and we linearize $\mathcal{Q}^{\Gamma_{\hat{u}}}(\tilde{u})$ as in the previous part. We obtain

$$\mathcal{Q}^{\Gamma_{\hat{u}}}(\tilde{u}) \approx -\Big(d_{\Gamma_u}\mathcal{Q}^{\Gamma_u}(\tilde{u});\ \frac{\tilde{u}}{|\vec{\nabla}u|}\vec{n}_{\Gamma_u}\Big), \tag{30}$$

with $\big(d_{\Gamma_u}\mathcal{Q}^{\Gamma_u}(\tilde{u});\ \frac{\tilde{u}}{|\vec{\nabla}u|}\vec{n}_{\Gamma_u}\big)$ given in (29). The following lemma gives the negativity of $\mathcal{Q}^{\Gamma_{\hat{u}}}(\tilde{u})$ by proving the positivity of $\big(d_{\Gamma_u}\mathcal{Q}^{\Gamma_u}(\tilde{u});\ \frac{\tilde{u}}{|\vec{\nabla}u|}\vec{n}_{\Gamma_u}\big)$ under assumptions. Let h_{z_u} be the wavelength of the front z_u defined by

$$\frac{1}{h_{z_u}} = \min_{\vec{x}\in\Gamma_u} \frac{|\vec{\nabla} z_u(\vec{x})\cdot\vec{n}|}{|\bar{C}_1 - \bar{C}_2|},$$

which means that the steeper z_u, the smaller h_{z_u}. An example of z_u in 1D is given in what follows.

Lemma 1 *If the condition*

$$\frac{1}{h_{z_u}} \geq \frac{1}{4}\Big(\frac{1}{|\Omega_{\hat{u}}^{in}|} + \frac{1}{|\mathcal{B}\setminus\overline{\Omega_{\hat{u}}^{in}}|}\Big)\Big(\int_{\Gamma_u}\alpha(|\vec{\nabla}\,u|)|\vec{\nabla}u|\,d\Gamma_u\ \int_{\Gamma_u}\frac{1}{\alpha(|\vec{\nabla}\,u|)|\vec{\nabla}u|}\,d\Gamma_u\Big)^{\frac{1}{2}}, \tag{31}$$

holds, then $\big(d_{\Gamma_u}\mathcal{Q}^{\Gamma_u}(\tilde{u});\ \frac{\tilde{u}}{|\vec{\nabla}u|}\vec{n}_{\Gamma_u}\big)$ *is positive.*

Proof Using the Cauchy-Schwarz inequality, we have

$$\begin{aligned}
&\Big|\int_{\Gamma_u}\frac{1}{|\vec{\nabla}u|}\tilde{u}\,d\Gamma_u\ \int_{\Gamma_u}\alpha(|\vec{\nabla}\,u|)\tilde{u}\,d\Gamma_u\Big| \\
&\quad= \Big|\int_{\Gamma_u}\sqrt{\frac{\alpha(|\vec{\nabla}\,u|)}{|\vec{\nabla}u|}}\tilde{u}\frac{1}{\sqrt{\alpha(|\vec{\nabla}\,u|)|\vec{\nabla}u|}}\,d\Gamma_u\ \int_{\Gamma_u}\sqrt{\frac{\alpha(|\vec{\nabla}\,u|)}{|\vec{\nabla}u|}}\tilde{u}\sqrt{\alpha(|\vec{\nabla}\,u|)|\vec{\nabla}u|}\,d\Gamma_u\Big| \\
&\quad\leq \int_{\Gamma_u}\frac{\alpha(|\vec{\nabla}\,u|)}{|\vec{\nabla}u|}\tilde{u}^2\,d\Gamma_u\Big(\int_{\Gamma_u}\alpha(|\vec{\nabla}\,u|)|\vec{\nabla}u|\,d\Gamma_u\ \int_{\Gamma_u}\frac{1}{\alpha(|\vec{\nabla}\,u|)|\vec{\nabla}u|}\,d\Gamma_u\Big)^{\frac{1}{2}}.
\end{aligned}$$

The lemma is then deduced by

$$
\begin{aligned}
&\big(d_{\Gamma_u}\mathcal{Q}^{\Gamma_u}(\tilde{u});\, \frac{\tilde{u}}{|\vec{\nabla}u|}\vec{n}_{\Gamma_u}\big) = 2\big(\bar{C}_1-\bar{C}_2\big)^2\frac{1}{h_{z_u}}\int_{\Gamma_u}\frac{\alpha(|\vec{\nabla}\,u|)}{|\vec{\nabla}u|}\tilde{u}^2\,d\Gamma_u\\
&-2\Big(\frac{\bar{C}_1-\bar{C}_2}{2}\Big)^2\Big(\frac{1}{|\Omega^{\rm in}_{\hat{u}}|}+\frac{1}{|\mathcal{B}\setminus\overline{\Omega^{\rm in}_{\hat{u}}}|}\Big)\int_{\Gamma_u}\frac{1}{|\vec{\nabla}u|}\,\tilde{u}\,d\Gamma_u\int_{\Gamma_u}\alpha(|\vec{\nabla}\,u|)\tilde{u}\,d\Gamma_u\\
&\geq 2\big(\bar{C}_1-\bar{C}_2\big)^2\frac{1}{h_{z_u}}\int_{\Gamma_u}\frac{\alpha(|\vec{\nabla}\,u|)}{|\vec{\nabla}u|}\tilde{u}^2\,d\Gamma_u\\
&-2\Big(\frac{\bar{C}_1-\bar{C}_2}{2}\Big)^2\Big(\frac{1}{|\Omega^{\rm in}_{\hat{u}}|}+\frac{1}{|\mathcal{B}\setminus\overline{\Omega^{\rm in}_{\hat{u}}}|}\Big)\Big(\int_{\Gamma_u}\alpha(|\vec{\nabla}\,u|)|\vec{\nabla}u|\,d\Gamma_u\int_{\Gamma_u}\frac{1}{\alpha(|\vec{\nabla}\,u|)|\vec{\nabla}u|}\,d\Gamma_u\Big)^{\frac{1}{2}}\\
&\qquad\times\int_{\Gamma_u}\frac{\alpha(|\vec{\nabla}\,u|)}{|\vec{\nabla}u|}\tilde{u}^2\,d\Gamma_u\\
&\geq\left(\frac{1}{h_{z_u}}-\frac{1}{4}\Big(\frac{1}{|\Omega^{\rm in}_{\hat{u}}|}+\frac{1}{|\mathcal{B}\setminus\overline{\Omega^{\rm in}_{\hat{u}}}|}\Big)\Big(\int_{\Gamma_u}\alpha(|\vec{\nabla}\,u|)|\vec{\nabla}u|\,d\Gamma_u\int_{\Gamma_u}\frac{1}{\alpha(|\vec{\nabla}\,u|)|\vec{\nabla}u|}\,d\Gamma_u\Big)^{\frac{1}{2}}\right)\\
&\qquad\times 2\big(\bar{C}_1-\bar{C}_2\big)^2\int_{\Gamma_u}\frac{\alpha(|\vec{\nabla}\,u|)}{|\vec{\nabla}u|}\tilde{u}^2\,d\Gamma_u
\end{aligned}
$$

□

Remark 2 *Using the Cauchy-Schwarz inequality, we have*

$$|\Gamma_u|^2=\Big(\int_{\Gamma_u}\frac{\sqrt{f}}{\sqrt{f}}d\Gamma_u\Big)^2\leq\int_{\Gamma_u}f\,d\Gamma_u\int_{\Gamma_u}\frac{1}{f}\,d\Gamma_u$$

and we deduce

$$\Big(\int_{\Gamma_u}\alpha(|\vec{\nabla}\,u|)|\vec{\nabla}u|\,d\Gamma_u\int_{\Gamma_u}\frac{1}{\alpha(|\vec{\nabla}\,u|)|\vec{\nabla}u|}\,d\Gamma_u\Big)^{\frac{1}{2}}\geq|\Gamma_u|. \tag{32}$$

The particular case

$$\alpha(|\vec{\nabla}\,u|)=\frac{1}{|\vec{\nabla}\,u|}, \tag{33}$$

verifies the case of equality in (32)*. This means that this choice imposes fewer restrictions on the data* z_u*.*

Theorem 1 *If* z_u *verifies* (31)*, the term*

$$\mathcal{Q}^{\Gamma_{\hat{u}}}(\chi)=\int_{\Gamma_{\hat{u}}}\alpha(|\vec{\nabla}\,\hat{u}|)\Big(\big(z_u-C_1(\Omega^{in}_{\hat{u}})\big)^2-\big(z_u-C_2(\Omega^{in}_{\hat{u}})\big)^2\Big)\,\chi\,d\Gamma_{\hat{u}}.$$

is a stabilizing term of the model of the error $\tilde{u}=u-\hat{u}$*.*

Proof We start with (25) and by using (30) which linearizes the correction term, and we have

$$
\begin{aligned}
\frac{1}{2}\int_{\mathcal{B}}\partial_t\tilde{u}^2d\vec{x} &= -\int_{\mathcal{B}}\vec{\bar{\sigma}}\cdot\vec{\nabla}\,\tilde{u}\cdot\vec{\nabla}\,\tilde{u}\,d\vec{x}+\int_{\mathcal{B}}k(f(u)-f(\hat{u}))\,\tilde{u}\,d\vec{x}\\
&\quad-\Big(d_{\Gamma_u}\mathcal{Q}^{\Gamma_u}(\tilde{u});\,\frac{\tilde{u}}{|\vec{\nabla}u|}\vec{n}_{\Gamma_u}\Big)+O(\|\tilde{u}\|^2).
\end{aligned}
$$

We can conclude using Lemma 1. □

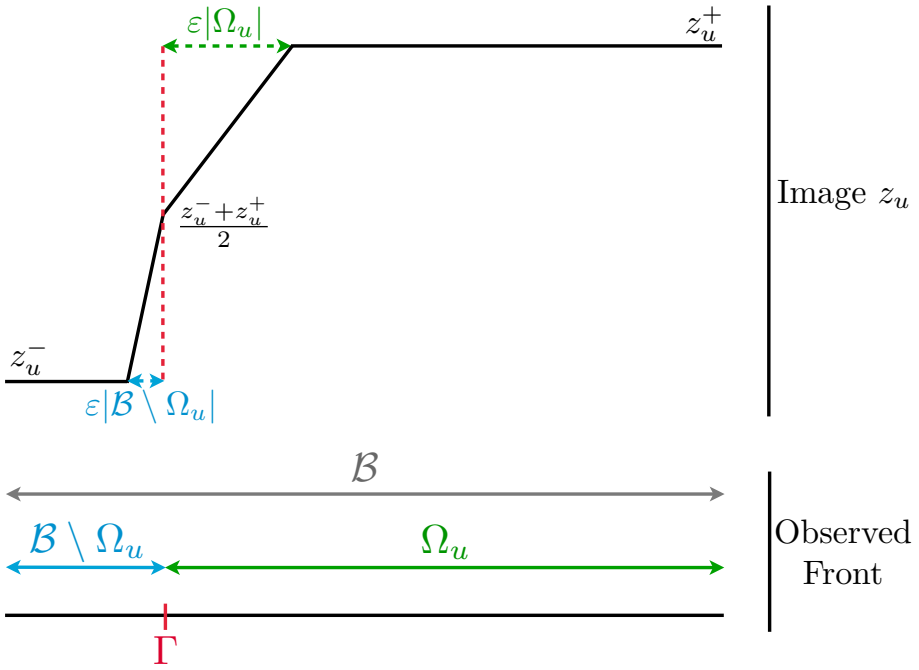

Figure 3: 1D object verifying Assumption 1

Example of a possible definition of z_u in 1D As previously explained, we need some assumptions on the definition of the object z_u constructed from the observed front to avoid consistency errors in the observation correction. Indeed, in the previous analysis, we suppose that this object verifies Assumption 1. We present here the construction of such an object in 1D. We consider a case with one front and we construct an associated object as seen in Figure 3, with z_u^+ and z_u^- two constants such that $z_u^+ > z_u^-$. The constant ε is strictly positive and is used in order to modify the wavelength of z_u. In Figure 3, the wavelength h_{z_u} equals $\varepsilon|\mathcal{B} \setminus \overline{\Omega_u^{\text{in}}}|$. We have

$$\bar{C}_1 = C_1(\Omega_u^{\text{in}}) = z_u^+ - \frac{1}{2}\frac{z_u^+ + z_u^-}{2}\varepsilon \text{ and } \bar{C}_2 = C_2(\Omega_u^{\text{in}}) = z_u^- + \frac{1}{2}\frac{z_u^+ + z_u^-}{2}\varepsilon$$

then $\frac{\bar{C}_1+\bar{C}_2}{2}$ equals $\frac{z_u^+ + z_u^-}{2}$, the front value of z_u. Assumption 1 holds.

Then we can study the value of $\vec{\nabla} z_u(\vec{x}) \cdot \vec{n}$ in this particular case,

$$\begin{aligned}
\vec{\nabla} z_u(\vec{x}) \cdot \vec{n} &= -(z_u)'(\vec{x}) \\
&= -\frac{1}{2}\left[\frac{1}{\varepsilon|\Omega_u^{\text{in}}|}\frac{z_u^+ - z_u^-}{2} + \frac{1}{\varepsilon|\mathcal{B} \setminus \overline{\Omega_u^{\text{in}}}|}\frac{z_u^+ - z_u^-}{2}\right] \\
&= -\frac{z_u^+ - z_u^-}{4\varepsilon}\left(\frac{1}{|\Omega_u^{\text{in}}|} + \frac{1}{|\mathcal{B} \setminus \overline{\Omega_u^{\text{in}}}|}\right) \\
&= \frac{\bar{C}_2 - \bar{C}_1}{4\varepsilon}\left(\frac{1}{|\Omega_u^{\text{in}}|} + \frac{1}{|\mathcal{B} \setminus \overline{\Omega_u^{\text{in}}}|}\right) - \frac{z_u^+ + z_u^-}{4}\left(\frac{1}{|\Omega_u^{\text{in}}|} + \frac{1}{|\mathcal{B} \setminus \overline{\Omega_u^{\text{in}}}|}\right).
\end{aligned}$$

In order to simplify this equation we can choose $z_u^- = -z_u^+$ and we have

$$-z_u'(\vec{x}) = \frac{\bar{C}_2 - \bar{C}_1}{4\varepsilon}\left(\frac{1}{|\Omega_u^{\text{in}}|} + \frac{1}{|\mathcal{B} \setminus \overline{\Omega_u^{\text{in}}}|}\right).$$

and

$$\big(d_{\Gamma_u}\mathcal{Q}^{\Gamma_u}(\tilde{u});\, \frac{\tilde{u}}{|\vec{\nabla}u|}\vec{n}_{\Gamma_u}\big) = 2\Big(\frac{\bar{C}_1-\bar{C}_2}{2}\Big)^2\Big(\frac{1}{|\Omega_u^{\rm in}|}+\frac{1}{|\mathcal{B}\setminus\overline{\Omega_u^{\rm in}}|}\Big)\frac{1}{\varepsilon}\int_{\Gamma_u}\frac{\alpha(|\vec{\nabla}\,u|)}{|\vec{\nabla}u|}\tilde{u}^2\,d\Gamma_u$$
$$-\,2\Big(\frac{\bar{C}_1-\bar{C}_2}{2}\Big)^2\Big(\frac{1}{|\Omega_{\hat{u}}^{\rm in}|}+\frac{1}{|\mathcal{B}\setminus\overline{\Omega_{\hat{u}}^{\rm in}}|}\Big)\int_{\Gamma_u}\frac{1}{|\vec{\nabla}u|}\,\tilde{u}\,d\Gamma_u\int_{\Gamma_u}\alpha(|\vec{\nabla}\,u|)\tilde{u}\,d\Gamma_u$$

In the case (33), the condition (31) becomes

$$\Big(\frac{1}{|\Omega_u^{\rm in}|}+\frac{1}{|\mathcal{B}\setminus\overline{\Omega_u^{\rm in}}|}\Big)\frac{1}{\varepsilon}\geq|\Gamma_{\hat{u}}|\Big(\frac{1}{|\Omega_{\hat{u}}^{\rm in}|}+\frac{1}{|\mathcal{B}\setminus\overline{\Omega_{\hat{u}}^{\rm in}}|}\Big)$$

and ε has to be chosen as

$$\varepsilon\geq\frac{1}{|\Gamma_{\hat{u}}|}\Big(\frac{1}{|\Omega_u^{\rm in}|}+\frac{1}{|\mathcal{B}\setminus\overline{\Omega_u^{\rm in}}|}\Big)/\Big(\frac{1}{|\Omega_{\hat{u}}^{\rm in}|}+\frac{1}{|\mathcal{B}\setminus\overline{\Omega_{\hat{u}}^{\rm in}}|}\Big).$$

This condition is very easy to verify. In practice the condition (31) given in Theorem 1 is not restrictive.

2.4.4 Gradient projection method

We denote

$$\mathcal{J}^{\Gamma_{\hat{u}}}=\int_{\Omega_{\hat{u}}^{\rm in}}(z_u-C_1(\Omega_{\hat{u}}^{\rm in}))^2d\vec{x}+\int_{\mathcal{B}\setminus\overline{\Omega_{\hat{u}}^{\rm in}}}(z_u-C_2(\Omega_{\hat{u}}^{\rm in}))^2d\vec{x}.$$

and we want to differentiate $\mathcal{J}^{\Gamma_{\hat{u}}}$ using the shape derivative. Using Proposition 3, we obtain

$$\begin{aligned}\big(d_{\Gamma_{\hat{u}}}\mathcal{J}^{\Gamma_{\hat{u}}};\,\vec{\psi}_{\Gamma_{\hat{u}}}\big)&=\int_{\Omega_{\hat{u}}^{\rm in}}\Big((z_u-C_1(\Omega_{\hat{u}}^{\rm in}))^2\Big)'d\vec{x}+\int_{\mathcal{B}\setminus\overline{\Omega_{\hat{u}}^{\rm in}}}\Big((z_u-C_2(\Omega_{\hat{u}}^{\rm in}))^2\Big)'d\vec{x}\\&+\int_{\Gamma_{\hat{u}}}(z_u-C_1(\Omega_{\hat{u}}^{\rm in}))^2\vec{\psi}_{\Gamma_{\hat{u}}}\cdot\vec{n}_{\Gamma_{\hat{u}}}\,d\Gamma_{\hat{u}}-\int_{\Gamma_{\hat{u}}}(z_u-C_2(\Omega_{\hat{u}}^{\rm in}))^2\vec{\psi}_{\Gamma_{\hat{u}}}\cdot\vec{n}_{\Gamma_{\hat{u}}}\,d\Gamma_{\hat{u}}\\&=-2C_1'(\Omega_{\hat{u}}^{\rm in})\int_{\Omega_{\hat{u}}^{\rm in}}(z_u-C_1(\Omega_{\hat{u}}^{\rm in}))d\vec{x}-2C_2'(\Omega_{\hat{u}}^{\rm in})\int_{\mathcal{B}\setminus\overline{\Omega_{\hat{u}}^{\rm in}}}(z_u-C_2(\Omega_{\hat{u}}^{\rm in}))d\vec{x}\\&+\int_{\Gamma_{\hat{u}}}\big[(z_u-C_1(\Omega_{\hat{u}}^{\rm in}))^2-(z_u-C_2(\Omega_{\hat{u}}^{\rm in}))^2\big]\vec{\psi}_{\Gamma_{\hat{u}}}\cdot\vec{n}_{\Gamma_{\hat{u}}}\,d\Gamma_{\hat{u}}\\&=\int_{\Gamma_{\hat{u}}}\big[(z_u-C_1(\Omega_{\hat{u}}^{\rm in}))^2-(z_u-C_2(\Omega_{\hat{u}}^{\rm in}))^2\big]\vec{\psi}_{\Gamma_{\hat{u}}}\cdot\vec{n}_{\Gamma_{\hat{u}}}\,d\Gamma_{\hat{u}}.\end{aligned}$$

Furthermore by definition of C_1 and C_2 we directly have

$$\int_{\Omega_{\hat{u}}^{\rm in}}(z_u-C_1(\Omega_{\hat{u}}^{\rm in}))d\vec{x}=0\ \text{and}\ \int_{\mathcal{B}\setminus\overline{\Omega_{\hat{u}}^{\rm in}}}(z_u-C_2(\Omega_{\hat{u}}^{\rm in}))d\vec{x}=0.$$

and then the shape derivative of $\mathcal{J}^{\Gamma_{\hat{u}}}$ is given by

$$\big(d_{\Gamma_{\hat{u}}}\mathcal{J}^{\Gamma_{\hat{u}}};\,\vec{\psi}_{\Gamma_{\hat{u}}}\big)\;=\;\int_{\Gamma_{\hat{u}}}\big[(z_u-C_1(\Omega_{\hat{u}}^{\rm in}))^2-(z_u-C_2(\Omega_{\hat{u}}^{\rm in}))^2\big]\vec{\psi}_{\Gamma_{\hat{u}}}\cdot\vec{n}_{\Gamma_{\hat{u}}}\,d\Gamma_{\hat{u}}.$$

Using the relation (28), we have

$$\big(d_{\hat{u}}\mathcal{J}^{\Gamma_{\hat{u}}};\,\psi_{\hat{u}-c_{\rm th}}\big)=\int_{\Gamma_{\hat{u}}}\frac{1}{|\vec{\nabla}\,\hat{u}|}\Big(\big(z_u-C_1(\Omega_{\hat{u}}^{\rm in})\big)^2-\big(z_u-C_2(\Omega_{\hat{u}}^{\rm in})\big)^2\Big)\,\psi_{\hat{u}-c_{\rm th}}\,d\Gamma_{\hat{u}}.$$

Formally, we define

$$\nabla \mathcal{J}^{\Gamma_{\hat{u}}} = \hat{\delta}(\Gamma_{\hat{u}}, \vec{x})\alpha(|\vec{\nabla}\, u|)|\vec{\nabla}\, u|\left(\frac{1}{|\vec{\nabla}\, \hat{u}|}\Big(\big(z_u - C_1(\Omega^{\text{in}}_{\hat{u}})\big)^2 - \big(z_u - C_2(\Omega^{\text{in}}_{\hat{u}})\big)^2\Big)\right).$$

The term α allows some flexibility in the choice of the identification norm. The observer model (4) can be rewritten as

$$\partial_t \hat{u} = \vec{\nabla} \cdot (\vec{\vec{\sigma}} \cdot \vec{\nabla}\, \hat{u}) + kf(\hat{u}) - \lambda \nabla \mathcal{J}^{\Gamma_{\hat{u}}}$$

and we recognize a gradient projection method. This makes a link with image processing and explain the choice of the discrepancy (5).

3 Parameter estimation

A very interesting aspect in data assimilation is the possibility to correct the uncertainties on the state and to also identify some parameters. The estimation of the parameters is a real issue. In this objective, we use the joint state and parameter strategy presented in [44] which has been applied for the data assimilation of bio-mechanical models. The strategy has been proposed to extend a possible Luenberger observer to a joint state and parameters estimation by combining the Luenberger filter and an optimal filter reduced to the remaining parameter space. Two different types of filter, one for the state and one for the parameters, are then used *i.e.*

$$\left\{\begin{array}{rcc} \dot{\hat{u}}(t) & = & \mathrm{A}(\hat{u}, \theta, t) + \mathrm{G}_u\big(\mathrm{D}(\mathrm{z}, \hat{u})\big) \\ \dot{\hat{\theta}}(t) & = & \mathrm{G}_\theta\big(\mathrm{D}(\mathrm{z}, \hat{u})\big) \\ \hat{u}(0) & = & \hat{u}_0 \\ \theta(0) & = & \theta_0. \end{array}\right.$$

The uncertainties on the state space are handled with the Luenberger observer presented in Section 2.2. In this work, the state correction term does not have the classical form $\mathrm{G}_u\big(\mathrm{D}(\mathrm{z}, \hat{u})\big)$ but a discrepancy defined in (5) can be extracted as explained in Section 2.4.4 and the joint state and parameter estimation model reads

$$\left\{\begin{array}{rcc} \dot{\hat{u}}(t) & = & \mathrm{A}(\hat{u}, \theta, t) + \mathcal{OBS}_u(z, \hat{u}) \\ \dot{\hat{\theta}}(t) & = & \mathrm{G}_\theta\big(\mathrm{D}(\mathrm{z}, \hat{u})\big) \\ \hat{u}(0) & = & x_0 \\ \hat{\theta}(0) & = & \theta_0 \end{array}\right.$$

A Reduced-order Unscented Kalman Filtering (RoUKF) [43] is used in order to handle the uncertainties of the parameters. As previously explained, a Kalman-like method for the parameter seems well-adapted due to the relatively small dimension of the parameters space.

4 Numerical simulations

4.1 Numerical methods

Algorithm for the state estimator This paragraph gives a space and time discretization of the observer model

$$\left\{\begin{array}{rcll} \partial_t \hat{u} - \vec{\nabla} \cdot (\vec{\vec{\sigma}} \cdot \vec{\nabla}\, \hat{u}) & = & kf(\hat{u}) & \\ +\lambda\delta(\hat{u} - c_{\text{th}})\Big(-\big(z_u - C_1(\Omega^{\text{in}}_{\hat{u}})\big)^2 + \big(z_u - C_2(\Omega^{\text{in}}_{\hat{u}})\big)^2\Big), & & & \mathcal{B} \times (0,T), \\ (\vec{\vec{\sigma}} \cdot \vec{\nabla}\, \hat{u}) \cdot \vec{n} & = & 0, & \partial\mathcal{B} \times (0,T), \\ \hat{u}(\vec{x}, 0) & = & \hat{u}_0(\vec{x}), & \mathcal{B}, \end{array}\right. \tag{34}$$

Note that this corresponds to (4) with

$$\alpha(|\vec{\nabla}\hat{u}|) = \frac{1}{|\vec{\nabla}\hat{u}|}.$$

We consider a space discretization – for example by finite difference or by finite element – denoted by $\hat{U} = (\hat{U}_k)_k$. Let $N \in \mathbb{N}^*$ be a given integer and consider a uniform partition $\{t_n, t_{n+1}\}_{0\leq n\leq N-1}$ with $t_n = nT/N = n\Delta t$ of the time interval $[0, T]$. We denote by $\hat{U}_n = (\hat{U}_{k,n})_k$ the approximation of $\hat{U}$ obtained at time t_n. For the sake of simplicity, we consider $\vec{\vec{\sigma}}$ constant in space and time. Denoting by Lp the Laplacian matrix, we define A^θ by $\mathrm{A}^\theta = \sigma\mathrm{Lp}$ with A^θ depending on the conductivity parameter (supposed here constant). And we denote by $\mathrm{F}(\cdot,\theta)$ the space discretization of $kf(\cdot)$. We consider an implicit discretization of the Laplacian and an explicit discretization of the reaction term and the data terms, which leads to the following prediction-correction time scheme

$$\left\{\begin{array}{rcl} \dfrac{\hat{U}^-_{n+1} - \hat{U}^+_n}{\Delta t} & = & \mathrm{A}^\theta \hat{U}^-_{n+1} + \mathrm{F}(\hat{U}^+_n, \theta) \\ \dfrac{\hat{U}^+_{n+1} - \hat{U}^-_{n+1}}{\Delta t} & = & \lambda\delta_\varepsilon(\hat{U}^-_{n+1} - c_{\mathrm{th}})\Big[-\Big(Z^u_{n+1} - \dfrac{\sum_k H(\hat{U}^-_{k,n+1} - c_{\mathrm{th}})Z^u_{k,n+1}}{\sum_k H(\hat{U}^-_{k,n+1} - c_{\mathrm{th}})}\Big)^2 \\ & & \qquad + \Big(Z^u_{n+1} - \dfrac{\sum_k \big(1 - H(\hat{U}^-_{k,n+1} - c_{\mathrm{th}})\big)Z^u_{k,n+1}}{\sum_k \big(1 - H(\hat{U}^-_{k,n+1} - c_{\mathrm{th}})\big)}\Big)^2\Big] \end{array}\right. \tag{35}$$

where the discretization of the Dirac function δ_ε is given as follows.

We introduce the real-valued continuous functions ψ by $\psi(x) = \frac{1}{2}(1+\cos(\pi x))$ and a possible discretization of the Dirac function in 1D is given by

$$\delta_\varepsilon(x) = \left\{\begin{array}{cc} \frac{1}{\varepsilon}\psi(\frac{x}{\varepsilon}) & \text{if } |x| \leq \varepsilon, \\ 0 & \text{if } |x| > \varepsilon \end{array}\right. ,$$

with here ε is a strictly positive constant. When considering the more general case

$$\hat{\delta}(\Gamma_{\hat{u}}, \vec{x})\alpha(|\vec{\nabla}\,\hat{u}|),$$

the discretization of $\hat{\delta}(\Gamma, \vec{x})$ for a boundary Γ defined from a level set denoted here by ϕ is necessary. There is a difficulty in the discretization of $\hat{\delta}(\Gamma, \vec{x})$ in the case where the level set is not a distance level set function, *i.e.* for example in our case because $\phi = \hat{u} - c_{\mathrm{th}}$. Following the discretization given in [18], the term ε becomes a function and is defined by

$$\varepsilon(\vec{v}, \varepsilon_0) = \frac{|\vec{v}|_1}{|\vec{v}|}\varepsilon_0,$$

where $|\vec{v}|_1$ is the 1-norm of the vector $\vec{v}$, $|\vec{v}|$ the 2-norm, and ε_0 a strictly positive constant. In [18], the following approximation of $\hat{\delta}(\Gamma, \vec{x})$ is considered,

$$\delta_{\varepsilon(|\vec{\nabla}\,\phi|,\varepsilon_0)}\Big(\frac{\phi}{|\vec{\nabla}\,\phi|}\Big) = \delta_{\bar{\varepsilon}}(\phi)|\vec{\nabla}\,\phi|,$$

with $\bar{\varepsilon} = \varepsilon_0|\vec{\nabla}\,\phi|_1$.

Application of RoUKF to parameter identification For the sake of simplicity, the state space is considered as finite dimensional. For the parameter identification, we use the RoUKF filter which is a non-linear filter adapted from the linear Kalman filter [32]. The gain operator of the Kalman filter is optimal in the sense of Bensoussan [4] and takes the following form

$$\mathrm{G}_\theta = \mathrm{P}\mathrm{D}^\intercal\mathrm{M},$$

where M is a metric on the observation space and P is called the covariance operator and verifies a Ricatti equation. Approximate Kalman filters are developed in non-linear cases as for example the *Extended Kalman filter* [51], the *Unscented Kalman filter* (UKF), [31] or the *Ensemble Kalman filter*, [19]. The UKF is a discrete-time estimator based on sampling points in the state space – called *sigma points* – whose role is to replace the tangent computations. We introduce the so-called p *unitary sampling points* $\mathrm{sp}^{[i]}$ and the p weights α^i as

$$\sum_{i=1}^{p} \alpha^i \mathrm{sp}^{[i]} = 0 \text{ and } \sum_{i=1}^{p} \alpha^i \mathrm{sp}^{[i]} \cdot (\mathrm{sp}^{[i]})^T = Id.$$

From a stochastic point of view, these sampling points can be generated around the estimated values based on the covariance estimation as explained for example in [30]. The optimal filters have an important drawback, namely the computation of the covariance P which is a full matrix. A classical strategy consists in using reduced-order for the covariance operators. For example, we consider the following decomposition for P,

$$\forall t,\ \mathrm{P}(t) = \mathrm{L}(t)\mathrm{U}^{-1}(t)\mathrm{L}(t)^T$$

with U an invertible matrix of small size r and L an extension operator. This strategy leads to the Reduced-order UKF (RoUKF). We denote by D_α, the diagonal matrix of the p weights α^i. We denote by $\mathrm{sp}^{[*]}$ the matrix concatenating the $\mathrm{sp}^{[i]}$ vectors side by side, and we do similarly for other vectors. The above idea has been applied in [44] for parameter identification by reducing the uncertainty space to the parameter space. This strategy is used and leads to the following prediction correction scheme used in our simulations.

Joint state (Luenberger observer) - parameters (RoUKF) scheme

1. Sampling
$$\left\{\begin{array}{rcl} \mathrm{C}_n & = & \sqrt{\mathrm{U}_n^{-1}} \\ \begin{pmatrix}\hat{U}_n^{[i]+}\\ \hat{\theta}_n^{[i]+}\end{pmatrix} & = & \begin{pmatrix}\hat{U}_n^{+}\\ \hat{\theta}_n^{+}\end{pmatrix} + \begin{pmatrix}\mathrm{L}_n^{u}\\ \mathrm{L}_n^{\theta}\end{pmatrix} \mathrm{C}_n^T \mathrm{sp}^{[i]} \end{array}\right.$$

2. Prediction
$$\left\{\begin{array}{rcl} \hat{U}_{n+1}^{[i]-} & = & \Big(\dfrac{1}{\Delta t} Id - \mathrm{A}^{\hat{\theta}_n^{[i]+}}\Big)^{-1}\Big(\dfrac{1}{\Delta t}\hat{U}_n^{[i]+} + \mathrm{F}(\hat{U}_n^{[i]+}, \hat{\theta}_n^{[i]+})\Big) \\ \hat{\theta}_{n+1}^{-} & = & \sum_{i=1}^{p} \alpha^i \hat{\theta}_n^{[i]+} = \hat{\theta}_n^{+} \text{ and } \hat{\theta}_{n+1}^{[i]-} = \hat{\theta}_n^{[i]+} \end{array}\right.$$

3. Solutions (of the observer model and using in the discrepancy)
$$\hat{U}_{n+1}^{-} = \sum_{i=1}^{p} \alpha^i \hat{U}_{n+1}^{[i]-}$$

4. State correction
$$\left\{\begin{array}{rcl} \hat{U}_{n+1}^{[i]+-} & = & \hat{U}_{n+1}^{[i]-} + \lambda\delta_\varepsilon(\hat{U}_{n+1}^{[i]-} - c_{\mathrm{th}})\Big[-\Big(Z_{n+1}^{u} - \dfrac{\sum_k H(\hat{U}_{k,n+1}^{[i]-} - c_{\mathrm{th}}) Z_{k,n+1}^{u}}{\sum_k H(\hat{U}_{k,n+1}^{[i]-} - c_{\mathrm{th}})}\Big)^2 \\ & & \qquad + \Big(Z_{n+1}^{u} - \dfrac{\sum_k \big(1 - H(\hat{U}_{k,n+1}^{[i]-} - c_{\mathrm{th}})\big) Z_{k,n+1}^{u}}{\sum_k \big(1 - H(\hat{U}_{k,n+1}^{[i]-} - c_{\mathrm{th}})\big)}\Big)^2\Big] \\ \hat{U}_{n+1}^{+-} & = & \sum_{i=1}^{p} \alpha^i \hat{U}_{n+1}^{[i]+-} \end{array}\right.$$

5. RoUKF correction
$$\left\{\begin{array}{rcl} \mathrm{L}_{n+1}^{u} & = & \hat{U}_{n+1}^{[*]+-} D_\alpha(\mathrm{sp}^{[*]})^T \\ \mathrm{L}_{n+1}^{\theta} & = & \hat{\theta}_{n+1}^{[*]-} D_\alpha(\mathrm{sp}^{[*]})^T \\ \Gamma_{n+1} & = & \mathrm{z}_{n+1}^{[*]-} D_\alpha(\mathrm{sp}^{[*]})^T \\ \mathrm{U}_{n+1} & = & 1 + \Gamma_{n+1}^T \mathrm{M}_{n+1} \Gamma_{n+1} \\ \hat{U}_{n+1}^{+} & = & \hat{U}_{n+1}^{+-} + \mathrm{L}_{n+1}^{u} \mathrm{U}_{n+1}^{-1} \Gamma_{n+1}^T \mathrm{M}_{n+1} \mathrm{D}_{n+1}(Z_{n+1}^{u}, \hat{U}_{n+1}^{-}) \\ \hat{\theta}_{n+1}^{+} & = & \hat{\theta}_{n+1}^{-} + \mathrm{L}_{n+1}^{\theta} \mathrm{U}_{n+1}^{-1} \Gamma_{n+1}^T \mathrm{M}_{n+1} \mathrm{D}_{n+1}(Z_{n+1}^{u}, \hat{U}_{n+1}^{-}) \end{array}\right.$$

4.2 Numerical Results in cardiac electrophysiology

4.2.1 Cardiac electrophysiology

As previously explained, this work can directly be used in cardiac electrophysiology. Cardiac electrophysiology describes the electrical phenomena which occur in the cardiac tissue and which trigger the cardiac mechanical contraction. Due to the frequent electrical pathologies – such as for example atrial fibrillation or ventricular bradycardia or tachycardia – which directly affect the origin of the heart contraction, the study of cardiac electrical mechanisms is very important. Over the past decades, cardiac electrophysiological modeling has made huge progresses and very realistic modeling and simulations are presented in the literature [12, 23, 11, 53, 25, 22], in particular in pathological cases [29, 5, 50]. Various electrophysiological models – based on a reaction-diffusion model – exist as for example the bidomain model [56, 49] and its first ap-

proximation the monodomain model [11, 50]. In order to obtain simulations adapted to each patient, we want to assimilate some data corresponding to the depolarization maps. These data are the isochrone maps of the transmembrane potential, called electrocardiographic imaging (ECGI). An electrode vest gives the potential at various points on the body surface and an inverse method allows to reconstruct the depolarization maps, see for example [47]. The bidomain equation can be written – in terms of u_e, the extracellular potential and $V_m = u_i - u_e$, the transmembrane potential with u_i, the intracellular potential – as

$$\begin{cases} A_m\Big(C_m\partial_t V_m + I_{ion}(V_m,\cdots)\Big) - \operatorname{div}\big(\vec{\vec{\sigma}}_i\cdot\vec{\nabla}V_m\big) \\ \qquad\qquad = \operatorname{div}\big(\vec{\vec{\sigma}}_i\cdot\vec{\nabla}u_e\big) + A_m I_{app}, \quad \text{in } \mathcal{B}\times(0,T), \\ \operatorname{div}\Big(\big(\vec{\vec{\sigma}}_i+\vec{\vec{\sigma}}_e\big)\cdot\vec{\nabla}u_e\Big) = -\operatorname{div}\big(\vec{\vec{\sigma}}_i\cdot\vec{\nabla}V_m\big), \quad \text{in } \mathcal{B}\times(0,T), \end{cases} \tag{36}$$

where $\mathcal{B}$ denotes the domain of interest, and with appropriate boundary conditions

$$\begin{cases} \big(\vec{\vec{\sigma}}_i\cdot\vec{\nabla}u_e\big)\cdot\vec{n} &=& -\big(\vec{\vec{\sigma}}_i\cdot\vec{\nabla}V_m\big)\cdot\vec{n}, & \text{in } \partial\mathcal{B}\times(0,T), \\ \big(\vec{\vec{\sigma}}_e\cdot\vec{\nabla}u_e\big)\cdot\vec{n} &=& 0, & \text{in } \partial\mathcal{B}\times(0,T), \end{cases} \tag{37}$$

where A_m is a positive constant denoting the ratio of membrane area per unit volume, C_m the membrane capacitance per unit surface and I_{app} a given applied stimulus current. The term $I_{ion}(V_m,\cdots)$ – called the reaction term – represents the ionic current across the membrane. It models the transmembrane potential evolution in the cell over time. Initially, the cell is at the resting potential around -80mV (0mV in rescaling cases). Due to a stimulus, the cell becomes depolarized very quickly and the value of the transmembrane potential is around 20mV (1mV in rescaling cases). During the depolarization phase – called the plateau phase – the mechanical contraction of the cell occurs. Then the repolarization of the cell begins. During this phase, the cell is in refractory phase, This means that a new stimulus cannot trigger a new depolarization. The reaction term $I_{ion}(V_m,\cdots)$ governs these different states of the cell. In order to describe this complex process which occurs across the membrane, one (or more) ordinary differential equation(s) can be coupled with the previous bidomain equations. There is a large variety of ionic models which represent the action potential. The extra-cellular potential u_e is defined up to a constant. In order to have a well-posed problem, the condition $\int_{\mathcal{B}} u_e = 0$ is added. The existence and uniqueness of the bidomain model has been studied for different ionic models in the literature, see for example [12, 3, 6, 57]. A particular case of the bidomain model is the monodomain model, see [46],

$$\begin{cases} A_m\Big(C_m\partial_t V_m + I_{ion}(V_m,w)\Big) - \operatorname{div}\big(\vec{\vec{\sigma}}_m\cdot\vec{\nabla}V_m\big) &=& A_m I_{app}, & \text{in } \mathcal{B}\times(0,T), \\ \big(\vec{\vec{\sigma}}_m\cdot\vec{\nabla}V_m\big)\cdot\vec{n} &=& 0, & \text{in } \partial\mathcal{B}\times(0,T). \end{cases} \tag{38}$$

We recall that we couple these monodomain or bidomain models with ionic models. There is a difficulty in 3D cases that we do not consider in this work. In fact, the ECGI allows to reconstruct the depolarization maps at the epicardium (outer surface) and sometimes at the endocardium (inner surface) but not through the thickness. In this part, only 1D and 2D cases are considered.

In a first section, we validate the Luenberger observer (4) in an electrophysiological framework. We start with simple 1D cases with the monodomain model coupled with the Mitchell-Schaeffer model [41]. Then we give a realistic medical case where a surface-based bidomain model specifically adapted to the atria – the upper chambers of the heart – able to take into account strong anisotropy in the atrial thickness, is used for the simulations. This model has been proposed and mathematically justified in [9] and simulations in realistic cases have been

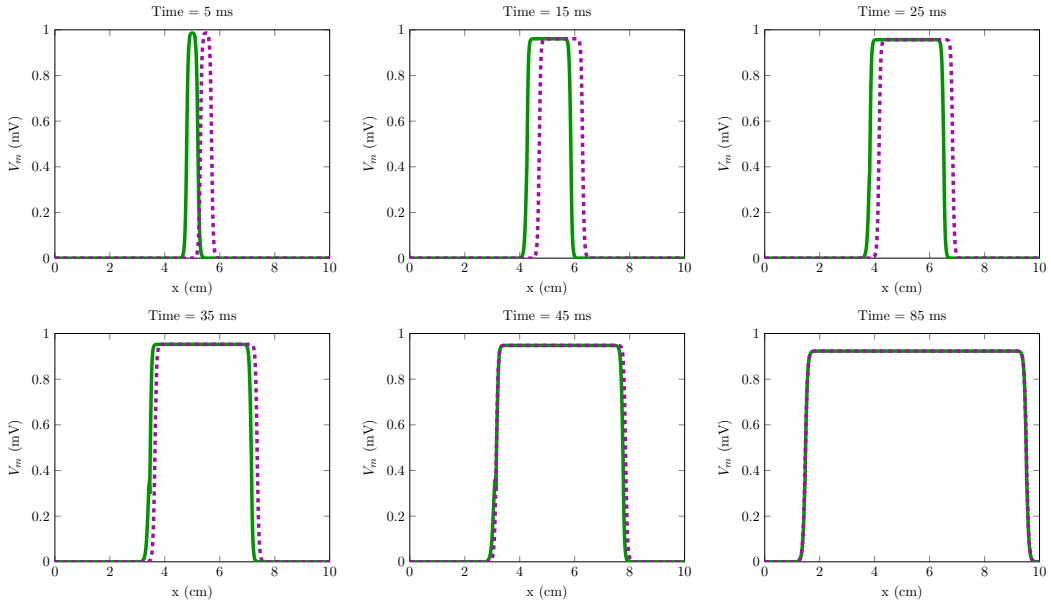

Figure 4: Solution of the observer model (in green) compared to the target solution (in dashed purple). Only an initial condition error is considered.

given in [13]. This surface-based model is coupled with the atrial physiological ionic model called the Courtemanche-Ramirez-Nattel model [16]. As we will see, the Luenberger state observer also gives very efficient results in this realistic situation. We also provide a simulation which represents a complex phenomenon, namely, spiral waves. The surface-based model coupled with the Mitchell-Schaffer is used. The excellent results confirm an excellent robustness of the observer model.

In a second section, we give simulations – always in 1D – which validate the joint state-parameter strategy previously introduced. We consider in this part the monodomain model coupled with the Mitchell-Schaeffer model. The parameter that we estimate is the conductivity parameter.

The 2D simulations are performed with the finite element library *FELiScE*[1], developed at Inria by REO and MΞDISIM teams.

4.2.2 Numerical simulations of the state observer

Simulations in 1D We recall that the monodomain model (38) coupled with the Mitchell-Schaeffer model is considered in this section. The parameters are given in Table 1 and we consider an isotropic and spacewise-constant diffusion tensor

$$\vec{\vec{\sigma}}_m = \sigma_m \vec{\vec{Id}},$$

with σ_m, the conductivity constant. The same assumptions and parameters are used for all 1D cases, except in the cases when otherwise specified. The initial conditions are $w = 1$ and $V_m = 0$mV. The stimulus is given by the term I_{app}. The observations are generated with the solution of the target model represented in the figures in dashed purple. When we assume an initial condition error, this means that the stimulus domain of the observer domain is different from the stimulus domain of the target model. The threshold constant equals 0.4mV. The

[1] `http://felisce.gforge.inria.fr`

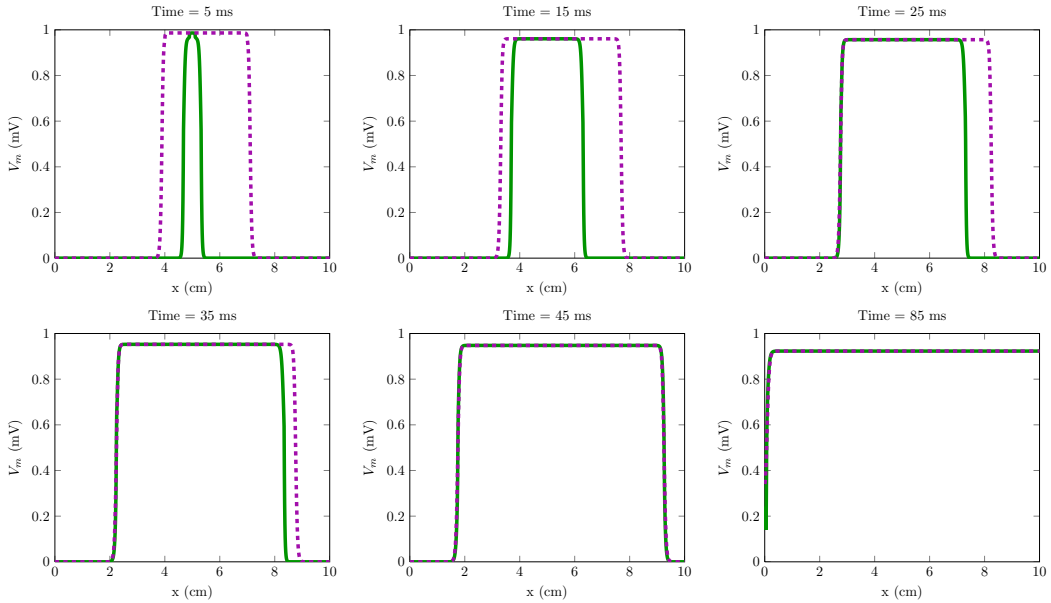

Figure 5: Solution of the observer model (in green) compared to the target solution (in dashed purple). Only an initial condition error is considered.

σ_m (S.cm^{-1})	A_m (cm^{-1})	C_m (mF.cm^{-2})	τ_{in} (cm^2.mA^{-1})	τ_{out} (cm^2.mA^{-1})
$2.0\,10^{-3}$	1	1	0.6	6
τ_{open} (ms)	τ_{close} (ms)	V_{min} (mV)	V_{max} (mV)	V_{gate} (mV)
120	300	0	1	0.13

Table 1: Model parameters

solution of the observer model is given in green. Here, we suppose that the parameters of the model are known and we only consider an error in the initial condition. The first simulation is given in Figure 4. The solution of the observer model adapts very well to the observations of the target model. Figure 5 shows a case where a large delay is caught.

Validation of the state estimator in a realistic case In this section, the state estimator is applied to a more realistic case. We consider a simulation of cardiac electrophysiology of the two atria, the upper chambers of the heart. The specificity of the atria is that they have very thin walls and a surface model can be considered. We use a surface-based bidomain model [9] – derived from the classical bidomain model (36) – able to take into account the strong anisotropy of the cardiac fibers in the atrial thickness. Indeed, the fibers vary through the wall of the atria and a linear variation of the fibers of a total angle 2θ between the endocardium (inner surface) and the epicardium (outer surface) can be considered. This implies that the diffusion tensors have the following form

$$\vec{\vec{\sigma}}_{i,e} = \sigma^t_{i,e}\vec{\vec{I}} + (\sigma^l_{i,e} - \sigma^t_{i,e})\big[I_0(\theta)\vec{\tau}_0 \otimes \vec{\tau}_0 + J_0(\theta)\vec{\tau}_0^{\perp} \otimes \vec{\tau}_0^{\perp}\big], \tag{39}$$

where $\vec{\vec{I}}$ denotes the identity tensor in the tangential plane, $\vec{\tau}_0$ is a unit vector parallel to the local fiber direction on the atria midsurface, and $\vec{\tau}_0^{\perp}$ such that $(\vec{\tau}_0, \vec{\tau}_0^{\perp})$ gives an orthonormal

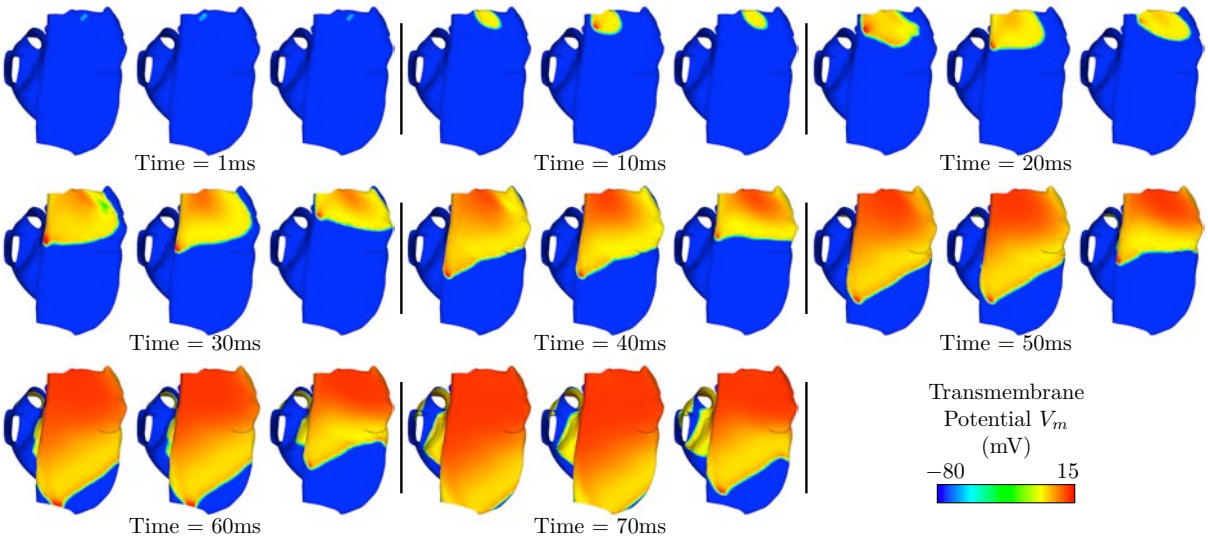

Figure 6: Right atrium view. In the middle, the target solution, on the right, the solution with a wrong initial condition without correction and on the left, the solution of the state observer model.

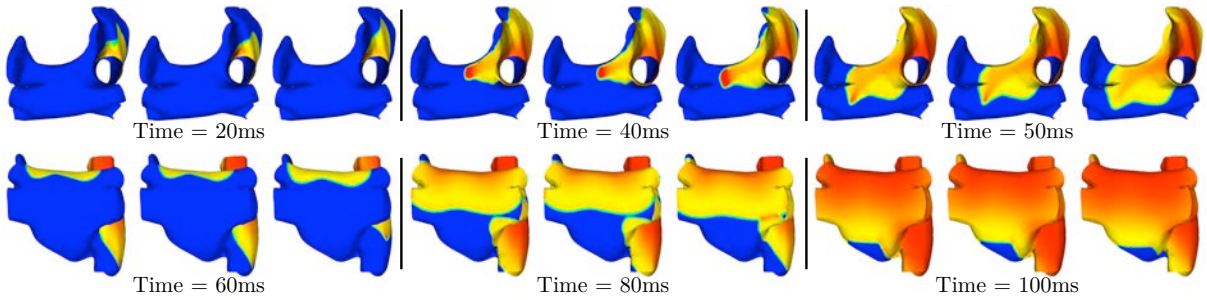

Figure 7: Left atrium view. In the middle, the target solution, on the right, the solution with a wrong initial condition without correction and on the left, the solution of the state observer model.

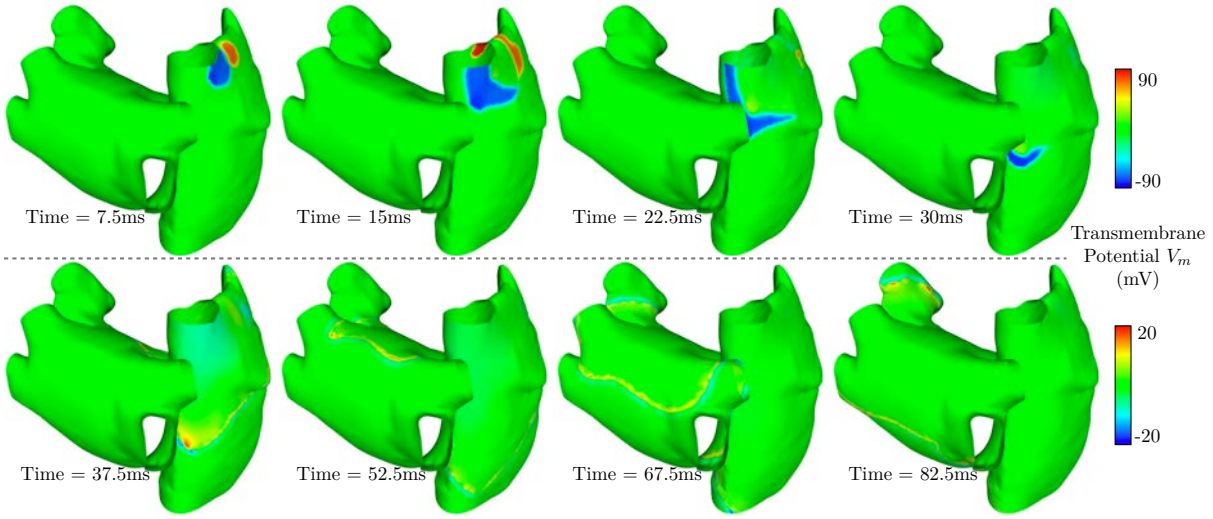

Figure 8: Atrial simulation – Point-wise difference between the observer model and the target model (to be compared with the typical variation range of the transmembrane potential, *i.e.* about 100 mV)

T(ms)	7.5	15	22.5	30	37.5	52.5	67.5	82.5
$\|\cdot\|_{l^2}$ (obs)	$8.14\,10^{-2}$	$1.15\,10^{-1}$	$9.47\,10^{-2}$	$6.11\,10^{-2}$	$1.66\,10^{-2}$	$1.41\,10^{-2}$	$1.69\,10^{-2}$	$1.59\,10^{-2}$
$\|\cdot\|_{l^2}$	$8.16\,10^{-2}$	$1.34\,10^{-1}$	$1.75\,10^{-1}$	$2.16\,10^{-1}$	$2.31\,10^{-1}$	$2.34\,10^{-1}$	$2.65\,10^{-1}$	$2.19\,10^{-1}$

Table 2: Atrial simulation – Error norm $|\cdot|_{l^2}$ between the target solution and the observer model (obs) on the second line and the model without correction on the third line.

basis of the tangential plane. An asymptotic analysis which validates this model is given in [9]. The ionic model used is the Courtemanche-Ramirez-Nattel model [16]. The position of the fibers, the parameters and the slow and fast conduction bundles considered are given in [14]. We only consider an error in the initial condition. This means that the sinus node, the natural pacemaker of the heart, is not well located. Figures 6 and 7 give three points of view at different times of the three following simulations: the target simulation in the middle, the simulation with the wrong initial condition on the right, and the simulation of the state observer on the left. The first remark that we can make is that even with a very small error in the initial condition the results obtained without correction (on the right) are very different from the target solution (in the middle). This shows the importance of a very efficient state observer. The second remark concerns the very good results of the state observer. As shown in Figure 6, at t= 40ms, the state observer is very close to the target solution for the right atria part despite a delay in the vertical direction and an advance in the horizontal direction at the origin. For the left part, we also obtain very quickly a very good approximation as shown in Figure 7.

We also report for this case in Table 2 the l^2 relative errors between the target model and the observer model (second line) *i.e.*

$$\frac{1}{V_{max}-V_{min}}\Big(\frac{1}{\sharp Nodes}\sum_{nodes}|u^{\text{obs}}-u^{\text{tar}}|^2\Big)^{\frac{1}{2}},$$

where $V_{max} = 15$mV and $V_{min} = -80$mV. Note that $V_{max} - V_{min}$ is a typical normalizing constant of the potential. We compare this with the l^2 difference between the target model and the model without correction (Table 2 third line). The l^2 error values between the target

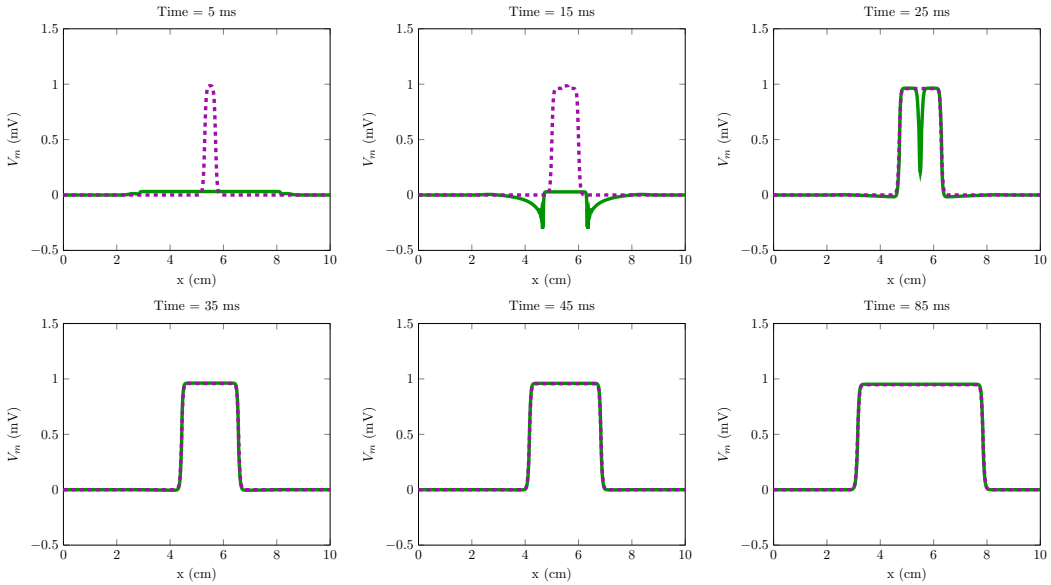

Figure 9: Solution of the observer model (in green) compared to the target solution (in dashed purple).

and the observer models are very small – inferior to 2% from $t = 35$ms – compared to the error between the target and the model – of the order of 20% – and this shows the efficiency of the observer. We also point out that these errors are quite stable over time. Figure 8 plots the difference $u^{\text{obs}} - u^{\text{tar}}$. At the beginning (first four times at the top), the absolute value of the maximum difference is around 90mV and we can see the early and the late areas. At the bottom of Figure 8, the absolute value of the difference is around 20mV. As seen in Figure 6, the solution of the observer model is very close to the target solution and the value of 20mV – which can be judged significant – is due to the discretization of the Dirac function. Indeed, as seen in Section 4.1, the approximate Dirac function is non-zero in an interval of the order of ε and this implies some errors in the observer model discretized. This is confirmed in Figure 8 by the narrow concentration of errors around the front. This can be improved by directly considering a Garlekin discretization of the surface term using an adaptive mesh which follows and determines the front [7].

Spiral waves A limitation occurs in the observer when the initial condition is not sufficient to trigger an actual propagation. This may be of concern for complex patterns such as spiral waves. We present a method which can be followed. First, we present this method in a 1D case. The idea is to play on the constant c_{th}, namely, the value above which the potential becomes depolarized. In Figure 9, we take a large interval of initial condition but a very weak stimulus. without correction, the solution returns to zero but with observations and a small value of $c_{\text{th}} = 0.01$mV, a depolarization is triggered as we will see in the figure. At $t = 1$ms, we change the value of c_{th} to a reasonable value equal to 0.4mV. The figure shows all the potential of this method.

This method is motivated by so-called spiral waves, a fascinating phenomenon in cardiac electrophysiology. The spiral waves are often considered as responsible for the atrial or ventricular fibrillation, [33]. We consider a surface mesh of an half cylinder. We use the surface model coupled with the Mitchell-Schaeffer model, the parameters and the direction of fibers given in [9]. The results are presented in Figure 10. The procedure which allows to initiate spiral waves

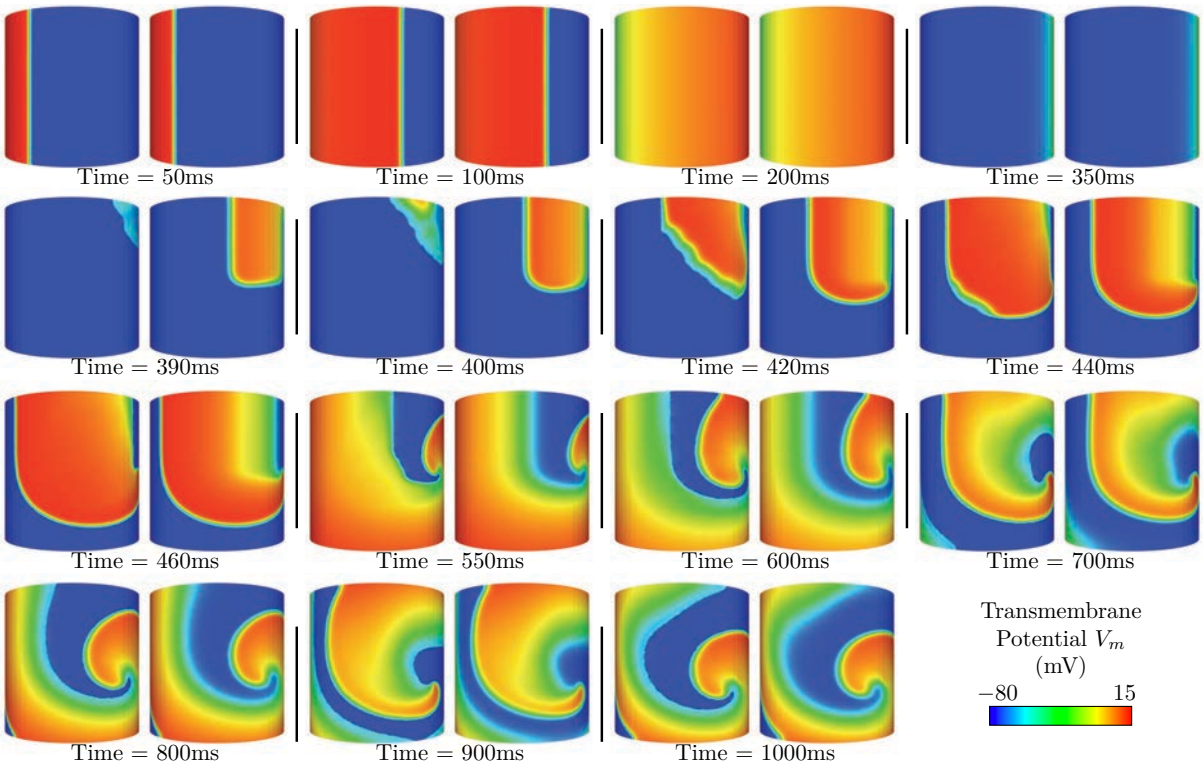

Figure 10: Spiral wave on a half cylinder. On the right, the target solution, and on the left, the solution of the state observer model.

is presented for example in [23]. A first wave in the horizontal direction is generated. Once the wave front has formed, it starts traveling along the cylinder (see the top line of Figure 10). A second region is depolarized at $t = 390$ms (see the target solution (right) at $t = 390$ms in Figure 10), just after the end of the repolarization. This depolarization signal rotates around the region that has just been repolarized and triggers a spiral wave. This new stimulus represents a pathological area which triggers a wrong signal which implies the apparition of a spiral wave. As it models a sick region of the heart, this region is unknown. That is why, for the observer model, we depolarize a very big region with a very small stimulus using the method presented just before in a 1D case. We play with the threshold, initially equal to -78mV, in order to create the spiral wave for the observer model. As seen in Figure 10, the spiral wave is initiated for the observer model and at $t = 450$ms, we change the value of c_{th} to a reasonable value equal to -67mV. The results obtained are excellent even if there is a small limitation for the observer. Indeed, Figure 10 shows that it is difficult for the observer to represent a smooth front (see the interior of the spiral wave).

A medical procedure called radio-frequency ablation consists in treating cardiac arrythmias (for example atrial fibrillation) by blocking these some abnormal conduction pathways [24]. Our personalized model may be used to optimize the procedure by detecting the position of the sick region. As a perspective here we could imagine to reconstruct the initial condition after choosing an appropriate parametrization of the initial condition and extending the joint state-parameter estimation to this case.

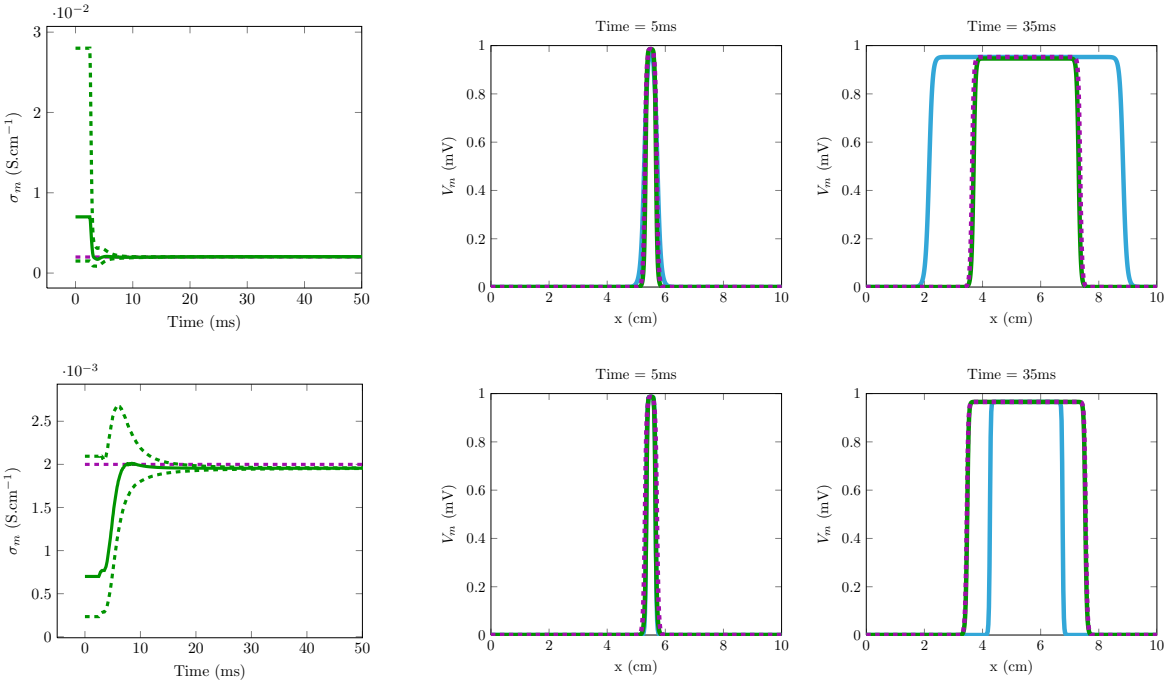

Figure 11: Left: Evolution of the estimated parameter $\sigma_m = 0.007\,\mathrm{S.cm^{-1}}$ (top) and $\sigma_m = 0.0007\,\mathrm{S.cm^{-1}}$ (bottom) toward the target parameter $\sigma_m = 0.002\,\mathrm{S.cm^{-1}}$. Right: Solution of the observer model with/without correction in green/blue compared to the target solution (in dashed purple). Only an error in conductivity parameter is considered.

4.2.3 Numerical simulations of the joint state and parameter estimator (only 1D cases)

Just for the sake of simplicity, we only consider 1D cases in this second part. The objective is to show that a joint state and parameter strategy – based on the *Luenberger* observer coupled with a RoUKF – allows to obtain the estimation of some parameters of the model.

Parameter estimator In this first case, we consider that the initial condition error is zero but we have an error in the conductivity parameter. Hence we only apply the parameter observer. We consider two different initial guesses for the conductivity parameter: $\sigma_m = 0.007\,\mathrm{S.cm^{-1}}$ and $\sigma_m = 0.0007\,\mathrm{S.cm^{-1}}$ instead of $\sigma_m = 0.002\,\mathrm{S.cm^{-1}}$. On the left of Figure 11 we plot the evolution of the conductivity parameters during the simulation runs of the data assimilation method, with the initial wrong parameter $\sigma_m = 0.007\,\mathrm{S.cm^{-1}}$ (resp. $\sigma_m = 0.0007\,\mathrm{S.cm^{-1}}$) at the top (resp. the bottom). We see that the mean value of the parameter (the green solid line) and the standard deviations of the parameter (the green dashed lines) converge in the two cases towards the correct value of the target model (the purple dashed line). As in the previous cases, the right of Figure 11 gives for the two cases the potential of the target model in dashed purple, and the solution of the observer model in green. In blue we represent the solution with the wrong conductivity parameters $\sigma_m = 0.007\,\mathrm{S.cm^{-1}}$ (top) and $\sigma_m = 0.0007\,\mathrm{S.cm^{-1}}$ (bottom) but without correction and this last curve shows that the parameter estimator is very efficient to retrieve the target parameter even from very inaccurate initial guesses.

Joint state and parameter estimator This second case is more complex. Indeed, we consider an initial condition error coupled with a wrong conductivity parameter equal to

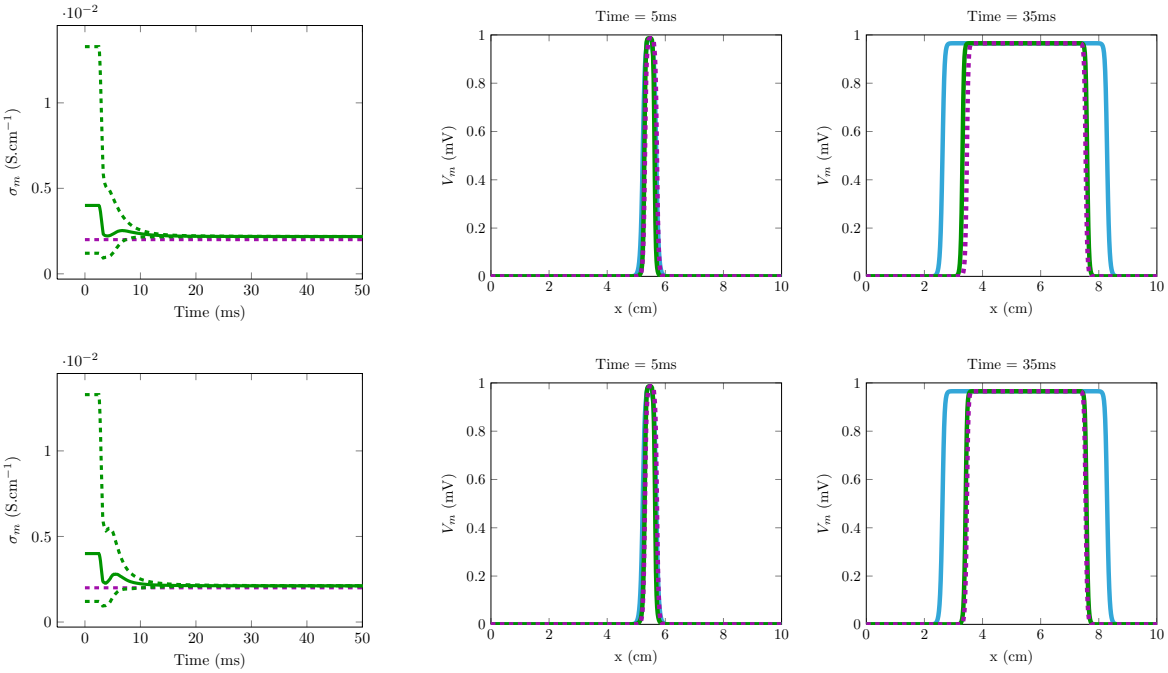

Figure 12: Left: Evolution of the estimated parameter $\sigma_m = 0.004\,\mathrm{S.cm^{-1}}$ toward the target parameter $\sigma_m = 0.002\,\mathrm{S.cm^{-1}}$ with only a parameter estimation (top) and a joint and state estimation (bottom). Right: Solution of the observer model with/without correction in green/blue compared to the target solution (in dashed purple). A small initial condition error is considered.

$0.004\,\mathrm{S.cm^{-1}}$ instead of the target conductivity value $0.002\,\mathrm{S.cm^{-1}}$. In a first simulation, we only try to perform a parameter identification without any state estimation in order to show the importance of the state observer. We consider a case where the initial condition error is very small, see Figure 12 (right-top) and a case with a larger initial condition error, see Figure 13 (right-top). In Figure 12 (left-top) and (resp. Figure 13 (left-top)), we plot the evolution of the parameter. With the small initial condition error, the mean value of the parameter (green solid line) converges to the target parameter but with a larger initial condition error, we do not have convergence toward the target parameter. In a second simulation, we use a first strategy of a joint state and parameter estimation, *i.e.* we couple the parameter estimator with the state estimator. The results are presented in Figure 12 (bottom) and Figure 13 (middle). In both cases, we obtain better results but in the case where the initial condition error is large, we are not able to estimate the target parameter. Then we use a second strategy. During a first time interval, we only use the state estimator in order to decrease the initial condition error which prevents the convergence of the parameter estimator. When the state estimation error is considered as sufficiently small, the parameter estimator is also applied. As shown in Figure 13 (bottom), we obtain a convergence toward the right parameter value with this method.

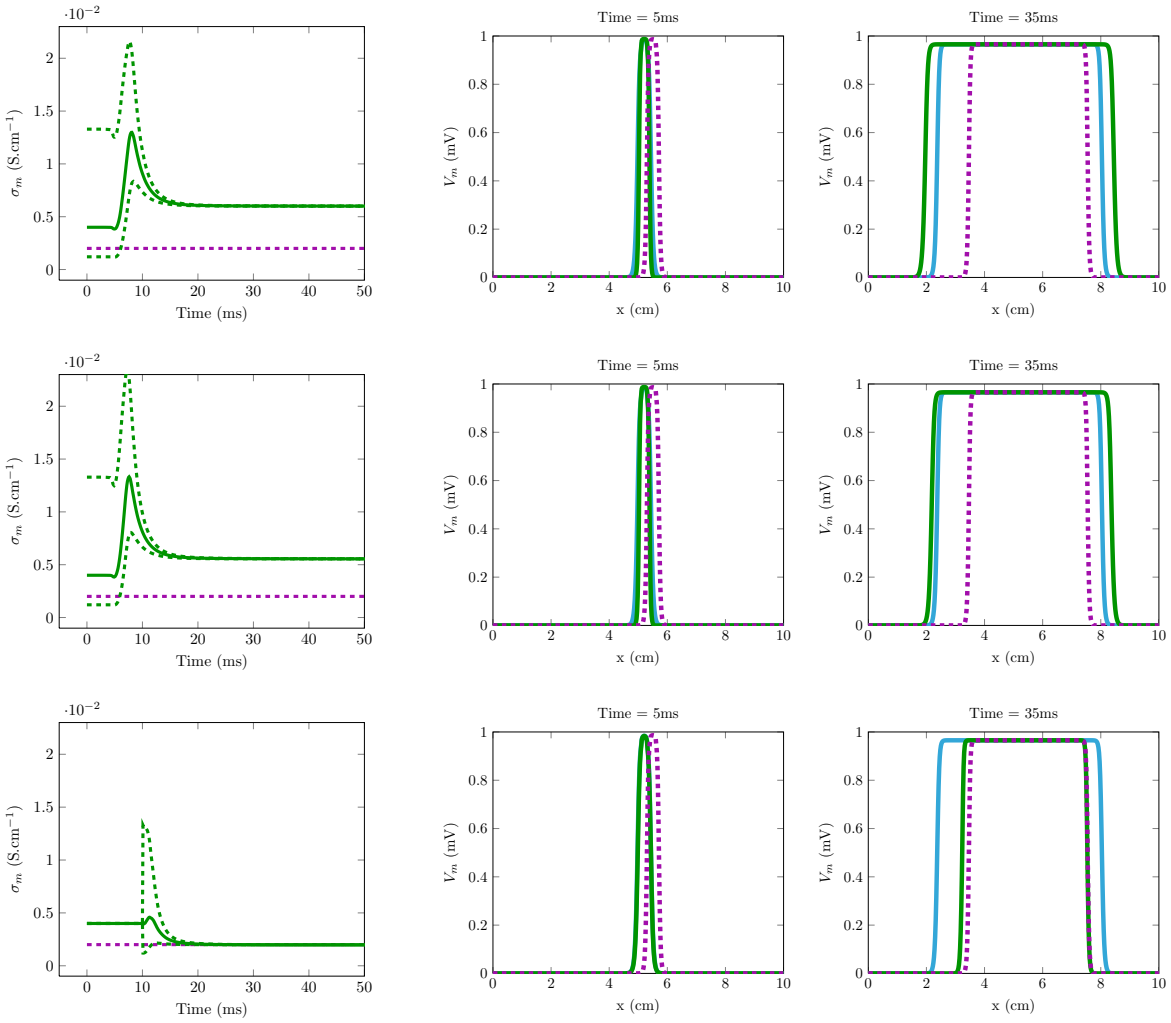

Figure 13: Left: Evolution of the estimated parameter $\sigma_m = 0.004\,\mathrm{S.cm}^{-1}$ with only a parameter estimation (top), a first joint and state estimation (middle) and a second joint and state estimation (bottom). Right: Solution of the observer model with/without correction in green/blue compared to the target solution (in dashed purple). A larger initial condition error is considered.

A Shape derivative theory

The first part of this appendix is an introduction – very much inspired from [17] and the appendix of [26] – on shape derivatives. In a second part, the proof of Proposition 1 is given.

A.1 Introduction

Let Γ the boundary of an open and bounded space $\Omega^{\text{in}} \in \mathbb{R}^N$ and we denote by $\vec{n}_\Gamma$, the exterior normal vector to Ω^{in}. We suppose that $\vec{\psi}$ is a given vector field with compact support in $\mathbb{R}^N$. We study the following problem

$$\left\{ \begin{array}{rcl} \vec{y}'(t) & = & \vec{\psi}(\vec{y}(t)),\ 0 \leq t \leq T \\ \vec{y}(0) & = & \vec{x}, \end{array} \right. \tag{40}$$

with $\vec{x} \in \mathbb{R}^N$ given. We define a mapping $T_t : \mathbb{R}^N \to \mathbb{R}^N$ with $T_t(\vec{x}) = \vec{y}(t)$ as the flow with respect to $\vec{\psi}$. We then define

$$\Gamma_t = T_t(\Gamma) \text{ and } \Omega^{\text{in}} = T_t(\Omega^{\text{in}}).$$

Remark 3 *If $\vec{\psi} \in C_0^k(\mathbb{R}^N, \mathbb{R}^N)$ then $T_t \in C^k(\mathbb{R}^N, \mathbb{R}^N)$, see [17].*

We want to calculate the derivatives for functionals of the following form

$$J^{\Omega^{\text{in}}} = \int_{\Omega^{\text{in}}} f(\Omega^{\text{in}}, \vec{x})\, d\vec{x} \text{ and } J^{\Gamma} = \int_{\Gamma} f(\Gamma, \vec{x})\, d\Gamma.$$

We start with the case where f is independent of Ω^{in} (and Γ) respectively and we have the Eulerian derivative of a domain integral.

Proposition 2 *Suppose $f \in W^{1,1}_{loc}(\mathbb{R}^N)$, then the functional*

$$J^{\Omega^{in}} = \int_{\Omega^{in}} f d\vec{x}$$

is shape differentiable for perturbation vector fields $\vec{\psi} \in C_0^1(\mathbb{R}^N; \mathbb{R}^N)$. The Eulerian derivative of $J^{\Omega^{in}}$ is given by

$$\left(dJ^{\Omega^{in}};\, \vec{\psi}\right) = \int_{\Omega^{in}} \operatorname{div}(f\vec{\psi}) d\vec{x} = \int_{\Gamma} f\, \vec{\psi} \cdot \vec{n}_\Gamma\, d\Gamma,$$

if Γ is of class C^1.
Suppose $f \in W^{2,1}_{loc}(\mathbb{R}^N)$ and Γ is a contour of class C^1, then the functional

$$J^{\Gamma} = \int_{\Gamma} f d\Gamma$$

is shape differentiable for perturbation vector fields $\vec{\psi} \in C_0^1(\mathbb{R}^N; \mathbb{R}^N)$ with

$$\left(dJ^{\Gamma};\, \vec{\psi}\right) = \int_{\Gamma} \left(\partial_{\vec{n}_\Gamma} f + f\kappa\right) \vec{\psi} \cdot \vec{n}_\Gamma\, d\Gamma,$$

where $\kappa = \vec{\nabla} \vec{n}_\Gamma$ is the curvature of Γ.

We are now interested in the general case, *i.e.* when f depends on Ω^{in} and Γ. We define two different derivatives.

Definition 1 *Suppose that $f(\Gamma)$ belongs to an appropriate Banach space $\mathcal{B}(\Gamma)$ of functions on Γ and let $\vec{\psi} \in C_0^k(\mathbb{R}^N, \mathbb{R}^N)$. We assume that $f(\Gamma_t) \circ T_t \in \mathcal{B}(\Gamma_t)$, $\forall t \in]0, T[$. If the limit*

$$\dot{f}(\Gamma; \vec{\psi}) = \lim_{t \to 0} \frac{1}{t} \big(f(\Gamma_t) \circ T_t - f(\Gamma) \big)$$

*exists in the strong (or weak) topology on $\mathcal{B}(\Gamma)$, then $\dot{f}(\Gamma; \vec{\psi})$ is called the strong (or weak) **material derivative** of f at Γ in direction $\vec{\psi}$.*

We define the analog definition for $f(\Omega^{\text{in}})$.

Definition 2 *We use the same hypothesis. If the limit*

$$\dot{f}(\Omega^{in}; \vec{\psi}) = \lim_{t \to 0} \frac{1}{t} \big(f(\Omega_t^{in}) \circ T_t - f(\Omega^{in}) \big)$$

exists in the strong (or weak) topology on $\mathcal{B}(\Omega^{in})$, then $\dot{f}(\Omega^{in}; \vec{\psi})$ is called the strong (or weak) material derivative of f at Ω^{in} in direction $\vec{\psi}$.

Remark 4 *In the case where the function $f : \Omega^{in} \to \mathbb{R}$ is independent of Ω^{in}, we have*

$$\lim_{t \to 0} \frac{1}{t} \big(f(\Omega_t^{in}) \circ T_t - f(\Omega^{in}) \big) = \lim_{t \to 0} \frac{1}{t} \big(f \circ T_t - f \big) = \vec{\nabla} f \cdot T_0' = \vec{\nabla} f \cdot \vec{\psi}.$$

For this type of functions – which do not depend on Ω^{in} – we wan to define a derivative with respect to Ω^{in} which must be zero. This is the goal of the following definition.

Definition 3 *We suppose that the weak material derivative of f and the term $\vec{\nabla} f \cdot \vec{\psi}$ exist in $\mathcal{B}(\Omega^{in})$. We then define*

$$f'(\Omega^{in}; \vec{\psi}) = \dot{f}(\Omega^{in}; \vec{\psi}) - \vec{\nabla} f \cdot \vec{\psi}$$

*and this new derivative is called the **domain shape derivative** of f at Ω^{in} in direction $\vec{\psi}$.*

Remark 5 *We have now*

$$f'(\Omega^{in}; \vec{\psi}) = f'(\vec{\psi}) = 0,$$

when f is independent of Ω^{in}.

For boundary functions $f(\Gamma) : \Gamma \mapsto \mathbb{R}$, the expression $\vec{\nabla} f \cdot \vec{\psi}$ does not exist. In this case, we define the **boundary shape derivative** as

$$f'(\Gamma; \vec{\psi}) = \dot{f}(\Gamma; \vec{\psi}) - \vec{\nabla}^{\Gamma} f \cdot \vec{\psi},$$

with

$$\vec{\nabla}^{\Gamma} f = \vec{\nabla} \bar{f} \Big|_{\Gamma} - \partial_{\vec{n}_\Gamma} \bar{f} \vec{n}_\Gamma,$$

where $\bar{f}$ is a smooth extension of f. In the case where $f(\Gamma) = f(\Omega^{\text{in}})\big|_{\Gamma}$, we have a relation between the **boundary shape derivative** and the **domain shape derivative** given by

$$f'(\Gamma; \vec{\psi}) = \big(\dot{f}(\Omega^{\text{in}}; \vec{\psi}) - \vec{\nabla} f \cdot \vec{\psi} + \partial_{\vec{n}_\Gamma} f \vec{n}_\Gamma \cdot \vec{\psi} \big) \Big|_{\Gamma} = \big(f'(\Omega^{\text{in}}; \vec{\psi}) + \partial_{\vec{n}_\Gamma} f \vec{n}_\Gamma \cdot \vec{\psi} \big) \Big|_{\Gamma}$$

Using this definition of the shape derivative of f, we are able to calculate the Eulerian derivatives for the shape functionals which concern us.

Proposition 3 *We suppose that the weak L^1-material derivation $\dot{f}$ and the shape derivation $f' \in L^1$ of f exist. Then*

$$J^{\Omega^{in}} = \int_{\Omega^{in}} f(\Omega^{in}, \vec{x})\, d\vec{x}$$

is shape differentiable and we have

$$\big(dJ^{\Omega^{in}};\, \vec{\psi}\big) = \int_{\Omega^{in}} f'(\Omega^{in};\, \vec{\psi}) d\vec{x} + \int_{\Gamma} f\, \vec{\psi} \cdot \vec{n}_{\Gamma}\, d\Gamma.$$

For boundary functions $f(\Gamma)$ we have

$$\big(dJ^{\Gamma};\, \vec{\psi}\big) = \int_{\Gamma} \Big(f'(\Gamma;\, \vec{\psi}) + f\kappa\, \vec{\psi} \cdot \vec{n}_{\Gamma} \Big)\, d\Gamma,$$

for

$$J^{\Gamma} = \int_{\Gamma} f(\Gamma, \vec{x})\, d\Gamma.$$

If $f(\Gamma) = f(\Omega^{in})\big|_{\Gamma}$, we obtain

$$\big(dJ^{\Gamma};\, \vec{\psi}\big) = \int_{\Gamma} \Big(f'(\Omega^{in};\, \vec{\psi})\Big|_{\Gamma} + \big(\partial_{\vec{n}_{\Gamma}} f + f\kappa\big)\vec{\psi} \cdot \vec{n}_{\Gamma} \Big)\, d\Gamma.$$

In our case, it is more complicated because f also depends on the variable t. Indeed we want to differentiate

$$\int_{\Gamma_t} f(t, \Gamma_t, \vec{x})\, d\Gamma_t.$$

Adapting the proof we can generalize the previous proposition.

Proposition 4 *We have*

$$\big(dJ^{\Gamma};\, \vec{\psi}\big) = \int_{\Gamma} \Big(f'(\Gamma;\, \vec{\psi}) + f\kappa\, \vec{\psi} \cdot \vec{n}_{\Gamma} \Big)\, d\Gamma \text{ for } J^{\Gamma} = \int_{\Gamma_t} f(t, \Gamma_t, \vec{x})\, d\Gamma_t.$$

If $f(\Gamma) = f(\Omega^{in})\big|_{\Gamma}$, we obtain

$$\big(dJ^{\Gamma};\, \vec{\psi}\big) = \int_{\Gamma} \Big(f'(\Omega^{in};\, \vec{\psi})\Big|_{\Gamma} + \big(\partial_{\vec{n}_{\Gamma}} f + f\kappa\big)\vec{\psi} \cdot \vec{n}_{\Gamma} \Big)\, d\Gamma.$$

Remark 6 *All this theory assumes a closed shape because we need to define the boundary Γ of the space Ω^{in}. Furthermore the boundary terms disappear thanks to the Green formula.*

A.2 Proof of Proposition 1

We want to prove Proposition 1 which gives the shape derivative of

$$\mathcal{Q}^{\Gamma_{\hat{u}}}(\chi) = \int_{\Gamma_{\hat{u}}} \alpha(|\vec{\nabla}\, \hat{u}|) \Big(\big(z_u - C_1(\Omega^{\text{in}}_{\hat{u}})\big)^2 - \big(z_u - C_2(\Omega^{\text{in}}_{\hat{u}})\big)^2 \Big)\, \chi\, d\Gamma_{\hat{u}}.$$

We assume that $\Gamma_{\hat{u}}$, $\Omega^{\text{in}}_{\hat{u}}$ and all the terms of $\mathcal{Q}^{\Gamma_{\hat{u}}}$ are as regular as needed for the use of the previous propositions. We want to use Proposition 4. We start with the boundary shape derivative of each term appearing in $\mathcal{Q}^{\Gamma_{\hat{u}}}$. In a second stage, we give the space derivative $\partial_{\vec{n}_{\Gamma}} f$ for each term and in a third paragraph, we gather the results.

Boundary shape derivative The subscript $(\cdot)'$ denotes in this part the boundary shape derivative. First, we want to compute the following derivative

$$\begin{aligned}\Big(\alpha(|\vec{\nabla}\,\hat{u}|)\Big(\big(z_u - C_1(\Omega^{\text{in}}_{\hat{u}})\big)^2 - \big(z_u - C_2(\Omega^{\text{in}}_{\hat{u}})\big)^2\Big)\Big)' \\ = \Big(\alpha(|\vec{\nabla}\,\hat{u}|)\Big)'\Big(\big(z_u - C_1(\Omega^{\text{in}}_{\hat{u}})\big)^2 - \big(z_u - C_2(\Omega^{\text{in}}_{\hat{u}})\big)^2\Big) \\ + \alpha(|\vec{\nabla}\,\hat{u}|)\Big(\big(z_u - C_1(\Omega^{\text{in}}_{\hat{u}})\big)^2 - \big(z_u - C_2(\Omega^{\text{in}}_{\hat{u}})\big)^2\Big)'.\end{aligned}$$

This term can be seen as of the form $f(t, \Gamma_{\hat{u}}, \vec{x})$. The variable t only appears in the term $\alpha(|\vec{\nabla}\,\hat{u}|)$ because of the function $\hat{u}$. The first step is to compute $\Big(\alpha(|\vec{\nabla}\,\hat{u}|)\Big)'$. We define the normal flow by

$$\vec{\psi}_{\Gamma_{\hat{u}}} = \frac{\psi_{\hat{u}-c_{\text{th}}}}{|\vec{\nabla}\hat{u}|}\vec{n}_{\Gamma_{\hat{u}}},$$

with $\hat{u}_t = \hat{u} + t\psi_{\hat{u}-c_{\text{th}}}$. Using the previous definition and Proposition 4, $\Big(\alpha(|\vec{\nabla}\,\hat{u}|)\Big)'$ can be written as

$$\Big(\alpha(|\vec{\nabla}\,\hat{u}|)\Big)' = \partial_t|_{\vec{x}}\,\alpha(|\vec{\nabla}\,\hat{u}|) + \vec{\nabla}\alpha(|\vec{\nabla}\,\hat{u}|)\cdot\vec{\psi}_{\Gamma_{\hat{u}}},$$

with

$$\begin{aligned}\partial_t|_{\vec{x}}\,\alpha(|\vec{\nabla}\,\hat{u}|) &= \lim_{t\to 0}\frac{1}{t}\big(\alpha(|\vec{\nabla}(\hat{u} + t\psi_{\hat{u}-c_{\text{th}}})|) - \alpha(|\vec{\nabla}\,\hat{u}|)\big) \\ &= \alpha'(|\vec{\nabla}\,\hat{u}|)\frac{\vec{\nabla}\,\hat{u}}{|\vec{\nabla}\,\hat{u}|}\cdot\vec{\nabla}\psi_{\hat{u}-c_{\text{th}}} \\ &= -\alpha'(|\vec{\nabla}\,\hat{u}|)\partial_{\vec{n}_\Gamma}\psi_{\hat{u}-c_{\text{th}}}\end{aligned}$$

using the Fréchet derivative of $\vec{x} \mapsto |\vec{\nabla}\,\hat{u}|$. In order to finish this computation, we study the following term

$$\begin{aligned}\Big(\big(z_u(\vec{x}) - C_1(\Omega^{\text{in}}_{\hat{u}})\big)^2 - \big(z_u(\vec{x}) - C_2(\Omega^{\text{in}}_{\hat{u}})\big)^2\Big)' \\ = -2C_1'(\Omega^{\text{in}}_{\hat{u}})\big(z_u(\vec{x}) - C_1(\Omega^{\text{in}}_{\hat{u}})\big) + 2C_2'(\Omega^{\text{in}}_{\hat{u}})\big(z_u(\vec{x}) - C_2(\Omega^{\text{in}}_{\hat{u}})\big).\end{aligned}$$

We recall that C_1 and C_2 are defined by

$$C_1(\Omega^{\text{in}}_{\hat{u}}) = \frac{\int_{\Omega^{\text{in}}_{\hat{u}}} z_u(\vec{x})d\vec{x}}{\int_{\Omega^{\text{in}}_{\hat{u}}} d\vec{x}}$$

and using Proposition 3, the derivative of C_1 is given by

$$\begin{aligned}C_1'(\Omega^{\text{in}}_{\hat{u}}) &= \frac{\Big(\int_{\Omega^{\text{in}}_{\hat{u}}} z_u(\vec{x})d\vec{x}\Big)' \times \int_{\Omega^{\text{in}}_{\hat{u}}} d\vec{x} - \Big(\int_{\Omega^{\text{in}}_{\hat{u}}} d\vec{x}\Big)' \times \int_{\Omega^{\text{in}}_{\hat{u}}} z_u(\vec{x})d\vec{x}}{(\int_{\Omega^{\text{in}}_{\hat{u}}} d\vec{x})^2}. \\ &= \frac{\int_{\Gamma_{\hat{u}}} z_u(\vec{x})\,\vec{\psi}_{\Gamma_{\hat{u}}}\cdot\vec{n}_{\Gamma_{\hat{u}}}\,d\Gamma_{\hat{u}}}{\int_{\Omega^{\text{in}}_{\hat{u}}} d\vec{x}} - \frac{\int_{\Gamma_{\hat{u}}}\vec{\psi}_{\Gamma_{\hat{u}}}\cdot\vec{n}_{\Gamma_{\hat{u}}}\,d\Gamma_{\hat{u}} \times C_1(\Omega^{\text{in}}_{\hat{u}})}{\int_{\Omega^{\text{in}}_{\hat{u}}} d\vec{x}} \\ &= \frac{1}{|\Omega^{\text{in}}_{\hat{u}}|}\int_{\Gamma_{\hat{u}}}(z_u(\vec{x}) - C_1(\Omega^{\text{in}}_{\hat{u}}))\,\vec{\psi}_{\Gamma_{\hat{u}}}\cdot\vec{n}_{\Gamma_{\hat{u}}}\,d\Gamma_{\hat{u}}.\end{aligned}$$

Using the same method, the derivative of C_2 is given by

$$C_2'(\Omega_{\hat{u}}^{\text{in}}) \;=\; -\frac{1}{|\mathcal{B}\setminus\overline{\Omega_{\hat{u}}^{\text{in}}}|}\int_{\Gamma_{\hat{u}}}\left(z_u(\vec{x})-C_2(\Omega_{\hat{u}}^{\text{in}})\right)\vec{\psi}_{\Gamma_{\hat{u}}}\cdot\vec{n}_{\Gamma_{\hat{u}}}\,d\Gamma_{\hat{u}}.$$

Space derivative In order to obtain the shape derivative of $\mathcal{Q}^{\Gamma_{\hat{u}}}$, we have to compute

$$\begin{aligned}
&\vec{\nabla}\Big(\alpha(|\vec{\nabla}\,\hat{u}|)\Big(\big(z_u(\vec{x})-C_1(\Omega_{\hat{u}}^{\text{in}})\big)^2-\big(z_u(\vec{x})-C_2(\Omega_{\hat{u}}^{\text{in}})\big)^2\Big)\Big)\cdot\vec{n}_{\Gamma_{\hat{u}}}\\
&\qquad=\vec{\nabla}\alpha(|\vec{\nabla}\,\hat{u}|)\cdot\vec{n}_{\Gamma_{\hat{u}}}\Big(\big(z_u(\vec{x})-C_1(\Omega_{\hat{u}}^{\text{in}})\big)^2-\big(z_u(\vec{x})-C_2(\Omega_{\hat{u}}^{\text{in}})\big)^2\Big)\\
&\qquad\qquad+\alpha(|\vec{\nabla}\,\hat{u}|)\,\vec{\nabla}\Big(\big(z_u(\vec{x})-C_1(\Omega_{\hat{u}}^{\text{in}})\big)^2-\big(z_u(\vec{x})-C_2(\Omega_{\hat{u}}^{\text{in}})\big)^2\Big)\cdot\vec{n}_{\Gamma_{\hat{u}}}.
\end{aligned}$$

Using the fact that $|\vec{\nabla}\,\hat{u}| = -\partial_{\vec{n}_{\Gamma_{\hat{u}}}}\hat{u}(\vec{x})$, we find

$$\begin{aligned}
\vec{\nabla}\alpha(|\vec{\nabla}\,\hat{u}|)\cdot\vec{n}_{\Gamma_{\hat{u}}} \;&=\; \alpha'(|\vec{\nabla}\,\hat{u}|)\vec{\nabla}\big(|\vec{\nabla}\,\hat{u}|\big)\cdot\vec{n}_{\Gamma_{\hat{u}}}=\alpha'(|\vec{\nabla}\,\hat{u}|)\vec{\nabla}\Big(-\partial_{\vec{n}_{\Gamma_{\hat{u}}}}\hat{u}(\vec{x})\Big)\cdot\vec{n}_{\Gamma_{\hat{u}}}\\
&=\; -\alpha'(|\vec{\nabla}\,\hat{u}|)\partial^2_{\vec{n}_{\Gamma_{\hat{u}}}}\hat{u}(\vec{x}).
\end{aligned}$$

and we finally obtain

$$\begin{aligned}
&\vec{\nabla}\Big(\big(z_u(\vec{x})-C_1(\Omega_{\hat{u}}^{\text{in}})\big)^2-\big(z_u(\vec{x})-C_2(\Omega_{\hat{u}}^{\text{in}})\big)^2\Big)\cdot\vec{n}_{\Gamma_{\hat{u}}}\\
&\qquad=2\vec{\nabla}z_u(\vec{x})\cdot\vec{n}_{\Gamma_{\hat{u}}}\Big(z_u(\vec{x})-C_1(\Omega_{\hat{u}}^{\text{in}})-\big(z_u(\vec{x})-C_2(\Omega_{\hat{u}}^{\text{in}})\big)\Big)\\
&\qquad\qquad=2\vec{\nabla}z_u(\vec{x})\cdot\vec{n}_{\Gamma_{\hat{u}}}\big(C_2(\Omega_{\hat{u}}^{\text{in}})-C_1(\Omega_{\hat{u}}^{\text{in}})\big).
\end{aligned}$$

By gathering the previous result, we obtain Proposition 1.

B Level Set Formulation

In this appendix, we want to present a second method for the differentiation of the shape estimator using the level set formulation, but in this case the calculus is done formally. We recall that

$$\mathcal{Q}^{\Gamma_{\hat{u}}}(\chi) = \int_{\mathcal{B}} \delta(\hat{u} - c_{\text{th}}) |\vec{\nabla}\,\hat{u}| \alpha(|\vec{\nabla}\,\hat{u}|) \Big(\big(z_u - C_1(\hat{u} - c_{\text{th}})\big)^2 - \big(z_u - C_2(\hat{u} - c_{\text{th}})\big)^2 \Big)\, \chi \, d\vec{x}.$$

We want to differentiate a functional of the following form

$$\mathcal{Q}^{\Gamma_{\hat{u}}} = \int_{\mathcal{B}} \delta(\hat{u} - c_{\text{th}}) |\vec{\nabla}\,\hat{u}| f(\vec{x},\, \hat{u}(\vec{x}),\, \hat{u}).$$

We have

$$\begin{aligned}
\Big(d\mathcal{Q}^{\Gamma_{\hat{u}}};\, \psi\Big) = \\
&\int_{\mathcal{B}} \delta'(\hat{u} - c_{\text{th}}) |\vec{\nabla}\,\hat{u}| f(\vec{x}, \hat{u}(\vec{x}), \hat{u})\, \psi \, d\vec{x} + \int_{\mathcal{B}} \delta(\hat{u} - c_{\text{th}}) |\vec{\nabla}\,\hat{u}| \left(\partial_{\hat{u}} f(\vec{x}, \hat{u}(\vec{x}), \hat{u}),\, \psi\right) d\vec{x} \\
&\qquad + \int_{\mathcal{B}} \delta(\hat{u} - c_{\text{th}}) \frac{\vec{\nabla}\,\hat{u}}{|\vec{\nabla}\,\hat{u}|} \cdot \vec{\nabla}\psi\, f(\vec{x}, \hat{u}(\vec{x}), \hat{u})\; d\vec{x} \\
&= \int_{\mathcal{B}} \delta'(\hat{u} - c_{\text{th}}) |\vec{\nabla}\,\hat{u}| f(\vec{x}, \hat{u}(\vec{x}), \hat{u})\, \psi \, d\vec{x} + \int_{\mathcal{B}} \delta(\hat{u} - c_{\text{th}}) |\vec{\nabla}\,\hat{u}| \left(\partial_{\hat{u}} f(\vec{x}, \hat{u}(\vec{x}), \hat{u}),\, \psi\right) d\vec{x} \\
&\qquad - \int_{\mathcal{B}} \vec{\nabla} \cdot \left(\delta(\hat{u} - c_{\text{th}}) f(\vec{x}, \hat{u}(\vec{x}), \hat{u}) \frac{\vec{\nabla}\,\hat{u}}{|\vec{\nabla}\,\hat{u}|} \right) \psi \, d\vec{x} \\
&= \int_{\mathcal{B}} \delta'(\hat{u} - c_{\text{th}}) |\vec{\nabla}\,\hat{u}| f(\vec{x}, \hat{u}(\vec{x}), \hat{u})\, \psi \, d\vec{x} + \int_{\mathcal{B}} \delta(\hat{u} - c_{\text{th}}) |\vec{\nabla}\,\hat{u}| \left(\partial_{\hat{u}} f(\vec{x}, \hat{u}(\vec{x}), \hat{u}),\, \psi\right) d\vec{x} \\
&- \int_{\mathcal{B}} \delta(\hat{u} - c_{\text{th}}) f(\vec{x}, \hat{u}(\vec{x}), \hat{u}) \vec{\nabla} \cdot \frac{\vec{\nabla}\,\hat{u}}{|\vec{\nabla}\,\hat{u}|}\, \psi \, d\vec{x} - \int_{\mathcal{B}} \delta(\hat{u} - c_{\text{th}}) \vec{\nabla} f(\vec{x}, \hat{u}(\vec{x}), \hat{u}) \cdot \frac{\vec{\nabla}\,\hat{u}}{|\vec{\nabla}\,\hat{u}|}\, \psi \, d\vec{x} \\
&\qquad - \int_{\mathcal{B}} f(\vec{x}, \hat{u}(\vec{x}), \hat{u}) \vec{\nabla} \delta(\hat{u} - c_{\text{th}}) \cdot \frac{\vec{\nabla}\,\hat{u}}{|\vec{\nabla}\,\hat{u}|}\, \psi \, d\vec{x} \\
&= \int_{\mathcal{B}} \delta'(\hat{u} - c_{\text{th}}) |\vec{\nabla}\,\hat{u}| f(\vec{x}, \hat{u}(\vec{x}), \hat{u})\, \psi \, d\vec{x} + \int_{\mathcal{B}} \delta(\hat{u} - c_{\text{th}}) |\vec{\nabla}\,\hat{u}| \left(\partial_{\hat{u}} f(\vec{x}, \hat{u}(\vec{x}), \hat{u}),\, \psi\right) d\vec{x} \\
&- \int_{\mathcal{B}} \delta(\hat{u} - c_{\text{th}}) f(\vec{x}, \hat{u}(\vec{x}), \hat{u}) \vec{\nabla} \cdot \frac{\vec{\nabla}\,\hat{u}}{|\vec{\nabla}\,\hat{u}|}\, \psi \, d\vec{x} - \int_{\mathcal{B}} \delta(\hat{u} - c_{\text{th}}) \vec{\nabla} f(\vec{x}, \hat{u}(\vec{x}), \hat{u}) \cdot \frac{\vec{\nabla}\,\hat{u}}{|\vec{\nabla}\,\hat{u}|}\, \psi \, d\vec{x} \\
&\qquad - \int_{\mathcal{B}} f(\vec{x}, \hat{u}(\vec{x}), \hat{u}) \delta'(\hat{u} - c_{\text{th}}) \vec{\nabla}\,\hat{u} \cdot \frac{\vec{\nabla}\,\hat{u}}{|\vec{\nabla}\,\hat{u}|}\, \psi \, d\vec{x} \\
&= \int_{\mathcal{B}} \delta(\hat{u} - c_{\text{th}}) |\vec{\nabla}\,\hat{u}| \left(\partial_{\hat{u}} f(\vec{x}, \hat{u}(\vec{x}), \hat{u}),\, \psi\right) d\vec{x} + \int_{\mathcal{B}} \delta(\hat{u} - c_{\text{th}}) f(\vec{x}, \hat{u}(\vec{x}), \hat{u}) \kappa\, \psi \, d\vec{x} \\
&\qquad + \int_{\mathcal{B}} \delta(\hat{u} - c_{\text{th}}) \vec{\nabla} f(\vec{x}, \hat{u}(\vec{x}), \hat{u}) \cdot \vec{n}_{\Gamma}\, \psi \, d\vec{x},
\end{aligned}$$

where κ is the curvature defined by

$$\kappa = -\vec{\nabla} \cdot \frac{\vec{\nabla}\,\hat{u}}{|\vec{\nabla}\,\hat{u}|}.$$

Remark 7 *We do not take into account the boundary term $\partial\mathcal{B}$*

$$\int_{\partial\mathcal{B}}\Big(\delta(\hat{u}-c_{th})f(\vec{x},\hat{u}(\vec{x}),\hat{u})\frac{1}{|\vec{\nabla}\,\hat{u}|}\Big)\vec{\nabla}\,\hat{u}\cdot\vec{n}\,\psi\,d\vec{x}$$

which appears when using the divergence formula in the previous calculus. For example, we can use this result

$$\vec{\nabla}\,\hat{u}\cdot\vec{n}=0,\ \textit{in}\ \partial\mathcal{B}.$$

In the shape derivative theory, this term does not appear because the contour is supposed to be closed and does not intersect the boundary of the domain.

In our case

$$f(\vec{x},\hat{u}(\vec{x}),\hat{u})=\alpha(|\vec{\nabla}\,\hat{u}|)\Big(\big(z_u-C_1(\hat{u}-c_{\rm th})\big)^2-\big(z_u-C_2(\hat{u}-c_{\rm th})\big)^2\Big)\,\chi.$$

Then the first term becomes

$$\begin{aligned}
&\int_{\mathcal{B}}\delta(\hat{u}-c_{\rm th})|\vec{\nabla}\,\hat{u}|\Big(\partial_{\hat{u}}\Big(\alpha(|\vec{\nabla}\,\hat{u}|)\Big(\big(z_u-C_1(\hat{u}-c_{\rm th})\big)^2-\big(z_u-C_2(\hat{u}-c_{\rm th})\big)^2\Big)\Big);\,\psi_{\hat{u}-c_{\rm th}}\Big)\,\chi\\
&=\int_{\mathcal{B}}\delta(\hat{u}-c_{\rm th})|\vec{\nabla}\,\hat{u}|\alpha'(|\vec{\nabla}\,\hat{u}|)\frac{\vec{\nabla}\,\hat{u}}{|\vec{\nabla}\,\hat{u}|}\cdot\vec{\nabla}\psi_{\hat{u}-c_{\rm th}}\Big(\big(z_u-C_1(\hat{u}-c_{\rm th})\big)^2-\big(z_u-C_2(\hat{u}-c_{\rm th})\big)^2\Big)\,\chi\\
&\qquad-2\Big(\partial_{\hat{u}}C_1(\hat{u}-c_{\rm th});\,\psi_{\hat{u}-c_{\rm th}}\Big)\int_{\mathcal{B}}\delta(\hat{u}-c_{\rm th})|\vec{\nabla}\,\hat{u}|\alpha(|\vec{\nabla}\,\hat{u}|)\big(z_u-C_1(\hat{u}-c_{\rm th})\big)\,\chi\,d\vec{x}\\
&\qquad+2\Big(\partial_{\hat{u}}C_2(\hat{u}-c_{\rm th});\,\psi_{\hat{u}-c_{\rm th}}\Big)\int_{\mathcal{B}}\delta(\hat{u}-c_{\rm th})|\vec{\nabla}\,\hat{u}|\alpha(|\vec{\nabla}\,\hat{u}|)\big(z_u-C_2(\hat{u}-c_{\rm th})\big)\,\chi\,d\vec{x}\\
&=-\int_{\mathcal{B}}\delta(\hat{u}-c_{\rm th})|\vec{\nabla}\,\hat{u}|\alpha'(|\vec{\nabla}\,\hat{u}|)\Big(\big(z_u-C_1(\hat{u}-c_{\rm th})\big)^2-\big(z_u-C_2(\hat{u}-c_{\rm th})\big)^2\Big)\,\chi\,\partial_{\vec{n}_{\Gamma_{\hat{u}}}}\psi_{\hat{u}-c_{\rm th}}\,d\vec{x}\\
&\qquad-2\Big(\partial_{\hat{u}}C_1(\hat{u}-c_{\rm th});\,\psi_{\hat{u}-c_{\rm th}}\Big)\int_{\mathcal{B}}\delta(\hat{u}-c_{\rm th})|\vec{\nabla}\,\hat{u}|\alpha(|\vec{\nabla}\,\hat{u}|)\big(z_u-C_1(\hat{u}-c_{\rm th})\big)\,\chi\,d\vec{x}\\
&\qquad+2\Big(\partial_{\hat{u}}C_2(\hat{u}-c_{\rm th});\,\psi_{\hat{u}-c_{\rm th}}\Big)\int_{\mathcal{B}}\delta(\hat{u}-c_{\rm th})|\vec{\nabla}\,\hat{u}|\alpha(|\vec{\nabla}\,\hat{u}|)\big(z_u-C_2(\hat{u}-c_{\rm th})\big)\,\chi\,d\vec{x}\\
&\qquad=-\int_{\Gamma_{\hat{u}}}\alpha'(|\vec{\nabla}\,\hat{u}|)\Big(\big(z_u-C_1(\hat{u}-c_{\rm th})\big)^2-\big(z_u-C_2(\hat{u}-c_{\rm th})\big)^2\Big)\,\chi\,\partial_{\vec{n}_{\Gamma_{\hat{u}}}}\psi_{\hat{u}-c_{\rm th}}\,d\Gamma_{\hat{u}}\\
&\qquad\qquad-2\Big(\partial_{\hat{u}}C_1(\hat{u}-c_{\rm th});\,\psi_{\hat{u}-c_{\rm th}}\Big)\int_{\Gamma_{\hat{u}}}\alpha(|\vec{\nabla}\,\hat{u}|)\big(z_u-C_1(\hat{u}-c_{\rm th})\big)\,\chi\,d\Gamma_{\hat{u}}\\
&\qquad\qquad+2\Big(\partial_{\hat{u}}C_2(\hat{u}-c_{\rm th});\,\psi_{\hat{u}-c_{\rm th}}\Big)\int_{\Gamma_{\hat{u}}}\alpha(|\vec{\nabla}\,\hat{u}|)\big(z_u-C_2(\hat{u}-c_{\rm th})\big)\,\chi\,d\Gamma_{\hat{u}}.
\end{aligned}$$

We recall that C_1 and C_2 are defined by

$$C_1(\hat{u}-c_{\rm th})=\frac{\int_{\mathcal{B}}z_uH(\hat{u}-c_{\rm th})d\vec{x}}{\int_{\mathcal{B}}H(\hat{u}-c_{\rm th})d\vec{x}}\ \text{and}\ C_2(\hat{u}-c_{\rm th})=\frac{\int_{\mathcal{B}}z_u(1-H(\hat{u}-c_{\rm th}))d\vec{x}}{\int_{\mathcal{B}}(1-H(\hat{u}-c_{\rm th}))d\vec{x}}.$$

We then have

$$\begin{aligned}
\Big(\partial_{\hat{u}}C_1(\hat{u}-c_{\rm th});\,\psi_{\hat{u}-c_{\rm th}}\Big)&=\frac{\int_{\mathcal{B}}z_u\delta(\hat{u}-c_{\rm th})\psi_{\hat{u}-c_{\rm th}}d\vec{x}\int_{\mathcal{B}}H(\hat{u}-c_{\rm th})d\vec{x}}{\big(\int_{\mathcal{B}}H(\hat{u}-c_{\rm th})d\vec{x}\big)^2}\\
&\quad-\frac{\int_{\mathcal{B}}z_uH(\hat{u}-c_{\rm th})d\vec{x}\int_{\mathcal{B}}\delta(\hat{u}-c_{\rm th})\psi_{\hat{u}-c_{\rm th}}d\vec{x}}{\big(\int_{\mathcal{B}}H(\hat{u}-c_{\rm th})d\vec{x}\big)^2}\\
&=\frac{1}{|\Omega_{\hat{u}}^{\rm in}|}\int_{\Gamma_{\hat{u}}}\frac{1}{|\vec{\nabla}\,\hat{u}|}\big(z_u-C_1(\hat{u}-c_{\rm th})\big)\psi_{\hat{u}-c_{\rm th}}d\Gamma_{\hat{u}}.
\end{aligned}$$

and also

$$\Big(\partial_{\hat{u}} C_2(\hat{u} - c_{\rm th});\, \psi_{\hat{u}-c_{\rm th}}\Big) = -\frac{1}{|\mathcal{B}\setminus\overline{\Omega^{\rm in}_{\hat{u}}}|}\int_{\Gamma_{\hat{u}}}\frac{1}{|\vec{\nabla}\,\hat{u}|}\big(z_u - C_2(\hat{u} - c_{\rm th})\big)\psi_{\hat{u}-c_{\rm th}} d\Gamma_{\hat{u}}.$$

We finally have for the first term

$$\begin{aligned}
&\int_{\mathcal{B}}\delta(\hat{u}-c_{\rm th})|\vec{\nabla}\,\hat{u}|\partial_{\hat{u}}\Big(\alpha(|\vec{\nabla}\,\hat{u}|)\Big(\big(z_u - C_1(\hat{u}-c_{\rm th})\big)^2 - \big(z_u - C_2(\hat{u}-c_{\rm th})\big)^2\Big)\Big)\,\chi\, d\vec{x}\\
&\quad = -\int_{\Gamma_{\hat{u}}}\alpha'(|\vec{\nabla}\,\hat{u}|)\Big(\big(z_u - C_1(\hat{u}-c_{\rm th})\big)^2 - \big(z_u - C_2(\hat{u}-c_{\rm th})\big)^2\Big)\,\chi\,\partial_{\vec{n}_{\Gamma_{\hat{u}}}}\psi_{\hat{u}-c_{\rm th}}\, d\Gamma_{\hat{u}}\\
&\quad - 2\frac{1}{|\Omega^{\rm in}_{\hat{u}}|}\int_{\Gamma_{\hat{u}}}\frac{1}{|\vec{\nabla}\,\hat{u}|}\big(z_u - C_1(\hat{u}-c_{\rm th})\big)\psi_{\hat{u}-c_{\rm th}} d\Gamma_{\hat{u}}\int_{\Gamma_{\hat{u}}}\alpha(|\vec{\nabla}\,\hat{u}|)\big(z_u - C_1(\hat{u}-c_{\rm th})\big)\,\chi\, d\Gamma_{\hat{u}}\\
&\quad - 2\frac{1}{|\mathcal{B}\setminus\overline{\Omega^{\rm in}_{\hat{u}}}|}\int_{\Gamma_{\hat{u}}}\frac{1}{|\vec{\nabla}\,\hat{u}|}\big(z_u - C_2(\hat{u}-c_{\rm th})\big)\psi_{\hat{u}-c_{\rm th}} d\Gamma_{\hat{u}}\int_{\Gamma_{\hat{u}}}\alpha(|\vec{\nabla}\,\hat{u}|)\big(z_u - C_2(\hat{u}-c_{\rm th})\big)\,\chi\, d\Gamma_{\hat{u}}.
\end{aligned}$$

We continue with the third term

$$\begin{aligned}
&\int_{\mathcal{B}}\delta(\hat{u}-c_{\rm th})\vec{\nabla}\Big(\alpha(|\vec{\nabla}\,\hat{u}|)\Big(\big(z_u - C_1(\hat{u}-c_{\rm th})\big)^2 - \big(z_u - C_2(\hat{u}-c_{\rm th})\big)^2\Big)\chi\Big)\cdot\vec{n}_{\Gamma_{\hat{u}}}\,\psi_{\hat{u}-c_{\rm th}}\, d\vec{x}\\
&\quad = \int_{\mathcal{B}}\delta(\hat{u}-c_{\rm th})\vec{\nabla}\alpha(|\vec{\nabla}\,\hat{u}|)\cdot\vec{n}_{\Gamma_{\hat{u}}}\Big(\big(z_u - C_1(\hat{u}-c_{\rm th})\big)^2 - \big(z_u - C_2(\hat{u}-c_{\rm th})\big)^2\Big)\chi\,\psi_{\hat{u}-c_{\rm th}}\, d\vec{x}\\
&\quad + \int_{\mathcal{B}}\delta(\hat{u}-c_{\rm th})\alpha(|\vec{\nabla}\,\hat{u}|)\vec{\nabla}\Big(\Big(\big(z_u - C_1(\hat{u}-c_{\rm th})\big)^2 - \big(z_u - C_2(\hat{u}-c_{\rm th})\big)^2\Big)\Big)\cdot\vec{n}_{\Gamma_{\hat{u}}}\chi\,\psi_{\hat{u}-c_{\rm th}}\, d\vec{x}\\
&\quad + \int_{\mathcal{B}}\delta(\hat{u}-c_{\rm th})\Big(\alpha(|\vec{\nabla}\,\hat{u}|)\Big(\big(z_u - C_1(\hat{u}-c_{\rm th})\big)^2 - \big(z_u - C_2(\hat{u}-c_{\rm th})\big)^2\Big)\Big)\vec{\nabla}\chi\cdot\vec{n}_{\Gamma_{\hat{u}}}\,\psi_{\hat{u}-c_{\rm th}}\, d\vec{x}\\
&\quad = -\int_{\mathcal{B}}\delta(\hat{u}-c_{\rm th})\alpha'(|\vec{\nabla}\,\hat{u}|)\partial^2_{\vec{n}_{\Gamma_{\hat{u}}}}\hat{u}\Big(\big(z_u - C_1(\hat{u}-c_{\rm th})\big)^2 - \big(z_u - C_2(\hat{u}-c_{\rm th})\big)^2\Big)\chi\,\psi_{\hat{u}-c_{\rm th}}\, d\vec{x}\\
&\qquad\quad + 2\int_{\mathcal{B}}\delta(\hat{u}-c_{\rm th})\alpha(|\vec{\nabla}\,\hat{u}|)\vec{\nabla}z_u\cdot\vec{n}_{\Gamma_{\hat{u}}}\big(C_2(\hat{u}-c_{\rm th}) - C_1(\hat{u}-c_{\rm th})\big)\chi\,\psi_{\hat{u}-c_{\rm th}}\, d\vec{x}\\
&\quad + \int_{\mathcal{B}}\delta(\hat{u}-c_{\rm th})\Big(\alpha(|\vec{\nabla}\,\hat{u}|)\Big(\big(z_u - C_1(\hat{u}-c_{\rm th})\big)^2 - \big(z_u - C_2(\hat{u}-c_{\rm th})\big)^2\Big)\Big)\vec{\nabla}\chi\cdot\vec{n}_{\Gamma_{\hat{u}}}\,\psi_{\hat{u}-c_{\rm th}}\, d\vec{x}\\
&\quad = -\int_{\Gamma_{\hat{u}}}\frac{\alpha'(|\vec{\nabla}\,\hat{u}|)}{|\vec{\nabla}\,\hat{u}|}\partial^2_{\vec{n}_{\Gamma_{\hat{u}}}}\hat{u}\Big(\big(z_u - C_1(\hat{u}-c_{\rm th})\big)^2 - \big(z_u - C_2(\hat{u}-c_{\rm th})\big)^2\Big)\chi\,\psi_{\hat{u}-c_{\rm th}}\, d\Gamma_{\hat{u}}\\
&\qquad\quad + 2\int_{\Gamma_{\hat{u}}}\frac{\alpha(|\vec{\nabla}\,\hat{u}|)}{|\vec{\nabla}\,\hat{u}|}\vec{\nabla}z_u\cdot\vec{n}_{\Gamma_{\hat{u}}}\big(C_2(\hat{u}-c_{\rm th}) - C_1(\hat{u}-c_{\rm th})\big)\chi\,\psi_{\hat{u}-c_{\rm th}}\, d\Gamma_{\hat{u}}\\
&\quad + \int_{\Gamma_{\hat{u}}}\frac{\alpha(|\vec{\nabla}\,\hat{u}|)}{|\vec{\nabla}\,\hat{u}|}\Big(\big(z_u - C_1(\hat{u}-c_{\rm th})\big)^2 - \big(z_u - C_2(\hat{u}-c_{\rm th})\big)^2\Big)\vec{\nabla}\chi\cdot\vec{n}_{\Gamma_{\hat{u}}}\,\psi_{\hat{u}-c_{\rm th}}\, d\Gamma_{\hat{u}}.
\end{aligned}$$

We finally obtain

$$
\begin{aligned}
&\Big(d\mathcal{Q}^{\Gamma_{\hat{u}}}(\chi);\, \psi_{\hat{u}-c_{\rm th}}\Big) \\
&\quad = -\int_{\Gamma_{\hat{u}}} \alpha'(|\vec{\nabla}\,\hat{u}|)\Big(\big(z_u - C_1(\hat{u}-c_{\rm th})\big)^2 - \big(z_u - C_2(\hat{u}-c_{\rm th})\big)^2\Big)\,\chi\,\partial_{\vec{n}_{\Gamma_{\hat{u}}}}\psi_{\hat{u}-c_{\rm th}}\,d\Gamma_{\hat{u}} \\
&\quad - 2\frac{1}{|\Omega^{\rm in}_{\hat{u}}|}\int_{\Gamma_{\hat{u}}}\frac{1}{|\vec{\nabla}\,\hat{u}|}\big(z_u - C_1(\hat{u}-c_{\rm th})\big)\psi_{\hat{u}-c_{\rm th}}d\Gamma_{\hat{u}}\int_{\Gamma_{\hat{u}}}\alpha(|\vec{\nabla}\,\hat{u}|)\big(z_u - C_1(\hat{u}-c_{\rm th})\big)\,\chi\,d\Gamma_{\hat{u}} \\
&\quad - 2\frac{1}{|\mathcal{B}\setminus\overline{\Omega^{\rm in}_{\hat{u}}}|}\int_{\Gamma_{\hat{u}}}\frac{1}{|\vec{\nabla}\,\hat{u}|}\big(z_u - C_2(\hat{u}-c_{\rm th})\big)\psi_{\hat{u}-c_{\rm th}}d\Gamma_{\hat{u}}\int_{\Gamma_{\hat{u}}}\alpha(|\vec{\nabla}\,\hat{u}|)\big(z_u - C_2(\hat{u}-c_{\rm th})\big)\,\chi\,d\Gamma_{\hat{u}} \\
&\qquad + \int_{\Gamma_{\hat{u}}}\kappa\,\frac{\alpha(|\vec{\nabla}\,\hat{u}|)}{|\vec{\nabla}\,\hat{u}|}\Big(\big(z_u - C_1(\hat{u}-c_{\rm th})\big)^2 - \big(z_u - C_2(\hat{u}-c_{\rm th})\big)^2\Big)\chi\,\psi_{\hat{u}-c_{\rm th}}\,d\Gamma_{\hat{u}} \\
&\qquad - \int_{\Gamma_{\hat{u}}}\frac{\alpha'(|\vec{\nabla}\,\hat{u}|)}{|\vec{\nabla}\,\hat{u}|}\partial^2_{\vec{n}_{\Gamma_{\hat{u}}}}\hat{u}\Big(\big(z_u - C_1(\hat{u}-c_{\rm th})\big)^2 - \big(z_u - C_2(\hat{u}-c_{\rm th})\big)^2\Big)\chi\,\psi_{\hat{u}-c_{\rm th}}\,d\Gamma_{\hat{u}} \\
&\qquad\quad + 2\int_{\Gamma_{\hat{u}}}\frac{\alpha(|\vec{\nabla}\,\hat{u}|)}{|\vec{\nabla}\,\hat{u}|}\vec{\nabla}z_u\cdot\vec{n}_{\Gamma_{\hat{u}}}\big(C_2(\hat{u}-c_{\rm th}) - C_1(\hat{u}-c_{\rm th})\big)\chi\,\psi_{\hat{u}-c_{\rm th}}\,d\Gamma_{\hat{u}} \\
&\quad + \int_{\Gamma_{\hat{u}}}\frac{\alpha(|\vec{\nabla}\,\hat{u}|)}{|\vec{\nabla}\,\hat{u}|}\Big(\big(z_u - C_1(\hat{u}-c_{\rm th})\big)^2 - \big(z_u - C_2(\hat{u}-c_{\rm th})\big)^2\Big)\vec{\nabla}\chi\,\cdot\vec{n}_{\Gamma_{\hat{u}}}\,\psi_{\hat{u}-c_{\rm th}}\,d\Gamma_{\hat{u}}.
\end{aligned}
$$

We find the results obtained with the shape derivatives given in Proposition 1 with

$$\vec{\psi}_{\Gamma_{\hat{u}}} = \frac{1}{|\vec{\nabla}\,\hat{u}|}\psi_{\hat{u}-c_{\rm th}}\vec{n}_{\Gamma_{\hat{u}}}.$$

References

[1] D. G. Aronso and H. F. Weinberger. Multidimensional nonlinear diffusion arising in population genetics. *Advances in Mathematics*, 30:33–76, 1978.

[2] M.I. Asensio and L. Ferragut. On a wildland fire model with radiation. *International Journal for Numerical Methods in Engineering*, 54(1):137–157, 2002.

[3] M. Bendahmane and H.K. Karlsen. Analysis of a class of degenerate reaction-diffusion systems and the bidomain model of cardiac tissue. *Networks and Heterogeneous Media*, 1:185–218, 2006.

[4] A. Bensoussan. *Filtrage optimal des systèmes linéaires*. Dunod, 1971.

[5] M. Boulakia, J.-F. Gerbeau, and E. Schenone. Reduced-order modeling for cardiac electrophysiology. Application to parameter identification. *International Journal for Numerical Methods in Biomedical Engineering*, 28:727–744, 2012.

[6] Y. Bourgault, Y. Coudière, and C. Pierre. Existence and uniqueness of the solution for the bidomain model used in cardiac electrophysiology. *Nonlinear Analysis-Real World Applications*, 10(1):458–482, 2009.

[7] C. Bui, C. Dapogny, and P. Frey. An accurate anisotropic adaptation method for solving the level set advection equation. *International Journal for Numerical Methods in Fluids*, 70(7):899–922, 2012.

[8] T.F. Chan and L.A. Vese. Active contours without edges. *IEEE Transactions on Image Processing*, 10(2):266–277, 1991.

[9] D. Chapelle, A. Collin, and J.-F. Gerbeau. A surface-based electrophysiology model relying on asymptotic analysis and motivated by cardiac atria modeling. *M3AS*, 23(14):2749–2776, 2013.

[10] T. Colin, A. Iollo, Lombardi. D, and O. Saut. System identification in tumor growth modeling using semi-empirical eigenfunctions. *M3AS*, 2(1):137–157, 2012.

[11] P. Colli Franzone, L.F. Pavarino, and B. Taccardi. Simulating patterns of excitation, repolarization and action potential duration with cardiac bidomain and monodomain models. *Mathematical Biosciences*, 197(1):35–66, 2005.

[12] P. Colli Franzone and G. Savaré. Degenerate evolution systems modeling cardiac electric field at micro and macroscopic. *Evolution Equations, Semigroups and Functional Analysis. Progress in Nonlinear Differential Equations and Their Applications*, 50:49–78, 2002.

[13] A. Collin, J.-F. Gerbeau, M. Hocini, M. Haïssaguerre, and D. Chapelle. Surface-based electrophysiology modeling and assessment of physiological simulations in atria. *FIMH 2013*, 7945:352–359, 2013.

[14] A. Collin, J.-F. Gerbeau, and E. Schenone. Simulations of full electrocardiogram cycles. 2014. In preparation.

[15] C. Corrado, J.-F. Gerbeau, and P. Moireau. Identification of weakly coupled multiphysics problems. application to the inverse problem of electrocardiology. *Journal of Computational Physics*, 2014. In review.

[16] M. Courtemanche, R.J. Ramirez, and S. Nattel. Ionic mechanisms underlying human atrial action potential properties: insights from a mathematical model. *American Journal of Physiology*, (275):H301–H321, 1998.

[17] M.C. Delfour and J.-P. Zolésio. *Shapes and Geometries: Analysis, Differential Calculus, and Optimization (Advances in Design and Control)*. SIAM, second edition, 2011.

[18] B. Engquist, A.-K. Tornberg, and R. Tsai. Discretization of Dirac delta functions in level set methods. *Journal of Computational Physics*, 207(1):28–51, 2005.

[19] G. Evensen. *Data Assimilation: The Ensemble Kalman Filter*. Springer, 2009.

[20] S.C. Ferreira, Jr. M. L. Martins, and M. J. Vilela. Reaction-diffusion model for the growth of avascular tumor. *Physical Review E*, 65:021907 (8pp), 2002.

[21] W. H. Fleming. Deterministic nonlinear filtering. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 25(3-4):435–454, 1997.

[22] P.C Franzone, L. Guerri, and S. Rovida. Wavefront propagation in an activation model of the anisotropic cardiac tissue: asymptotic analysis and numerical simulations. *Journal of Mathematical Biology*, 28:121–176, 1990.

[23] S. Göktepe and E. Kuhl. Computational modeling of cardiac electrophysiology: A novel finite element approach. *International Journal for Numerical Methods in Engineering*, 79(2):156–178, 2009.

[24] M. Haïssaguerre, D. C. Shah, Jaïs P., M. Hocini, T. Yamane, I. Deisenhofer, S. Garrigue, and J. Clémenty. Mapping-guided ablation of pulmonary veins to cure atrial fibrillation. *The American Journal of Cardiology*, 86(9):K9–K19, 2000.

[25] D.M. Harrild and S.H. Craig. A computer model of normal conduction in the human atria. *Circulation Research*, (87):e25–e36, 2000.

[26] M. Hintermüller and W. Ring. An inexact Newton-C-G-type active contour approach for the minimization of the Mumford-Shah functional. *Journal of Mathematical Imaging and Vision*, 20:19–42, 2004.

[27] C. Hogea, C. Davatzikos, and G. Biros. An image-driven parameter estimation problem for a reaction–diffusion glioma growth model with mass effects. *Journal of Mathematical Biology*, 56(6):793–825, 2008.

[28] J. E. Hoke and R. A. Anthes. The initialization of numerical models by a dynamic-initialization technique. *Monthly Weather Review*, 104(12):1551–1556, 1976.

[29] V. Jacquemet. An eikonal approach for the initiation of reentrant cardiac propagation in reaction-diffusion models. *IEEE Transactions On Biomedical Engineering*, 57(9):2090–2098, 2010.

[30] S. J. Julier and J. K. Uhlmann. Reduced sigma point filters for the propagation of means and covariances through nonlinear transformations. In *In Proceedings of the 2002 American Control Conference*, volume 2, pages 887–892, 2002.

[31] S.J. Julier and J.K. Uhlmann. A new extension of the Kalman filter to nonlinear systems. In *Proc. of AeroSense: The 11th Int. Symp. on Aerospace/Defence Sensing, Simulation and Controls*, 1997.

[32] R. Kalman and R. Bucy. New results in linear filtering and prediction theory. *Trans. ASME J. Basic. Eng.*, 83:95—108, 1961.

[33] A. Karma. Electrical alternans and spiral wave breakup in cardiac tissue. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 4(3):461–472, 1994.

[34] J. Keener and J. Sneyd. *Mathematical Physiology*. Springer, 2004.

[35] J.P. Keener. An eikonal-curvature equation for action potential propagation in myocardium. *Journal of Mathematical Biology*, 29:629–651, 1991.

[36] E. Konukoglu, O. Clatz, B. H. Menze, B. Stieltjes, M.-A. Weber, E. Mandonnet, H. Delingette, and N. Ayache. Image guided personalization of reaction-diffusion type tumor growth models using modified anisotropic eikonal equations. *Medical Imaging, IEEE Transactions*, 29(1):77–95, 2010.

[37] D.G. Luenberger. *Determining the State of a Linear with Observers of Low Dynamic Order*. PhD thesis, Stanford University, 1963.

[38] D.G. Luenberger. An introduction to observers. *IEEE Transactions on Automatic Control*, 16:596–602, 1971.

[39] J. Malmivuo and R. Plonsey. *Bioelectromagnetism - Principles and Applications of Bioelectric and Biomagnetic Fields*. Oxford University Press, 1995.

[40] J. Mandel, L.S. Bennethum, J.D. Beezley, J.L. Coen, C.C Douglas, L.P. Franca, M. Kim, and A. Vodacek. A wildland fire model with data assimilation. *Mathematics and Computers in Simulation*, 79(3):584–606, 2008.

[41] C.C. Mitchell and D.G. Schaeffer. A two-current model for the dynamics of cardiac membrane. *Bulletin Math. Bio.*, 65:767–793, 2003.

[42] P. Moireau and D. Chapelle. Erratum of article "Reduced-order Unscented Kalman Filtering with application to parameter identification in large-dimensional systems". *ESAIM: Control, Optimisation and Calculus of Variations*, 17(2):406–409, 2011.

[43] P. Moireau and D. Chapelle. Reduced-order Unscented Kalman Filtering with application to parameter identification in large-dimensional systems. *ESAIM: Control, Optimisation and Calculus of Variations*, 17(2):380–405, 2011.

[44] P. Moireau, D. Chapelle, and P. Le Tallec. Joint state and parameter estimation for distributed mechanical systems. *Computer Methods in Applied Mechanics and Engineering*, 1987(6–8):659–677, 2008.

[45] S. Osher and R. Fedkiw. *Level Set Methods and Dynamic Implicit Surfaces*. Applied Mathematical Sciences, 2002.

[46] M. Potse, B. Dubé, J. Richer, A. Vinet, and R.M. Gulrajani. A comparison of monodomain and bidomain reaction-diffusion models for action potential propagation in the human heart. *IEEE Transactions on Biomedical Engineering*, 53(12):2425–2435, 2006.

[47] C. Ramanathan, R.J. Ghanem, P. Jia, K. Ryu, and Y. Rudy. Noninvasive electrocardiographic imaging for cardiac electrophysiology and arrhythmia. *Nature Medicine*, 10:422–428, 2004.

[48] C.M. Rochoux, B. Cuenot, S. Ricci, A. Trouvé, B. Delmotte, Se. Massart, R Paoli, and R. Paugam. Data assimilation applied to combustion. *Comptes Rendus Mécanique*, 341:266–276, 2013.

[49] F.B. Sachse. *Computational Cardiology: Modeling of Anatomy, Electrophysiology and Mechanics*. Springer-Verlag, 2004.

[50] M. Sermesant, K. Rhode, G.I. Sanchez-Ortiz, 0. Camara, R. Andriantsimiavona, S. Hegde, D. Rueckert, P. Lambiase, C. Bucknall, E. Rosenthal, H. Delingette, D.L.G Hill, N. Ayache, and R. Razavi. Simulation of cardiac pathologies using an electromechanical biventricular model and {XMR} interventional imaging. *Medical Image Analysis*, 9(5):467–480, 2005.

[51] D. Simon. *Optimal State Estimation: Kalman, H^∞, and Nonlinear Approaches*. Wiley-Interscience, 2006.

[52] D. R. Stauffer and N. L. Seaman. Use of four-dimensional data assimilation in a limited-area mesoscale model. Part I: Experiments with synoptic-scale data. *Monthly Weather Review*, 118(6):1250–1277, 1990.

[53] J. Sundnes, G.T. Lines, X. Cai, B.F. Nielsen, K.A. Mardal, and A. Tveito. *Computing the Electrical Activity in the Heart*, volume 1 of *Monographs in Computational Science and Engineering*. Springer-Verlag, 2006.

[54] K.R. Swanson, R.C. Rostomily, and E.C. Alvord. A mathematical modelling tool for predicting survival of individual patients following resection of glioblastoma: a proof of principle. *British Journal of Cancer*, 98(1):113–119, 2008.

[55] P. Tracqui and M. Mendjeli. Modelling three-dimensional growth of brain tumours from time series of scans. *Mathematical Models and Methods in Applied Sciences*, 9(04):581–598, 1999.

[56] L. Tung. *A bi-domain model for describing ischemic myocardial d-c potentials*. PhD thesis, Massachusetts Institute of Technology. Dept. of Electrical Engineering and Computer Science, 1978.

[57] M. Veneroni. Reaction–diffusion systems for the macroscopic bidomain model of the cardiac electric field. *Nonlinear Analysis: Real World Applications*, 10(2):849–868, 2009.

[58] H.K. Zhao, T. Chan, B. Merriman, and S. Osher. A variational level set approach to multiphase motion. *Journal of Computational Physics*, 127(1):179–195, 1996.

Conclusion et Perspectives

De nombreuses pathologies cardio-vasculaires sont dues à des troubles électrophysiologiques qui perturbent le rythme cardiaque. Les problèmes mathématiques qui se posent sur ce sujet sont divers, ce qui explique que la littérature soit si abondante. Nous avons essayé dans cette thèse d'apporter des contributions à ce vaste projet de modélisation dont le but final est de produire des simulations – avec des temps de calcul raisonnables et adaptées à chaque patient – pouvant aider les médecins. Les principaux enseignements et les conclusions que nous avons tirés de ce travail sont donnés dans ce dernier chapitre. Les outils innovants développés dans cette thèse amènent de nouvelles questions, et des perspectives seront proposées. Nous commençons par une liste de conclusions et de perspectives par chapitre. Les équations qui sont référencées renvoient vers celles données dans l'introduction.

Conclusions et perspectives par chapitre

Chapitre 2

Conclusions Une compréhension très profonde du modèle bidomaine posé sur un domaine fixé est nécessaire si on veut pouvoir étudier l'impact des déformations mécaniques et proposer une extension rigoureuse de ce modèle en domaine mobile. C'est pourquoi nous avons proposé dans une première partie de ce chapitre une étude détaillée s'appuyant sur la littérature de la mise en équation du modèle bidomaine. Nous avons vu aussi que le processus d'homogénéisation conduit à un modèle qui peut être directement formulé dans le cadre de la théorie des mélanges. Cependant, cette méthode donne moins d'informations dans le cas périodique (pas de problème cellule identifié). Cette théorie nous a permis d'obtenir un modèle bidomaine (4) prenant en compte les déformations mécaniques. Les simulations que nous avons faites sous une hypothèse d'incompressibilité du cœur ont montré que l'incidence des déplacements est du même ordre que le déplacement lui-même.

Perspectives Les *gap junctions* qui relient électriquement les cellules entre elles ne sont pas considérées dans le processus d'homogénéisation. Une nouvelle analyse prenant en compte ce qui est peut-être une deuxième échelle asymptotique est une première perspective à ce travail. Dans [2], une homogénéisation des *gap junctions* est proposée mais c'est le modèle bidomaine déjà homogénéisé, et non pas le modèle intra-cellulaire, qui est écrit dans chaque cellule. Les *gap junctions* ne sont donc pas étudiées au niveau microscopique et une double asymptotique n'est donc pas considérée.

Dans les simulations du modèle mécano-électrique (4), nous avons négligé certains termes sous des hypothèses d'incompressibilité du cœur. De plus, la déformation mécanique considérée manque de réalisme. Une prise en compte dans les simulations de tous les termes devrait être faite, afin de réellement mesurer leur impact un à un. Enfin, le couplage avec un modèle mécanique cardiaque performant [8] permettrait d'obtenir des simulations réalistes que nous pourrions alors comparer avec les simulations obtenues dans la littérature [6, 3].

Chapitre 3 et Chapitre 5

Conclusions L'objectif du Chapitre 3 était de trouver un modèle bidomaine réduit capable de prendre en compte des fortes anisotropies dans l'épaisseur, situation rencontrée dans les oreillettes cardiaques. En utilisant une analyse asymptotique, nous avons justifié et obtenu un modèle bidomaine surfacique (12). La comparaison des simulations de ce modèle avec le modèle 3D nous a montré que cette analyse était nécessaire et qu'un modèle naïf qui moyenne l'anisotropie n'est pas suffisant, en particulier dans des cas pathologiques. La mise en place des simulations sur une géométrie atriale réelle données dans le Chapitre 5 montre la complexité des oreillettes. En plus de la variation des fibres dans l'épaisseur, la modélisation des oreillettes nécessite la prise en compte des faisceaux rapides et lents ainsi qu'un modèle ionique adapté si l'on souhaite obtenir des simulations réalistes.

Perspectives Grâce au modèle surfacique bidomaine (12) proposé dans le Chapitre 3, nous sommes donc en mesure d'obtenir des simulations réalistes mais toujours à domaine fixé. Un modèle mécanique surfacique des oreillettes pourrait aussi être dérivé. Comme la contraction se fait principalement dans la direction des fibres, les mêmes questions de prise en compte de l'anisotropie sont posées. C'est un projet ambitieux puisque les modèles de coques actuels ne prennent pas en compte ce type d'anisotropie dans l'épaisseur. L'objectif global serait d'avoir un jour un modèle surfacique d'oreillettes complet capable de prendre en compte les phénomènes électriques et mécaniques ainsi que leurs interactions.

Chapitre 6

Conclusions Les excellents résultats que nous obtenons dans la simulation d'électrocardiogrammes – que ce soit dans un cas sain ou dans des cas pathologiques – donnent confiance dans le modèle bidomaine couplé, et en particulier dans le modèle bidomaine asymptotique que nous avons établi dans le Chapitre 3. Sans réglage supplémentaire sur le modèle d'oreillettes, les ondes P ont le bon signe ce qui signifie que la direction du signal à chaque instant dans nos simulations atriales est réaliste.

Ces simulations ont aussi été l'occasion de mieux comprendre l'interaction entre les oreillettes et les ventricules. En effet bien qu'ils soient électriquement isolés, nous ne pouvons pas séparer chacune des deux simulations puisque le potentiel extra-cellulaire est défini à une constante près. En séparant les deux simulations et donc en réglant la constante séparément, un dipôle artificiel serait créé.

Perspectives Toutes les simulations des électrocardiogrammes (ECGs) que nous proposons sont effectuées sur une géométrie fixe. Il est intéressant de se demander quelles sont les effets de la contraction cardiaque sur les ECGs. En effet les ventricules se contractent avec un effet de torsion, ce qui peut avoir un impact sur la direction de propagation de l'onde et induire une variation dans l'ECG.

Une deuxième perspective sur ce sujet est la prise en compte d'un vrai couplage oreillettes-ventricules avec la modélisation du nœud atrio-ventriculaire ainsi que des fibres

de Purkinje. En effet dans ce chapitre, les ventricules sont stimulés manuellement en appliquant un courant directement dans toute la zone proche des fibres de Purkinje.

Une dernière perspective serait de voir s'il est possible d'améliorer les conditions de couplage que nous avons utilisées, par exemple avec une prise en compte du péricarde. L'idée serait de mieux comprendre ce qui induit une amplitude de l'onde P si petite par rapport à l'onde R des ventricules (de l'ordre de 1/5) alors que la surface externe des oreillettes est de taille comparable à la surface externe des ventricules (de l'ordre de 2/3). La littérature médicale suggère que c'est dû à l'épaisseur du muscle cardiaque plus fin dans les oreillettes que dans les ventricules. Peut-être qu'un couplage fort serait une solution à ce problème puisque pour le moment avec le couplage faible, seul le potentiel sur l'épicarde se diffuse dans le reste du corps.

Chapitre 8

Conclusions Dans ce chapitre, nous proposons un estimateur d'état (15) de type *Luenberger* [4] pour un modèle de réaction-diffusion et des données de type position de fronts. Notre estimation d'état est compatible avec l'ajout d'un filtre de paramètres de type RoUKF [5] afin d'estimer les paramètres du modèle. La mise en place de l'estimateur d'état nous a montré qu'il était difficile de comparer une donnée de type ligne de niveau avec la solution d'un modèle de réaction diffusion. En jonglant avec plusieurs théories mathématiques – comme l'analyse asymptotique, les techniques de traitement d'image ou encore les dérivées de forme – nous avons pu dériver et justifier cet estimateur. Les résultats sur données synthétiques que nous obtenons sont excellents.

Perspectives Les simulations 1D et 2D de ce chapitre sur des données synthétiques montrent tout le potentiel de l'observateur que nous proposons. Cependant, pour pouvoir le valider définitivement, des simulations sur un cas réel – que ce soit dans le domaine de l'électrophysiologie ou encore dans celui de la propagation de feu – est nécessaire. Une méthode permettant d'identifier la condition initiale dans certains cas – comme les pathologies où une onde spirale apparait – devrait être proposée. Concernant l'estimation paramétrique, des simulations permettant de retrouver d'autres paramètres – comme par exemple ceux du terme de réaction – ou encore le même paramètre mais décomposé dans plusieurs zones sont envisagées. Pour finir, des cas où l'incertitude concerne plusieurs paramètres permettraient d'étudier leur interaction les uns avec les autres.

Du point de vue numérique, des difficultés de discrétisation du terme de correction – qui est un terme surfacique qui agit uniquement sur le front – ont été rencontrées. En effet, la méthode qui consiste à discrétiser la fonction Dirac multidimensionnelle entraîne des oscillations près du front. Une décomposition de Garlekin directement du terme surfacique – en utilisant des maillages s'adaptant au front au cours du temps [1] – permettrait de répondre à cette difficulté numérique. C'est une des perspectives de ce travail.

Pour finir, la dernière perspective concerne la difficulté du cas 3D en électrophysiologie. En effet, la méthode qui permet de reconstruire les cartes de dépolarisation d'un patient – c'est à dire la résolution d'un problème inverse à partir de signaux obtenus par une veste d'électrode [7] – ne donne la position du front qu'à l'épicarde, surface externe du cœur

(et parfois à l'endocarde, surface interne). Développer des techniques mathématiques permettant d'appliquer l'observateur uniquement sur la surface externe du cœur est alors un véritable enjeu.

Conclusion générale

Pour conclure, des outils innovants s'appuyant sur des théories mathématiques variées ont été développés dans cette thèse. L'approche mathématique a toujours été appréhendée dans le but d'améliorer la modélisation en électrophysiologie cardiaque.

Cette thèse a donc été l'occasion, à partir d'une problématique initiale principalement médicale d'introduire, de développer ou d'utiliser des techniques, des outils ou encore des théories appartenant au domaine mathématique. Nous avons dû aussi nous appuyer sur des modélisations poussées développées dans le domaine biologique et parfois les formaliser du point de vue mathématique.

Cette interdisciplinarité représente une difficulté pour ces sujets aux carrefours de plusieurs sciences puisque cela demande des connaissances dans chacun des domaines. L'enjeu de médiation qui permettra d'améliorer la communication entre spécialistes de domaines différents est aussi très important. Cependant, cela apporte aux mathématiques purement théoriques sous-jacentes un sens concret et un objectif stimulant. Nous espérons avec cette thèse avoir offert des résultats novateurs et des réflexions pertinentes qui trouveront leur place dans l'ensemble des outils médicaux – innovants et prometteurs – développés grâce à la modélisation en médecine.

Bibliographie

[1] C. Bui, C. Dapogny, and P. Frey. An accurate anisotropic adaptation method for solving the level set advection equation. *International Journal for Numerical Methods in Fluids*, 70(7) :899–922, 2012.

[2] P. E. Hand and C. S. Peskin. Homogenization of an electrophysiological model for a strand of cardiac myocytes with gap-junctional and electric-field coupling. *Bulletin of Mathematical Biology*, 72 :1408–1424, 2010.

[3] R.H. Keldermann, M.P. Nash, and A.V. Panfilov. Modeling cardiac mechano-electrical feedback using reaction-diffusion-mechanics systems. *Physica D*, 238 :1000–1007, 2008.

[4] D.G. Luenberger. An introduction to observers. *IEEE Transactions on Automatic Control*, 16 :596–602, 1971.

[5] P. Moireau and D. Chapelle. Reduced-order Unscented Kalman Filtering with application to parameter identification in large-dimensional systems. *ESAIM : Control, Optimisation and Calculus of Variations*, 17(2) :380–405, 2011.

[6] M.P. Nash and A.V. Panfilov. Electromechanical model of excitable tissue to study reentrant cardiac arrhythmias. *Progress in Biophysics and Molecular Biology*, 85(2-3) :501–522, 2004.

[7] C. Ramanathan, R.J. Ghanem, P. Jia, K. Ryu, and Y. Rudy. Noninvasive electrocardiographic imaging for cardiac electrophysiology and arrhythmia. *Nature Medicine*, 10 :422–428, 2004.

[8] J. Sainte-Marie, D. Chapelle, R. Cimrman, and M. Sorine. Modeling and estimation of the cardiac electromechanical activity. *Computers and Structures*, 84 :1743–1759, 2006.