On Boundary Stimulation and Optimal Boundary Control of the Bidomain Equations

Chamakuri Nagaiah[1*], Karl Kunisch[1,2†], Gernot Plank[3‡]

[1] Radon Institute for Computational and Applied Mathematics,
Altenbergerstr. 69, Linz, A-4040 Austria.
[2] Institute of Mathematics and Scientific Computing,
University of Graz, Heinrichstr. 36, Graz, A-8010 Austria.
[3] Institute of Biophysics, Medical University of Graz,
Harrachgasse 21, Graz, A-8010 Austria.

September 26, 2012

Abstract

The bidomain equations with Neumann boundary stimulation and optimal control of these stimuli are investigated. First an analytical framework for boundary control is provided. Then a parallel finite element based algorithm is devised and its efficiency is demonstrated not only for the direct problem but also for the optimal control problem. The computations realize a model configuration corresponding to optimal boundary defibrillation of a reentry phenomenon by applying current density stimuli.

1 Introduction

The heart is an electrically-controlled mechanical pump which drives blood through the circulatory system with remarkable efficiency. Under healthy conditions its electrical activation is highly organized, in disease, however, disturbances in the formation and/or propagation of electrical impulses may induce reentrant activation patterns which precipitate its rhythm significantly. Ultimately, such fast rhythms may transition to highly disorganized almost chaotic activation patterns, an electrical state referred to as fibrillation. Under such conditions the heart looses its capacity to pump a sufficiently large mass of blood through the circulatory system. Without therapy, death would ensue within minutes. The only reliable therapy to terminate this otherwise lethal condition is the delivery of a strong electrical shock. This therapy, referred to as electrical defibrillation, is nowadays reliably achieved in a large patient population via the implantation of devices, so-called implantable cardioverters defibrillators (ICD), which monitor the heart rate and, if needed, deliver a discharge to restore a normal rhythm.

Although ICD therapy has proved to be efficient and reliable in preventing sudden cardiac death [2], it is far from ideal. There are several known adverse effects secondary to the administration of strong electrical shocks which are caused by the high field strengths required to terminate fibrillation with a sufficiently high probability. Moreover, psychological effects play an important role as well since shock delivery is perceived as extremely painful by conscious patients, leading to traumatization and a reduced quality of life. The

*nagaiah.chamakuri@ricam.oeaw.ac.at
†karl.kunisch@uni-graz.at
‡gernot.plank@medunigraz.at

link between the high shock strengths required and adverse effects provides the motivation for posing the defibrillation process as an optimization problem, where one aims to achieve defibrillation with minimal energy and, consequently, with minimal detrimental side effects.

The optimal control approach to defibrillation is to determine an applied electrical field in such a way that it optimizes a given design objective, which is, in our case, the restoration of a tissue state in which fibrillatory propagation cannot be maintained. This can be achieved by driving the whole tissue to a resting state, or equivalently, to an excited state. In both cases the main ingredients for maintaining fibrillation, namely the presence of both propagating wavefronts and a sufficient mass of excitable tissue at rest, referred to as "excitable gap", in which these wavefronts can travel, are missing. Achieving these objectives is challenging since, on biophysical grounds, shock-induced changes in polarization of both polarities are always present during shock delivery [23, 19].

In previous work [13, 14, 12] we addressed these points by modeling the controller action representing the current delivered by the electrodes as distributed forces. One of the main objectives of the current work consists in analyzing the case when the action of the electrodes is modeled as Neumann boundary conditions.

From a methodological point of view, in most, if not all, recent finite element modeling studies the effect of extracellularly applied electric fields has been accounted for either by imposing inhomogeneous Dirichlet boundary conditions to model extracellular potential stimuli, or, by using current volume sources to model current stimuli.

The use of inhomogeneous Neumann boundary conditions for modeling current stimuli has been the method of choice in early pioneering monodomain modeling work where the finite difference method was employed to model impulse propagation in 1D strands or 2D sheets [24]. Surprisingly, to the best of our knowledge, in the bidomain literature Neumann boundary conditions for modeling current injection via electrodes have not been rigorously stated yet, neither in bidomain forward models nor in the context of optimal control. While equivalence between both formulations, i.e. Neumann boundary conditions and volume sources, can be achieved for any given setup, in the latter case where currents are injected via shape functions into 2D or 3D elements, the total injected current depends on spatial discretization and choice of weighting function. Thus, in the present work we aim to investigate the suitability of of using inhomogeneous Neumann boundary conditions in a bidomain model, specifically, the feasibility of optimal boundary control for the bidomain equations.

A second important issue of the current work consists in comparing the nonlinear conjugate gradient and the Newton method as iterative solution processes to solve the resulting optimization problems. Here we point out that due to the complicated dynamical systems behavior of the bidomain equations and the scale differences between its components it is prima vista not evident that the Newton method is applicable.

Finally any numerical optimization approach requires repeated solution of the bidomain equations and the associated adjoint equations. While an efficient solution strategy is already of paramount importance for the direct numerical simulation of bidomain equation it is indispensable in the context of optimal control. For this reason our numerical realization relies on parallelization. We report on the parallel efficiency both for the direct simulation and for the optimization algorithms.

The optimal pacing of the cardiac tissue is expressed by optimal control with partial differential equations as constraints. Let $\Omega \subset \mathbb{R}^d$, $d \in \{2, 3\}$, denote a bounded connected domain with Lipschitz continuous boundary $\partial\Omega$. The space-time domain and its lateral boundary are denoted by $Q = \Omega \times (0, T]$ and $\Sigma = \partial\Omega \times (0, T]$, respectively. Also we denote

the observation domain by $\Omega_{obs} \subset \Omega$. The standard form control problem is expressed as:

$$\text{(P)} \qquad \begin{cases} \min J(v, I_e)\,, \\ \\ e(u, v, w, I_e) = 0, \end{cases} \tag{1.1}$$

where u, v and w are the state variables, and I_e is the extracellular current which is utilized as a control variable in the optimal control problem and $e = 0$ stands formally for the dynamical system constraint. The dynamical behavior of the intra- and extracellular potentials is described by the coupled system of reaction-diffusion equations which can be expressed as follows

$$0 = \nabla \cdot (\sigma_i + \sigma_e)\nabla u + \nabla \cdot \sigma_i \nabla v \quad \text{in } Q \tag{1.2}$$

$$\frac{\partial v}{\partial t} = \nabla \cdot \sigma_i \nabla v + \nabla \cdot \sigma_i \nabla u - I_{ion}(v, w) + I_{tr}(x, t) \quad \text{in } Q \tag{1.3}$$

$$\frac{\partial w}{\partial t} = G(v, w) \quad \text{in } Q, \tag{1.4}$$

where $u\colon Q \to \mathbb{R}$ is the extracellular potential, $v\colon Q \to \mathbb{R}$ is the transmembrane voltage, $w\colon Q \to \mathbb{R}^n$ represents the ionic current variables, $\sigma_i\colon \Omega \to \mathbb{R}^{d\times d}$ and $\sigma_e\colon \Omega \to \mathbb{R}^{d\times d}$ are respectively the intracellular and extracellular conductivity tensors. The term $I_{tr}(x, t)$ is the transmembrane current density stimulus as delivered by an intracellular electrode. The $I_{ion}(v, w)$ is the current density flowing through the ionic channels and the function $G(v, w)$ determines the evolution of the gating variables, which is determined by an electrophysiological cell model, see e.g. [1] for more description on these models. Eq. (1.2) above is an elliptic type equation, Eq. (1.3) is a parabolic type equation and Eq. (1.4) is a set of ordinary differential equations which can be solved independently for each node. Typically, the conductivity tensors, which were considered in our computations, are expressed in the following form,

$$\sigma_c = \begin{pmatrix} \sigma_{cl} & 0 \\ 0 & \sigma_{ct} \end{pmatrix}, \quad \text{where } c = i, e\,, \tag{1.5}$$

where σ_{cl} and σ_{ct} are longitudinal and transverse fiber conductivities, respectively.

The membrane model for the ionic activity is described by a set of ordinary differential equations. The dimension of the ODE system is a consequence of the ionic model. In our numerical computations, we used a modified FitzHugh-Nagumo (FHN) model, called Rogers-McCulloch model [21], which consists of only two state variables and has a cubic non-linearity in the transmembrane potential

$$I_{ion}(v, w) = gv(1 - \frac{v}{v_{th}})(1 - \frac{v}{v_p}) + \eta_1 v w\,. \tag{1.6}$$

$$G(v, w) = \eta_2(\frac{v}{v_p} - \eta_3 w) \tag{1.7}$$

where $g, \eta_1, \eta_2, \eta_3$ are prescribed positive real coefficients, $v_{th} > 0$ is the threshold potential and $v_p > v_{th}$ is the peak potential.

In the absence of a conductive bath both intracellular and extracellular domains are electrically isolated along the tissue boundaries and homogeneous Neumann boundary conditions are appropriate to reflect this fact, except for those parts of the boundary where extracellular stimuli are applied. The initial values of the transmembrane voltage and state variables are prescribed by given values $v_0 \in L^2(\Omega)$ and $w_0 \in L^2(\Omega)$. The initial and boundary conditions are therefore prescribed as

$$\eta \cdot (\sigma_i \nabla v + \sigma_i \nabla u) = 0 \quad \text{on } \Sigma \tag{1.8}$$

$$\eta \cdot \sigma_e \nabla u = I_e \quad \text{on } \Gamma_{12} \times (0, T] \tag{1.9}$$

$$\eta \cdot \sigma_e \nabla u = 0 \quad \text{on } \Gamma_3 \times (0, T] \tag{1.10}$$

$$v(x, 0) = v_0 \quad \text{and } w(x, 0) = w_0 \quad \text{on } \Omega\,, \tag{1.11}$$

where η denotes the outwards normal to the boundary of Ω. Here I_e is the extracellular current density stimulus which acts as control along the boundary $\Gamma_{12} = \Gamma_1 \cup \Gamma_2$, where Γ_i, $I = 1, 2, 3$ are mutually disjoint and satisfy $\Gamma_1 \cup \Gamma_2 \cup \Gamma_3 = \partial\Omega$. For compatibility reasons it is assumed throughout that

$$\int_{\partial\Omega} I_e(t, \cdot)\, d\, s = 0 \tag{1.12}$$

for almost every $t \in (0, T)$. In the numerical experiments I_e will be only temporally dependent and will be of the form

$$I_e = \hat{I}_e(t)(\chi_{\Gamma_1} - \chi_{\Gamma_2})\,,$$

where χ_{Γ_i} is the characteristic function of the set Γ_i, $i = 1, 2$. Then condition (1.12) is satisfied if $|\Gamma_1| = |\Gamma_2|$. The support regions Γ_1 and Γ_2 can be considered to represent a cathode and an anode, respectively.

In terms of optimal control we recall that the optimality system involves the primal as well as the adjoint equations. Each of these two systems has similar complexity and must be solved within an iterative solution process. Also, the linearized primal and dual equations need to be resolved frequently in each iteration of Newton-Krylov methods. The computational costs involved in this solution process are significant, rendering the use of efficient numerical approaches a key ingredient to obtain results within reasonable time frames. Parallelization techniques, as they are routinely used in virtually any bidomain modeling study, suggest themselves quite naturally also as a means to speed up the solution process of the optimal control problem. In this regard, we parallelized our optimization codes based on the public domain package DUNE [4]. In the numerical section we compare the improvements in terms of computational cost that can be offered by Newton-Krylov optimization algorithm over gradient type algorithms.

The remainder of the paper is organized as follows. In the next section we will show the well-posedness of the bidomain boundary value problem. In Section 3, the boundary control formulation of the bidomain model is presented. There we discuss the necessary first order optimality for the boundary control problem. Also, we lay out the necessary steps to implement Newton's method in our numerical realization. The numerical discretization of the optimality system is described in Section 4. At the end of that section a brief overview of parallelization results is given. In Section 5 numerical results are presented for the different test cases. A summary is given then in the last section.

2 Existence for the bidomain system

Here we provide a brief analytical setting for well-posedness of the state equation (1.2)-(1.4) together with initial and boundary conditions (1.8)-(1.11). For the most part we can rely on the analysis in [8] where a related problem with distributed forcing and homogenous boundary conditions was treated. In the variational setting we only need to replace a distributed by a boundary forcing term. For the sake of consistency we provide the main steps here. Throughout we assume that the conductivity tensors satisfy $\sigma_i \in L^\infty(\Omega)$, $\sigma_e \in L^\infty(\Omega)$ and that they are uniformly elliptic, i.e. there exist $0 < m < M < \infty$ such that

$$m|\xi|^2_{\mathbb{R}^d} \leq \xi^T \sigma_{i,e} \xi \leq M|\xi|^2_{\mathbb{R}^d} \text{ for } \xi \in \mathbb{R}^d. \tag{2.1}$$

Definition 2.1. *A triple* (u, v, w) *with* $u \in L^2(0, T; H^1(\Omega))$, $v \in C([0, T],\, L^2(\Omega)) \cap L^2(0, T; H^1(\Omega)) \cap L^4(Q)$, $w \in C([0, T], L^2(\Omega))$ *is called a weak solution to* (1.2)-(1.4), (1.8)-(1.11)*, if it sat-*

isfies for a.e. $t \in (0,T)$ *the equations*

$$\int_\Omega (\sigma_i \nabla v(t) \nabla \varphi + (\sigma_i + \sigma_e) \nabla u(t) \nabla \varphi)\, dx = \int_{\Gamma_{12}} I_e(t) \varphi \, ds,$$
$$\int_\Omega (v_t(t)\psi + \sigma_i(\nabla u(t) + \nabla v(t))\nabla \psi)\, dx + \int_\Omega I_{ion}(v(t), w(t))\psi \, dx = \int_\Omega I_{tr}(t)\psi \, dx,$$
$$\int_\Omega (v_t(t) + G(v(t), w(t)))w \, dx = 0,$$

for all $(\varphi, \psi, w) \in H^1(\Omega)/_{\mathbb{R}} \times H^1(\Omega) \times L^2(\Omega)$, *and* $(v(0), w(0)) = (v_0, w_0)$.

For the following result we need to introduce the space

$$\mathcal{X} = L^2(0,T,H^1(\Omega)) \times (C([0,T], L^2(\Omega)) \cap L^2(0,T;H^1(\Omega)) \cap L^4(Q) \cap W^{1,\frac{4}{3}}(0,T;H^1(\Omega)^*)$$
$$\times (C([0,T), L^2(\Omega)) \times W^{1,2}(0,T;H^1(\Omega)^*)),$$

which is endowed with the natural norm.

Theorem 2.1. *Assume that* $I_{tr} \in L^2(0,T;H^1(\Omega)^*)$, $I_e \in L^2(0,T;\Gamma_{12})$ *and that* (1.12), (2.1) *hold. Then* (1.2)-(1.4), (1.8)-(1.11) *admits a unique weak solution and there exists a constant* $K > 0$ *such that the following a-priori bound holds:*

$$|(u,v,w)|_{\mathcal{X}} \le K(1 + |v_0|_{L^2(\Omega)} + |w_0|_{L^2(\Omega)} + |I_{tr}|_{L^2(0,T;H^1(\Omega)^*)} + |I_e|_{L^2(0,T;L^2(\Gamma_{12}))}).$$

For the proof we can follow the verification of Theorems 2.5 and 2.6 in [8] noting that $\psi \to \int_{\Gamma_{12}} I_e(t,s)\psi(t,s)\, ds\, dt$ defines a continuous functional on $L^2(0,T;H^1(\Omega))$.

For the following stability estimate, which in particular implies uniqueness of the solutions to (1.2)-(1.4),(1.8)-(1.11), we use $\tilde{\mathcal{X}} = \{(u,v,w) \in \mathcal{X} : w \in W^{1,2}(0,T;L^2(\Omega))\}$, endowed with the natural norm.

Theorem 2.2. *If in addition to the assumptions of Theorem 2.1 we have* $w_0 \in L^4(\Omega)$, *then for any* R *there exists a constant* C_2 *such that for any two pairs* (I'_{tr}, I'_e) *and* $(\tilde{I}_{tr}, \tilde{I}_e)$ *in* $L^\infty(0,T;H^1(\Omega)^*) \times L^\infty(0,T;L^2(\Gamma_{12}))$ *whose norms are bounded by* R *we have*

$$|(u',v',w') - (\tilde{u},\tilde{v},\tilde{w})|_{\tilde{\mathcal{X}}} \le C_R(|I'_{tr} - \tilde{I}_{tr}|_{L^\infty(0,T;H^1(\Omega)^*)} + |I'_{tr} - \tilde{I}_{tr}|^2_{L^\infty(0,T;H^1(\Omega)^*)}$$
$$+ |I'_e - \tilde{I}_e|_{L^\infty(0,T;L^2(\Gamma_{12}))} + |I'_e - \tilde{I}_e|^2_{L^\infty(0,T;L^2(\Gamma_{12}))}),$$

where (u',v',w') *and* $(\tilde{u},\tilde{v},\tilde{w})$ *denote the solutions corresponding to* (I'_{tr}, I'_e) *and* $(\tilde{I}_{tr}, \tilde{I}_e)$ *respectively.*

3 The optimal control problem

We return to the optimal control problem (P). The cost functional J will be chosen in such a manner that the controlled trajectory of the transmembrane voltage is driven to a desired state v_d by applying extracellular current I_e. Application of I_e has an averse effect on the tissue and hence it enters into the cost functional as penalization term.

We consider

$$\begin{cases} \min J(v, I_e) = \dfrac{1}{2}\displaystyle\int_0^T (\int_{\Omega_{obs}} |v - v_d|^2 \, dx + \alpha \int_{\Gamma_{12}} I_e^2 \, ds)\, dt \\ \text{subject to (1.2)-(1.4),(1.8)-(1.11) and } I_e \in U, \end{cases} \tag{3.1}$$

where $\alpha > 0$ is the weight for the control cost, $\Omega_{obs} \subset \Omega$ is the observation domain, $v_d \in L^2(0,T;L^2(\Omega_{obs}))$ and

$$U = \{I_e - \frac{1}{|\Gamma_{12}|}\int_{\Gamma_{12}} I_e \, ds : \, I_e \in L^2(0,T;L^2(\Gamma_{12})), |I_e(t,x)| < R \text{ for a.e. } (t,x) \in (0,T) \times \Gamma_{12}\}.$$

The set U is a closed, convex and weakly* sequentially compact subset of $L^\infty(0,T;L^2(\Gamma_{12}))$. This fact together with the well-posedness results of section 2 allow to deduce the existence of a solution (u^*, v^*, v^*, I_e^*) to (3.1) by standard arguments. For the case of distributed control action the details are provided in [9]. For computational purposes the first order necessary conditions are of paramount importance. We give a formal derivation of these conditions here. They can be made rigorous by rather minor modifications of the proofs given in [9].

To obtain these first order conditions we introduce the Lagrangian associate to (3.1) by defining

$$\begin{aligned}
&\mathcal{L}(u,v,w,I_e,p,q,r) = J(v,I_e)\\
&-\int_0^T\int_\Omega(\sigma_i\nabla v\nabla p+(\sigma_i+\sigma_e)\nabla u\nabla p)\,dx\,dt+\int_0^T\int_{\Gamma_{12}} I_e p\,ds\,dt\\
&-\int_0^T\int_\Omega(v_t q+\sigma_i(\nabla u+\nabla v)\nabla q+I_{ion}(v,w)q-I_{tr}q)\,dx\,dt\\
&-\int_0^T\int_\Omega(w_t-G(v,w))r\,dx\,dt,
\end{aligned}$$

where p, q, r are the Lagrange multipliers associated to the equations (1.2)-(1.4). Taking variations with respect to the state variables, invoking integration by parts with respect to the temporal variable and Green's formula with respect to the spatial variable, and keeping in mind the initial- and boundary conditions we obtain the first order optimality system which we express in operator form:

$$\nabla\left((\sigma_i+\sigma_e)\nabla p\right)+\nabla(\sigma_i\nabla q)=0 \text{ in } Q\,, \tag{3.2}$$

$$(v-v_d)|_{\Omega_{obs}}+\frac{\partial}{\partial t}q+\nabla(\sigma_i\nabla p)+\nabla(\sigma_i\nabla q)-(I_{ion}(v,w))_v q+G_v(v,w)r=0 \text{ in } Q\,, \tag{3.3}$$

$$-(I_{ion}(v,w))_w q+\frac{\partial}{\partial t}r+G_w(v,w)^T r=0 \text{ in } Q\,, \tag{3.4}$$

together with the terminal conditions

$$q(T)=0,\ r(T)=0,$$

and the boundary conditions for the adjoint states

$$\begin{aligned}
(\sigma_i\nabla q+\sigma_i\nabla p)\eta&=0 \qquad \text{on } \Sigma,\\
\sigma_e\nabla p\, n&=0 \qquad \text{on } \Sigma,
\end{aligned}$$

and $\int_\Omega p(t)\,dx=0$, for a.e. $t\in(0,T)$. Furthermore for any optimal control I_e^* the following variational inequality must be satisfied:

$$\int_0^T\int_{\Gamma_{12}}(\alpha I_e^*+Qp)(I_e-I_e^*)\,ds\,dt\geq 0, \text{ for all } I_e\in U, \tag{3.5}$$

$$\tag{3.6}$$

where $(Qp)(t)=p(t)-\frac{1}{|\Gamma_{12}|}\int_{\Gamma_{12}}p(t,s)\,ds$ on Γ_{12}. The pointwise formulation of (3.5) is found to imply a.e. in Q

$$\begin{aligned}
&\alpha I_e^*(t,x)+Qp(t,x)>0\Rightarrow \tilde{I}_e^*(t,x)=-R\\
&\alpha I_e^*(t,x)+Qp(t,x)<0\Rightarrow \tilde{I}_e^*(t,x)=R\\
&\qquad -R<\tilde{I}_e^*(t,x)<R\Rightarrow \alpha I_e^*(t,x)+Qp(t,x)=0,
\end{aligned}$$

where $\tilde{I}_e^*$ is such that $Q\tilde{I}_e^* = I_e^*$.

For the numerical realization of (3.1) we shall require the reduced cost functional

$$\hat{J}(I_e) = J((v(I_e), I_e)),$$

where $v(I_e)$ is the solution component to (1.2)-(1.4). From (3.5) the gradient of $\hat{J}$ at some I_e is given by

$$\hat{J}(I_e)' = \alpha I_e^* + Qp. \tag{3.7}$$

In the numerical experiments we consider the practically relevant case that I_e is only spatially dependent and $I_e(t)$ is of the form $I_e(t,x) = \hat{I}_e(t)(\chi_{\Gamma_1} - \chi_{\Gamma_2})$ where $\Gamma_{12} = \Gamma_1 \cup \Gamma_2$, $|\Gamma_1| = |\Gamma_2|$, and

$$\hat{I}_e \in U = \{\hat{I}_e \in L^2(0,T) : |\hat{I}_e(t)| \leq R \text{ for a.e. } T \in (0,T)\}.$$

For this case the optimality condition (3.5) turns out to be

$$(2\alpha \hat{I}_e^*(t)|\Gamma_1| + \int_{\Gamma_1} p(t,s)\,ds - \int_{\Gamma_2} p(t,s)\,ds\,)\,(\hat{I}_e - \hat{I}_e^*(t)) \geq 0, \text{ for all } |\hat{I}_e| \leq R, \text{ and a.e. } t \in (0,T).$$

From the point of view of numerical optimization we shall not only consider a nonlinear conjugate gradient, but also a Newton method, which, without considering line-search steps, is based on the iteration procedure $I_e^{k+1} = I_e^k + \delta I_e$ where δI_e is the solution to

$$\hat{J}''(I_e^k)\delta I_e = -\hat{J}'(I_e^k), \tag{3.8}$$

and k denotes the iteration level. Here $\hat{J}''(I_e)$ denotes the Hessian of the reduced cost functional. It is numerically infeasible to set up $\hat{J}''(I_e)$ as a (discretized) matrix operator, rather one needs to resort to an iterative procedure to solve (3.8) approximately for δI_e, e.g. by a Krylov solver. Below we summarize the necessary steps to realize the Newton step (3.8) by an iterative procedure:

1. Compute the first derivative $\hat{J}'(I_e^k)$ according to (3.7), which requires one primal and one adjoint solve.

2. In each iteration step for solving (3.8), evaluate the action of $\hat{J}''(I_e^n)$ on δI by means of the following sequence of computation:

   (a) solve the linearized primal equation for $(\delta u, \delta v, \delta w)$ using δI

$$\begin{pmatrix} \nabla \cdot (\sigma_i + \sigma_e)\,\nabla \delta u + \nabla \cdot (\sigma_i \nabla \delta v) \\ \nabla \cdot (\sigma_i \nabla \delta v) - (\delta v_t + [I_{ion}]_v\,\delta v + [I_{ion}]_w\,\delta w) \\ \delta w_t - \frac{\eta_2}{v_p}\delta v + \eta_2\eta_3 \delta w \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

   with the following initial and boundary conditions

$$\begin{aligned} \eta \cdot (\sigma_i \nabla \delta v + \sigma_i \nabla \delta u) &= 0 \quad \text{on} \quad \Sigma \\ \eta \cdot \sigma_e \nabla \delta u &= \left(\chi_{c\Gamma_1} - \chi_{c\Gamma_2}\right)\delta I(t) \quad \text{on} \quad \Gamma_1 \cup \Gamma_2 \\ \eta \cdot \sigma_e \nabla \delta u &= 0 \quad \text{on} \quad \Gamma_3 \\ \delta v(x,0) = \delta v_0 \qquad & \text{and} \quad \delta w(x,0) = \delta w_0 \quad \text{on} \quad \Omega\,, \end{aligned}$$

   (b) evaluate $(\xi_1, \xi_2, \xi_3, \xi_4) := \mathcal{L}_{yy}(y^k, z^k)(\delta u, \delta v, \delta w, \delta I_e)$ from (3.9),

   (c) solve the adjoint equation with (ξ_1, ξ_2, ξ_3) as r.h.s., i.e.

$$\begin{pmatrix} \nabla \cdot (\sigma_i + \sigma_e)\,\nabla w_1 + \nabla \cdot \sigma_i \nabla w_2 \\ \nabla \cdot \sigma_i \nabla w_2 + w_{2_t} - [I_{ion}]_v w_2 - \frac{\eta_2}{v_p} w_3 \\ -[I_{ion}]_w w_3 - w_{3_t} + \eta_2 \eta_3\, w_3 \end{pmatrix} = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix}$$

   by using homogenous initial and boundary conditions,

(d) compute the action $\hat{J}''(I_e^k)\delta I = \alpha\,\delta I|(\Gamma_1 \cup \Gamma_2) + w_1|\Gamma_1 - w_1|\Gamma_2$.

Above we use that the second derivative of the operator $\mathcal{L}$ at $y^k = (u^k, v^k, w^k, I_e^k)$, $z^k = (p^k, q^k, r^k)$, with k denoting the iteration level, applied to the vector $(\delta u, \delta v, \delta w, \delta I_e)$ can be expressed as

$$\mathcal{L}_{yy}(y^k, z^k)(\delta u, \delta v, \delta w, \delta I_e) = \begin{pmatrix} 0 \\ \begin{pmatrix} \delta v|_{\Omega obs} - [I_{ion}]_{vv}\, q^k\, \delta v - \eta_1 q^k\, \delta w \\ -\eta_1 q^k\, \delta v \end{pmatrix} \\ \alpha\,\delta I_e \end{pmatrix} \tag{3.9}$$

$$\text{where } [I_{ion}]_{vv} \;=\; \frac{2g}{v_p v_{th}}\left(3v^k - v_p - v_{th}\right).$$

4 Numerical approach

In this section we describe the numerical treatment of the complete optimality system. For the spatial discretization of the primal and dual equations we used a piecewise linear finite element method and linearly implicit Runge-Kutta methods for the temporal discretization.

4.1 Semi-discretization in space

Here we give an overview of the spatial discretization of the primal problem by a finite element method based on the weak formulation. An analogous discretization is used for the dual equations, we refer to our previous paper [13, Section 3.1.2] in this respect.

A weak solution triple (u, v, w) satisfies for a.e. $t \in (0, T)$ and for all $\varphi \in H^1(\Omega)$

$$0 = \langle \nabla \cdot (\sigma_i + \sigma_e)\nabla u + \nabla \cdot \sigma_i \nabla v, \varphi \rangle\,, \tag{4.1}$$

$$\left\langle \frac{\partial v}{\partial t}, \varphi \right\rangle = \langle \nabla \cdot \sigma_i \nabla v + \nabla \cdot \sigma_i \nabla u - I_{ion}(v, w) + I_{tr}(x, t), \varphi \rangle\,, \tag{4.2}$$

$$\left\langle \frac{\partial w}{\partial t}, \varphi \right\rangle = \langle G(v, w), \varphi \rangle\ , \tag{4.3}$$

together with initial and boundary conditions (1.8)-(1.11). Let $V_h \subset H^1(\Omega)$ be the finite dimensional subspace of piecewise linear basis functions with respect to the spatial grid. The approximate solutions $\mathbf{u}, \mathbf{v}$ and $\mathbf{w}$ are expressed in the form $\mathbf{u}(t) = \sum_{i=1}^N u_i(t)\omega_i$, $\mathbf{v}(t) = \sum_{i=1}^N v_i(t)\omega_i$ and $\mathbf{w}(t) = \sum_{i=1}^N w_i(t)\omega_i$, respectively, where $\{\omega_i\}_{i=1}^N$ denote the basis functions. The semi-discretization of primal equations in space results in the differential algebraic system as follows:

$$\mathbf{A_{ie}}\mathbf{u} + \mathbf{A_i}\mathbf{v} = \mathbf{I}_e \tag{4.4}$$

$$\mathbf{M}\frac{\partial \mathbf{v}}{\partial t} = -\mathbf{A_i}\mathbf{v} - \mathbf{A_i}\mathbf{u} - \mathbf{I}_{ion}(\mathbf{v}, \mathbf{w}) + \mathbf{I}_{tr} \tag{4.5}$$

$$\frac{\partial \mathbf{w}}{\partial t} = \mathbf{G}(\mathbf{v}, \mathbf{w}), \tag{4.6}$$

together with initial conditions for $\mathbf{v}$ and $\mathbf{w}$, where $\mathbf{A_{ie}} = \{\langle (\sigma_i + \sigma_e)\nabla\omega_i, \nabla\omega_j \rangle\}_{i,j=1}^N$ and $\mathbf{A_i} = \{\langle \sigma_i \nabla\omega_i, \nabla\omega_j \rangle\}_{i,j=1}^N$ are the stiffness matrices, $\mathbf{M} = \{\langle \omega_i, \omega_j \rangle\}_{i,j=1}^N$ is the mass matrix, and the vectors $\mathbf{I}_e$, $\mathbf{I}_{itr}$ are defined by $\mathbf{I}_e = \{\langle (\chi_{\Gamma_1} I_e - \chi_{\Gamma_2} I_e)\,, \omega_j \rangle\}_{j=1}^{N_\Gamma}$ and $\mathbf{I}_{tr} = \{\langle I_{tr}, \omega_j \rangle\}_{j=1}^N$, respectively. The expression $(\mathbf{I}_{ion})(\mathbf{v}, \mathbf{w})$ is defined by

$$(\mathbf{I}_{ion})(\mathbf{v}, \mathbf{w}) = (\mathbf{I}_{ion})\left(\sum_{i=0}^N v_i\omega_i, \sum_{i=0}^N w_i\omega_i\right).$$

The same spatial discretization technique is applied for the linearized primal and dual equations.

4.2 Full discretization and solution procedure

After the semi discretization of primal and dual equations one has to apply a time discretization technique to get a fully discretized system. For this purpose we applied a third order Rosenbrock method called ROS3PL [10]. Here we avoid the time discretization details for this problem which was explained in [13, Section 3.2].

The solution of the singular linear systems which arise after the full discretization of Eqs (1.2) and (3.2) are defined up to an additive constant. We mentioned earlier that to fix this constant we impose a zero mean condition. For the numerical realization of this condition we adopted a stabilized saddle point formulation from the work of Bochev and Lehoucq [5]. See [13] for the discussion and implementation details of this technique for the current problem. After the full discretization of the PDEs we obtain a system of linear algebraic equations. To solve this linear system we employed a BiCGSTAB [26] method with Jacobi preconditioning.

To solve the complete optimality system the Hager-Zhang variant of the non-linear conjugate gradient (NCG) [16], and the Newton method as explained in the previous section, were developed. Their performances will be compared in the following section.

4.3 Parallel implementation

To solve the optimality system numerically, many PDE solves are required to obtain an approximate solution to the optimal control problem. The solution of the bidomain equation system itself, due different time-scales that are present in the dynamical system, as well as the wave formation and the resulting necessity for fine scale spatio-temporal simulations, puts substantial demands on the computation see eg. [18, 17, 15, 25, 20]. This suggests to apply parallelization techniques. For this purpose, the public domain FEM package DUNE [4] is used and the internal parallel Yasp grid in DUNE is employed for parallel grid constructions. It supports different levels of overlapping grids for parallel simulations. We used zero level overlapping, i.e. non overlapping, grids. The Yasp grid uses the simple coordinate bisection method for the domain decomposition of the spatial grid. The finite element discretization module *dune-pdelab* in the DUNE package [4] was used.

Let us recall the number of PDE solves that are need for optimization. One iteration of the NCG algorithm needs the solution of the primal and of dual equations, plus one additional primal evaluation for the line search method. The Newton method, moreover requires evaluation of the Hessian of the reduced cost functional in direction of the increment of the control δI_e which is required for the iterative Krylov solver. Each inner iteration is achieved by one linearized primal, see 2(a) at the end of Section 3, and a dual equation solver 2(c), which together achieve the matrix-vector product of the Hessian on δI_e. The solution of these linearized equations in parallel was done similarly to the solution procedure for primal and dual equations.

To achieve overall parallel scalability in the optimization problem we require first and foremost efficient realization of the communication in the finite element context of the solution of the direct problem in every iteration of the linear solver at each time step. For optimization this communication must also be considered for the adjoint equations, the computation of the Jacobian update and for the efficient parallel implementation of the Newton-Krylov method. Finally communication is required for the evaluation of the norm of the gradient of the cost functional, and for the cost itself.

5 Numerical results

Numerical results on the basis of two different test cases are discussed in this section. First numerical results based on boundary control of the bidomain equations are presented where the control acts on the left and right boundary of the computational domain, as shown in Figure 1. The second test case deals with a similar situation but now there are four smaller electrodes placed on either side of the domain, as shown in right hand side of Figure 1. In all cases the computational domain $\Omega = [0,2] \times [0,2] \subset \mathbb{R}^2$ of size $2 \star 2\ cm^2$ is fixed and a 256×256 uniform quadrilateral spatial grid is used which consists of $65,536$ elements and $66,049$ nodes. Thus the computation involved $198,147$ dofs for one PDE solve.

Now we turn to the stopping criteria for the optimization algorithm. The termination of the optimization algorithm is based on the following condition:

$$\left\| \nabla J(\mathbf{I}_e^k) \right\|_{L^2} \leq 10^{-3} \cdot \left| J(\mathbf{I}_e^k) \right| \text{ or } \left| J(\mathbf{I}_e^k) - J(\mathbf{I}_e^{k-1}) \right| \leq 10^{-4} \tag{5.1}$$

If this condition was not satisfied within a prescribed number of 300 iterations, the algorithm was terminated. An Armijo type condition is imposed in the line search algorithm.

The presented numerical results as well as the parallel efficiency of the optimization algorithms are done on a Linux cluster (GHOST) with two nodes, where each node consists of 8 quad-core AMD Opteron processors 8356 clocked at 2.3 GHz and equipped with 256 GB RAM. For our computations we used only one node (overall 32 cpus).

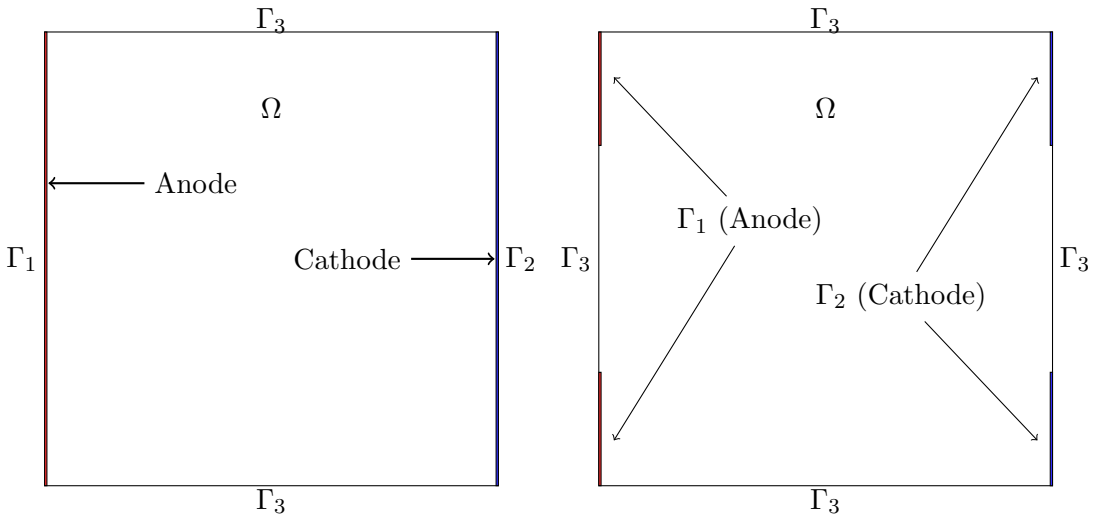


Figure 1: Computational domain with two and four stimulation boundaries.

As presented previously [13], the conductivity tensors were chosen, on one hand, to arrive at a good match with physiological conduction velocities with the simplified FitzHugh-Nagumo model, and, on the other hand, to keep anisotropy ratios within the range of values reported in experimental studies [22]. This led to the following choice of simulation parameters, $\sigma_{il} = 2.0 \cdot 10^{-3}\ S/cm$, $\sigma_{it} = 3.1 \cdot 10^{-4}\ S/cm$, $\sigma_{el} = 2.0 \cdot 10^{-3}\ S/cm$, $\sigma_{et} = 1.3 \cdot 10^{-3}\ S/cm$ $g = 1.5\ S/cm^2$, $v_{th} = 13\ mV$, $v_p = 100\ mV$, $\eta_1 = 4.4\ S/cm^2$, $\eta_2 = 0.012$, $\eta_3 = 1$. To account for the statistical discontinuous variation of conductivities which are omnipresent in cardiac tissue due to the discrete micro-structure, the intracellular conductivity tensors were modified by multiplying with random numbers in the range $(0,1)$. Moreover, those random numbers which were below a threshold of 0.34 were set equal to 10^{-12}, for details of the underlying rationale see [6, 13].

A standard $S1 - S2$ stimulation protocol was applied to induce a reentrant activation pattern. An initial $S1$ stimulus of $I_{tr} = 100\ \mu A/cm^2$ and 5 $msec$ duration was applied

at time $t = 0$ $msec$ along the bottom edge of the tissue sheet. At time $t = 183$ $msec$, when the critical recovery isoline crossed the center of the sheet, a second $S2$ stimulus of $I_{tr} = 100$ $\mu A/cm^2$ and 5 $msec$ duration was applied in a small region of 0.3 cm radius at the center of the domain. This $S2$ stimulus generated two phase singularities at the intersections between critical recovery isoline and the boundary of the $S2$ stimulus region, leading to a so-called Figure of Eight reentrant pattern. The solution at $t = 435$ $msec$ was then chosen as the initial state for simulating the delivery of electrical shocks and the post-shock evolution following at the end of the shock.

The three temporal horizons are illustrated in Figure 2. The solution for the transmembrane voltage v in absence of any control is shown in Figure 3 at different instances of time to verify that the reentry is sustained for a sufficiently long time. During the shock, which constitutes the time period within which optimization takes place, the size of the time step was kept constant at $\Delta t = 0.04$ $msec$, while during the pre- and post-shock phases adaptive time stepping features of the ROS3PL method were used to speed up computation.

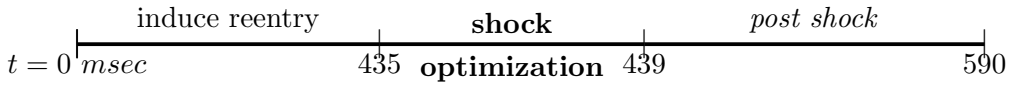
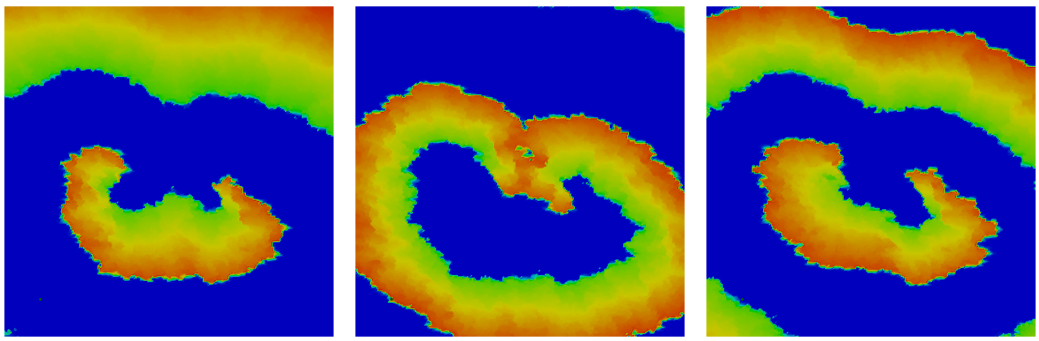


Figure 2: Different time horizons considered in the computations.



(a) t= 437 msec (b) t= 507 msec (c) t= 590 msec

Figure 3: 2D visualization of uncontrolled solution (v) at different times of simulation.

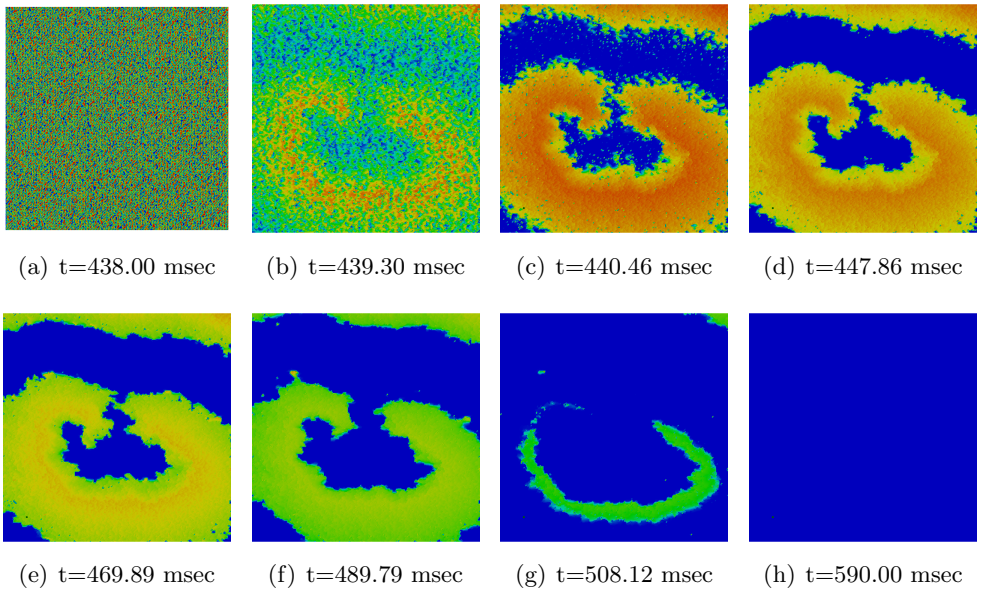
5.1 Boundary control with two stimulation boundaries

Here we present the numerical results for the bidomain model with Neumann boundary control. The anode is modeled as a Neumann boundary along the surface Γ_1 and the cathode along the surface Γ_2 of the computational domain, as shown in Figure 1. The desired trajectory of the transmembrane potential (v_d) is obtained by solving once the primal problem using a prescribed time course of a stimulation current, $I_e(t) = 200mA/s^2$. Here the desired trajectory ensures that optimized states attain a steady state during the post shock period. The optimization algorithm constructs the best optimal control while keeping the energy low, due to the term $\int I_e^2 \, dt$ in cost functional.

The phase plane analysis of the modified FHN model, see for instance [11, 13], explains that if a sufficient perturbation is given to the excitable state, then it tends rapidly to a maximal value and subsequently gradually goes to a stable resting point. In optimal control we influence a sufficient portion of the cardiac tissue to reach the maximal value by providing the optimum extracellular current. Then the excitable gap over the cardiac

tissue is sufficiently small so that no further excitation wave fronts can arise. After turning off the external stimulus in the post shock simulation the cardiac tissue goes to a resting state. This post shock simulation time takes approximately 150 msec.

The 2D spatial representation of the optimized transmembrane voltage is shown in Figure 4 at different time instances. Observe that during the shock period, at time $t =$



(a) t=438.00 msec (b) t=439.30 msec (c) t=440.46 msec (d) t=447.86 msec

(e) t=469.89 msec (f) t=489.79 msec (g) t=508.12 msec (h) t=590.00 msec

Figure 4: 2D visualization of controlled solution (v) at different times of simulation.

$438msec$, most of the computational domain is depolarized. Due to the applied external stimulus strength the cardiac tissue responds strongly and produces a large number of virtual electrodes at the microscopic size scale during the shock period. Eventually, these virtual electrodes over the tissue effectively block further wave propagation of the reentry wave and this leads to a successful defibrillation if the optimum external stimulus. The results depicted in Figure 4 were obtained by the Newton method, and those obtained by NCG are graphically indistinguishable. But algorithmically these methods behave differently as we discuss next.

The norm of the gradients of the cost functionals and the values of the cost functionals for the NCG and Newton methods are shown in the leftmost and middle panel of Figure 5, over the optimization iterations. We observe that both algorithms achieve approximately at same minimization values. The corresponding optimal controls are shown on the right hand side of Figure 5. The optimal controls are graphically indistinguishable.

Turning to the computational cost, we report that the Newton method takes 12 outer iterations to achieve converge. Per outer iteration an average of 9 inner CG iterations were taken. The accepted step length is 0.5 for the first three optimization iterations and full step lengths for subsequent ones. Overall, the complete Newton optimization algorithm needed 240 PDE solves and additionally 15 PDE solves for the line search method to converge the optimal solution. The NCG method took 300 iterations. It accepted the step length 1.0 right from beginning. Thus the NCG required 600 PDE solves and additionally 300 PDE solves for the line search algorithm. In conclusion the Newton method was significantly more effective than the NCG algorithm for this configuration.

Let us finally discuss the total extracellular current that is required for defibrillation. It is computed by the following expression:

$$\int_0^T \int_\Gamma ((\sigma \nabla u) \cdot \eta)^2 \; ds \; dt = |\Gamma| \int_0^T I_e(t)^2 \; ds \; dt.$$

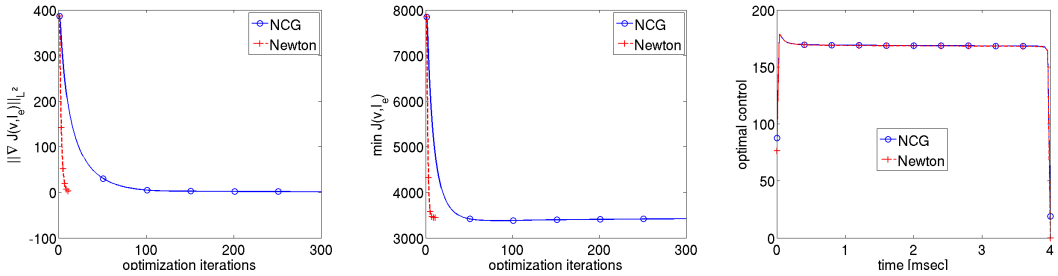


Figure 5: The gradient, minimum value of the cost functional and the optimal control values are shown on the left, middle and right figures respectively.

The extracellular current for desired trajectory v_d on the basis of a constant value for the stimulus required 3.2 A/cm^3. The total extracellular current required to drive the system from arrhythmia pattern to steady state pattern by the optimized control based on Newton's algorithm is 2.6788 A/cm^3 and that for NCG is 2.6844 A/cm^3 . Thus both methods give comparable results and both require less total current to obtain the successful defibrillation than that chosen on heuristic consideration to compute v_d. In Section 5.2 below we shall consider the effect of reducing the control domain of the electrodes on the total current required for defibrillation.

Parallel scalability for direct problem

Owing to the fact that the solution of the direct problem dominates overall computational cost in an optimal control solver, it is key to optimize solvers in terms of strong scalability to achieve significant reduction in execution time. To assess the strong scaling properties of the solvers used in this study benchmark simulations were performed using a high resolution mesh ($h = 3.9063\ \mu m$) of a square 2D domain of size $2 \times 2\ cm$ which consisted of 26,214,400 elements and 26,224,641 nodes (5120×5120). In this benchmarks the bottom edge of the 2D tissue sheet was stimulated to initiate a planar propagating wavefront. In the benchmark runs execution times were recorded then for the first 5ms of wavefront propagation, which is, typically, most expensive in terms of solver time since the fastest transients in the solution arise there. To be able to cover a wide range of number of cores benchmarks were performed at the UK national supercomputer HECToR phase 2b service, where strong scalability was tested up to 8192 cores, measuring parallel efficiency against the run using 64 cores, i.e. the smallest setup which provided enough memory to accommodate the whole model. Parallel efficiency was computed as

$$e = \frac{N_r}{N} \frac{T_r}{T_N}$$

where N_r and T_r are number of cores and total cpu time of a reference simulation, respectively, and N and T_N are number of cores and total cpu time in the scalability experiments where N is increased in multiples of N_r. Typically, $N_r = 1$ is chosen, however, in our case $N_r = 64$ was chosen since the model could not be fitted into memory otherwise.

Benchmark results are illustrated in Fig. 6. Up to 4096 cores adequate parallel efficiency of $e = 0.89$ could be achieved. Increasing N beyond 4096 cores e dropped significantly, for instance, with $N = 8192$ efficiency was quite poor with $e = 0.61$. This can be attributed to the unfavorable surface-to-volume ratio of local domains. With increasing N, the relation between local compute work performed on inner nodes of the domain and the data communication which are proportional to the size of shared domain interfaces, deteriorates, thus impeding any further efficient scaling. The domain sizes with 4096 and 8192 cores in terms of nodes were 6,561 and 3321 only.

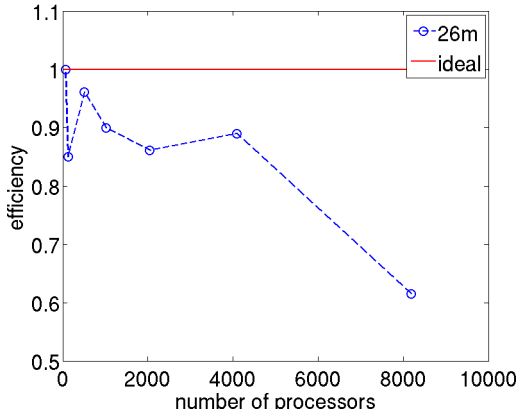


Figure 6: Parallel efficiency for the direct simulation.

Parallel scalability for optimization

For benchmarking parallel scalability of the optimization algorithm the computational setup which is demonstrated in Section 5.1 is used. In this case the mesh size is $h = 7.8\ \mu m$ and consists of $65,536$ elements and $66,049$ nodes. Compared to the direct simulation benchmarks this setup was smaller to keep computations tractable. Strong scalability experiments were performed on the 32 core GHOST cluster where also the boundary control simulations were performed. Parallel efficiency e of the first primal solve during the optimization with NCG and Newton optimization algorithms are shown in Figure 7. The parallel efficiency e is calculated as before with the reference simulation against which we compare is a run on a single core, i.e. $N_r = 1$. We achieved 82% of cpu gain on 32 processors for the first primal solve which is a crucial step in the optimization algorithms. For the overall optimization by the NCG algorithm 75% cpu time gain and by the Newton algorithm 73% cpu time gain were achieved. Absolute execution times on a single core for one primal solve was 8 min, while the execution of NCG and Newton algorithm lasted approximately 84 hours and 41 hours, respectively. Evidently, Newton's method performed better, being roughly twice as fast as NCG. These numbers highlight the pivotal importance of using scalable parallel algorithms for the optimization of the application under consideration.

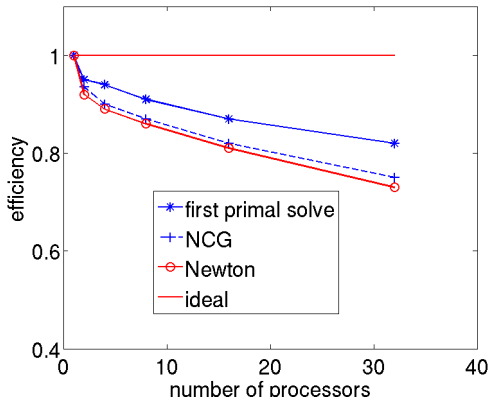


Figure 7: Parallel efficiency for different optimization algorithms.

5.2 Numerical results for multiple boundary stimulations

Here we discuss the numerical results for the set up of multiple stimulation electrodes which are placed at the left and right boundary of the computational domain, see right hand side of Figure 1. The smaller support patches of the control domains for the anode region are chosen to be $\Omega_{c_1} = [0, 0] \times [0, 0.5]$ and $\Omega_{c_2} = [0, 1.5] \times [0, 2.0]$ of size 1 cm as well as for the cathode region to be $\Omega_{c_3} = [2, 0] \times [2, 0.5]$ and $\Omega_{c_3} = [2, 1.5] \times [2, 2]$ of size 1 cm. From the view of applications, the main goal is to consider the control domains as small as possible. For this case the desired transmembrane solution trajectory is constructed using the stimulus strength $I_e = 400\ mA/cm^3$.

We again conducted the numerical experiments based on the NCG and Newton's optimization algorithms and both led to a successful defibrillation. Since the images for the optimized transmembrane voltage at different time instances are quite similar to those in Figure 4 we do not depict them here.

The gradient value of the cost functionals, the values of the cost themselves as well as the optimal controls are depicted in Figure 8. Again the NCG does not reach the stopping criterion within the first 300 iterations, and it is terminated there. At this point the algorithm decreases only very slowly. Comparable numbers as in the case of two electrodes are required as in the case of two electrodes. Here we only report that the overall computational time for (parallelized) NCG algorithm is 8174 seconds and that for the Newton method takes about 68% of the NCG CPU time. Thus clearly for this set-up the Newton optimization outperforms the NCG method.

The computed total optimum extracellular current for NCG is 2.0146 A/cm^3 and that for Newton is quite close with value 2.0134 A/cm^3. This optimized applied current is distinctly less than that which is used to design the desired trajectory which uses 2.4 A/cm^3.

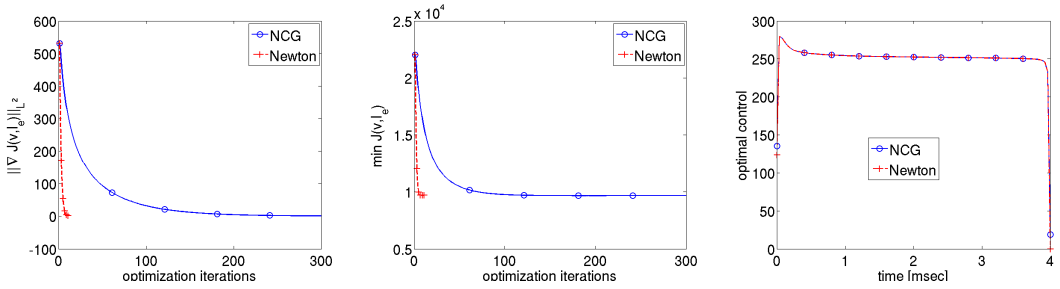


Figure 8: The gradient, minimum value of the cost functional and the optimal control values are shown on the left, middle and right respectively.

6 Discussion

In practice, defibrillation is achieved either externally by two large paddle electrodes placed on the chest, or, internally, via ICDs. In the latter scenario, both the size of shock electrodes as well as their number are small. Typically, in concurrent transvenous ICDs the shocking coil through which current is injected, is placed in the right ventricular apex, and the case of the ICD serves as reference electrode. Very recent devices offer more flexibility in terms of number of electrodes and placement options. Alternative subcutaneous placement strategies have been proposed recently [3], which allow a much wider range of feasible lead placements. However, the rationalization of lead placements is challenging [7]. In all scenarios which are currently considered, currents are injected via a sparse set of small electrodes. Hence, the control regions are sparse and have local support. This is well in line with the boundary formulation underlying this study. To be

of practical relevance for optimal control approaches to the cardiac defibrillation problem, it has to be demonstrated first that reductions in required energy can be achieved by finding optimal pulses as well as to enable the investigation of various electrode sizes and placements. In this study two electrode configurations were considered where for the setup with 4 smaller electrodes it could be shown that defibrillation could be achieved with less energy.

We use Neumann boundary conditions for modeling current injection. This is advantageous over the standard approaches typically used in bidomain modeling where currents are injected via nodes weighted with the corresponding nodal hat functions. The total injected current is then readily computed as the total length/area of the Neumann boundary multiplied by the imposed current density and does not depend on discretization and choice of weighting function.

The optimization of the applied current problem was achieved by the NCG and Newton algorithms, and it was found that the latter is advantageous in terms of computational efficiency as well as achieving local minima.

The optimization algorithm rely on frequent calls to the primal and adjoint equations and hence parallel efficiency is essential. In our computations we observed good parallel efficiency up to 4096 CPU cores for the primal solve. We also demonstrate good parallel efficiency of the NCG and Newtons optimization algorithms.

Acknowledgment

The authors gratefully acknowledge the Austrian Science Foundation (FWF) for financial support under SFB 032, "Mathematical Optimization and Applications in Biomedical Sciences".

References

[1] CellML Model Repository. `http://models.cellml.org/cellml`.

[2] G. H. Bardy, B. Hofer, G. Johnson, P. J. Kudenchuk, J. E. Poole, G. L. Dolack, M. Gleva, R. Mitchell, and D. Kelso. Implantable transvenous cardioverter-defibrillators. *Circulation*, 87(4):1152–68, 1993.

[3] G. H. Bardy, W. M. Smith, M. A. Hood, I. G. Crozier, I. C. Melton, L. Jordaens, D. Theuns, R. E. Park, D. J. Wright, D. T. Connelly, S. P. Fynn, F. D. Murgatroyd, J. Sperzel, J. Neuzner, S. G. Spitzer, A. V. Ardashev, A. Oduro, L. Boersma, A. H. Maass, I. C. Van Gelder, A. A. Wilde, P. F. van Dessel, R. E. Knops, C. S. Barr, P. Lupo, R. Cappato, and A. A. Grace. An entirely subcutaneous implantable cardioverter-defibrillator. *N Engl J Med*, 363(1):36–44, 2010.

[4] P. Bastian, M. Blatt, A. Dedner, C. Engwer, R. Klöfkorn, R. Kornhuber, M. Ohlberger, and O. Sander. A generic grid interface for parallel and adaptive scientific computing. Part II: implementation and tests in DUNE. *Computing*, 82(2):121–138, July 2008.

[5] P. Bochev and R. B. Lehoucq. On the finite element solution of the pure neumann problem. *SIAM Rev.*, 47:50–66, January 2005.

[6] J. M. T. de Bakker and H. M. V. van Rijen. Continuous and discontinuous propagation in heart muscle. *J Cardiovasc Electrophysiol*, 17(5):567–573, May 2006.

[7] M. Jolley, J. Stinstra, S. Pieper, R. Macleod, D. H. Brooks, F. Cecchin, and J. K. Triedman. A computer modeling tool for comparing novel ICD electrode orientations in children and adults. *Heart Rhythm*, 5(4):565–72, 2008.

[8] K. Kunisch and M. Wagner. Optimal control of the bidomain system (ii): uniqueness and regularity theorems for weak solutions. *Annali di Matematica Pura ed Applicata*, pages 1–36. 10.1007/s10231-012-0254-1.

[9] K. Kunisch and M. Wagner. Optimal control of the bidomain system (iii): Existence of minimizers and first-order optimality conditions. *appear in ESAIM: M2AN*.

[10] J. Lang and D. Teleaga. Towards a fully space-time adaptive FEM for magnetoquasistatics. *IEEE Transactions on Magnetics*, 44(6):1238–1241, 2008.

[11] J. Murray. *Mathematical biology: I. An introduction*. Springer, Heidelberg, 2002.

[12] C. Nagaiah and K. Kunisch. Higher order optimization and adaptive numerical solution for optimal control of monodomain equations in cardiac electrophysiology. *Appl. Numer. Math.*, 61:53–65, 2011.

[13] C. Nagaiah, K. Kunisch, and G. Plank. Optimal control approach to termination of re-entry waves in cardiac electrophysiology. *Journal of Mathematical Biology*, pages 1–30. 10.1007/s00285-012-0557-2.

[14] C. Nagaiah, K. Kunisch, and G. Plank. Numerical solution for optimal control of the reaction-diffusion equations in cardiac electrophysiology. *Computational Optimization and Applications*, 49:149–178, 2011. 10.1007/s10589-009-9280-3.

[15] S. Niederer, L. Mitchell, N. Smith, and G. Plank. Simulating human cardiac electrophysiology on clinical time-scales. *Front Physiol*, 2:14, 2011.

[16] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer Verlag, New York, second edition edition, 2006.

[17] L. F. Pavarino and S. Scacchi. Multilevel additive schwarz preconditioners for the bidomain reaction-diffusion system. *SIAM J. Sci. Comput.*, 31:420–443, October 2008.

[18] G. Plank, M. Liebmann, R. W. dos Santos, E. Vigmond, and G. Haase. Algebraic multigrid preconditioner for the cardiac bidomain model. *IEEE Trans Biomed Eng.*, 54(4):585–596, 2007.

[19] G. Plank, A. Prassl, E. Hofer, and N. A. Trayanova. Evaluating intramural virtual electrodes in the myocardial wedge preparation: simulations of experimental conditions. *Biophys J*, 94(5):1904–1915, Mar 2008.

[20] M. Potse, B. Dube, J. Richer, A. Vinet, and R. 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, Dec. 2006.

[21] J. M. Rogers and A. D. McCulloch. A collocation-Galerkin finite element model of cardiac action potential propagation. *IEEE Trans. Biomed. Eng.*, 41:743–757, 1994.

[22] B. J. Roth. Electrical conductivity values used with the bidomain model of cardiac tissue. *IEEE Trans Biomed Eng*, 44(4):326–328, Apr 1997.

[23] N. G. Sepulveda, B. J. Roth, and J. P. Wikswo, Jr. Current injection into a two-dimensional anisotropic bidomain. *Biophys J*, 55(5):987–99, 1989.

[24] M. S. Spach and J. M. Kootsey. Relating the sodium current and conductance to the shape of transmembrane and extracellular potentials by simulation: effects of propagation boundaries. *IEEE Trans Biomed Eng*, 32(10):743–755, Oct 1985.

[25] N. A. Trayanova. Whole-heart modeling: applications to cardiac electrophysiology and electromechanics. *Circ Res*, 108(1):113–128, Jan 2011.

[26] H. A. van der Vorst. Bi-CGSTAB: A fast and smoothly converging variant of bi-cg for the solution of nonsymmetric linear systems. *SIAM J. Sci. Stat. Comput.*, 13:631–644, 1994.