Time optimal control of the monodomain model in cardiac electrophysiology

K. Kunisch A. Rund

This article is a preprint. It was published in IMA Journal of Applied Mathematics by Oxford University Press. The version of record is available at https://doi.org/10.1093/imamat/hxv010.

Time optimal control of the monodomain model in cardiac electrophysiology

Karl Kunisch*, Armin Rund†

Institute of Mathematics and Scientific Computing
University of Graz
Heinrichstr. 36
8010 Graz, Austria

Keywords: time optimal control, reaction diffusion equation, monodomain model, PDE constrained optimization, trust region semismooth Newton method

Abstract

An optimal control approach to a simplified reaction diffusion system describing cardiac defibrillation is proposed that allows for joint optimization of shape and duration of defibrillation pulses. Within the framework, optimized multiphasic pulses with low energy, short duration and/or low amplitude can be designed according to specific needs. The approach is based on a novel time optimal control formulation for the monodomain model, which takes into consideration the dynamical system properties of the uncontrolled equation. The highly complex dynamics requires a consistent discretization of first and second order information to guarantee effective optimization schemes leading to successful defibrillation. Numerical examples underline the efficiency of the proposed method.

1 Introduction and problem formulation

Over the last decades significant progress was made in the numerical treatment of open loop optimal control problems governed by distributed parameter systems. The techniques that were developed were adapted for a wide range of important equations, including wave and diffusion equations, the equations of fluid mechanics and fluid-structure interaction models. In contrast very little attention was paid to reaction diffusion systems, whose dynamical systems behavior is significantly different from those of the systems mentioned before. In this paper we continue our efforts on one particular reaction diffusion system, which describes the electrical activity of the heart. Compared to our earlier work we propose a new choice of cost functionals, which allows a much wider class of optimized trajectories. This is only possible by using well conceived numerical optimal control techniques. Due to the rich dynamical systems behavior for the problems under consideration, ad hoc techniques will simply fail, especially for second order methods.

Let us briefly describe the physiological background for the problem to be investigated. The heart supplies all organs with blood by rhythmic contractions that are triggered electrically. Disturbances in the formation and/or propagation of electrical impulses may induce reentrant activation patterns which lead to a noticeable increase in the hearts activation rate. Such

*karl.kunisch@uni-graz.at
†armin.rund@uni-graz.at

fast rhythms may lead to fibrillation. To restore a healthy rhythm the delivery of electrical shocks, referred to as defibrillation, is a reliable therapy. It can be administered by means of implantable cardioverters defibrillators (ICDs), which monitor the heart rate and deliver a discharge, which acts as a control, to restore a normal rhythm.

The bidomain model is a well-accepted continuous and macroscopic description of the electrical activity of cardiac muscle cells. The model consists of two coupled reaction-diffusion equations together with an ODE describing the ionic currents associated with the reaction terms, see e.g. [14]. Assuming the intracellular and extracellular conductivity tensors to be linearly dependent, the model can be simplified to the monodomain model, which results in a substantial reduction of the computational effort [20, 24]. Once a model for the physiological phenomena and their dependence on a control input are fixed, an optimal control approach can be utilized to decide on the optimal shock delivery.

Due to severe physiological constraints, involving time scales, geometry and multi-physics aspects, the current optimal control techniques certainly fall short of addressing all relevant aspects. But the medical technology itself is still changing rapidly, so that certain assumptions, as for instance the availability of observations or actuator support which is not too small relative to the overall tissue size, may become reality. Current technological advances include, for instance, the development of a new type of ICDs, see e.g. [21]. They consist of flexible arrays of leads which act as sensors, gathering information on the electrical state of the heart, and as actuator-electrodes, delivering a defibrillation shock when arrhythmias are detected. In case of a defibrillation therapy, each lead is provided with an defibrillation pulse that has to be designed appropriately, based on the measured data.

Within the optimal control approach to cardiac defibrillation defibrillation, pulses are designed by solving an optimal control problem constrained by a reaction diffusion system. The aims of effective defibrillation and minimal detrimental side effects to the patient are modeled within the control objective. By adapting the objective and its parameters, a wide range of goals can be achieved, which makes the optimal control approach a powerful and flexible tool for defibrillation pulse design. The design of the objective is of paramount importance and, together with an efficient numerical realization, are the main innovation of this paper. For the choice of the control objective, several conflicting interests need to be taken into consideration. They include the behavior of the unforced dynamical system, which for the simplified ODE-FitzHugh Nagumo model states that once the state is sufficiently excited it must necessarily reach a plateau value before it can return to the stable equilibrium, see e.g. [17, pg 241f]. For the infinite dimensional system (1) this behaviour can occur at different times at any point in the spatial domain. This suggests to use a control objective (defibrillation) which only involves the terminal time of the control horizon. This leads to a highly ill-conditioned optimal control problem, making exact computation of gradient and Hessian information indispensable.

The topic of numerical simulation of the electrical activity of the heart has inspired much research, so that we can only quote selected references [9, 27]. The optimal control approach to cardiac defibrillation was previously investigated in [18, 10] for the monodomain model, and in [19] for the bidomain model. Differently from the present paper, these papers consider the case where the shock length is fixed. Moreover the cost functional for the optimal control formulation involves a reference trajectory. As a consequence the number of phases of the optimal pulse is determined a-priori. - The optimal control of reaction diffusion systems involving wave phenomena was also the focus of the research in [3, 6].

The article is organized as follows: the monodomain model is described in Section 2. Section 3 is devoted to the formulation of the optimal control problem. The necessary conditions are obtained in Section 4. In Section 5 the optimization method is presented, which is based on a bilevel formulation together with a trust region semismooth Newton method. Section 6 introduces the numerical framework which is chosen in such a manner that discretization

before or after deriving the necessary optimality conditions commute, and lead to a Galerkin discretization with the exact discrete derivatives, see Section 6. The proposed techniques are tested by numerical experiments on termination of reentry waves in Section 7. One of the examples also addresses robustness of the computed controls.

2 The controlled state equation

We investigate a sample of heart tissue described by the domain Ω. The electrophysiology is modeled by the monodomain equation using the cell model of Rogers-McCulloch [22], which is a modified FitzHugh-Nagumo model. For simplicity we do not consider a conductive bath and therefore model the heart tissue to stay electrically isolated leading to homogeneous Neumann boundary conditions. Thus, the dynamical system is given by

$$v_t + I(v,w) - \nabla\cdot(\bar\sigma_i \nabla v) = I_\mathrm{e} \qquad \text{a. e. in } Q := (0,t_\mathrm{f})\times\Omega\,, \tag{1a}$$

$$w_t + G(v,w) = 0 \qquad \text{a. e. in } Q\,, \tag{1b}$$

$$\nu\cdot\bar\sigma_i\nabla v = 0 \qquad \text{on } \Sigma := (0,t_\mathrm{f})\times\partial\Omega\,, \tag{1c}$$

$$v(x,0) = v_0(x)\,,\quad w(x,0) = w_0(x) \qquad \text{a. e. in } \Omega\,. \tag{1d}$$

The independent variables are $x\in\Omega\subset\mathbb{R}^d$, $d=2$, and time $t\in(0,t_\mathrm{f})$ with the terminal time $t_\mathrm{f}>0$. Ω is a bounded domain with Lipschitz continuous boundary $\partial\Omega$ and unit outer normal ν. The functions $v(t,x)$, $w(t,x)$ denote the transmembrane electric potential and the gating or recovery variable. The intercellular conductivity tensor $\bar\sigma_i\in L^\infty(\Omega,\mathbb{R}^{d\times d})$ is assumed to be symmetric and uniformly elliptic. The extracellular stimulation current I_e depends on the defibrillation pulse, which has to be controlled. The ionic current $I(v,w)$ and $G(v,w)$ are given as

$$I(v,w) = \eta_0\, v\left(1-\frac{v}{v_\mathrm{th}}\right)\left(1-\frac{v}{v_\mathrm{pk}}\right) + \eta_1\, v\, w\,, \tag{2a}$$

$$G(v,w) = \eta_2\left(\eta_3\, w - \frac{v}{v_\mathrm{pk}}\right)\,, \tag{2b}$$

with $\eta_0,\eta_1,\eta_2,\eta_3\in\mathbb{R}^+$. A cell is excited, if the transmembrane potential exceeds the threshold potential $v_\mathrm{th}>0$. Further $v_\mathrm{pk}>v_\mathrm{th}$ is the peak potential. The initial conditions $v_0(x)\in L^2(\Omega)$, $w_0(x)\in L^4(\Omega)$ describe a fibrillatory situation.
The geometric setting represents a layer of heart muscle tissue modeled as a 2D domain Ω. On top of it, a finite number of electrode plates $\Omega_{\mathrm{con},k}$, $k=1,\dots,N_\mathrm{e}$ are pasted. In the common setting this would just be a pair, alternatively it can be an array of plates in case of a flexible sensor array. For the monodomain model, each electrode is assigned an independent defibrillation pulse $u_k(t)$ whereas a compatibility condition would be needed for bidomain modeling. The extracellular stimulation current I_e is modeled as

$$I_\mathrm{e}(t,x) = \sum_{k=1}^{N_\mathrm{e}} u_k(t)\chi_{\Omega_{\mathrm{con},k}}(x)\,, \tag{3}$$

where $\chi_{\Omega_{\mathrm{con},k}}(x)$ denotes the characteristic function of electrode plate k, $u_k(t)$ the corresponding pulse and $u(t)=(u_1(t),\dots,u_{N_\mathrm{e}}(t))$ the control vector.

3 The optimal control problem

Here defibrillation will be posed as an optimal control problem. The aim consists in influencing the extracellular stimulation current $I_\mathrm{e}(t,x)$ in such a way, that the tissue changes to a state

where fibrillatory propagation is hindered. Additionally, side effects on the tissue should be kept small. While this description is clear, its particular modeling is involved. We first discuss the choice of the time horizons and then define the optimal control problem.

3.1 Modeling the time horizon

After a defibrillation shock has been applied successfully, the heart muscle tissue needs a certain amount of time to reach a non fibrillatory state, especially in the presence of complicated patterns of reentry waves. Therefore, a successful defibrillation can only be confirmed at a time t_f with $t_\mathrm{f} \gg T$, where T is the end time of the defibrillation shock. There are several ways how one might incorporate this fact into the optimal control problem.

Nagaiah et al. propose in [19] a formulation with a short fixed time horizon $[0, T]$ and enforce the defibrillation on the basis of a tracking functional using a desired trajectory given by an a-priori known defibrillation pulse, which brings the tissue to a non-excited state at $t_\mathrm{f} \gg T$. Post-optimally, the simulation on (T, t_f) is continued to confirm successful defibrillation.
Here we propose a formulation which is different in several ways. First, we do not rely on a desired trajectory, secondly the shock duration itself is optimized. Thirdly, defibrillation is quantified in the cost by demanding that at some final time of simulation t_f the electric potential is small throughout Ω. Thus, the optimization problem is posed on some fixed horizon $[0, t_\mathrm{f}]$ at the end of which defibrillation must be achieved. The defibrillation pulse is applied on the first part $[0, T]$, with T being part of the optimization. Compared to [19] this gives an increased flexibility in the way, how the defibrillation is achieved. In particular the number of phase changes is part of the optimization and is not aligned with some desired trajectory. In addition, for successful defibrillation the system is monitored throughout the time interval $[0, t_\mathrm{f}]$, rather than only on $[0, T]$.
From the point of view of numerical optimization this problem is significantly more challenging, since the elimination of the use of a desired trajectory leads to a drastically reduced coercivity of the optimal control formulation. Our approach will lead to different optimal controls, that deliver less energy to the tissue, since the optimal control formulation is more flexible in choosing the pulses.

3.2 Defibrillation as optimal control problem

For effective defibrillation at time t_f we aim at bringing as much tissue to the resting state as possible. Then, the next natural activation given by the sinoatrial node or by a pacemaker should be able to reestablish the normal heart rhythm. How to model this terminal condition? The goal is realized by a terminal penalty term.

To model negative side effects of the applied shock I_e three different quantities are considered: the duration, the energy and the amplitude of the pulse. Since the exposure of the patient is related to the duration of the electrical shock, we aim at minimizing the duration T. Moreover, the energy of the pulses $\|u\|_2^2$ has to be minimized. Additionally, we restrict the amplitudes by imposing inequality constraints $u_\mathrm{min} \leq u_k(t) \leq u_\mathrm{max}$, since too large amplitudes would result in a local damage to the tissue adjacent to the electrodes.

These considerations suggest the following optimal control problem

$$\min_{0\le T\le t_\mathrm{f},\, u(t)\in U_\mathrm{ad}} J(v,u,T) := T + \frac{\mu}{2}\|v(\cdot,t_\mathrm{f})\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\sum_{k=1}^{N_\mathrm{e}} \|u_k\|^2_{L^2(0,T)}\,, \tag{4a}$$

$$\text{subject to (1) with } I_\mathrm{e} = \sum_{k=1}^{N_\mathrm{e}} u_k(t)\chi_{\Omega_{\mathrm{con},k}}(x)\chi_{(0,T)}(t)\,, \tag{4b}$$

with weighting parameters $\mu > 0, \alpha > 0$. The amplitude of the controls are bounded via the set of admissible controls

$$U_\mathrm{ad} := \{u \in U \,:\, u_\mathrm{min}(t) \le |u_k(t)| \le u_\mathrm{max}(t) \text{ for a. a. } t \in (0,t_\mathrm{f})\,, k = 1,\dots,N_\mathrm{e}\}\,, \tag{4c}$$

where u_min, $u_\mathrm{max} \in L^\infty(0,t_\mathrm{f})$ and $U := L^2(0,t_\mathrm{f};\mathbb{R}^{N_\mathrm{e}})$. Eq. (4) constitutes a time optimal control problem with a nonlinear ODE-PDE system as constraints. The objective (4a) is a scalarized multi-objective formulation favoring successful defibrillation for large μ, small energy inputs for large α and short pulses for small α and μ.

3.3 Existence

At first, we recall the existence and regularity results for the solutions of the monodomain equations, which are defined next. We introduce $Q = (0,t_\mathrm{f}) \times \Omega$ and the Sobolev space $V := H^1(\Omega)$ with its dual V^*. The duality pairing between V and V^* is denoted by $\langle\cdot,\cdot\rangle_{V^*,V}$.

Definition 3.1 *For $I_e \in L^2(0,t_\mathrm{f},V^*)$ and $(v_0,w_0) \in L^2(\Omega)\times L^2(\Omega)$, a pair (v,w) is called weak solution to (1), if $(v,w) \in L^2(0,t_\mathrm{f};V)\cap C([0,t_\mathrm{f}];L^2(\Omega))\cap L^4(Q)\times C^1([0,t_\mathrm{f}];L^2(\Omega))$, $v_t \in L^2(0,t_\mathrm{f};V^*) + L^{\frac{4}{3}}(Q)$, and for a.a. $t \in (0,t_\mathrm{f})$ and all $\varphi \in V$*

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\displaystyle\int_\Omega v(t)\varphi\,\mathrm{d}x + \int_\Omega \bar\sigma_i \nabla v(t)\nabla\varphi\;\mathrm{d}x + \int_\Omega I(v(t),w(t))\varphi\;\mathrm{d}x = \langle I_e(t),\varphi\rangle_{V^*,V}\,, \\ w_t(t) + G(v(t),w(t)) = 0 \qquad \text{a.e. in } \Omega\,, \end{cases}$$

where the time derivative is to be understood in the distributional sense.
Existence and uniqueness results for the bidomain equation are considered in e.g. [4, 18]. Since we restrict ourselves to the monodomain equation here and since we use a simple form for G, only minor modifications in the proof of these results imply the following proposition, which holds in dimensions 2 and 3, see also [16] for the monodomain equation.

Proposition 3.1 *Let $I_e \in L^2(0,t_\mathrm{f},V^*)$ and $(v_0,w_0) \in L^2(\Omega)\times L^2(\Omega)$,. Then System (1) admits a weak solution. Furthermore, there exists a constant C, such that*

$$\begin{aligned}\|v\|^2_{C([0,t_\mathrm{f}];L^2)} + \|v\|^2_{L^2(0,t_\mathrm{f};V)} + \|v\|^4_{L^4(Q)} + \|v_t\|^{\frac{4}{3}}_{L^{\frac{4}{3}}(0,t_\mathrm{f};V^*)} + \|w\|^2_{C^1([0,t_\mathrm{f}];L^2)} \\ \le C(1 + \|v_0\|^2_{L^2(\Omega)} + \|w_0\|^2_{L^2(\Omega)} + \|I_e\|^2_{L^2(V^*)})\,.\end{aligned}$$

If additionally $I_e \in L^\infty(0,t_\mathrm{f};V^*)$ and $w_0 \in L^4(\Omega)$ holds, then the weak solution is unique.

This proposition applies in particular to the choice of I_e made in (4b).
In the following, we prove the existence of a global minimizer of the time optimal control problem (4).

Proposition 3.2 *Problem (4) admits a solution* (v^*, w^*, u^*, T^*).

Proof. Let $\{(u^n, T^n)\}_{n=1}^{\infty}$ denote a minimizing sequence. This sequence is bounded and hence there exists a subsequence, denoted by the same symbols, and (u^*, T^*) such that $(u^n, T^n) \rightharpoonup (u^*, T^*)$ weakly in $L^2(0, t_\mathrm{f}; \mathbb{R}^{N_\mathrm{e}}) \times \mathbb{R}$ with $u^* \in U_\mathrm{ad}$.
Let $(v^n, w^n) = (v(u^n), w(u^n))$ denote the associated states of the monodomain equation. By Proposition 3.1 they are bounded in $\mathbb{X} := L^2(0, t_\mathrm{f}; V) \cap W^{1,\frac{4}{3}}(0, t_\mathrm{f}; V^*) \times W^{1,2}(0, t_\mathrm{f}; L^2(\Omega))$. In particular, there exists a weakly convergent subsequence of $\{(v^n, w^n)\}$ in $\mathbb{X}$ on which we can pass to the limit in the state equations so that $(v(u^*), w(u^*))$ satisfy (1). Since $\{v^n\}$ is bounded in $L^2(0, t_\mathrm{f}; V)$ and $\{v_t^n\}$ is bounded in $L^{\frac{4}{3}}(V^*)$ it follows that, possibly on a further subsequence, $v^n(t_\mathrm{f}) \to v^*(t_\mathrm{f})$ strongly in V^*, see e.g. [7], pg. 71. Since $\{v^n(t_\mathrm{f})\}$ is bounded in $L^2(\Omega)$ we also have that $v^n(t_\mathrm{f}) \rightharpoonup v^*(t_\mathrm{f})$ weakly in $L^2(\Omega)$. Now we can pass to the limes inferior in

$$\begin{aligned}\inf_{0\le T\le t_\mathrm{f}, u\in U_{ad}} J(v,u,T) &= \varliminf_{n\to\infty} (T^n + \frac{\mu}{2}\|v(\cdot;u^n,t_\mathrm{f})\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\sum_{k=1}^{N_\mathrm{e}} \|u_k^n\|^2_{L^2(0,T^n)}) \\ &\ge T^* + \frac{\mu}{2}\|v(\cdot;u^*,t_\mathrm{f})\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\sum_{k=1}^{N_\mathrm{e}} \varliminf_{n\to\infty} \|u_k^n\|^2_{L^2(0,T^n)}).\end{aligned}$$

To treat the last term we define

$$\tilde{u}_k^n = \begin{cases} u_k^n & \text{on } (0,T^n) \\ 0 & \text{on } (T^n, t_\mathrm{f}) \end{cases}, \qquad \tilde{u}_k^* = \begin{cases} u_k^* & \text{on } (0,T^*) \\ 0 & \text{on } (T^*, t_\mathrm{f}). \end{cases}$$

It is simple to verify that $\tilde{u}_k^n \rightharpoonup u_k^*$ weakly in $L^2(0, t_\mathrm{f})$. Therefore,

$$\varliminf_{n\to\infty} \int_0^{T_n} |u_k^n|^2 = \varliminf_{n\to\infty} \int_0^{t_\mathrm{f}} |\tilde{u}_k^n|^2 \ge \int_0^{t_\mathrm{f}} |\tilde{u}_k^*|^2 = \int_0^{T^*} |u_k^*|^2,$$

consequently

$$\inf_{0\le T\le t_\mathrm{f}, u\in U_{ad}} J(v,u,T) \ge J(v^*,u^*,T^*),$$

and thus, (u^*, T^*) is a solution to (4). ∎

4 Necessary conditions

The numerical realization of (4) is based on first order necessary optimality conditions that an optimal solution $(\bar{u}, \bar{v}, \bar{w}, \bar{T})$ has to fulfil. Applying a formal Lagrangian approach with $p(t,x)$ and $q(t,x)$ as the Lagrange multipliers associated to the parabolic PDE and the ODE, one can proceed in a by now standard manner to obtain the first order necessary system, see e.g. [25, 16], for problems with fixed time horizon and [13] for time optimal control problems. The first order necessary system consists of the state equations (1), the adjoint equations (5), the optimality conditions (6), and a transversality condition (7) for the optimal free time $\bar{T}$.

$$-p_t - \nabla\cdot(\bar{\sigma}_i \nabla p) + I_v(\bar{v},\bar{w})\, p + G_v\, q = 0 \qquad \text{in } Q\,, \tag{5a}$$

$$-q_t + I_w(\bar{v},\bar{w})\cdot p + G_w \cdot q = 0 \qquad \text{in } Q\,, \tag{5b}$$

$$\nu\cdot\bar{\sigma}_i\nabla p = 0 \qquad \text{on } \Sigma\,, \tag{5c}$$

$$p(t_\mathrm{f}) = \mu\,\bar{v}(t_\mathrm{f})\,, \quad q(t_\mathrm{f}) = 0 \qquad \text{in } \Omega\,. \tag{5d}$$

$$(\alpha\,\bar{u}(t) + B^*p(t)) \cdot (u(t) - \bar{u}(t)) \geq 0 \qquad \text{a.a. } t \in (0,\bar{T})\,, \forall u \in U_{\text{ad}}\,. \tag{6}$$

$$\begin{aligned} 0 = \frac{1}{\bar{T}} \int_0^{\bar{T}} \big(1 + \frac{\alpha}{2}\|\bar{u}\|_2^2 + \langle I_{\text{e}}(\bar{u}) + \nabla\cdot(\bar{\sigma}_i \nabla v) - I, p\rangle - \langle G, q\rangle\big)\,\mathrm{d}t \\ - \frac{1}{t_{\text{f}} - \bar{T}} \int_{\bar{T}}^{t_{\text{f}}} \big(\langle \nabla\cdot(\bar{\sigma}_i\nabla v) - I, p\rangle - \langle G, q\rangle\big)\,\mathrm{d}t\,. \end{aligned} \tag{7}$$

Here I_v, I_w, G_v, G_w denote the partial derivatives of the model functions (2) and $B^* : L^2(Q) \to U$, $B^*p := (\chi_{(0,\bar{T})}(t) \int_{\Omega_{\text{con},k}} p(t,x)\,\mathrm{d}x)_{k=1,\dots,N_{\text{e}}}$. For the derivation of the transversality condition by a time transformation we refer to [15].
To apply the semismooth Newton method later on, we first reformulate (6) using the projection operator $P_{\text{ad}} : L^2(0,t_{\text{f}};\mathbb{R}^{N_{\text{e}}}) \to L^2(0,t_{\text{f}};\mathbb{R}^{N_{\text{e}}})$, $P_{\text{ad}}(y) = \min(u_{\max}, \max(u_{\min}, y))$ resulting in

$$\bar{u}(t) = P_{\text{ad}}\left(-\frac{1}{\alpha}B^*p\right) \qquad \text{a.a. } t \in (0,\bar{T})\,. \tag{8}$$

Secondly, we introduce artificial optimization variables

$$z \in U\,, \qquad z := (z_k) = -\frac{1}{\alpha}B^*p$$

and parametrize the controls as $u = P_{\text{ad}}(z)$. Hence, we shift the non-smooth projection operator to the state equation and the objective. Thus, the first order necessary conditions are equivalent to (1), (5), (7) with eliminated control $u = P_{\text{ad}}(z)$ and

$$0 = F(z) := \alpha z + B^*p \qquad \text{a.a. } t \in (0,\bar{T})\,. \tag{9}$$

5 Methods

Time optimal control problems are challenging numerically. To partially appreciate this fact we note that by means of a time transformation, time optimal problems can be transformed to a fixed time interval, at the expense of an additional nonlinearity in the dynamical system. We want to avoid such a new nonlinearity since already (1) is known to be rich in structure, allowing wave-like- and reentry phenomena, for example.
Therefore we propose a bilevel approach for solving (4), separating T and the controls u by treating T as parameter in the lower level problem:

$$\min_{0<T\leq t_{\text{f}}} \left(\min_{\substack{u\in U_{\text{ad}} \\ \text{s.t. } (4b)}} J(v,u;T) \right)\,. \tag{10}$$

Obviously, this problem has the same solution as the time optimal control problem (4). For each fixed T the lower level problem (LLP) constitutes a terminal tracking problem for a coupled ODE-PDE system with controls acting on a fixed part of the time interval. An alternative all-at-once approach will be developed in [15].
The bilevel problem will be solved by an iterative method, where the LLP is solved by a semismooth Newton method (TR-SN). It consists of a combination of the reduced Newton method of [11] with a globalization based on Steihaug-CG [23]. The extension to semismooth Newton methods to allow for the control constraints $u(t) \in U_{\text{ad}}$ will be explained in the next section. The method is matrix-free i.e. the Hessians are not set up explicitly, but we compute only the action of the Hessians and resort to Krylov methods. All forward and backward

systems are solved efficiently with time-stepping methods, see Sect. 6. A globally convergent (derivative-free) direct search method is used for the upper level minimization problem avoiding the transversality condition (7), which is checked a-posteriori.
Before we describe the TR-SN, we note that with the technique of the proof of Proposition 3.2 it is simple to argue the following result.

Lemma 5.1 (Existence of an optimal solution of the LLP)
The lower level problem has an optimal solution for every $T > 0$.

The optimality conditions for the LLP consist of (1), (5), and (8) with a fixed current guess for $\bar{T}$, and follow from the results in [16].

5.1 Trust region semismooth Newton method for solving the LLP

In the following, we describe the solution of the LLP with a matrix-free semismooth Newton method. Therefore we treat all state and adjoint variables as functions of z (as solutions of (1) and (5)), and we define the reduced objective w.r.t. u as $j(u) = J(v(u), u; T)$. Consequently, the reduced optimality condition is $0 = F(z)$ with F from (9). Here F is non-smooth, but it allows for the application of a semismooth Newton method according to [15]. Using the semismoothness calculus in Banach spaces from e.g. [12, 26], we introduce the generalized differential of the projection operator $P_{\mathrm{ad}}(y)$

$$DP_{\mathrm{ad}}(z)(h) = \chi_{\mathcal{I}} h\,, \tag{11}$$

where $\chi_{\mathcal{I}} h := (\chi_{\mathcal{I}^k} h_k)_k$ and $\chi_{\mathcal{I}^k}$ denotes the indicator function of the inactive set $\mathcal{I}^k = \{t \in (0,T) \,|\, u_{\min}(t) < z_k(t) < u_{\max}(t)\}$ of component u_k. The generalized derivative of F at point z^n in the direction δz is then given by

$$H(z^n)(\delta z) = \alpha \delta z + B^* \delta p(\delta z)\,. \tag{12}$$

To compute $\delta p(\delta z)$, first the tangent equation depending on δz and incorporating $\chi_{\mathcal{I}^k}$ is solved for δv, δw, and then the second adjoint equation is solved for δp, δq, see the end of appendix A. Together we can formulate the semismooth Newton iteration

$$H(z^n)(\delta z) = -F(z^n)\,, \qquad z^{n+1} = z^n + \delta z\,. \tag{13}$$

While the Hessian H is in general non symmetric, it is symmetric with respect to the L^2-inner product of the inactive set $(a,b)_{\mathcal{I}} := \sum_{k=1}^{N_{\mathrm{e}}} \int_0^T \chi_{\mathcal{I}^k} a_k b_k \,\mathrm{d}t$. Therefore we compute d by solving (13) with the CG method using $(\cdot,\cdot)_{\mathcal{I}}$ as inner product. By this we obtain a solution of (13) on the inactive set i.e. $\chi_{\mathcal{I}}(Hd + F) = 0$. Afterwards, a solution of the full system (13) is obtained by updating the components on the active set according to

$$\delta z = d - \frac{1}{\alpha}(F(z^n) + H(z^n)d)\,. \tag{14}$$

We note that for $U_{\mathrm{ad}} = U$ the semismooth Newton method coincides with the well-known matrix-free Newton method of [11].

Since Newton methods are generally only locally convergent, we embed the method into a trust region framework following the lines of [23], which in the unconstrained case is proven to be globally convergent. Therefore we note that the CG method with $(\cdot,\cdot)_{\mathcal{I}}$ computes a particular solution of the quadratic problem

$$\min_{h \in U} \varphi_{z^n}(h) := (h, F(z^n))_{\mathcal{I}} + \frac{1}{2}(h, H(z^n)h)_{\mathcal{I}}\,.$$

We replace this problem by the trust region problem

$$\min_{h\in U} \varphi_{z^n}(h) \quad \text{s.t. } \|h\|_{\mathcal{I}} \leq \Delta_n \,,$$

with trust region radius $\Delta_n > 0$ and $\|h\|_{\mathcal{I}} = \sqrt{(h,h)_{\mathcal{I}}}$. It is solved with Steihaug-CG [23, Sect. 2] using the inner product $(\cdot,\cdot)_{\mathcal{I}}$. The update (14) is done only for a fully converged CG method, hence not for the cases when negative curvature or a large step is encountered. For practical realization the update (14) should be replaced by minimizing the residual in direction of $r = -F - Hd$ according to

$$\delta z = d + \theta r \qquad \text{with } \theta \in \mathbb{R}\,, \quad \theta = \arg\min(\|H(d+\theta r) + F\|_{L^2(0,T;\mathbb{R}^{N_e})})\,, \tag{15}$$

in order to make the procedure more robust w.r.t. rounding errors.
The update of the trust region radius Δ_n and the decision of accepting or rejecting a step are done analogously to [23], see the full algorithm in Appendix A. Additionally, we modify the trust region method to be monotone, i.e. accepted steps will always yield a decrease in the objective.

5.2 Direct search method for the upper level problem

The upper level problem is solved with a globally convergent derivative-free optimization method based on bisection. It is assumed that the optimal values $G(T)$ of the LLP are continuous w.r.t T. We start from a triple $L < M < R$ with $G(M) < G(L)$ and $G(M) < G(R)$, i.e. we assume that a minimizer is contained in $[L, R]$. Then both intervals are bisected by $P := (L+M)/2$ and $Q := (M+R)/2$ and $G(P)$, $G(Q)$ are computed. Next we chose M as minimizer in $\{M, P, Q\}$, tighten both intervals and iterate. Additionally, we skip the computation of $G(Q)$ if $G(P) < G(M)$ holds.

6 Discretization

We give a brief description of the discretization of the LLP. To combine the advantages of First-Discretize-Then-Optimize methods (FDTO) and First-Optimize-Then-Discretize (FOTD) methods, we choose a FE-Galerkin method in space together with a Petrov-Galerkin method in time, which allows for exact discrete derivatives and a natural translation of the optimality conditions from the continuous to the discrete level, see [2]. Hence, FDTO and FOTD commute and coincide within our framework, which is very important for trust region Newton methods.
In particular, we choose Lagrange Q1 elements on a quadrilaterally structured grid for spatial and the Crank-Nicolson method in the cG(1)-scheme for temporal discretization, for the latter see e.g. [8, 2]. Since the spatial discretization is straightforward, we defer it to Appendix B. However, the time discretization is important to gain exact discrete derivatives and decoupling. Therefore the essential parts are presented in the following, concentrating on the semidiscretization in time.
We aim for an efficient decoupling method to solve the ODE and PDE variables independently per time step. Therefore we utilize a decoupling of the ODE from the PDE by taking the gating variable explicitly in the PDE. By working thoroughly through the Lagrangian calculus, we reestablish the exact discrete derivatives respecting the decoupling.
A time grid $t_0 < \cdots < t_N$ with stepsizes $\tau_m := t_m - t_{m-1}$ is chosen. The state variables are semidiscretized in time as continuous piecewise linear functions with values $V^m(x) = v(t_m, x)$, $m = 0, \ldots, N$ and analogously for w, see Fig. 1. The adjoint and control variables are piecewise constant in time with values $P^m(x)$. Hence we have $p(t,x) = \sum_{m=1}^{N} P^m(x)\chi_{(t_{m-1},t_m]}(t)$, and analogously for $q(t,x)$ and $u_k(t) = \sum_{m=1}^{N} u_k^m \chi_{(t_{m-1},t_m]}(t)$.

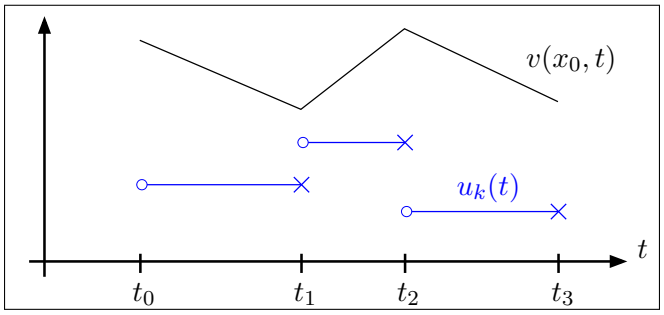

Figure 1: Ansatz space in time for state (black), control and adjoint variables (blue).

Therefore, the semidiscrete Lagrangian $\mathcal{L}$ can be expressed as

$$\begin{aligned}\mathcal{L}(\dots) := T + \frac{\mu}{2}\int_\Omega (V^N)^2\,\mathrm{d}x + \frac{\alpha}{2}\sum_{k=1}^{N_\mathrm{e}}\sum_{m=1}^{N}\tau_m (u_k^m)^2 - \sum_{m=1}^{N}\int_\Omega \frac{\tau_m}{2}\nabla P^m\cdot\bar{\sigma}_i\nabla(V^m + V^{m-1}) \\ + P^m\Big[V^m - V^{m-1} - \tau_m\sum_{k=1}^{N_\mathrm{e}}\chi_{\Omega_{\mathrm{con},k}}u_k^m + \frac{\tau_m}{2}I(V^m, W^{m-1}) + \frac{\tau_m}{2}I(V^{m-1}, W^{m-1})\Big]\,\mathrm{d}x \\ - \sum_{m=1}^{N}\int_\Omega Q^m\left[W^m - W^{m-1} + \frac{\tau_m}{2}G\left(V^m + V^{m-1}, W^m + W^{m-1}\right)\right]\mathrm{d}x\,,\end{aligned} \tag{16}$$

where we leave the inequality constraints as explicit constraints. We again emphasize the decoupling of w at $I(V^m, W^{m-1})$, which later results also in an adapted decoupling in the adjoint and tangent equations. Therefore, the ODE can generally be solved efficiently in a matrix-free manner.
Next, the well-known Lagrange formalism yields a consistent semidiscretization of tangent, adjoint and second adjoint equation. A subsequent spatial discretization with FE is straightforward and results in the equations in Appendix B.
The FE calculations are done with `deal.II` [1]. The nonlinear systems in each time step of the state equation are solved with Newton's method, and the linear systems are solved directly with UMFPACK.

7 Numerical experiments

In the following, the proposed formulation and method are tested on several examples. The choice of parameters is inspired by [9], where one can also find the aforegoing nondimensionalization. The following parameters are fixed throughout all examples:

η_0	η_1	η_2	η_3	v_th	v_pk	$\bar{\sigma}_i$
1.5	4.4	0.012	1.0	13	100	$\mathrm{diag}(3\cdot 10^{-3}, 3.1525\cdot 10^{-4})$

The geometry is set to be a rectangle $\Omega = (0,2)\times(0,0.8)$ of size 2cm $\times$ 0.8cm, which is discretized into 128×64 cells. All computations were done with an equidistant time discretization with step size $\tau = 0.04$ (msec). The stopping criteria are set to $\|F_n\| < \min(10^{-5}, 10^{-5}\|F_0\|)$ for the (trust region) Newton method – where the gradient F_n is the discretization of (9) – and $\|r_k\| < 10^{-5}\|r_0\|^{1.3}$ for the residual of the Steihaug-CG method.
The initial condition (v_0, w_0) describes a reentry wave of the type "figure of eight". It is constructed by the usual S1-S2-protocol as follows. Starting by exciting the lower edge $v(x,0) =$

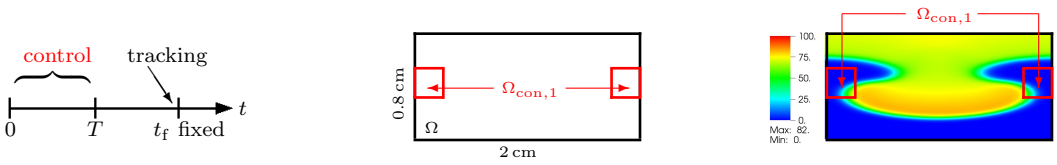

Figure 2: Time domain, geometry and initial condition of the 1st example.

101 if $x_2 \leq 1/160$ and 0 otherwise, $w(x,0) = 0$, we integrate the uncontrolled solution until $t = 130$ using a fixed step size $\tau = 0.1$. The solution describes a planar wave front travelling from the bottom up. As soon as the center gets excitable again, a second stimulus is based on a circle around the midpoint with radius 0.3 for 2 ms, i.e. $I_e = 200\,\chi_{\Omega_{S2}}(x)\,\chi_{[130,132]}(t)$ with $\Omega_{S2} = B_{0.3}(1, 0.4)$. We carry on the simulation without any further stimulus up to $t = 217$ and save both states $v(x, 217), w(x, 217)$ as future initial conditions for the optimization. The timing and radius of the second stimulus are crucial. For different domains or parameters, one has to adapt it by trial and error, otherwise a reentry wave will not evolve.

In the examples which follow, we address the different demands for optimized pulses, looking for: a short pulse with restricted amplitude in Example 1, a low norm $\|u\|$ in Example 2, and a robust optimized pulse w.r.t. the tensor data in Example 3.

7.1 1. Example: symmetric reentry wave

We start with an axially symmetric problem, where it is possible to defibrillate with just one control pulse, i.e. $N_e = 1$. The geometry of the control domain is $\Omega_{con,1} = [0, 0.25] \times [0.3, 0.55] \cup [1.75, 2] \times [0.3, 0.55]$, see Fig. 2. The bilevel method was started on the interval $[L, R] = [30, 40]$ and convergence was reported for $|R - L| < 4 \cdot 10^{-2}$. The parameters were $t_f = 64$, $\alpha = 10^{-3}$, $\mu = 1000$ and $u_{max} = -u_{min} = 100$. The initial control is set to $u_0^1 = u_0 = -50$ for the first LLP with $T^1 = 40$. All other LLP for $k \geq 2$ were warm-started with the optimal control of the former LLP $\bar{u}^{k-1}$ restricted to the current interval $[0, T^k]$ resp. expanded with zero, e.g. $u_0^k(t) = \bar{u}^{k-1}(t)\chi_{[0,\min(T^k,T^{k-1})]}(t)$. An alternative procedure to obtain the new initial control $u_0^k(t)$ is to linearly map $\bar{u}^{k-1}$ from $[0, T^{k-1}]$ to $[0, T^k]$ by $u_0^k(t) = \bar{u}^{k-1}(t\,\frac{T^{k-1}}{T^k})$.
The direct search method in the upper level needs 16 function evaluations to converge at $\bar{T} = 34.12$ with $\bar{J} = 238.786$, i.e. 16 LLP were solved in total. We note, that this is not the shortest pulse that effectively defibrillates, since we are facing a multi-objective formulation with 3 goals. It is an optimal compromise between short duration and low energy. The total number of state, gradient and Hessian evaluations throughout the bilevel run are 78, 71, and 658, respectively. 7 of the 62 TR-Newton steps are rejected. The total number of 559 CG steps yields ≈ 9 CG steps per Newton step; excluding the globalization steps, we observe ≈ 14 CG steps per fully converged CG call.
Typically, the most CPU work is required for the first LLP with $T = R = 40$, since it is not warm-started (see the left part of Tab. 1). The TR-SN method needs 22 steps to converge, reducing the objective from $j(u_0) = 38118$ to $j(u_{22}) = 244$ and reducing the first order optimality $\|F(u_n)\|$ significantly. The last column shows the number of CG iterations. All subsequent LLP solves show a fast convergence of the TR-SN method, see e.g. the second LLP solve with $T = L = 30$ on the right of Table 1. Due to the warm-start, only a few globalization steps are needed, where Steihaug-CG is stopped due to too large steps (flag 1) or negative curvatures (flag 2). Afterwards, the CG is fully solved (flag 0) and the number of inactive time points $|\mathcal{I}|$ converges. Superlinear convergence of the objective j is observed from $s_n := \frac{j(u_{n+1})-j(u_n)}{j(u_n)-j(u_{n-1})}$ in the last column.

n	$j(u_n)$	$\|F(u_n)\|$	#CG	$\|\mathcal{I}\|$
0	38118	$8.3 \cdot 10^{1}$		1000
5	538	$6.9 \cdot 10^{0}$	2	555
10	327	$1.8 \cdot 10^{0}$	1	738
15	262	$6.5 \cdot 10^{-1}$	2	837
20	244	$1.6 \cdot 10^{-2}$	13	915
21	244	$1.9 \cdot 10^{-3}$	13	915
22	244	$1.3 \cdot 10^{-4}$	14	915

n	$j(u_n)$	$\|F(u_n)\|$	#CG	flag	$\|\mathcal{I}\|$	s_n
0	255.744	$6.2 \cdot 10^{-1}$	0		665	
1	254.084	$4.3 \cdot 10^{-1}$	2	1	663	
2	254.084	$4.3 \cdot 10^{-1}$	8	2	663	
3	253.574	$4.7 \cdot 10^{-2}$	7	1	577	0.31
4	253.533	$2.0 \cdot 10^{-3}$	14	0	570	0.08
5	253.533	$1.9 \cdot 10^{-5}$	14	0	569	0.00
6	253.533	$7.3 \cdot 10^{-11}$	15	0	569	0.00

Table 1: TR-SN method for the 1st LLP with $T = 40$ (left) and the 2nd LLP with $T = 30$ (right).

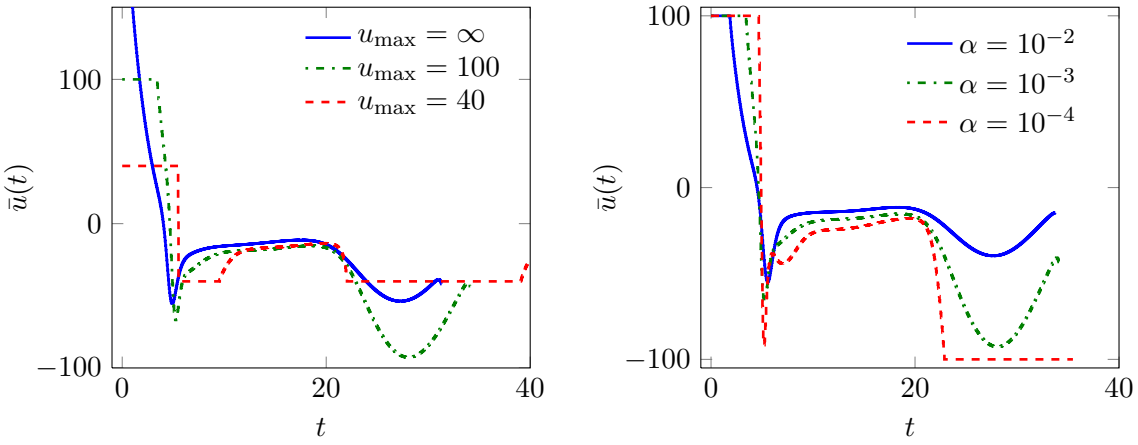

Figure 3: Time optimal controls for different $u_{\max} = -u_{\min}$ with $\alpha = 10^{-3}$ (left) and different α with $u_{\max} = 100$ (right).

The time optimal control $\bar{u}(t)$ is depicted in green in both graphs of Fig. 3. Additional curves show the time optimal controls for different control bounds $u_{\max}$ (left) and different cost parameters α (right). Apparently, all time optimal controls differ to a large extent from the initial control $u_0(t) = -50\chi_{[0,40]}(t)$, in particular the shape, the duration and the switching structure. Consequently, the corresponding trajectories behave qualitatively different. While the initial control only counteracts the wave propagation due to $u_0 \leq 0$, we observe a speed up of the wave propagation at certain points for the time optimal control, since it features positive values, too. According to the left plot and Tab. 2, a lower bound $u_{\max}$ leads to an increase in the optimal pulse length $\bar{T}$ and the optimal value $\bar{J}$, since the effectivity of the control decreases. On the other hand, reducing the cost parameter α results in a smaller optimal value, a slightly increased pulse length and a larger energy of the optimal pulse.

$u_{\max}$	$\bar{J}$	$\bar{T}$	$\|\bar{u}(t)\|_{L^2(0,\bar{T})}$
∞	130	31	329
100	239	34	334
40	2167	39	217

α	$\bar{J}$	$\bar{T}$	$\|\bar{u}(t)\|_{L^2(0,\bar{T})}$
10^{-2}	501	33.9	206
10^{-3}	239	34.1	334
10^{-4}	173	35.1	436

Table 2: Optimal value, pulse length and norm of the time optimal pulse for different $u_{\max}$ with $\alpha = 10^{-3}$ (left) and for different α with $u_{\max} = 100$ (right).

For a verification we compute the transversality condition (7) both for the initial guess (u_0^1, T^1) and the optimal pair $(\bar{u}, \bar{T})$, which yields -1660 and -0.1, respectively. The comparison shows a relative decrease of $6 \cdot 10^{-5}$ in this optimality condition, which underlines the optimality of the computed time optimal control.

7.2 2. Example: asymmetric reentry wave

The 2nd example considers two independent electrode plates with $I_e = \chi_{[0,T]}(t)\left(u_1(t)\chi_{\Omega_{con,1}}(x) + u_2(t)\ \chi_{\Omega_{con,2}}(x)\right)$ in an asymmetric setting $\Omega_{con,1} = [0, 0.25] \times [0.4, 0.55]$, $\Omega_{con,2} = [1.75, 2] \times [0.35, 0.4]$, see Fig. 4. The parameters are $t_f = 65$, $\alpha = 1 \cdot 10^{-5}$, $\mu = 100$ and $U_{ad} = U$ i.e. the LLP method coincides with a trust region Newton method.

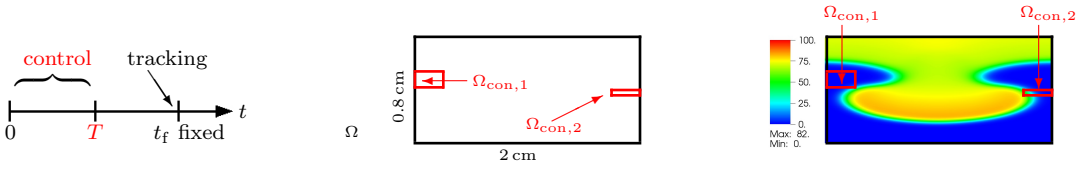

Figure 4: Time domain, geometry and initial condition of the 2nd example.

The bilevel method was started on the interval $[L, R] = [27.5, 37.5]$ with $u_0 = -50$ and convergence was reported for $|R - L| < 4 \cdot 10^{-2}$. We observe again global convergence of the bilevel method and locally superlinear convergence of each LLP. Fig. 5 depicts snapshots of the time optimal transmembrane voltage $\bar{v}(t, x)$ for six different times t, both for the optimally controlled (above the line) resp. for the uncontrolled case. At the very beginning of the time horizon the positive part of the pulses heavily influence the excitable part of the tissue adjacent to the wave front, bringing it to a non excitable state (parts in color red). Thus, the wave can not progress upwards, falls apart and leaves the domain. At the terminal time, not a single part of the tissue is excited, which confirms a successful defibrillation.

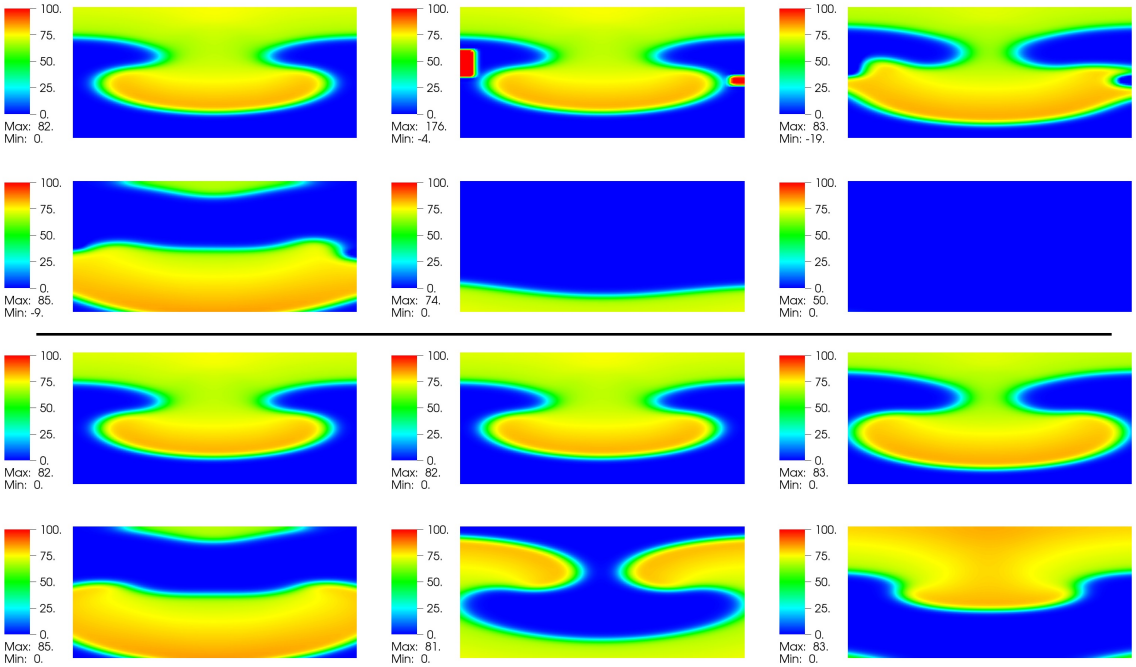

Figure 5: Above line: snapshots of the optimal state $\bar{v}(t, x)$ at $t = 0$, 0.12, 6 (upper row) and $t = 16$, 48, 65 (lower row). Below line: analogous snapshots for the uncontrolled reentry wave $u \equiv 0$.

For checking the gradient and Hessian consistency, we verify the derivatives given by the adjoint calculus via a comparison with finite differences in Tab. 3 using the initial control $u = -50$ and $d = -F(z)$. For the gradient, the absolute difference abs $= C - (g, d)$ and the relative difference rel $= \frac{\text{abs}}{C}$ are computed using the central difference $C = \frac{j(u+\varepsilon d) - j(u-\varepsilon d)}{2\varepsilon}$. For the

Hessian, the differences are abs $= C - (d, Hd)$, rel $= \frac{\text{abs}}{C}$ with $C = \frac{j(u+\varepsilon d)-2j(u)+j(u-\varepsilon d)}{\varepsilon^2}$. All columns confirm a quadratic convergence of the finite differences to the adjoint-based values of the first and second derivatives, as well as a very high precision of the gradient and Hessian code. This is crucial for the success of the optimization since optimal control problems with only terminal observation are known to be highly ill-conditioned.

	Gradient		Hessian	
ε	abs	rel	abs	rel
$1.0e+01$	$7.9e+03$	$1.0e+00$	$1.6e+03$	$1.0e+00$
$1.0e+00$	$5.4e-01$	$1.7e-02$	$2.0e-01$	$2.9e-02$
$1.0e-01$	$5.2e-03$	$1.6e-04$	$1.9e-03$	$2.8e-04$
$1.0e-02$	$5.2e-05$	$1.6e-06$	$1.9e-05$	$2.8e-06$
$1.0e-03$	$5.8e-07$	$1.8e-08$	$8.8e-06$	$1.3e-06$
$1.0e-04$	$1.2e-07$	$3.6e-09$	$1.4e-03$	$2.1e-04$
$1.0e-05$	$4.6e-07$	$1.4e-08$	$1.3e-01$	$2.0e-02$

Table 3: Verification of the gradient and Hessians against finite differences with $U_{\text{ad}} = U$.

To find time optimal control pulses with consideration for small energy, we successively increase α and depict the corresponding time optimal controls and their energy in Fig. 6. The required energy decreases from 2195 to 121 while maintaining an effective defibrillation pulse. For increasing α the optimal duration increases as well.

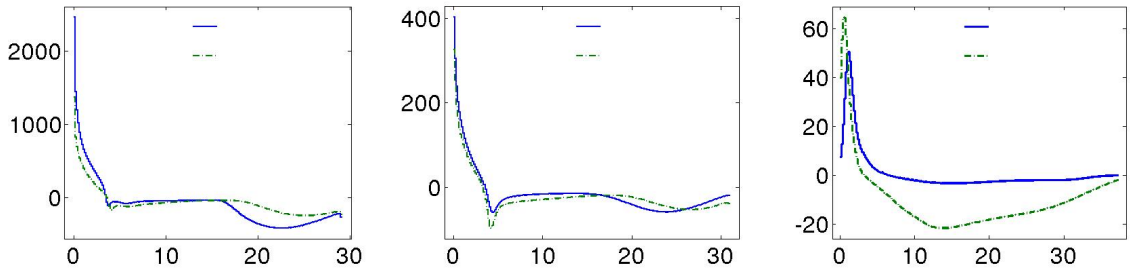

Figure 6: Time optimal controls $\bar{u}_1(t)$ (blue) and $\bar{u}_2(t)$ (green) for different $\alpha = 10^{-5}$, 10^{-3}, 10^0 with corresponding norms $\|u\|_U = 2195$, 442, 121.

7.3 3. Example: a robust design

In the next example, we take into account some uncertainty in the conductivity tensor data, reflecting the fact that they may vary heavily between different settings. As an example, we set $\bar{\sigma}_i = \text{diag}(\sigma \cdot 10^{-3}, 3.1525 \cdot 10^{-4})$ and assume that $\sigma \in \mathbb{R}^+$ is a random variable. By extending [5, Sect. 6.4] to optimal control problems, the expectation value of the tracking term enters the objective. Thus we replace J by

$$J_{\mathbb{E}} = T + \frac{\mu}{2}\mathbb{E}\left(\|v(x, t_{\text{f}}; \sigma)\|^2\right) + \frac{\alpha}{2}\|u\|_U^2 \,. \tag{17}$$

Together with the constraints (1) and $u \in U_{\text{ad}}$, this constitutes a stochastic robust control problem, which in general is computationally demanding. Therefore we restrict ourselves to the case where σ takes only a finite number of values $\{\sigma^1, \ldots, \sigma^r\}$ with probabilities $P_1, \ldots, P_r \geq 0$,

$\sum_{k=1}^{r} P_k = 1$. Consequently, the objective, the reduced gradient and Hessian change to

$$J_r = T + \frac{\mu}{2}\sum_{k=1}^{r} P_k \|v(x, t_\mathrm{f}; \sigma^k)\|^2 + \frac{\alpha}{2}\|u\|^2 \;= \sum_{k=1}^{r} P_k J(v, u, T; \sigma^k)\,, \tag{18a}$$

$$F_r = \sum_{k=1}^{r} P^k F(z; \sigma^k)\,, \qquad H_r(z)\delta z = \sum_{k=1}^{r} P^k H(z; \sigma^k)\delta z\,. \tag{18b}$$

We see that each call to the objective, the gradient and the Hessian has to be split into r calls to the existing solvers (with different $\bar{\sigma}_i$) followed by a weighted mean. This would allow a parallelization of the code, which, however, is not pursued here.
To investigate the effect of the robustness approach, we compare an optimal control (for fixed $\sigma = 3$) to a robust optimal control in the following. To facilate this comparison, we fix the pulse length $T = t_\mathrm{f}$, i.e. we compute only one LLP for both settings. Thus we compute the solution u_1 of the LLP with fixed $\sigma = 3$ on the one hand, and the robust counterpart u_r, that minimizes the LLP incorporating the changes from (18), on the other hand.
We investigate the reentry setting with an electrode placement different from above: $\Omega_{\mathrm{con},1} = [0.05, 0.5] \times [0.45, 0.55]$, $\Omega_{\mathrm{con},2} = [1.8, 1.9] \times [0, 0.45]$. The parameters are set to $\alpha = 10^{-2}$, $\mu = 1000$ and $t_\mathrm{f} = 84$. As an example, we test a uniform distribution for $\sigma \in \{2, 4, 6, 8, 10\}$, i.e. $p_j = 1/r \;\forall j$ with $r = 5$.
The optimization yields a robust pulse at the expense of a higher norm: $\|u_\mathrm{r}\| = 713$ compared to $\|u_1\| = 189$. To inspect the robustness of the two pulses, we test them for different values of $\sigma = 1 + n/20$, $n = 0, \ldots, 200$. For each value of σ, the monodomain equation is solved and successful defibrillation is confirmed at t_f and a later time $t = t_\mathrm{f} + 4$, to exclude regeneration of a reentry wave. Fig. 7 shows the norm $\|v(x, t_\mathrm{f} + 4)\|_{L^2(\Omega)}$ over σ. The zero set of the curves corresponds to a successful defibrillation. While u_r defibrillates for all $\sigma \in [1.6, 10]$, the pulse u_1 is found to be successful only for $\sigma \in [2.4, 3]$, and by chance also for $\sigma \in [9.8, 11]$, see Fig. 7.

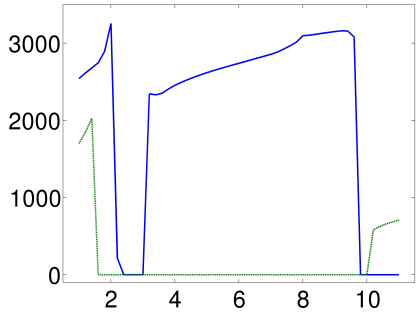

Figure 7: $\|v(t_\mathrm{f} + 4, x)\|_{L^2(\Omega)}$ for different values of σ, both for the optimal pulse u_1 (blue) and the robost optimal control u_r (green, dotted).

8 Conclusion and outlook

It was demonstrated that the choice of cost functional reflecting the system dynamics and incorporating time-optimality for the joint optimization of the shape and the duration of defibrillation pulses is effective for the optimal control of the monodomain equation. Certainly it would be of interest to extended the proposed methods to the bidomain equations, to realistic geometries and to more complex ionic models.

Acknowledgment

The authors gratefully acknowledge the Austrian Science Fund (FWF) for financial support under SFB F32, "Mathematical Optimization and Applications in Biomedical Sciences". The authors thank Konstantin Pieper, TU Munich for many fruitful discussions, especially concerning Section 5.1.

A Optimization algorithm TR-SN

1. Initialize z^0, maximal radius $\Delta_{\max} > 0$, initial radius $0 < \Delta_0 \leq \Delta_{\max}$ and set $n = 0$.
2. Solve state and adjoint equations for z^n , set up gradient $F(z^n)$ from (9) and determine inactive sets $\mathcal{I}^k = \{t \in (0,T) \,|\, u_{\min}(t) < z_k^n(t) < u_{\max}(t)\}$.
3. Compute d from (13) by Steihaug-CG using the L^2-inner product on the inactive set $(\cdot,\cdot)_{\mathcal{I}}$.
4. If Steihaug-CG is fully converged (i.e. [23, (2.3)] is fulfilled), then compute δz according to (14). Otherwise set $\delta z = d$.
5. Calculate $j(P_{\mathrm{ad}}(z^n + \delta z))$ and $\varrho_n := \frac{\varrho^{\mathrm{act}}}{\varrho^{\mathrm{pred}}} = \frac{j(P_{\mathrm{ad}}(z^n)) - j(P_{\mathrm{ad}}(z^n+\delta z))}{-\varphi_{z^n}(\delta z)}$.
6. Update z:
$$z^{n+1} := \begin{cases} z^n + \delta z\,, & \text{if } \varrho_n > \alpha_2 \text{ and } \varrho^{\mathrm{act}} > \epsilon \quad \text{(accept)}\,, \\ z^n\,, & \text{otherwise} \qquad\qquad\qquad\quad \text{(reject)}. \end{cases}$$
7. update radius Δ_n:
$$\Delta_{n+1} = \begin{cases} \min(2\|\delta z\|_{\mathcal{I}}, \Delta_{\max})\,, & \text{if } \varrho_n \in [0.7, 1.3] \quad \text{(model good)} \\ 0.25\Delta_n\,, & \text{elseif } \varrho^{\mathrm{act}} \leq \epsilon \quad \text{(no decrease)} \\ 0.5\|\delta z\|_{\mathcal{I}}\,, & \text{elseif } \varrho_n \notin [0.25, 1.75] \quad \text{(model bad)} \\ \Delta_n\,, & \text{else}\,. \end{cases}$$
8. If stopping criteria are not fulfilled, set $n = n + 1$ and goto 2.

Each Hessian evaluation in 3. is carried out by the following steps.

1. Solve the tangent equation with δz and corresponding states (v^n, w^n) for $u^n = P_{\mathrm{ad}}(z^n)$
$$\begin{aligned} \delta v_t - \nabla\cdot(\bar\sigma_i \nabla \delta v) + I_v(v_n, w_n)\,\delta v + I_w(v_n)\,\delta w &= \sum_{k=1}^{N_{\mathrm{e}}} \chi_{\Omega_{\mathrm{con},k}}(x)\chi_{\mathcal{I}^k}(t)\delta z_k(t)\chi_{(0,T)}(t) && \text{in } Q\,, \\ \delta w_t + G(\delta v, \delta w) &= 0 && \text{in } Q\,, \\ \nu\cdot\bar\sigma_i\nabla\delta v &= 0 && \text{on } \Sigma\,, \\ \delta v(x,0) = 0\,, \quad \delta w(x,0) &= 0 && \text{in } \Omega\,. \end{aligned}$$
2. Solve the second adjoint equation with p^n the adjoint to (v^n, w^n)
$$\begin{aligned} -\delta p_t - \nabla\cdot(\bar\sigma_i\nabla\delta p) + I_v(v_n,w_n)\,\delta p + G_v\,\delta q &= -I_{vv}(v_n)\,p_n\,\delta v - I_{vw}\,p_n\,\delta w && \text{in } Q\,, \\ -\delta q_t + I_w(v_n)\,\delta p + G_w\,\delta q &= -I_{vw}\,p_n\,\delta v && \text{in } Q\,, \\ \nu\cdot\bar\sigma_i\nabla\delta p &= 0 && \text{on } \Sigma\,, \\ \delta p(x,t_{\mathrm{f}}) = \mu\,\delta v\,, \quad \delta q(x,t_{\mathrm{f}}) &= 0 && \text{in } \Omega\,. \end{aligned}$$
3. Evaluate (12) .

B Discretization formulas for state, adjoint and second-order solvers

The space is discretized with a FE-Galerkin method using Lagrange-Q1 elements $\{\varphi_i(x)\,, i = 1,\dots,N_x\}$. Hence, we search for FE-coordinates $v_m := (v_m^i)_{i=1,\dots,N_x}$ with $v(t_m,x) = V^m(x) = \sum_{i=1}^{N_x} v_m^i \varphi_i(x)$ and analogously for w, δv, δw, p, q, δp, δq.
As matrices we define the mass matrix $M := (\int_\Omega \varphi_i\varphi_j \,\mathrm{d}x)_{i,j}$, the negative stiffness matrix $\Delta_\sigma := -(\int_\Omega \nabla\varphi\bar\sigma_i\nabla\varphi_j \,\mathrm{d}x)_{i,j}$ and the Jacobian $J_{m,n} = \left(\int_\Omega \frac{\partial I}{\partial v}(v_m(x),w_n(x))\varphi_i(x)\varphi_j(x)\,\mathrm{d}x\right)_{i,j}$. Further we define the vectors $\vec\chi_k := (\int_{\Omega_{\mathrm{con},k}} \varphi_j \,\mathrm{d}x)_j$ and $I_{m,n} := (\int_\Omega I(v_m(x),w_n(x))\varphi_j(x)\,\mathrm{d}x)_j$. v_0, w_0 are the FE coordinates of the initial states $v_0(x)$, $w_0(x)$. The index m passes through $1,\dots,N$ for primal and tangent equations, and through $1,\dots,N-1$ for adjoint and second adjoint equation.

$$
\begin{aligned}
\text{state:}\quad & [M - \tfrac{\tau_m}{2}\Delta_\sigma]v_m + \tfrac{\tau_m}{2} I_{m,m-1} = [M + \tfrac{\tau_m}{2}\Delta_\sigma]v_{m-1} - \tfrac{\tau_m}{2} I_{m-1,m-1} \\
&\qquad + \tau_m \sum_{k=1}^{N_\mathrm{e}} u_k^m \vec\chi_k \chi_{(0,T)}(t_m), \\
& [1 + \tfrac{\tau_m}{2}G_w]Mw_m = [1 - \tfrac{\tau_m}{2}G_w]Mw_{m-1} - \tfrac{\tau_m}{2}G_v M(v_m + v_{m-1})\,, \\
\text{adj.:}\quad & q_N = 0\,, \qquad [M - \tfrac{\tau_N}{2}\Delta_\sigma + \tfrac{\tau_N}{2}J_{N,N-1}]p_N = \mu M v_N\,, \\
& [M - \tfrac{\tau_m}{2}\Delta_\sigma + \tfrac{\tau_m}{2}J_{m,m-1}]p_m = [M + \tfrac{\tau_{m+1}}{2}\Delta_\sigma - \tfrac{\tau_{m+1}}{2}J_{m,m}]p_{m+1} \\
&\qquad - \frac{G_v}{2} M(\tau_m q_m + \tau_{m+1} q_{m+1}), \\
& [1 + \tfrac{\tau_m}{2}G_w]Mq_m = [1 - \tfrac{\tau_{m+1}}{2}G_w]Mq_{m+1} - \tfrac{\tau_{m+1}}{2}\int_\Omega I_w(V^{m+1}+V^m)\,P^{m+1}\varphi_j \,\mathrm{d}x\,, \\
\text{tang.:}\quad & \delta v_0 = 0\,, \qquad \delta w_0 = 0\,, \\
& [M - \tfrac{\tau_m}{2}\Delta_\sigma + \tfrac{\tau_m}{2}J_{m,m-1}]\delta v_m = [M + \tfrac{\tau_m}{2}\Delta_\sigma - \tfrac{\tau_m}{2}J_{m-1,m-1}]\delta v_{m-1} \\
&\qquad - \frac{\tau_m}{2}\int_\Omega I_w(V^m + V^{m-1})\delta W^{m-1}\varphi_j \,\mathrm{d}x + \tau_m \sum_{k=1}^{N_\mathrm{e}} \chi_{\mathcal{I}^k}(t_m)\chi_{(0,T)}(t_m)\delta z_k^m \vec\chi_k\,, \\
& [1 + \tfrac{\tau_m}{2}G_w]M\delta w_m = [1 - \tfrac{\tau_m}{2}G_w]M\delta w_{m-1} - \tfrac{\tau_m}{2}G_v M(\delta v_m + \delta v_{m-1})\,, \\
\text{2nd adj.:}\quad & \delta q_N = 0\,, \\
& [M - \tfrac{\tau_N}{2}\Delta_\sigma + \tfrac{\tau_N}{2}J_{N,N-1}]\delta p_N = -\tfrac{\tau_N}{2}\int_\Omega P^N[I_{vv}(V^N)\delta V^N + I_{vw}\delta W^{N-1}]\varphi_j \,\mathrm{d}x + M\delta v_N\,, \\
& [M - \tfrac{\tau_m}{2}\Delta_\sigma + \tfrac{\tau_m}{2}J_{m,m-1}]\delta p_m = [M + \tfrac{\tau_{m+1}}{2}\Delta_\sigma - \tfrac{\tau_{m+1}}{2}J_{m,m}]\delta p_{m+1} \\
&\qquad - \frac{1}{2}G_v M(\tau_m\delta q_m + \tau_{m+1}\delta q_{m+1}) - \frac{1}{2}\int_\Omega \Big\{\tau_m P^m[I_{vv}(V^m)\delta V^m + I_{vw}\delta W^{m-1}] \\
&\qquad + \tau_{m+1}P^{m+1}[I_{vv}(V^m)\delta V^m + I_{vw}\delta W^m]\Big\}\varphi_j \,\mathrm{d}x\,, \\
& [1 + \tfrac{\tau_m}{2}G_w]M\delta q_m = [1 - \tfrac{\tau_{m+1}}{2}G_w]M\delta q_{m+1} \\
&\qquad - \frac{\tau_{m+1}}{2}\int_\Omega \big[\delta P^{m+1} I_w(V^{m+1}+V^m) + P^{m+1}(\delta V^{m+1}+\delta V^m)I_{vw}\big]\varphi_j \mathrm{d}x.
\end{aligned}
$$

All solves with a pure mass matrix are avoided by directly updating w_m resp. δw_m and by storing Mq_m resp. $M\delta q_m$.

References

[1] W. Bangerth, R. Hartmann, and G. Kanschat. deal.II – a general purpose object oriented finite element library. *ACM Trans. Math. Softw.*, 33(4):24/1–24/27, 2007.

[2] R. Becker, D. Meidner, and B. Vexler. Efficient numerical solution of parabolic optimization problems by finite element methods. *Optim. Methods Softw.*, 22(5):813–833, 2007.

[3] A. Borzì and R. Griesse. Distributed optimal control of lambda-omega systems. *J. Numer. Math.*, 14(1): 17–40, 2006.

[4] Y. Bourgault, Y. Coudière, and C. Pierre. Existence and uniqueness of the solution for the bidomain model used in cardiac electrophysiology. *Nonlinear Anal. Real World Appl.*, 10:458–482, 2009.

[5] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2009.

[6] E. Casas, C. Ryll, and F. Tröltzsch. Sparse optimal control of the Schlögl and FitzHugh-Nagumo systems. *Comput. Meth. in Appl. Math.*, 13(4):415–442, 2013.

[7] P. Constantin and C. Foias. *Navier-Stokes Equations*. The University of Chicago Press, 1988.

[8] K. Eriksson, D. Estep, P. Hansbo, and C. Johnson. *Computational Differential Equations*. Cambridge University Press, 1996.

[9] P. Franzone, P. Deuflhard, B. Erdmann, J. Lang, and L. Pavarino. Adaptivity in space and time for reaction-diffusion systems in electrocardiology. *SIAM J. Sci. Comput.*, 28(3):942–962, 2006.

[10] S. Götschel, N. Chamakuri, K. Kunisch, and M. Weiser. Lossy compression in optimal control of cardiac defibrillation. *J. Sci. Comput.*, 1–25, Springer US, 2013.

[11] M. Hinze and K. Kunisch. Second order methods for optimal control of time-dependent fluid flow. *SIAM J. Control Optim.*, 40(3):925–946, 2001.

[12] K. Ito and K. Kunisch. *Lagrange Multiplier Approach to Variational Problems and Applications*. Advances in Design and Control, SIAM, Philadelphia, 2008.

[13] K. Ito and K. Kunisch. Semismooth Newton methods for time-optimal control for a class of ODEs *SIAM J. Control Optim.*, 48(6):3997–4013, 2010.

[14] J. Keener and J. Sneyd. *Mathematical Physiology II: Systems Physiology*. Springer New York, 2009.

[15] K. Kunisch, K. Pieper, and A. Rund. Time-optimal control for the monodomain equations — a monolithic approach. In preparation, 2014.

[16] K. Kunisch and M. Wagner. Optimal control of the bidomain system (I): the monodomain approximation with the Rogers-McCulloch model. *Nonlinear Anal. Real World Appl.*, 13(4):1525–1550, 2011.

[17] J.D. Murray. *Mathematical Biology: I. An Introduction*. Springer New York, 2002.

[18] C. Nagaiah, K. Kunisch, and G. Plank. Numerical solution for optimal control of the reaction-diffusion equations in cardiac electrophysiology. *Comput. Optim. Appl.*, 49:149–178, 2011.

[19] C. Nagaiah, K. Kunisch, and G. Plank. Optimal control approach to termination of re-entry waves in cardiac electrophysiology. *J. Math. Biol.*, 67(2):359–388, 2013.

[20] 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 Trans. Biomed. Eng.*, 53(12):2425–2435, 2006.

[21] M. Puri, K.C. Chapalamadugu, A. Miranda, S. Gelot, W. Moreno, P.C. Adithya, C. Law, and S.M. Tipparaju. Integrated approach for smart implantable cardioverter defibrillator (ICD) device with real time ECG monitoring: use of flexible sensors for localized arrhythmia sensing and stimulation. *Front. Physiol.*, 4(300):1–4, 2013.

[22] J.M. Rogers and A.D. McCulloch. A collocation-Galerkin finite element model of cardiac action potential propagation. *IEEE Trans. Biomed. Eng.*, 41(8):743–757, 1994.

[23] T. Steihaug. The conjugate gradient method and trust regions in large scale optimization. *SIAM J. Numer. Anal.*, 20:626–637, 1983.

[24] J. Sundnes, B.F. Nielsen, K.A. Mardal, X. Cai, G.T. Lines, and A. Tveito. On the computational complexity of the bidomain and the monodomain models of electrophysiology. *Ann. Biomed. Eng.*, 34(7):1088–1097, 2006.

[25] F. Tröltzsch. *Optimal Control of Partial Differential Equations.* Graduate Studies in Mathematics 112, American Mathematical Society, Providence, Rhode Island, 2010.

[26] M. Ulbrich. *Semismooth Newton Methods for Variational Inequalities and Constrained Optimization Problems in Function Spaces.* MOS-SIAM Series on Optimization, SIAM, Philadelphia, 2011.

[27] E.J. Vigmond, R. Weber dos Santos, A.J. Prassl, M. Deo, and G. Plank. Solvers for the cardiac bidomain equations. *Prog. Biophys. Mol. Bio.*, 96(1-3):3–18, 2008.