Computational Methods in Applied Mathematics
Vol. 1 (2001), No. 1, pp. 1–29

Sparse optimal control of the Schlögl and FitzHugh-Nagumo systems

Eduardo Casas · Christopher Ryll · Fredi Tröltzsch

Abstract — We investigate the problem of sparse optimal controls for the so-called Schlögl model and the FitzHugh-Nagumo system. In these reaction-diffusion equations, traveling wave fronts occur that can be controlled in different ways. The L^1-norm of the distributed control is included in the objective functional so that optimal controls exhibit effects of sparsity. We prove the differentiability of the control-to-state mapping for both dynamical systems, show the well-posedness of the optimal control problems and derive first-order necessary optimality conditions. Based on them, the sparsity of optimal controls is shown. The theory is illustrated by various numerical examples, where wave fronts or spiral waves are controlled in a desired way.

2010 Mathematical subject classification: 49K20, 49M05, 35K57, 37N25.

Keywords: optimal control; reaction-diffusion system; Schlögl model; FitzHugh-Nagumo system; sparse control; traveling wave fronts; spiral waves.

Dedicated to Prof. Dr. Rolf Grigorieff on the occasion of his 75th birthday

1. Introduction

In this paper, we investigate optimal control problems for two reaction-diffusion problems, namely the so-called Schlögl or Nagumo model and the FitzHugh-Nagumo equations. Under appropriate initial conditions, the uncontrolled solutions of these systems behave like a traveling wave or a spiral wave. It is a natural task to control such wave type solutions in an optimal way. Often, it is desired to apply controls only in small parts of the spatial domain, since it is not always realistic to apply distributed controls in the whole spatial domain. This is a typical situation, where the theory of sparse optimal control can be applied that has recently been studied actively for elliptic partial differential equations. For example [Stadler (2009)], [Wachsmuth and Wachsmuth (2010)], and [Casas et al.(2012)] deal with

This work was supported by DFG in the framework of the Collaborative Research Center SFB 910, project B6. The first author was also supported by Spanish Ministerio de Economía y Competitividad under project MTM2011-22711.

Eduardo Casas
Departmento de Matemática Aplicada y Ciencias de la Computación, E.T.S.I. Industriales y de Telecomunicación, Universidad de Cantabria, Av. Los Castros s/n, 39005 Santander, Spain
E-mail: `eduardo.casas@unican.es`.

Christopher Ryll
Institut für Mathematik, Technische Universität Berlin, D-10623 Berlin, Germany
E-mail: `ryll@math.tu-berlin.de`.

Fredi Tröltzsch
Institut für Mathematik, Technische Universität Berlin, D-10623 Berlin, Germany
E-mail: `troeltzsch@math.tu-berlin.de`.

this topic. In [Casas et al.(2012)], first-order necessary conditions along with second-order necessary and second-order sufficient optimality conditions are derived for sparse optimal controls of semilinear elliptic equations. Moreover, [Casas et al.(2013)] study this problem for a general class of semilinear parabolic equations with monotone nonlinearity.

Our state equations belong to the class of reaction-diffusion equations. They can also be classified as semilinear parabolic equations or systems, but they are of non-monotone type. Therefore, the theory of existence, uniqueness and regularity of associated solutions is more delicate than for equations of monotone type. Existence and uniqueness theorems for the Schlögl and FitzHugh-Nagumo system have already been proved by several authors. In particular, we mention the paper [Jackson (1990)] on the FitzHugh-Nagumo system with nonsmooth data. We also refer to the books [Smoller (1994)] and [Murray (1993)].

The optimal control theory for such equations requires additional investigations. To our knowledge, [Brandão et al.(2008)] published the first paper on the optimal control theory for the FitzHugh-Nagumo system. They derived first-order necessary optimality conditions for optimal controls and discuss also aspects of controllability. Their approach is based on an existence theorem for the FitzHugh-Nagumo equations that is proved by the Leray-Schauder principle in an L^2-setting for domains with smooth boundary. The authors do not discuss the differentiability of the control-to-state mapping. Moreover, we mention the paper [Kunisch and Wang (2012)] on time-optimal control of a linear version of the FitzHugh-Nagumo system. The authors solve the existence problem by a semigroup technique and derive optimality conditions.

Further contributions to the optimal control of reaction-diffusion equations that admit wave-type solutions were published by [Borzì and Griesse (2006)] and in the sequence of papers [Kunisch and Wagner (I)] - [Kunisch and Wagner (III)].

Our paper contains the following novelties: In the first part, we prove existence and uniqueness of a solution to the FitzHugh-Nagumo equations in a new way that also works in Lipschitz domains. By an L^∞-approach, we show the second-order Fréchet differentiability of the control-to-state mapping. Based on this foundation, we derive first-order necessary optimality conditions for sparse optimal controls. Here, we mainly follow the lines of [Casas et al.(2012)], where the sparsity of optimal controls was derived in the elliptic case. By the presence of the L^1-norm of the control, the objective function is not differentiable. This needs special techniques, which were developed in [Casas et al.(2012)] and can be extended to our parabolic case in a direct way. We do not exploit the proved second-order differentiability to set up second-order optimality conditions. However, we state the associated differentiability as a basis for later applications to second-order optimization methods.

In the second part of our paper we study various numerical examples in 1D and 2D spatial domains. We control traveling wave fronts as solutions to the Schlögl-equation. In the case of the FitzHugh-Nagumo-system, spiral waves occur. Controlling such patterns is geometrically impressive but numerically fairly demanding. To our best knowledge, sparse optimal controls for such equations were not yet discussed in literature. However, there is a rich literature on feedback control problems in the community of Physics. We refer exemplarily to [Zykov and Engel(2004)], [Breuer (2006)], [Mantel et al.(1996)], or [Schrader et al.(1995)], and to the survey volume [Schöll and Schuster (2007)].

We consider optimal control problems for the following two reaction-diffusion equations: The first one, in Physics known as *Schlögl model* and in Neurology as *Nagumo equation*, has the form

$$\begin{aligned}
\frac{\partial}{\partial t}y(x,t) - \Delta y(x,t) + R(y(x,t)) &= u(x,t) \quad \text{in } Q_T \\
\partial_\nu y(x,t) &= 0 \quad \text{in } \Sigma_T \\
y(x,0) &= y_0(x) \quad \text{in } \Omega,
\end{aligned} \tag{1}$$

where R is the cubic polynomial

$$R(y) = k\,(y-y_1)(y-y_2)(y-y_3)$$

with given real numbers $k > 0$ and $y_1 < y_2 < y_3$.

In this setting, Ω is a bounded open Lipschitz domain of $\mathbb{R}^n$, $n \in \{1,2,3\}$, $T > 0$ is a fixed time, and we use the notation $Q_T := \Omega \times (0,T)$ and $\Sigma_T := \partial\Omega \times (0,T)$. Moreover, an initial function $y_0 \in L^\infty(\Omega)$ is given. By ν and ∂_ν, we denote the outward unit normal vector and the associated outward normal derivative on $\partial\Omega$, respectively.

The second reaction-diffusion equation, the *FitzHugh-Nagumo system*, is given by

$$\begin{aligned}
\frac{\partial}{\partial t}y(x,t) - \Delta y(x,t) + R(y(x,t)) + z(x,t) &= u(x,t) \quad \text{in } Q_T \\
\partial_\nu y(x,t) &= 0 \quad \text{in } \Sigma_T \\
y(x,0) &= y_0(x) \quad \text{in } \Omega \\
\frac{\partial}{\partial t}z(x,t) + \beta\, z(x,t) - \gamma\, y(x,t) + \delta &= 0 \quad \text{in } Q_T \\
z(x,0) &= z_0(x) \quad \text{in } \Omega.
\end{aligned} \tag{2}$$

Here, more real constants β, γ, δ, and initial data $z_0 \in L^\infty(\Omega)$ are given.

In this system, the partial differential equation for y is said to be the *activator equation*, while the one for z is called the *inhibitor equation*. The function y is the *state* that is to be controlled, while the inhibitor z has only some auxiliary character with respect to the control.

To handle the analysis for both equations at once, we shall discuss the more general model

$$\begin{aligned}
\frac{\partial}{\partial t}y(x,t) - \Delta y(x,t) + R(y(x,t)) + \alpha\, z(x,t) &= u(x,t) \quad \text{in } Q_T \\
\partial_\nu y(x,t) &= 0 \quad \text{in } \Sigma_T \\
y(x,0) &= y_0(x) \quad \text{in } \Omega \\
\frac{\partial}{\partial t}z(x,t) + \beta\, z(x,t) - \gamma\, y(x,t) + \delta &= 0 \quad \text{in } Q_T \\
z(x,0) &= z_0(x) \quad \text{in } \Omega,
\end{aligned} \tag{3}$$

where α is a real number. For the choice $\alpha = 0$, both equations decouple and the state function y has to solve the Schlögl equation. Here, the inhibitor equation is meaningless. For $\alpha = 1$, the FitzHugh-Nagumo system is obtained. These values $\alpha \in \{0,\ 1\}$ are the values of our interest, but the analysis for the system (3) is also true for arbitrary real α. We shall consider optimal control problems, where the objective functional

$$f(y_u, z_u, u) = I(u) + \mu\, j(u) =: J(u) \tag{4}$$

is to be minimized, where

$$I(u) := \frac{1}{2} \int_0^T \int_\Omega c_Q^Y(x,t)\,(y_u(x,t) - y_Q(x,t))^2 + c_Q^Z(x,t)\,(z_u(x,t) - z_Q(x,t))^2\,dx\,dt$$

$$+ \frac{1}{2} \int_\Omega c_T^Y(x)\,(y_u(x,T) - y_T(x))^2 + c_T^Z(x)\,(z_u(x,T) - z_T(x))^2\,dx + \frac{\kappa}{2} \int_0^T \int_\Omega u^2(x,t)\,dx\,dt,$$

$$j(u) := \int_0^T \int_\Omega |u(x,t)|\,dx\,dt,$$

and the state (y_u, z_u) is the unique solution of equation (3) for the given control u. Existence and uniqueness of such a solution are proved in Theorem 2.1, respectively in Corollary 2.1. Moreover, the functions c_T^Y, $c_T^Z \in L^\infty(\Omega)$, c_Q^Y, $c_Q^Z \in L^\infty(Q_T)$ and the constants κ, μ are non-negative weights, and y_T, $z_T \in L^\infty(\Omega)$ and y_Q, $z_Q \in L^\infty(Q_T)$ are given target functions. In control problems for the Schlögl model, we fix $c_Q^Z = 0$ and $c_T^Z = 0$, since only the function y is of interest.

The *control functions* u are taken from the set of *admissible controls* defined by

$$\mathcal{U}_{ad} := \{u \in L^\infty(Q_T) \mid u(x,t) \in [a,b] \text{ for a.a. } (x,t) \in Q_T\} \tag{5}$$

with real constants $a \leqslant 0 < b$. We also allow for the case $a = 0$, because in some applications $u = 0$ (no control) stands for the smallest value. Then the objective functional is differentiable so that the later use of a subdifferential is not needed in this case.

The optimal control problems for the Schlögl model and the FitzHugh-Nagumo system are both covered by the following one:

Minimize the objective functional

$$\min_{u \in \mathcal{U}_{ad}} f(y_u, z_u, u), \tag{6}$$

where the pair (y_u, z_u) denotes the solution of the general system (3) that is associated to the control u. To make this well defined, we have to show that to each $u \in \mathcal{U}_{ad}$ there exists a unique solution (y_u, z_u) of (3) and that the mapping $u \mapsto (y_u, z_u)$ is continuous. This is the subject of the next section.

2. Well-posedness of the state equation

To prove the existence and uniqueness of a solution (y,z) of (3), we proceed as follows: By eliminating the linear ordinary differential equation for z, we transform (3) to an integro-differential equation for y. Next, following [Engel et al.(2013)], we substitute $y = e^{\eta t} v$ with sufficiently large $\eta > 0$ to get a new equation with monotone nonlinearity. Finally, based on standard estimates, we invoke the Schauder fixed point theorem to show the existence of a solution. The uniqueness follows then by standard energy estimates.

2.1. Transformation of the state equation

Let us first perform the transformation of (3) to an integro-differential equation. The last two equations of (3) can be resolved by

$$z(x,t) = e^{-\beta t} z_0(x) + \int_0^t e^{-\beta(t-s)}(\gamma\, y(x,s) - \delta) ds = e^{-\beta t} z_0(x) + \frac{\delta}{\beta}\big(e^{-\beta t} - 1\big) + (K\, y)(x,t), \quad (7)$$

where the integral operator K is defined by

$$(K\, y)(x,t) = \int_0^t \gamma\, e^{-\beta(t-s)} y(x,s)\, ds.$$

We have $K \in \mathcal{L}(L^p(Q_T))$ for all $1 \leqslant p \leqslant \infty$, where $\mathcal{L}(X)$ denotes the space of all linear and continuous operators acting in a Banach space X. Inserting (7) in the first equation of (3), the integro-differential equation

$$\frac{\partial}{\partial t} y - \Delta y + R(y) + \alpha\, K y = u - \alpha \left(e^{-\beta t} z_0 + \frac{\delta}{\beta}\big(e^{-\beta t} - 1\big) \right) \quad \text{in } Q_T$$

is obtained. Since R is not monotone, we follow [Engel et al.(2013)] and substitute

$$y(x,t) := e^{\eta\, t}\, v(x,t)$$

with a sufficiently large real parameter η. This leads to a new equation for v,

$$\frac{\partial}{\partial t} v - \Delta v + e^{-\eta\, t} R\big(e^{\eta\, t} v\big) + \eta v + \alpha\, K_\eta v = e^{-\eta\, t}\left(u - \alpha \left(e^{-\beta t} z_0 + \frac{\delta}{\beta}\big(e^{-\beta t} - 1\big) \right)\right) \text{ in } Q_T, \quad (8)$$

with the given initial and boundary conditions, where the operator K_η is defined by

$$(K_\eta\, v)(x,t) := \int_0^t \gamma\; ds e^{-(\beta+\eta)(t-s)}\, v(x,s).$$

We discuss equation (8) for convenience with simplified right-hand side u, i.e. in the form

$$\begin{aligned} \frac{\partial}{\partial t} v - \Delta v + e^{-\eta\, t} R(e^{\eta\, t} v) + \eta v + \alpha\, K_\eta v &= u \quad \text{in } Q_T \\ \partial_\nu v &= 0 \quad \text{in } \Sigma_T \\ v(0) &= y_0 \quad \text{in } \Omega. \end{aligned} \qquad (9)$$

2.2. A priori estimates

Next, preparing a fixed point argument, we derive a bound for any weak solution $y \in W(0,T) \cap L^\infty(Q_T)$ of (9), where

$$W(0,T) = \left\{ y \in L^2\big(0,T; H^1(\Omega)\big) \;\middle|\; \frac{\partial y}{\partial t} \in L^2\big(0,T; H^1(\Omega)^*\big) \right\}.$$

A function $y \in W(0,T) \cap L^\infty(Q_T)$ is said to be a *weak solution* of (9), if

$$\int_0^T \langle v'(t)\,,\, \varphi(t)\rangle\, dt + \int_0^T\!\!\int_\Omega \left\{\nabla v(t)\cdot \nabla\varphi(t) + [e^{-\eta t}R(e^{\eta t}v(t)) + \eta v(t)]\,\varphi(t)\right\}\, dx\, dt$$
$$+ \int_0^T\!\!\int_\Omega \alpha\,(K_\eta v)(t)\,\varphi(t)\, dx\, dt = \int_0^T\!\!\int_\Omega u(t)\,\varphi(t)\, dx\, dt$$

holds for all $\varphi \in W(0,T)$ and the initial condition $v(0) = y_0$ is satisfied in Ω. Here, $\langle\cdot\,,\,\cdot\rangle$ denotes the pairing between $H^1(\Omega)$ and $H^1(\Omega)^*$.

To this aim, let us first estimate the norm of the operator K_η. We get

$$\left|\int_0^t e^{-(\beta+\eta)(t-s)}\, v(x,s)\, ds\right| \leqslant \left(\int_0^t e^{-2(\beta+\eta)(t-s)}\, ds\right)^{\frac{1}{2}} \left(\int_0^t v^2(x,s)\, ds\right)^{\frac{1}{2}}$$
$$= \left(\frac{1}{2(\beta+\eta)}\big(1 - e^{-2(\beta+\eta)t}\big)\right)^{\frac{1}{2}} \left(\int_0^t v^2(x,s)\, ds\right)^{\frac{1}{2}} \leqslant \frac{1}{\sqrt{2(\beta+\eta)}} \left(\int_0^t v^2(x,s)\, ds\right)^{\frac{1}{2}}$$

provided that $\eta > |\beta|$. Therefore, we have that

$$\begin{aligned}\int_0^T\!\!\int_\Omega (K_\eta\, v)^2\, dx\, dt &= \gamma^2 \int_0^T\!\!\int_\Omega \left|\int_0^t e^{-(\beta+\eta)(t-s)}\, v(x,s)\, ds\right|^2 dx\, dt \\ \leqslant \frac{\gamma^2}{2(\beta+\eta)} \int_0^T\!\!\int_\Omega &\left(\int_0^T v^2(x,s)\, ds\right) dx\, dt \leqslant \frac{T\,\gamma^2}{2(\beta+\eta)}\, \|v\|^2_{L^2(Q_T)}\end{aligned}$$

for $\eta > |\beta|$. It follows that

$$\|K_\eta\|_{\mathcal{L}(L^2(Q_T))} \leqslant |\gamma| \sqrt{\frac{T}{2(\beta+\eta)}} \qquad \forall \eta > |\beta|\,, \tag{10}$$

hence K_η tends to zero in $\mathcal{L}(L^2(Q_T))$ as $\eta \to \infty$.

Let us mention that the derivative of R is a convex quadratic polynomial, hence its derivative is bounded from below. There is some constant c_R such that

$$R'(v) \geqslant c_R \quad \forall v \in \mathbb{R}. \tag{11}$$

Lemma 2.1 (L^2-a-priori estimate)**.** *There exist positive constants C_2 and η_0 with the following properties: If $\eta \geqslant \eta_0$ and $v \in W(0,T) \cap L^\infty(Q_T)$ is any weak solution of the system (9), then there holds for all $u \in L^2(Q_T)$ and $y_0 \in L^2(\Omega)$*

$$\|v\|_{L^2(0,T;H^1(\Omega))} \leqslant C_2 \Big(\|u\|_{L^2(Q_T)} + \|y_0\|_{L^2(\Omega)} + |R(0)|\Big). \tag{12}$$

Proof. Define $R_\eta(t,v) := e^{-\eta t} R(e^{\eta t} v) + \frac{\eta}{3}\, v$. If $\frac{\eta}{3} \geqslant c_R$, then

$$\frac{\partial}{\partial v} R_\eta(t,v) \geqslant 0 \quad \forall v, t \in \mathbb{R}. \tag{13}$$

We write the parabolic PDE of (9) in the form

$$\frac{\partial}{\partial t} v - \Delta v + \underbrace{e^{-\eta t} R\big(e^{\eta t} v\big) + \frac{\eta}{3} v}_{R_\eta(t,v)} + \frac{\eta}{3} v + \Big(\frac{\eta}{3} v + \alpha\, K_\eta v\Big) = u. \tag{14}$$

The first $\eta v/3$ is added to the term with R to get a monotone function R_η, the second one contributes to an $L^2(Q_T)$-estimate for the solution v, and the third $\eta v/3$ is to compensate the operator K_η.

Next, we subtract $R_\eta(t,0) = e^{-\eta t} R(0)$ from both sides of (14) and test this equation by v, associated with an integration over Q_T. After integrating by parts w.r.t. to t in the first item and w.r.t. to x in the second, we obtain

$$\frac{1}{2}\, \|v(T)\|^2_{L^2(\Omega)} + \int_0^T\!\!\int_\Omega (|\nabla v|^2 + \frac{\eta}{3}\, v^2)\, dx\, dt + \int_0^T\!\!\int_\Omega (R_\eta(t,v) - R_\eta(t,0))(v-0)\, dx\, dt$$
$$+ \left(\frac{\eta}{3} - |\alpha\,\gamma| \sqrt{\frac{T}{2(\beta+\eta)}}\right) \|v\|^2_{L^2(Q_T)} \;\leqslant\; \int_0^T\!\!\int_\Omega |u - R_\eta(t,0)|\; |v|\; dx\, dt + \frac{1}{2}\, \|v(0)\|^2_{L^2(\Omega)}\,.$$

The first and second terms in the left-hand side are obviously non-negative. The third is non-negative by monotonicity for $\eta \geqslant |c_R|$, while the fourth is non-negative provided that

$$\eta \geqslant 3\, |\alpha\,\gamma| \sqrt{\frac{T}{2(\beta+\eta)}}.$$

We define

$$\eta_0 = \max\left\{ |c_R|\,,\; 3\, |\alpha\,\gamma|\, \sqrt{T},\; \frac{1}{2} - \beta \right\};$$

then we have $2(\beta+\eta) \geqslant 1$ and

$$\eta \geqslant 3\, |\alpha\,\gamma|\, \sqrt{T} \geqslant 3\, |\alpha\,\gamma| \sqrt{\frac{T}{2(\beta+\eta)}} \qquad \forall \eta \geqslant \eta_0.$$

Young's inequality yields for all $\eta \geqslant \eta_0$

$$\int_0^T\!\!\int_\Omega \Big(|\nabla v|^2 + \frac{\eta}{3} v^2\Big)\, dx\, dt \leqslant c\Big(\big\|u - e^{-\eta\,\cdot}\, R(0)\big\|^2_{L^2(Q_T)} + \|v(0)\|^2_{L^2(\Omega)}\Big).$$

An application of the triangle inequality and $e^{-\eta t} \leqslant 1$ finally yield the estimate (12). □

Let us complement this L^2-estimate by an L^∞-estimate. Obviously, the operator K_η maps continuously $L^2(0,T;H^1(\Omega))$ into $C([0,T],H^1(\Omega))$ and we easily verify

$$\left\| \int_0^t e^{-(\beta+\eta)(t-s)} v(x,s)\,ds \right\|_{H^1(\Omega)} \leqslant \int_0^t e^{-(\beta+\eta)(t-s)} \|v(s)\|_{H^1(\Omega)}\,ds \leqslant \frac{1}{\sqrt{2(\beta+\eta)}} \|v\|_{L^2(0,T;H^1(\Omega))}.$$

The continuous embedding of $H^1(\Omega)$ in $L^6(\Omega)$ for $n \leqslant 3$ yields

$$\|K_\eta v\|_{L^6(Q_T)} \leqslant C\, \|K_\eta v\|_{C([0,T],L^6(\Omega))} \leqslant C\, \|K_\eta v\|_{C([0,T],H^1(\Omega))} \leqslant \frac{C}{\sqrt{2(\beta+\eta)}} \|v\|_{L^2(0,T;H^1(\Omega))}. \tag{15}$$

Assume now that u belongs to $L^p(Q_T)$ with $p > 5/2$ and set $q := \min\{p,6\}$. In (14), we shift the term $\alpha\, K_\eta v$ to the right hand side and consider the associated semilinear equation

$$\frac{\partial}{\partial t} v - \Delta v + R_\eta(t,v) + \frac{2}{3}\eta v = u - \alpha\, K_\eta v$$

subject to the given initial and boundary conditions $v(0) = y_0$, $\partial_\nu v = 0$. The nonlinearity R_η is monotone increasing, hence we can invoke known L^∞-estimates for semilinear parabolic equations for the given q, cf. the treatment of quasilinear equations in [Ladyzhenskaya et al.(1968)], or the discussion of the semilinear case in [Casas (1997)], or [Tröltzsch (2010)]. We obtain from (15) and (12) with a generic constant $c > 0$

$$\begin{aligned}
\|v\|_{L^\infty(Q_T)} &\leqslant c\Big(\|u - \alpha\, K_\eta v - R_\eta(\cdot,0)\|_{L^q(Q_T)} + \|y_0\|_{L^\infty(\Omega)}\Big) \\
&\leqslant c\Big(\|u\|_{L^p(Q_T)} + |\alpha|\ \|K_\eta v\|_{L^6(Q_T)} + |R(0)| + \|y_0\|_{L^\infty(\Omega)}\Big) \\
&\leqslant c\Big(\|u\|_{L^p(Q_T)} + |R(0)| + \|y_0\|_{L^\infty(\Omega)}\Big) + \frac{c\ |\alpha|}{\sqrt{2(\beta+\eta)}} \|v\|_{L^2(0,T;H^1(\Omega))} \\
&\leqslant c\Big(\|u\|_{L^p(Q_T)} + |R(0)| + \|y_0\|_{L^\infty(\Omega)}\Big) + \frac{c\ |\alpha|\ C_2}{\sqrt{2(\beta+\eta)}}\Big(\|u\|_{L^2(Q_T)} + \|y_0\|_{L^2(\Omega)} + |R(0)|\Big) \\
&\leqslant c\left(1 + \frac{1}{\sqrt{2(\beta+\eta)}}\right)\Big(\|u\|_{L^p(Q_T)} + \|y_0\|_{L^\infty(\Omega)} + |R(0)|\Big) \\
&\leqslant 2c\Big(\|u\|_{L^p(Q_T)} + \|y_0\|_{L^\infty(\Omega)} + |R(0)|\Big)
\end{aligned}$$

provided that $\eta \geqslant \eta_0$. In this way, we have proved the following result:

Lemma 2.2 (L^∞-a-priori estimate)**.** *Assume $u \in L^p(Q_T)$ with $p > 5/2$ and $y_0 \in L^\infty(\Omega)$. If $\eta \geqslant \eta_0$ and $v \in W(0,T) \cap L^\infty(Q_T)$ is any weak solution to* (9)*, then there holds*

$$\|v\|_{L^\infty(Q_T)} \leqslant C_\infty \Big(\|u\|_{L^p(Q_T)} + \|y_0\|_{L^\infty(\Omega)} + |R(0)|\Big) \tag{16}$$

with some constant $C_\infty > 0$ that does not depend on α, η, y_0, u, and R.

2.3. Solvability of the state equation

Now we keep the given control u, together with y_0, fixed and set

$$M_\infty := C_\infty \Big(\|u\|_{L^p(Q_T)} + \|y_0\|_{L^\infty(\Omega)} + |R(0)|\Big).$$

By M_∞, we define the following auxiliary function cutting off R_η:

$$\hat{R}_\eta(t,v) = \begin{cases} R_\eta(t, M_\infty) & \text{if } \ v \ \geqslant \ M_\infty \\ R_\eta(t,v) & \text{if } \ |v| \ < \ M_\infty \\ R_\eta(t,-M_\infty) & \text{if } \ v \ \leqslant \ -M_\infty. \end{cases}$$

Theorem 2.1 (Existence and uniqueness). *For all $\eta \geqslant \eta_0$, $u \in L^p(Q)$ with $p > 5/2$, and $y_0 \in L^\infty(\Omega)$, the integro-differential system (9) has a unique solution $v \in W(0,T) \cap L^\infty(Q_T) \cap C\big(\overline{\Omega} \times (0,T]\big)$. There is a constant $C_\infty > 0$ such that*

$$\|v\|_{L^\infty(Q_T)} + \|v\|_{W(0,T)} \leqslant C_\infty \Big(\|u\|_{L^p(Q_T)} + \|y_0\|_{L^\infty(\Omega)} + |R(0)| \Big).$$

If y_0 is continuous in $\overline{\Omega}$, then the solution v belongs to $C\big(\overline{Q_T}\big)$.

Proof. (i)*Existence of a solution.* For given $w \in L^2(Q_T)$, we consider the equation

$$\frac{\partial}{\partial t} v - \Delta v + \hat{R}_\eta(t,v) + \frac{2}{3}\eta\, v = u - \alpha\, K_\eta w \tag{17}$$

subject to the given initial and boundary conditions $v(\cdot,0) = y_0$ and $\partial_\nu v = 0$. Though the right-hand side does possibly not belong to $L^p(Q_T)$ with $p > 5/2$, there is a unique solution of this system in $W(0,T)$, because the function $(x,t) \mapsto \hat{R}_\eta(t, v(x,t))$ is bounded in $L^\infty(Q_T)$ independently of $v \in L^2(Q_T)$. Therefore, the existence of a solution v follows by a simple application of Schauder's theorem. The uniqueness is an immediate consequence of the monotonicity of $\hat{R}$ w.r. to v.

Let us denote by F the mapping $F : w \mapsto v$ acting in $L^2(Q_T)$.

The a-priori estimate (12) for solutions of equation (9) holds in particular for $\alpha = 0$, hence it can be applied to equation (17), too. We define

$$M := C_2 \Big(\|u\|_{L^2(Q_T)} + \|y_0\|_{L^2(\Omega)} + |R(0)| \Big)$$

and assume that $\|w\|_{L^2(Q_T)} \leqslant 2\,M$. Then we obtain from (12), applied to the right-hand side with u substituted by $u - \alpha\, K_\eta w$,

$$\begin{aligned} \|F\,w\|_{L^2(Q_T)} &= \|v\|_{L^2(Q_T)} \leqslant \|v\|_{L^2(0,T;H^1(\Omega))} \\ &\leqslant C_2 \Big(\|u\|_{L^2(Q_T)} + |\alpha|\ \|K_\eta w\|_{L^2(Q_T)} + \|y_0\|_{L^2(\Omega)} + |R(0)| \Big) \\ &\leqslant M + C_2\ |\alpha|\ \|K_\eta w\|_{L^2(Q_T)} \leqslant M + \frac{C_2\ |\alpha|\ C}{\sqrt{2(\beta+\eta)}}\, \|w\|_{L^2(Q_T)} \leqslant 2\,M, \end{aligned}$$

if η is sufficiently large. Therefore, F maps $B_{2M}(0)$, the closed ball of $L^2(Q_T)$ around zero with radius $2\,M$, into itself. Moreover, considering equation (17) in the form

$$\frac{\partial}{\partial t} v - \Delta v + \frac{2}{3}\eta\, v = u - \alpha\, K_\eta w - \hat{R}_\eta(t,v)$$

we obtain from standard estimates for linear parabolic equations that

$$\|F\,w\|_{W(0,T)} \leqslant c$$

holds with some constant c for all w in $B_{2M}(0)$. Notice that $\hat{R}_\eta$ is uniformly bounded in $L^\infty(Q_T)$. By Aubin's Lemma, bounded sets of $W(0,T)$ are relatively compact in $L^2(Q_T)$, hence the mapping F is compact. By Schauder's theorem, F has a fixed point in $B_{2M}(0)$; this is a solution of (9) when R_η is replaced by $\hat{R}_\eta$.

(ii) *Uniqueness of the solution.* Suppose that v_1 and v_2 are solutions of (9) and set $v := v_1 - v_2$. Subtracting the associated equations and applying the mean value theorem to the appearing difference $\hat{R}_\eta(t, v_1) - \hat{R}_\eta(t, v_2)$, we see that v solves

$$\frac{\partial}{\partial t} v - \Delta v + \left(\frac{\partial}{\partial v}\hat{R}_\eta(t, v_\vartheta) + \frac{2}{3}\eta\right) v + \alpha\, K_\eta v = 0$$

subject to homogeneous initial and boundary conditions, where $v_\vartheta = v_1 + \vartheta\,(v_2 - v_1)$ with some measurable ϑ taking values in $(0,1)$. This is a linear equation with non-negative coefficient $\frac{\partial}{\partial v}\hat{R}_\eta(t, v_\vartheta) + \frac{2}{3}\eta$. Applying the same technique as in the proof of Lemma 2.1, we find a constant $\tilde{C}_2$ such that $\|v\|_{L^2(0,T;H^1(\Omega))} \leqslant \tilde{C}_2\, \|0\|_{L^2(Q_T)} = 0$, hence $v = v_1 - v_2 = 0$ showing the uniqueness.

(iii) *The solution v obeys* (9). By Lemma 2.2, the solution v satisfies the L^∞-estimate (16) provided that η is taken sufficiently large. In this case, $\hat{R}_\eta(t, v) = R_\eta(t, v)$ is satisfied so that v is a solution of (9).

(iv) *Continuity properties of v.* As a solution to (17) with bounded initial function and right-hand side in $L^p(Q_T)$, v belongs to the space $L^\infty(Q_T) \cap C(\bar{\Omega} \times (0,T])$, we refer to [Raymond and Zidani (1999)]. If v_0 is even continuous in $\bar{\Omega}$, then $v \in C(\bar{\Omega} \times [0,T]) = C(\overline{Q_T})$ follows from [Casas (1997)] and the references therein. □

Corollary 2.1. *For all $u \in L^p(Q_T)$ with $p > 5/2$, $y_0 \in L^\infty(\Omega)$, $z_0 \in L^\infty(\Omega)$ and all $\alpha \in \mathbb{R}$, the equation* (3) *has a unique solution $(y,z) \in \big(W(0,T) \cap L^\infty(Q_T) \cap C(\overline{\Omega} \times (0,T])\big)^2$. If y_0 and z_0 are continuous on $\overline{\Omega}$, then y and z belong to $C\big(\overline{\Omega} \times [0,T]\big)$.*

Proof. We apply Theorem 2.1 to the equation (8) and obtain a unique solution v. Next, we return to the original quantities y and z by $y = e^{\eta t} v$ and $z = e^{-\beta t} z_0 + \frac{\delta}{\beta}\big(e^{-\beta t} - 1\big) + K_\eta\, v$. □

2.4. Differentiability of the control-to-state mapping

To show the differentiability of the control-to-state mapping $u \mapsto y$, we first prove an analog of Theorem 2.1 for a linear system.

Lemma 2.3. *If η is taken sufficiently large, $c_0 \in L^\infty(Q_T)$ is almost everywhere non-negative, $u \in L^2(Q_T)$ and $y_0 \in L^2(\Omega)$, then the linear integro-differential system*

$$\begin{aligned} \frac{\partial}{\partial t} v - \Delta v + c_0(x,t)\, v + \eta v + \alpha\, K_\eta\, v &= u && \text{in } Q_T \\ \partial_\nu v &= 0 && \text{in } \Sigma_T \\ v(x,0) &= y_0(x) && \text{in } \Omega \end{aligned} \tag{18}$$

has a unique solution $v \in W(0,T)$. There is $C_2 > 0$ depending neither on y_0 nor on c_0 such that

$$\|v\|_{W(0,T)} \leqslant C_2 \Big(\|u\|_{L^2(Q_T)} + \|y_0\|_{L^2(\Omega)}\Big) \qquad \forall u \in L^2(Q_T),\ y_0 \in L^2(\Omega). \tag{19}$$

Proof. The result can be shown completely analogously to Lemma 2.1 and Theorem 2.1 by formally substituting $c_0(x,t)v$ for $R_\eta(t,v)$. If η is taken sufficiently large, then the term $\alpha\, K_\eta v$ can be absorbed by ηv in the estimation. Since the equation (18) is linear, a cut-off function like $\hat{R}_\eta$ is not needed. □

Remark 2.1. In the same way, we show for $u \in L^p(Q_T)$, $p > 5/2$, and $y_0 \in L^\infty(\Omega)$, that the solution v of (18) belongs to $L^\infty(Q_T) \cap C\big(\overline{\Omega} \times (0,T]\big)$ and there exists a constant C_∞ such that

$$\|v\|_{L^\infty(Q_T)} \leqslant C_\infty \Big(\|u\|_{L^p(Q_T)} + \|y_0\|_{L^\infty(\Omega)} \Big). \tag{20}$$

We reduce the problem of differentiability of the control-to state mapping for the equation (3) to known results for semilinear parabolic equations with monotone nonlinearity.

Lemma 2.4. *For all $p > 5/2$ and all sufficiently large η, the solution mapping $\mathcal{G}_\eta : u \mapsto v$ for equation (9) is of class C^2 from $L^p(Q_T)$ to $W(0,T) \cap L^\infty(Q_T) \cap C(\overline{\Omega} \times (0,T])$.*

Proof. At first, we consider the semilinear parabolic differential equation of monotone type

$$\begin{aligned} \frac{\partial}{\partial t} v - \Delta v + R_\eta(t,v) + \frac{2}{3}\eta\, v &= u && \text{in } Q_T \\ \partial_\nu v &= 0 && \text{in } \Sigma_T \\ v(x,0) &= y_0(x) && \text{in } \Omega. \end{aligned} \tag{21}$$

For each $u \in L^p(Q_T)$, $y_0 \in L^\infty(\Omega)$ and $\eta \geqslant \eta_0$, this equation has a unique solution $v_u \in V_\infty := W(0,T) \cap L^\infty(Q_T) \cap C\big(\overline{\Omega} \times (0,T]\big)$. By G_η we denote the associated solution mapping,

$$G_\eta : u \mapsto v_u, \qquad G_\eta : L^p(Q_T) \to V_\infty.$$

It is known that G_η is twice continuously Fréchet-differentiable. For this differentiability property and the concrete form of the first- and second-order derivatives, we refer to [Casas et al.(2008)] or to Thm. 5.9 and Thm. 5.16 in [Tröltzsch (2010)].

By Lemma 2.3 and Remark 2.1, there exists a constant C_∞ such that

$$\|v_u\|_{L^\infty(Q_T)} \leqslant C_\infty \Big(\|u\|_{L^p(Q_T)} + \|y_0\|_{L^\infty(\Omega)} \Big) \qquad \forall u \in L^p(Q_T)$$

holds, no matter how large $\eta \geqslant \eta_0$ or $\frac{\partial}{\partial v} R_\eta$ are. With this prerequisite at hand, we return to the nonlinear equation (9) in the form

$$\begin{aligned} \frac{\partial}{\partial t} v - \Delta v + R_\eta(t,v) + \frac{2}{3}\eta\, v + \alpha\, K_\eta v &= u && \text{in } Q_T \\ \partial_\nu v &= 0 && \text{in } \Sigma_T \\ v(x,0) &= y_0(x) && \text{in } \Omega. \end{aligned} \tag{22}$$

Let us denote the solution mapping for this equation by $\mathcal{G}_\eta$, $\mathcal{G}_\eta : L^p(Q_T) \to V_\infty$; then we have $v = \mathcal{G}_\eta(u)$. Obviously, v solves (22) if and only if, using the mapping G_η for (21),

$$v - G_\eta(u - \alpha\, K_\eta\, v) =: \mathcal{F}(v,u) = 0. \tag{23}$$

Suppose that the pair (v_0, u_0) is a solution to (23). We know that G_η is of class C^2, hence also the mapping $(v,u) \mapsto \mathcal{F}(v,u)$, where $\mathcal{F}$ defined above is considered for simplicity as $\mathcal{F}$:

$L^\infty(Q_T) \times L^p(Q_T) \to L^\infty(Q_T)$. To apply the implicit function theorem, we need continuous invertibility of the derivative

$$\frac{\partial}{\partial v}\mathcal{F}(v_0, u_0) = I + \alpha\, G_\eta'(u_0 - \alpha\, K_\eta\, v_0)\, K_\eta.$$

The norm of $\|K_\eta\|_{\mathcal{L}(L^\infty(Q_T))}$ tends to zero as $\eta \to \infty$, because it holds

$$\|K_\eta\, v\|_{L^\infty(Q_T)} \leqslant \frac{c}{\sqrt{2(\beta+\eta)}}\, \|v\|_{L^\infty(Q_T)}\,,$$

hence

$$\left\|\alpha\, G_\eta'(u_0 - \alpha\, K_\eta\, v_0)\, K_\eta\right\|_{L^\infty(Q_T)} < 1$$

holds for all sufficiently large η. Here we exploit that, by Lemma 2.3 with $c_0(x,t) := \frac{\partial}{\partial v} R_\eta(t, v_0(x,t))$, the norm of the operator $G_\eta'(u_0 - \alpha\, K_\eta\, v_0)$ remains bounded for $\eta \to \infty$. Therefore, $\frac{\partial}{\partial v}\mathcal{F}(v_0, u_0)$ is continuously invertible for sufficiently large η. By the implicit function theorem, the mapping $\mathcal{G}_\eta : u \mapsto v_u$ is also of class C^2 from $L^p(Q_T)$ to $L^\infty(Q_T)$ provided that η is sufficiently large.

Let us finally verify this property for the original range V_∞ of $\mathcal{G}_\eta$. We re-write (23) as

$$v = \mathcal{G}_\eta(u) = G_\eta(u - \alpha K_\eta \mathcal{G}_\eta(u)).$$

Since G_η is of class C^2 from $L^p(Q_T)$ to V_∞ and $\mathcal{G}_\eta(u)$ is of class C^2 from $L^p(Q_T)$ to $L^\infty(Q_T)$, we obtain by the chain rule that $\mathcal{G}_\eta$ is also of class C^2 from $L^p(Q_T)$ to V_∞. □

Remark 2.2. Since the function $y \mapsto R(y)$ is of class C^∞, the implicit function theorem yields immediately that $\mathcal{G}_\eta$ is even of class C^∞.

By this result, it is easy to prove our main result with respect to differentiability.

Theorem 2.2 (Differentiability of the control-to-state mapping)**.** *The solution mapping $G : u \mapsto (y_u, z_u)$ associated with the system* (3) *is twice continuously Fréchet differentiable from $L^p(Q_T)$ to $\big(W(0,T) \cap L^\infty(Q_T) \cap C(\overline{\Omega} \times (0,T])\big)^2$. The derivative $(y_h, z_h) := G'(u)h$ is equal to the pair (y,z) solving the system*

$$\begin{array}{rcll}
\dfrac{\partial}{\partial t} y - \Delta y + R'(y_u) y + \alpha\, z &=& h & \mathit{in}\ Q_T \\
\partial_\nu y &=& 0 & \mathit{in}\ \Sigma_T \\
y(x,0) &=& 0 & \mathit{in}\ \Omega \\
\dfrac{\partial}{\partial t} z + \beta\, z - \gamma\, y &=& 0 & \mathit{in}\ Q_T \\
z(x,0) &=& 0 & \mathit{in}\ \Omega.
\end{array} \tag{24}$$

The second derivative $(y_{h_1 h_2}, z_{h_1 h_2}) := G''(u)[h_1, h_2]$ in the directions h_1, $h_2 \in L^p(Q_T)$ is equal to the pair (y,z) that solves the equation

$$\frac{\partial}{\partial t} y - \Delta y + R'(y_u) y + \alpha\, z \;=\; -R''(y_u)\, y_{h_1} y_{h_2} \quad \mathit{in}\ Q_T \tag{25}$$

and the last 4 equations of (24)*. Here, y_{h_i}, $i = 1,2$, are the first components of the derivatives $G'(u)h_i$ defined in* (24)*.*

Proof. We apply Lemma 2.4 to (8). By affine linearity, the associated mapping $u \mapsto v$ is of class C^2, hence also the mapping $u \mapsto y_u = e^{\eta t}\, v$. The same is true for the mapping $u \mapsto z_u$, because z_u is obtained from (7), hence $G : u \mapsto (y_u, z_u)$ is of class C^2 in the indicated spaces. Let us write $G(u) = (G_y(u), G_z(u))$. Inserting this in (3) and differentiating in a direction $h \in L^p(Q_T)$, we find

$$\begin{aligned}
\frac{\partial}{\partial t} G_y'(u)h - \Delta G_y'(u)h + R'(G_y(u))\, G_y'(u)h + \alpha\, G_z'(u)h &= h \quad \text{in } Q_T\\
\partial_\nu G_y'(u)h &= 0 \quad \text{in } \Sigma_T\\
(G_y'(u)h)(x,0) &= 0 \quad \text{in } \Omega\\
\frac{\partial}{\partial t} G_z'(u)h + \beta\, G_z'(u)h - \gamma\, G_y'(u)h &= 0 \quad \text{in } Q_T\\
(G_z'(u)h)(x,0) &= 0 \quad \text{in } \Omega.
\end{aligned}$$

This system is equivalent to (24). Differentiating the first partial differential equation above again with respect to another direction $k \in L^p(Q_T)$, we obtain

$$\begin{aligned}
&\frac{\partial}{\partial t} G_y''(u)[h,k] - \Delta G_y''(u)[h,k] + R'(G_y(u))\, G_y''(u)[h,k]\\
&\quad + R''(G_y(u))[G_y'(u)h, G_y'(u)k] + \alpha\, G_z''(u)[h,k] = 0,
\end{aligned}$$

i.e. the first equation of (25) with $h_1 := h$ and $h_2 = k$. The other equations follow easily. □

3. Well-posedness of the optimal control problems and first-order necessary optimality conditions

3.1. Solvability of the general optimal control problem

Let us introduce the *reduced objective functional.*

$$J(u) := f(y_u, z_u, u) = I(u) + \mu\, j(u).$$

Then we consider for both model equations (Schlögl for $\alpha = 0$, FitzHugh-Nagumo for $\alpha = 1$) the optimal control problem

$$\min_{u \in \mathcal{U}_{ad}} J(u) := f(y_u, z_u, u), \tag{26}$$

where the state (y_u, z_u) is the unique solution of equation (3) for the given control u.

Theorem 3.1 (Existence of an optimal solution)**.** *The optimal control problem* (26) *has at least one optimal solution* $\bar{u}$ *with associated optimal state* $\bar{y} := G(\bar{u})$*.*

Proof. The set $\mathcal{U}_{ad}$ is non-empty and weakly compact in $L^p(Q_T)$. Moreover, the reduced objective functional $J : u \mapsto f(y_u, z_u, u)$ is weakly lower semicontinuous in $L^p(Q_T)$ for $p > 2$ because of the compactness of the mapping $u \in L^p(Q_T) \to (y, z) \in L^2(Q_T)^2$ and the convexity of the terms involving the control. Notice also that the mapping $G : u \mapsto (y_u, z_u)$ is of class C^2. The result follows now by standard arguments. □

In view of this general result, the optimal control problems defined upon the Schlögl or the FitzHugh-Nagumo equations are solvable.

3.2. First-order necessary optimality conditions

Again, we deal with both problems at once by considering the general state equation (3). Let $\bar u \in \mathcal{U}_{ad}$ be a locally optimal control with associated state $(\bar y, \bar z)$. Since any global solution is also a local one, we formulate the optimality conditions for local solutions. The triple $(\bar y, \bar z, \bar u)$ has to satisfy a variational inequality including the subdifferential $\partial j(\bar u)$. We recall that

$$\partial j(\bar u) = \Big\{ \lambda \in L^\infty(Q_T) \,\Big|\, j(u) \geqslant j(\bar u) + \int_0^T\!\!\int_\Omega \lambda\,(u - \bar u)\,dx\,dt \quad \forall u \in L^\infty(Q_T) \Big\}.$$

In our case, where j is the norm of $L^1(Q_T)$, this means almost everywhere

$$\lambda(x,t) \in \begin{cases} \{1\}, & \text{if } \bar u(x,t) > 0 \\ [-1,1], & \text{if } \bar u(x,t) = 0 \\ \{-1\}, & \text{if } \bar u(x,t) < 0. \end{cases}$$

Lemma 3.1. *If $(\bar y, \bar z, \bar u)$ is a local solution to the optimal control problem* (26)*, then there exists a function $\bar\lambda \in \partial j(\bar u)$ such that, with μ introduced in* (4)*,*

$$I'(\bar u)(u - \bar u) + \int_0^T\!\!\int_\Omega \mu\,\bar\lambda(x,t)(u(x,t) - \bar u(x,t))\,dx\,dt \geqslant 0 \qquad \forall u \in \mathcal{U}_{ad}. \tag{27}$$

This result is obtained from the standard variational inequality

$$I'(\bar u)(u - \bar u) + \mu\, j'(\bar u, u - \bar u) \geqslant 0 \qquad \forall u \in \mathcal{U}_{ad},$$

where $j'(\bar u, u - \bar u)$ denotes the directional derivative of j. Then (27) follows from

$$j'(\bar u, u - \bar u) = \max_{\lambda \in \partial j(\bar u)} \langle \lambda\,,\, u - \bar u \rangle,$$

cf. [Casas et al.(2012)]. The derivative $I'(\bar u)$ is given by

$$\begin{aligned} I'(\bar u)(u - \bar u) = & \int_\Omega c_T^Y(\bar y(x,T) - y_T(x))\,v_{u-\bar u}(x,T) + c_T^Z(\bar z(x,T) - z_T(x))\,w_{u-\bar u}(x,T)\,dx \\ & + \int_0^T\!\!\int_\Omega c_Q^Y(\bar y - y_Q)\,v_{u-\bar u} + c_Q^Z(\bar z - z_Q)\,w_{u-\bar u}\,dx\,dt + \kappa \int_0^T\!\!\int_\Omega \bar u\,(u - \bar u)\,dx\,dt, \end{aligned} \tag{28}$$

where $v_{u-\bar u} \in C(\overline{Q_T})$ denotes the y-component and $w_{u-\bar u} \in C(\overline\Omega \times [0,T])$ the z-component of the derivative $G'(\bar u)(u - \bar u)$. Thanks to Theorem 2.2 and equation (24), $(v_{u-\bar u}, w_{u-\bar u})$ solves the linear equation

$$\frac{\partial}{\partial t} v - \Delta v + R'(\bar y)\,v + \alpha\,w = u - \bar u \qquad \text{in } Q_T \tag{29}$$

$$\partial_\nu v = 0 \qquad \text{in } \Sigma_T \tag{30}$$

$$v(x,0) = 0 \qquad \text{in } \Omega \tag{31}$$

$$\frac{\partial}{\partial t} w + \beta\,w - \gamma\,v = 0 \qquad \text{in } Q_T \tag{32}$$

$$w(x,0) = 0 \qquad \text{in } \Omega. \tag{33}$$

Introducing adjoint states, we are able to get rid of the functions $v_{u-\bar u}$ and $w_{u-\bar u}$ in (28). We define the following adjoint system for a pair of adjoint states $(\varphi_1,\varphi_2)\in W(0,T)\times W(0,T)$:

$$-\frac{\partial}{\partial t}\varphi_1-\Delta\varphi_1+R'(\bar y)\,\varphi_1-\gamma\,\varphi_2=c_Q^Y\,(\bar y-y_Q)\qquad \text{in } Q_T \tag{34}$$

$$\partial_\nu\varphi_1=0\qquad \text{in } \Sigma_T \tag{35}$$

$$\varphi_1(x,T)=c_T^Y(x)\,(\bar y(x,T)-y_T(x))\qquad \text{in } \Omega \tag{36}$$

$$-\frac{\partial}{\partial t}\varphi_2+\beta\,\varphi_2+\alpha\,\varphi_1=c_Q^Z\,(\bar z-z_Q)\qquad \text{in } Q_T \tag{37}$$

$$\varphi_2(x,T)=c_T^Z(x)\,(\bar z(x,T)-z_T(x))\qquad \text{in } \Omega. \tag{38}$$

Existence, uniqueness and regularity of the adjoint state $(\varphi_1,\,\varphi_2)$ follows from Theorem (2.1) after the transformation $\tilde\varphi_i(t):=\varphi_i(T-t)$, $i=1,2$. In this way, the (well-posed) backward problem (34)-(38) is transformed to a forward one, where Theorem 2.1 can be applied.

Remark 3.1. If $y_0,\ z_0$ belong to $C(\overline\Omega)$, then $(\bar y,\bar z)$ and $(\bar\varphi_1,\bar\varphi_2)$ are continuous on $\overline\Omega\times[0,T]$. If in addition $y_T,\ z_T$ belong to $C(\overline\Omega)$ and $y_Q,\ z_Q$ to $L^\infty(Q_T)$, the data $\bar y-y_Q$, $\bar z-y_Q$, $\bar y(\cdot,T)-y_T$ and $\bar z(\cdot,T)-z_T$ given in the adjoint system belong to $L^\infty(Q_T)$ and $C(\overline\Omega)$, respectively. In this case, Theorem 2.1 yields the regularity $\bar\varphi_i\in C\big(\overline\Omega\times[0,T]\big)$.

Lemma 3.2. *Let $(\bar\varphi_1,\bar\varphi_2)\in W(0,T)\times W(0,T)$ be the unique solution of the adjoint system* (34)-(38)*. Then it holds*

$$\begin{aligned}\int_0^T\!\!\int_\Omega \bar\varphi_1\,(u-\bar u)\,dx\,dt &= \int_0^T\!\!\int_\Omega c_Q^Y\,(\bar y-y_Q)\,v_{u-\bar u}+c_Q^Z\,(\bar z-z_Q)\,w_{u-\bar u}\,dx\,dt\\ &+\int_\Omega c_T^Y(\cdot)\,(\bar y(\cdot,T)-y_T(\cdot))\,v_{u-\bar u}(\cdot,T)+c_T^Z(\cdot)\,(\bar z(\cdot,T)-z_T(\cdot))\,w_{u-\bar u}(\cdot,T)\,dx.\end{aligned} \tag{39}$$

Proof. Let us write for short $v:=v_{u-\bar u}$, $w:=w_{u-\bar u}$. We multiply (29) with $\bar\varphi_1$ and (32) with $\bar\varphi_2$, integrate both equations over Q_T and integrate by parts in the term containing Δv. Next, we add both equations to obtain

$$\begin{aligned}\int_0^T\!\!\int_\Omega (u-\bar u)\,\bar\varphi_1\,dx\,dt &= \int_0^T \langle w'\,,\,\bar\varphi_2\rangle+\langle v'\,,\,\bar\varphi_1\rangle\ dt\\ &+\int_0^T\!\!\int_\Omega \beta\,w\,\bar\varphi_2-\gamma\,v\,\bar\varphi_2+\nabla v\cdot\nabla\bar\varphi_1+R'(\bar y)\,v\,\bar\varphi_1+\alpha\,w\,\bar\varphi_1\,dx\,dt.\end{aligned} \tag{40}$$

Next, we multiply (34) with v and (37) with w and perform the same operations as above. In addition, we integrate by parts with respect to t in the terms containing $\bar\varphi_i'$, $i=1,2$. We arrive at

$$\begin{aligned}\int_0^T\!\!\int_\Omega c_Q^Y\,(\bar y-y_Q)\,v+c_Q^Z\,(\bar z-z_Q)\,w\,dx\,dt &= -\int_\Omega \bar\varphi_1(\cdot,T)\,v(\cdot,T)+\bar\varphi_2(\cdot,T)\,w(\cdot,T)\,dx\\ +\int_0^T \langle v'\,,\,\bar\varphi_1\rangle+\langle w'\,,\,\bar\varphi_2\rangle\ dt &+\int_0^T\!\!\int_\Omega \nabla v\cdot\nabla\bar\varphi_1+R'(\bar y)\,v\,\bar\varphi_1-\gamma\,v\,\bar\varphi_2+\beta\,w\,\bar\varphi_2+\alpha\,w\,\bar\varphi_1\,dx\,dt.\end{aligned} \tag{41}$$

Now we insert the initial and final conditions $\bar\varphi_1(\cdot,T) = c_T^Y(\cdot)\,(\bar y(\cdot,T) - y_T(\cdot))$ and $\bar\varphi_2(\cdot,T) = c_T^Z(\cdot)\,(\bar z(\cdot,T) - z_T(\cdot))$ and subtract (41) from (40) to find

$$\begin{aligned}
&\int\limits_\Omega c_T^Y(\cdot)\,(\bar y(\cdot,T) - y_T(\cdot))\,v(\cdot,T) + c_T^Z(\cdot)\,(\bar z(\cdot,T) - z_T(\cdot))\,w(\cdot,T)\,dx \\
&= \int\limits_0^T\!\!\int\limits_\Omega \bar\varphi_1\,(u-\bar u)\,dx\,dt - \int\limits_0^T\!\!\int\limits_\Omega c_Q^Y\,(\bar y - y_Q)\,v + c_Q^Z\,(\bar z - z_Q)\,w\,dx\,dt.
\end{aligned} \tag{42}$$

This is equivalent to the statement of the lemma. □

The next theorem summarizes our findings.

Theorem 3.2 (Necessary optimality conditions)**.** *If $\bar u$ is a local solution to the optimal control problem* (26) *and $(\bar y,\bar z)$ is the associated state, then there exists a unique pair $(\bar\varphi_1,\bar\varphi_2) \in W(0,T)^2$ of adjoint states solving the adjoint system* (34)-(38) *and a function $\bar\lambda \in L^\infty(Q_T)$ such that*

$$\int\limits_0^T\!\!\int\limits_\Omega \big(\bar\varphi_1(x,t) + \kappa\,\bar u(x,t) + \mu\,\bar\lambda(x,t)\big)\big(u(x,t) - \bar u(x,t)\big)\,dx\,dt \geqslant 0 \qquad \forall u \in \mathcal{U}_{ad}. \tag{43}$$

Proof. The theorem follows from the relations (27), (28), and Lemma 3.2. □

In the case $\kappa > 0$, from the variational inequality (43) the following standard projection formula is obtained, cf. [Casas (1997)] or the exposition in [Tröltzsch (2010)]:

$$\bar u(x,t) = \mathbb{P}_{[a,b]}\Big\{-\frac{1}{\kappa}\big(\bar\varphi_1(x,t) + \mu\,\bar\lambda(x,t)\big)\Big\} \qquad \text{for a.a. } (x,t) \in Q_T, \tag{44}$$

where $\mathbb{P}_{[a,b]} : \mathbb{R} \to [a,b]$ is defined by

$$\mathbb{P}_{[a,b]}(u) = \max\{a, \min\{v,b\}\}.$$

A further discussion, related to the element $\bar\lambda$ of the subdifferential $\partial j(\bar u)$, reveals the sparsity of $\bar u$. The associated analysis is completely analogous to the one of the elliptic case discussed in [Casas et al.(2012)]. We recall this approach for convenience [Casas et al.(2013)].

Theorem 3.3. *Assume that κ and μ are positive. Then, for almost all $(x,t) \in Q_T$, there holds*

$$\bar u(x,t) \;=\; 0, \quad \textit{if and only if} \quad \begin{cases} |\bar\varphi_1(x,t)| \leqslant \mu, & \textit{if } a < 0 \\ \bar\varphi_1(x,t) \geqslant -\mu, & \textit{if } a = 0, \end{cases} \tag{45}$$

$$\bar\lambda(x,t) \;=\; \mathbb{P}_{[-1,1]}\Big\{-\frac{1}{\mu}\bar\varphi_1(x,t)\Big\}. \tag{46}$$

Proof. The proof is completely analogous to [Casas et al.(2012)]. Let us show this result for convenience of the reader. In the proof, we write $\bar\varphi := \bar\varphi_1$ to avoid the repeated use of the index 1. At first, we take $a > 0$. By the projection formula (44), it is obvious that

$$\bar u(x,t) = 0 \;\Leftrightarrow\; \bar\varphi(x,t) + \mu\,\bar\lambda(x,t) = 0 \tag{47}$$

holds for almost all $(x,t) \in Q_T$. Assume now the right-hand part of (47) holds true. Then we have $\bar\varphi(x,t) = -\mu\,\bar\lambda(x,t)$ and on the other hand it holds $|\bar\lambda(x,t)| \leqslant 1$. Together, this yields

$$|\bar\varphi(x,t)| = \left|\mu\,\bar\lambda(x,t)\right| = \mu\,\left|\bar\lambda(x,t)\right| \leqslant \mu.$$

Consider now the points $(x,t) \in Q_T$ with $\bar u(x,t) \neq 0$, say first that $\bar u(x,t) > 0$. Then (44) gives

$$0 < -\frac{1}{\kappa}\big(\bar\varphi(x,t) + \mu\,\bar\lambda(x,t)\big).$$

Because $\bar\lambda$ is equal to 1 where $\bar u > 0$, this inequality implies $0 > \bar\varphi(x,t) + \mu$, hence $\bar\varphi(x,t) < 0$ and finally

$$|\bar\varphi(x,t)| = -\bar\varphi(x,t) > \mu > 0.$$

A similar argumentation shows $|\bar\varphi(x,t)| > \mu$ also in the case $\bar u(x,t) < 0$. Therefore, for almost all $(x,t) \in Q_T$, nonzero values of $\bar u(x,t)$ can only occur if and only if $|\bar\varphi(x,t)| > \mu$. This shows the first case of (45).

If $a = 0$, it is easy to see that $\lambda(x,t) = 1$, because j is differentiable with derivative 1. For this reason, a simple computation and (44) lead to the second case of (45),

$$\bar u(x,t) = 0 \;\Leftrightarrow\; \bar\varphi(x,t) \geqslant -\mu$$

Next, we derive the projection formula (45) for $\bar\lambda$. In a.a. points with $\bar u(x,t) > 0$, it holds $\bar\lambda(x,t) = 1$ by the definition of the subdifferential. Then the projection formula (44) implies $-(\bar\varphi(x,t) + \mu\cdot 1) > 0$. This is equivalent to $-\bar\varphi(x,t)/\mu > 1$, hence

$$\bar\lambda(x,t) = 1 = \mathbb{P}_{[-1,1]}\Big\{-\frac{\bar\varphi(x,t)}{\mu}\Big\}$$

holds true as claimed. Analogously, (45) is proved, if $\bar u(x,t) < 0$. Consider finally the case $\bar u(x,t) = 0$. By formula (44), we have then $0 = \bar\varphi(x,t) + \mu\bar\lambda(x,t)$ and hence

$$\bar\lambda(x,t) = -\frac{\bar\varphi(x,t)}{\mu} = \mathbb{P}_{[-1,1]}\Big\{-\frac{\bar\varphi_1(x,t)}{\mu}\Big\},$$

because we also know that $\bar\lambda(x,t)$ belongs to $[-1,1]$ for $\bar u(x,t) = 0$. $\square$

From this theorem, we are able to draw several important conclusions about the behavior of the optimal control.

Corollary 3.1 (Sparsity of optimal controls)**.** *Assume that y_0, y_T, z_0, $z_T \in L^\infty(\Omega)$ and y_Q, $z_Q \in L^\infty(Q_T)$.*

(i) Then there exists a constant $M_\varphi > 0$ such that there holds

$$|\varphi_1(x,t)| \leqslant M_\varphi \quad \text{for a.a. } (x,t) \in Q_T \tag{48}$$

for all adjoint states φ_1 associated with any admissible control $u \in \mathcal{U}_{ad}$. Any locally optimal control $\bar u$ of the problem (26) vanishes in almost all points $(x,t) \in Q_T$, where $|\varphi(x,t)| > \mu$ is satisfied. For $\mu > M_\varphi$, problem (26) has the unique (locally) optimal control $\bar u = 0$.

(ii) The functions $\bar\lambda$ and $\bar u$ have the regularity $L^2(0,T;H_0^1(\Omega)) \cap C(\bar\Omega \times (0,T])$. If in addition y_T, z_T belong to $C(\overline{\Omega})$, then any locally optimal control $\bar u$ belongs to $C(\overline{\Omega} \times [0,T])$.

Proof. The set $\mathcal{U}_{ad}$ is bounded in $L^\infty(Q_T)$, therefore, Lemma 2.2 provides a bound for the norm of all admissible states v_u in $L^\infty(Q_T)$. Obviously, we have the same for all states z_u. Therefore, thanks to the boundedness of the data y_0, y_T, z_0, z_T, y_Q, $z_Q \in L^\infty(Q_T)$, the right-hand sides of the adjoint system (34)–(38) are uniformly bounded in $L^\infty(Q_T)$ or $L^\infty(\Omega)$, respectively. This turns over to the adjoint states, hence the existence of the constant M_φ in (48) is clear. The other statements of (i) are only a reformulation of property (45) in 3.3.

(ii) The adjoint state $\bar\varphi_1$ belongs to $L^2(0,T;H_0^1(\Omega))\cap C(\bar\Omega\times(0,T])$. By the representation (46), $\bar\lambda$ has the same regularity. The projection formula (44) yields this regularity for $\bar u$.

If y_T and z_T are continuous, then the continuity of the adjoint state $\bar\varphi_1$ follows from an application of Remark 3.1 to the adjoint system. Therefore, (46) and (44) ensure the continuity of $\bar\lambda$ and $\bar u$. □

Remark 3.2. If $\bar u$ is continuous in $\overline{Q_T}$, then the set $\{(x,t) \in Q_T \mid \bar u(x,t) \neq 0\}$ is open and hence the union of countably many disjoint open and connected subsets (so-called components) of Q_T, cf. [Alexandroff (2001)]. Therefore, the optimal control has a fairly regular shape.

4. Numerical examples

In this section, we present various numerical examples of optimal control problems for the Schlögl and the FitzHugh-Nagumo model, where traveling wave fronts or spiral waves are controlled. The sparse optimal controls turn out to be concentrated at the front region of the traveling waves. We shall also see in some of our examples that sparse controls improve the convergence of the CG method compared to the case of $\mu = 0$ without sparsity.

For the numerical treatment of the forward problem (3), respectively the adjoint system (34) - (38), we use a semi-implicit Euler-method with respect to the time and linear continuous finite elements for the spatial discretization. As optimization procedure for (6), respectively (26), a projected conjugate gradient method with nonlinear CG-step (namely the one by *Hestenes-Stiefel*) along with the strong *Wolfe-Powell* step-size rule for the line-search is used, cf. [Engel et al.(2013)].

We recall the steps of this method for the convenience of the reader:

1. Initialization: Select an initial control u^0 and an initial step size s_0. Compute $(y,z)^0 = (y_{u^0}, z_{u^0})$ (states), $\varphi_1^0 = \varphi_{1,y^0,z^0}$ (adjoint state), $\lambda^0 = \lambda_{u^0,\varphi_1^0}$ (subgradient of j), $g^0 = \kappa\, u^0 + \varphi_1^0 + \mu\,\lambda^0$ (subgradient of f), $d^0 = -g^0$ (anti-subgradient of f); set $k := 0$.

2. New subgradient:

$$
\begin{array}{lcll}
u^{k+1} & = & u^k + s_k d^k & \textit{(new control)},\\
(y,z)^{k+1} & = & (y_{u^{k+1}}, z_{u^{k+1}}) & \textit{(new states)},\\
\varphi_1^{k+1} & = & \varphi_{1,y^{k+1},z^{k+1}} & \textit{(new adjoint state)},\\
\lambda^{k+1} & = & \lambda_{u^{k+1},\varphi_1^{k+1}} & \textit{(new subgradient of } j\textit{)},\\
g^{k+1} & = & \kappa\, u^{k+1} + \varphi_1^{k+1} + \mu\,\lambda^{k+1} & \textit{(new subgradient of } f\textit{)}.
\end{array}
$$

3. Stop, if $\left\|g^{k+1}\right\| < \varepsilon$.

4. Direction of descent: Compute the conjugate direction according to Hestenes und Stiefel
$$d^{k+1} = -g^{k+1} + \frac{\left(g^{k+1}\right)^T\left(g^{k+1}-g^k\right)}{\left(d^k\right)^T\left(g^{k+1}-g^k\right)}.$$

5. Select the step size s_{k+1} by the *strong Wolfe-Powell rule*, cf. [Engel et al.(2013)], set $k := k+1$ and go to (2.).

The notation $\varphi_{1,y,z}$ is used for the adjoint state that corresponds to the solution φ_1 of (34) - (38) with y and z given instead of $\bar y$ and $\bar z$ in the right hand side. Moreover, λ_{u,φ_1} denotes the subgradient of $j(u)$ that satisfies formula (46) with respect to φ_1 in the points (x,t), where $u(x,t)=0$ holds. We remark that we modified (3.) such that the algorithm also terminates, if the controls and the states of two consecutive iterations differ only very marginally.

Since we also have to deal with a subgradient, we should remark that the projection formula (46) is used to set λ, if $|u|$ is below a certain threshold. Further, in all of our numerical optimization runs, we selected the weight $\kappa = 10^{-10}$ and started with the admissible initial control $u \equiv 0$.

4.1. The Schlögl model in $\Omega \subset \mathbb{R}$

In this subsection, we consider the model described by (1) and therefore, we set $c_Q^Z = c_T^Z = z_T = z_Q = 0$. The nonlinearity R is taken as $R(y) = \frac{1}{3}y^3 - y$. We discuss the problem of extinguishing a wave after a certain time and that of reaching a desired target state y_T in a one-dimensional spatial domain.

Example 1: Extinction of waves. We discuss the following setting: the domain is given by $\Omega = (0,40)$, the final time by $T=5$, and the initial state by

$$y_0(x) := \sum_{k=1}^{4}(-1)^k k^{-2}(1-\cos(\psi(x,\ell_k))),$$

where $\ell_1 = 7$, $\ell_2 = 16$, $\ell_3 = 24$, $\ell_4 = 33$, and ψ is defined as

$$\psi(x,\ell) := \begin{cases} (x-\ell+1)\frac{\pi}{7}, & \text{if } x \in [\ell-7,\ \ell+7] \\ 0, & \text{else.} \end{cases}$$

Without any control, i.e. for $u \equiv 0$, traveling wave fronts appear that expand outwards and

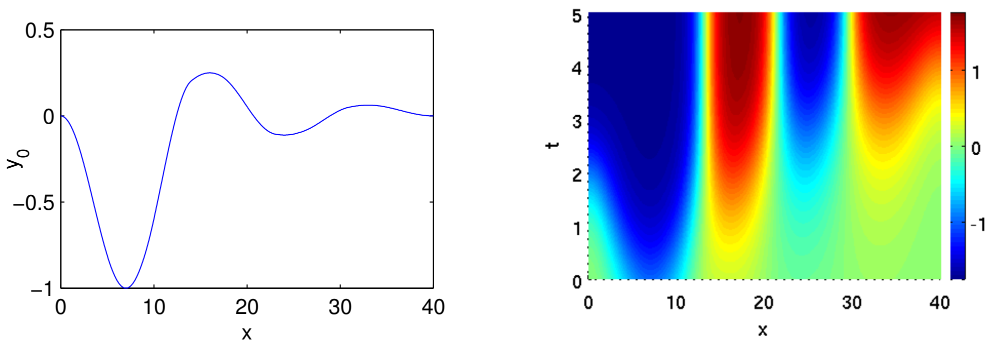

Figure 1. Example 1: Initial state y_0 (left) and state y_{nat} (right) for $u \equiv 0$.

seem to lock each other where they collide. In Fig. 1, the initial state y_0 and the development of the state y_{nat} as the solution of (1) for $u \equiv 0$ are presented. Notice that, in the figures displaying wave fronts, the vertical axis is the time axis.

Our task is to extinguish the wave fronts at $t = 1$ and keep the state at zero until $t = T$. Therefore, let $c_Q^Y(t) = 1$, if $t \in [1, T]$, and $c_Q^Y(t) = 0$, else, as well as $y_Q \equiv 0$. Further, we set $c_T^Y = 0$ and $y_T \equiv 0$. We split $(0, 40)$ and $(0, T)$ by partitions with $m = 601$ and $n = 500$ equidistant node points, respectively, and select $a = -1$ and $b = 1$. The optimal control and its associated state behave as shown in Fig. 2. Due to the constraints on u, the desired

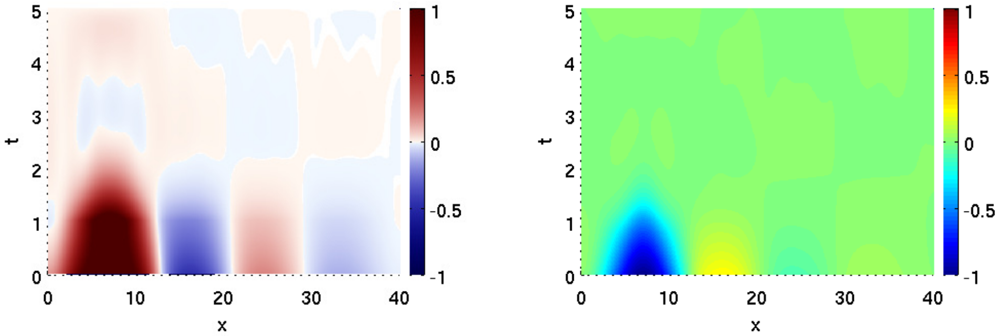

Figure 2. Example 1: Optimal control $\bar{u}$ (left) and associated state $\bar{y}$ (right) for $\mu = 0$.

state cannot be approximated that well. However, the objective functional value $J = 0.094$ is acceptable.

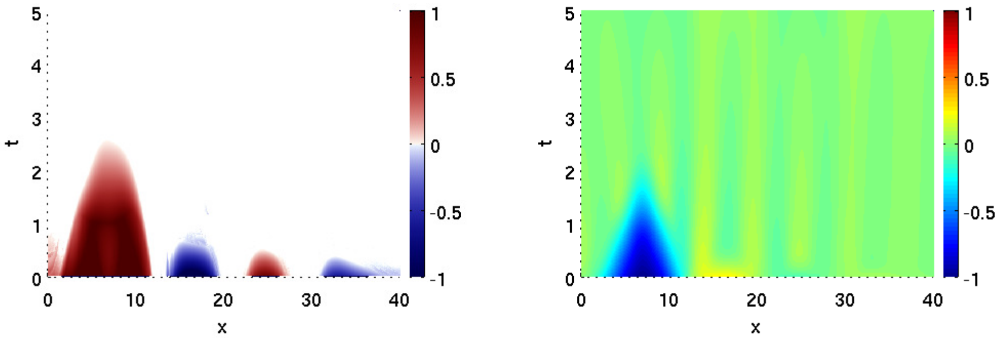

Figure 3. Example 1: Sparse optimal control $\bar{u}$ (left) and associated state $\bar{y}$ (right) for $\mu = 1$.

It is obvious that the optimal value increases as $\mu > 0$ increases. For $\mu = 1$, it is $J = 13.954$. Also the value $J(\bar{u}) - \mu j(\bar{u}) = 0.23$ (the objective functional value without the additional costs $\mu j(\bar{u})$) is larger than before. But the computed control is sparse, indeed. Fig. 3 presents the result for this case.

We remark that, in addition to the sparsity of the optimal control, the CPU-time for optimization decreases with increasing μ. For $\mu = 1$ it is half the time than for $\mu = 0$.

Example 2: Reaching a desired state. The second task is to reach a desired state y_T at the final time $t = T$. To this aim, we take $c_Q^Y = 0$, $y_Q \equiv 0$, and $c_T^Y = 1$. Moreover, let y_T be defined as y_0 of the first example. With $y_0 \equiv 0$, Q_T as before, $a = -\frac{1}{20}$, and $b = \frac{1}{20}$, an optimal objective functional value of $J = 5.07 \cdot 10^{-7}$ is reached.

Fig. 4 displays the optimal control for $\mu = 0$ and for $\mu = \frac{1}{4}$. The sparsity of $\bar{u}$ is quite obvious for $\mu = \frac{1}{4}$. In this case, the costs $J = 0.0396$ ($J(\bar{u}) - \mu j(\bar{u}) = 0.0032$) are much higher but still acceptable. The desired trajectory is qualitatively well approximated, since $\|\bar{y}(\,\cdot\,, T) - y_T\|_{L^\infty(\Omega)} = 3.12 \cdot 10^{-2}$.

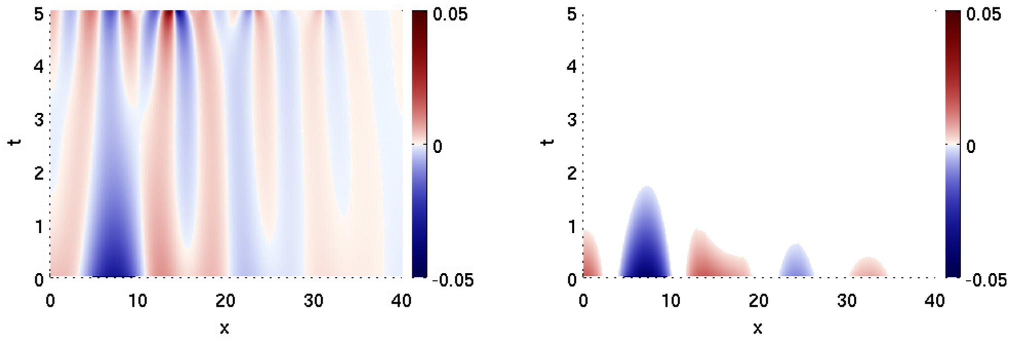

Figure 4. Example 2: Optimal control $\bar{u}$ for $\mu = 0$ (left) and sparse optimal control $\bar{u}$ for $\mu = \frac{1}{4}$ (right).

4.2. The Schlögl model in $\Omega \subset \mathbb{R}^2$

In this subsection, we consider the model (1) as before, but with the nonlinearity $R(y) = y(y-0.25)(y+1)$. Therefore, we set $c_Q^Z = c_T^Z = z_T = z_Q = 0$, again. Since the spatial dimension is $N = 2$, the optimization method turned out to be much slower than for $N = 1$, though we selected only 141×141 node points in Ω. Due to this, reducing the calculation time is even more important.

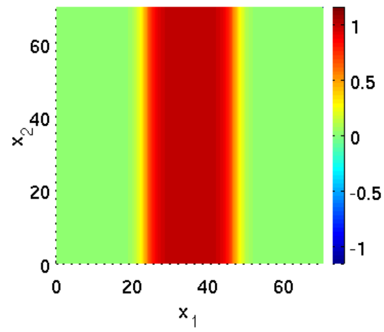

Figure 5. Initial state y_0.

For the examples that are discussed in this subsection, let $\Omega = (0, 70) \times (0, 70)$ and the initial state be given by two parallel wave fronts as shown in Fig. 5. We define the initial state y_0 by

$$y_0(x) := \left(\exp\left(\frac{\frac{70}{3} - x_1}{\sqrt{2}}\right)\right)^{-1} + \left(\exp\left(\frac{x_1 - \frac{140}{3}}{\sqrt{2}}\right)\right)^{-1} - 1.$$

Without any control, i.e. for $u \equiv 0$, the waves expand outwards in positive respectively negative x_1-direction and cover the whole spatial domain after approximately $t = 65$.

Example 3: Extinction of wave fronts. The first task is to extinguish both wave fronts at the final time $T = 20$. With $c_T^Y = 1$, $y_T \equiv 0$, $c_Q^Y = 0$, $y_Q \equiv 0$, $a = -1$, and $b = 1$, an optimal objective functional value of $J = 0.08$ is reached. Since the CG method was quite fast for $\mu = 0$, a positive parameter μ did not essentially reduce the needed CPU-time noticeable. Because the optimal control $\bar{u}$ shows similar behaviour along the x_2-axis, we only present it in the x_1-t-plane for $x_2 = 35$. Fig. 6 presents this view of $\bar{u}$ for $\mu = 0$ and $\mu = \frac{1}{10}$.

Although the use of $\mu = \frac{1}{10}$ causes an optimal functional value of $J = 242.19$ ($J(\bar{u}) - \mu j(\bar{u}) = 9.13$), we have $\|\bar{y}(.,T)\|_{L^\infty(\Omega)} = 0.097$ and therefore, the reached state is close enough to zero such that $\bar{y}(\cdot\,,t)$ tends rapidly to zero for increasing $t > T$ and $\bar{u}(\cdot\,,t) \equiv 0$ for all $t > T$. Notice that zero is a stable fixed point here.

Example 4: Turning and stopping wave fronts. Our second task is to rotate the vertical initial wave fronts of Example 3, displayed in Fig. 5, counter-clockwise in a horizontal position until $t = 60$ and to keep them like this for another 20 time-units. Therefore, we

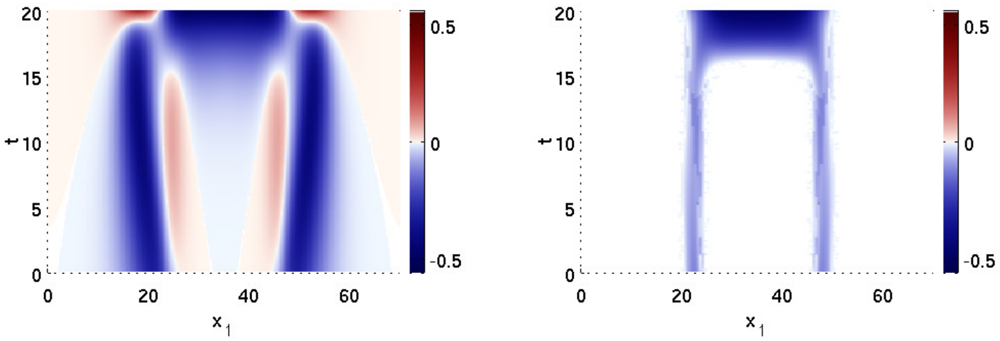

Figure 6. Example 3: Optimal control $\bar{u}$ for $\mu = 0$ (left) and sparse optimal control $\bar{u}$ for $\mu = \frac{1}{10}$ (right), both shown in the x_1-t-plane at $x_2 = 35$.

take $T = 80$, $a = -1$, $b = 1$, $c_Q^Y = 1$, $c_T^Y = 0$, $y_T \equiv 0$, and define the desired state y_Q as

$$y_Q(x,t) := \left(\exp\left(\frac{\cos(\psi(t))(\frac{70}{3} - x_1) + \sin(\psi(t))(\frac{70}{3} - x_2)}{\sqrt{2}}\right)\right)^{-1} + \left(\exp\left(\frac{\cos(\psi(t))(x_1 - \frac{140}{3}) + \sin(\psi(t))(x_2 - \frac{140}{3})}{\sqrt{2}}\right)\right)^{-1} - 1,$$

where $\psi(t) := \frac{\pi}{2} \min\left(1, \frac{4t}{3T}\right)$. Fig. 7 displays the behaviour of this desired trajectory.

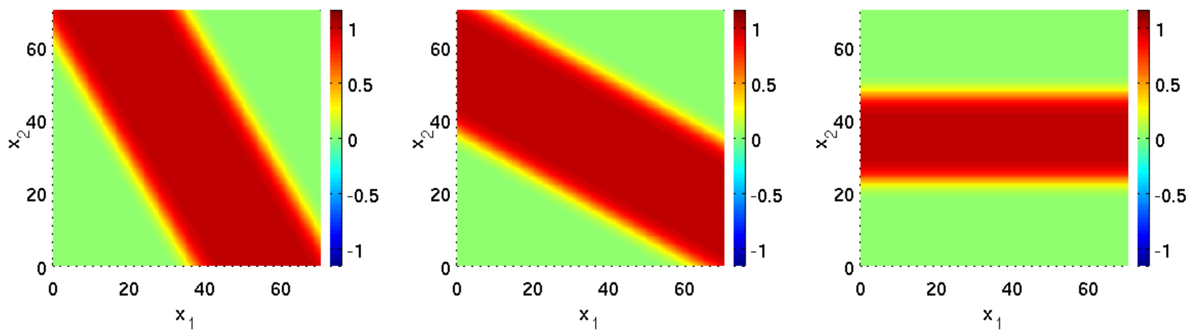

Figure 7. Example 4: Desired trajectory y_Q at $t = 20$ (left), $t = 40$ (middle) and $t = 60$ (right).

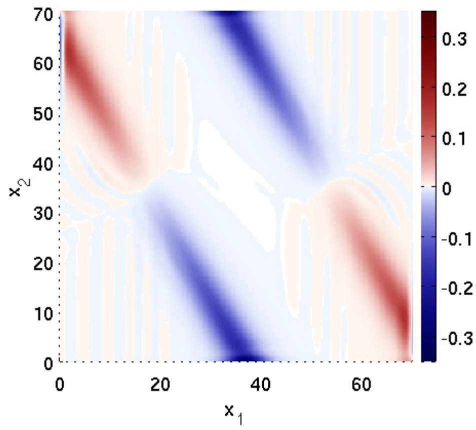

Figure 8. Example 4: Optimal control $\bar{u}$ at $t = 20$ for $\mu = 0$.

For $\mu = 0$, the desired state is approximated quite well and there is no visible difference between $\bar{y}$ and y_Q. The associated optimal control is presented in Fig. 8. We remark that the CG method was very slow for $\mu = 0$. After 440 iterations, it stopped at an objective functional value of $J = 0.083$. However, the method terminated much faster for a positive sparse-parameter $\mu > 0$. For instance, the method terminated after 59 iterations in the case $\mu = 1$. Indeed, the optimal costs $J = 4.95 \cdot 10^3$ are much higher, even if we take into account that J contains the additional cost $\mu\, j(\bar{u})$ and consider $J(\bar{u}) - \mu j(\bar{u}) = 911.62$, but the desired trajectory is qualitatively approximated well. The associated optimal controls and the corresponding states are presented in Fig. 9, exemplarily at $t = 20$.

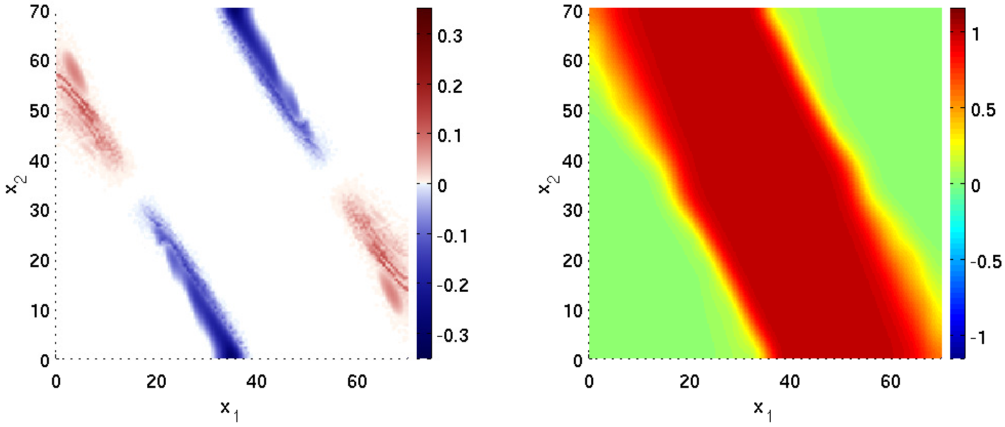

Figure 9. Example 4: Sparse optimal control $\bar{u}$ (left) and associated state $\bar{y}$ (right) at $t = 20$ for $\mu = 1$.

Considering the result of the CG method for $\mu = 0$ for the same fixed CPU-time, namely after 60 iterations, the desired state is not approximated as well as for $\mu = 1$. Fig. 10 displays the calculated control and associated state.

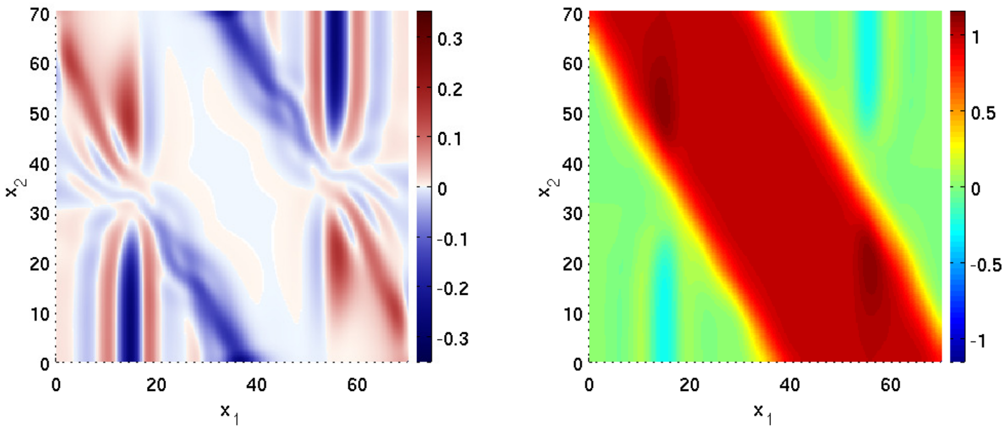

Figure 10. Example 4: Control u (left) and associated state y (right) at $t = 20$ for $\mu = 0$ after 60 iterations of the CG method.

It is worth mentioning that not only a better approximation of the terminal target is achieved by sparse controls but also a tremendous reduction of the running time of the CG method by more than 85% in this example.

4.3. The FitzHugh-Nagumo-model in $\Omega \subset \mathbb{R}^2$

As mentioned in the first section, the FitzHugh-Nagumo system (2) has solutions which form patterns of spiral waves. These patterns do not occur for all settings of the parameters β, γ, δ, k, y_1, y_2, and y_3. Since a spiral wave is needed as initial state in some of the following examples, we briefly introduce one possible way of exciting spiral wave solutions.

Let Ω be rectangular and $u = 1$ close to the bottom boundary of Ω in a certain short period of time and $u = 0$ elsewhere. As result, a traveling wave appears that propagates to the upper boundary of the spatial domain. After a short period of time, when the wave front is located between the upper and the bottom boundary, we set the state (y, z) equal to zero in the left half of Ω. Then the wave starts to curl up and forms a spiral pattern.

In the following examples, the considered spatial domain Ω is discretized by 97×97 up to 121×121 nodes. Further, we select $\beta = \frac{1}{100}$, $\delta = 0$, and $k = 1$. Moreover, we discuss the case of given trajectories y_Q and z_Q in the objective functional and we therefore set $c_T^Y = c_T^Z = y_T = z_T = 0$.

Example 5: Acceleration of a spiral wave. The first task is to accelerate an initially given spiral wave. We set $\Omega = (-150, 150)^2$, $\gamma = \frac{1}{500}$, $y_1 = 0$, $y_2 = \frac{1}{20}$, and $y_3 = 1$. With the above mentioned procedure, an initial state (y_0, z_0) is generated as shown in Fig. 11.

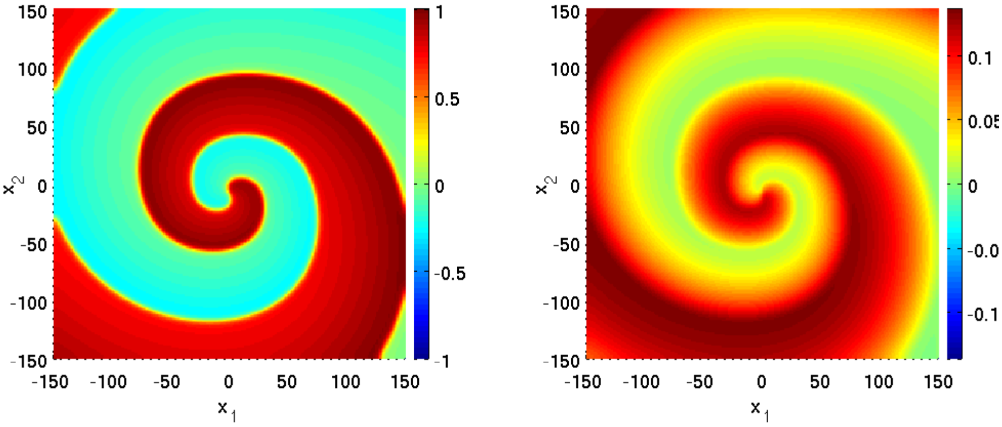

Figure 11. Example 5: Initial states y_0 (left) and z_0 (right) for the acceleration of the spiral wave.

Let $T = 50$ and $(y_{\text{nat}}, z_{\text{nat}})$ denote the natural development of (y, z) in the uncontrolled case $u \equiv 0$ starting with (y_0, z_0). We define

$$y_Q(x,t) := y_{\text{nat}}\left(x, \frac{1}{5}t^2 + t\right) \qquad \text{and} \qquad z_Q(x,t) := z_{\text{nat}}\left(x, \frac{1}{5}t^2 + t\right),$$

where the term $t^2/5$ accounts for the acceleration. The admissible controls are restricted by $a = -5$ and $b = 5$. After the optimization, an objective functional value of $J = 0.664$ is reached, since the calculated state follows the desired trajectory very well. Fig. 12 presents the optimal control and its associated state, exemplarily at $t = 40$. Graphically, the computed optimal state $\bar{y}$ coincides with y_Q.

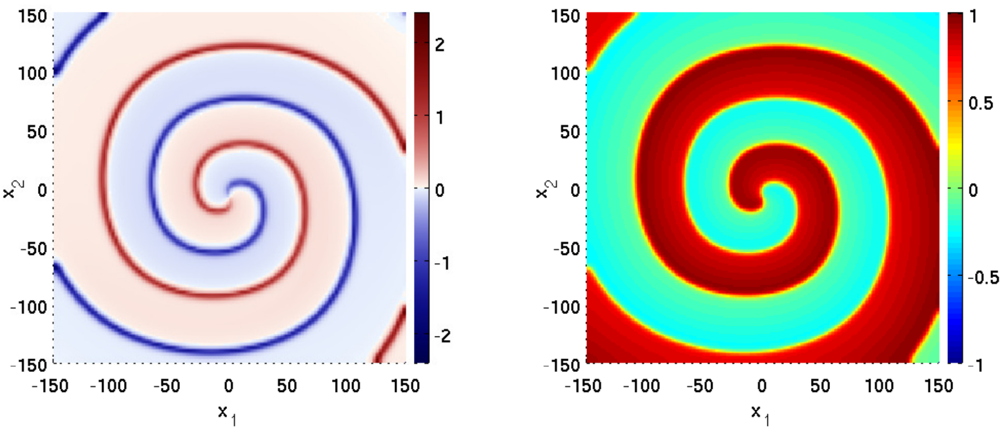

Figure 12. Example 5: Optimal control $\bar{u}$ (left) and associated state $\bar{y}$ (right) at $t = 40$ for $\mu = 0$.

Again, a positive parameter $\mu > 0$ causes sparsity of the optimal control and accelerates the CG method considerably. Instead of 850 iterations in the case of $\mu = 0$, the CG method stopped after only 59 iterations for $\mu = \frac{1}{3}$. Due to the high costs of $J = 1.351 \cdot 10^5$

($J - \mu j = 4.471 \cdot 10^4$), the result seems not to be acceptable. Nevertheless, the goal of accelerating the natural spiral wave is achieved fairly well. In Fig. 13, the calculated optimal state is displayed at $t = 40$. Comparing it to $\bar{y}$ in Fig. 12, the patterns are qualitatively well approximated. For any other time $0 \leqslant t \leqslant T$, a similar result is obtained.

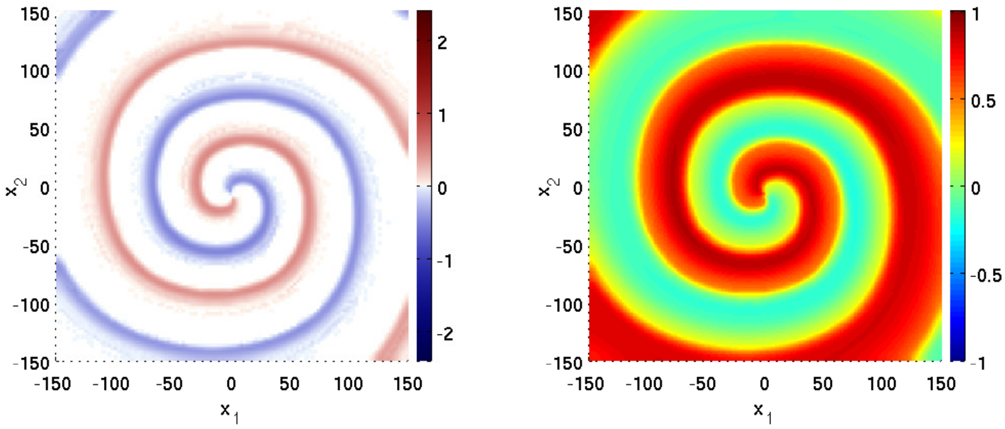

Figure 13. Example 5: Optimal control u (left) and associated state y (right) at $t = 40$ for $\mu = \frac{1}{3}$.

Example 6: Extinction of a spiral wave. The next example deals with the extinction of a spiral wave. This goal might be achieved by reaching the desired final state $(y_T, z_T) = (0, 0)$ at the final time $t = T$ instead of approaching a desired trajectory (y_Q, z_Q). But in this case, the target of extinction is not achieved. Therefore, we followed the idea of moving the tip of the spiral across the boundary of Ω, which was already known in Physics, cf. [Yang and Zhang (2006)]. In addition, referring to [Breuer (2006)], it was also known that a translation of the tip of the spiral wave is sufficient to move the whole wave. Determining the (approximated) tip numerically is very easy. The intersection of a level curve (isocline) of y with a level curve of z leads to a unique intersection point, the approximated tip, cf. [Breuer (2006)]. Based on this knowledge, we define the desired trajectory (y_Q, z_Q) and the coefficients $\left(c_Q^Y, c_Q^Z\right)$ in a way that only the area around a moved tip is considered in the objective functional.

Therefore, let $\overline{x} \in \Omega$ be the center of the tip trajectory of y_{nat}, the solution of (2) that develops in the uncontrolled case $u = 0$. The trajectory y_{nat} is a spiral that is turning around the fixed center $\overline{x}$. We define

$$\tilde{c}_Q^Y(x) := \begin{cases} 1, & \text{if } |x - \overline{x}| \leqslant r \\ 0, & \text{else} \end{cases} \quad \text{and} \quad \tilde{c}_Q^Z(x) := \begin{cases} 10, & \text{if } |x - \overline{x}| \leqslant r \\ 0, & \text{else} \end{cases} \tag{49}$$

for some $r > 0$, where $|\cdot|$ denotes the Euclidean norm. Thus, $\tilde{c}_Q^Y$ and $\tilde{c}_Q^Z$ only differ from zero in the circled area around $\overline{x}$ with radius r. To move the tip to a desired boundary point $\underline{x} \in \partial\Omega$, let us define the distance $\ell := |\underline{x} - \overline{x}|$ and the direction $d := (\underline{x} - \overline{x})/\ell$. With the velocity $c > 0$ for the desired movement, the supports of

$$c_Q^Y(x,t) := \tilde{c}_Q^Y(x - \min\{\ell, ct\}\, d), \qquad c_Q^Z(x,t) := \tilde{c}_Q^Z(x - \min\{\ell, ct\}\, d)$$

describe the translated circled area from $\overline{x}$ to $\underline{x}$ along a straight line in direction d. Analogously, we define the moved trajectories

$$\tilde{y}_Q(x,t) := c_Q^Y(x,t)\, y_{\text{nat}}(x - \min\{\ell, ct\}\, d), \qquad \tilde{z}_Q(x,t) := c_Q^Z(x,t)\, z_{\text{nat}}(x - \min\{\ell, ct\}\, d).$$

Notice, that y_{nat} and z_{nat} are spirals whose tip is turning around the fixed center $\overline{x}$. Since $(\tilde{y}_Q, \tilde{z}_Q)$ only accounts for the translation to the boundary but not for the extinction after it, we introduce

$$\psi(t) := \begin{cases} 1, & \text{if } t < T_1 \\ ((T_2 - t)/(T_2 - T_1))^2, & \text{if } t \in [T_1, T_2] \\ 0, & \text{if } t > T_2, \end{cases}$$

where $0 < \ell/c \leqslant T_1 < T_2$. Finally, the desired trajectories are given by

$$y_Q(x,t) := \psi(t)\, \tilde{y}_Q(x,t), \qquad z_Q(x,t) := \psi(t)\, \tilde{z}_Q(x,t).$$

Remark 4.1. Defining c_Q^Y and c_Q^Z in the way explained above, we restrict the *observation* of the states to a small moving region. This does not include that the control u is restricted to this region. However, we will observe that the support of the optimal control is indeed close to this region of observation. This confirms the knowledge of physicists.

In our example, let $\Omega = (-120, 120)^2$, $\gamma = \frac{3}{400}$, $y_1 = 0$, $y_2 = \frac{1}{200}$, $y_3 = 1$, $r = 20$, $\overline{x} = (0,0)$, $\underline{x} = (0, 120)$, $c = \frac{1}{16}$, $T = 2500$, $T_1 = 2000$, and $T_2 = 2120$. Further, we set

$$a = a(t) := \begin{cases} -5, & \text{if } t \leqslant T_1 \\ 0, & \text{else} \end{cases} \qquad b = b(t) := \begin{cases} 5, & \text{if } t \leqslant T_1 \\ 0, & \text{else} \end{cases}$$

such that the control is able to act only in $[0, T_1]$.

For $\mu = 0$, the optimization method did not lead to a satisfying result. Close to $t = 900$, the spiral wave starts to split in two parts. Fig. 14 presents the calculated optimal control and its associated state exemplarily at $t = 1200$. As result, a new spiral wave occurs whose tip is located around the point $(50, 95)$.

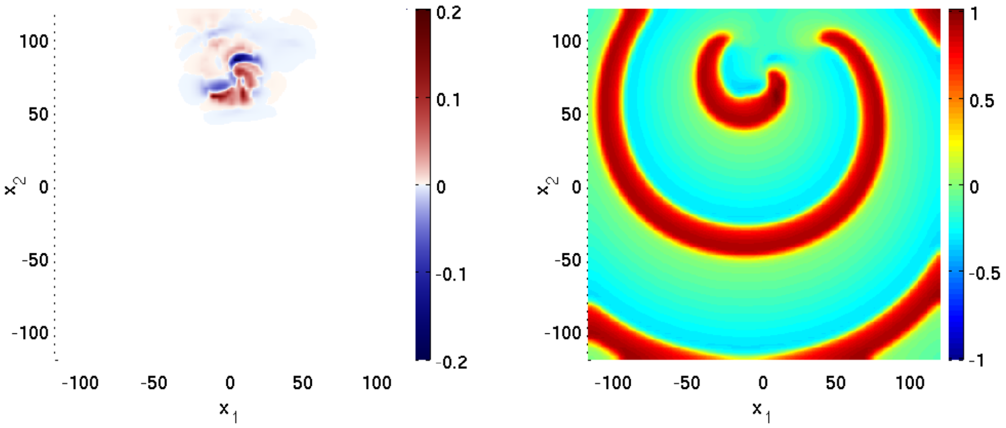

Figure 14. Example 6: Optimal control $\bar{u}$ (left) and associated state $\bar{y}$ (right) at $t = 1200$ for $\mu = 0$.

It is remarkable that, for sufficiently large μ close to 1, the spiral behaves more smoothly. Finally, the tip moves across the upper boundary of Ω such that the formed patterns disappear till $t = T$. This is exactly the desired development. Figure 15 displays the described behaviour in the case $\mu = 1$, exemplarily at $t = 1200$. With 169 iterations of the CG method for $\mu = 0$ and 161 iterations for $\mu = 1$, we cannot observe a further reduction of CPU time for increasing μ.

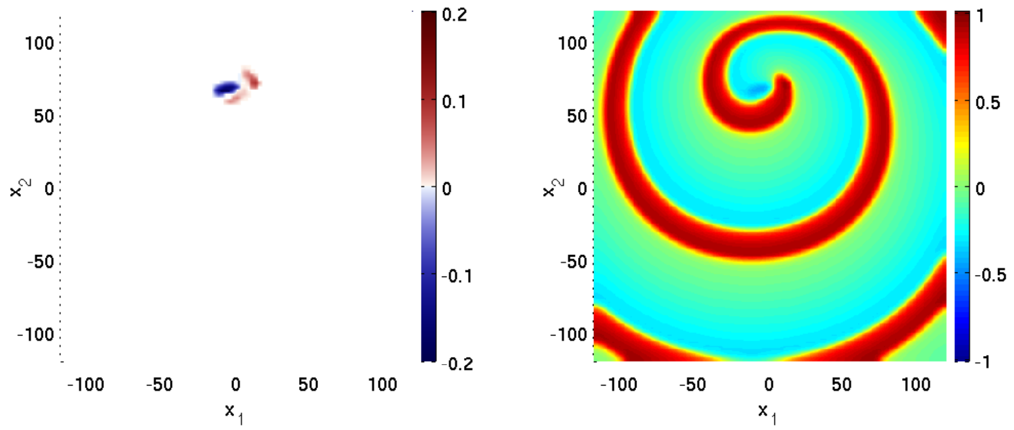

Figure 15. Example 6: Sparse optimal control $\bar{u}$ (left) and associated state $\bar{y}$ (right) at $t = 1200$ for $\mu = 1$.

Example 7: Exciting a spiral wave. Let us finally consider a different situation. Starting from an initial state that does not show a spiral pattern, a spiral should be excited by the control. For convenience we take $(y_0, z_0) = (0, 0)$. The use of only a desired final state (y_T, z_T) fails here again. However, we profit from our experience from the last example and define with the use of (49), $\big(c_Q^Y, c_Q^Z\big)(x,t) := \big(\tilde{c}_Q^Y, \tilde{c}_Q^Z\big)(x,t)$, where this time $r = r(t) := \frac{4}{7}t$. Moreover, let $(y_Q, z_Q)(x,t) := (y_{\text{nat}}, z_{\text{nat}})(x,t)$ and $(y_{\text{nat}}, z_{\text{nat}})$ denote the development of (y, z) in the uncontrolled case $u \equiv 0$.

Let $\Omega = (-150, 150)^2$, $\gamma = \frac{1}{200}$, $y_1 = 0$, $y_2 = \frac{1}{10}$, $y_3 = 1$, $T = 800$, $a = -5$, and $b = 5$. Though the objective functional value $J = 1.299 \cdot 10^5$ is quite large, the optimal control forces the state to develop as desired. A spiral wave is excited not only for $\mu = 0$ a but also

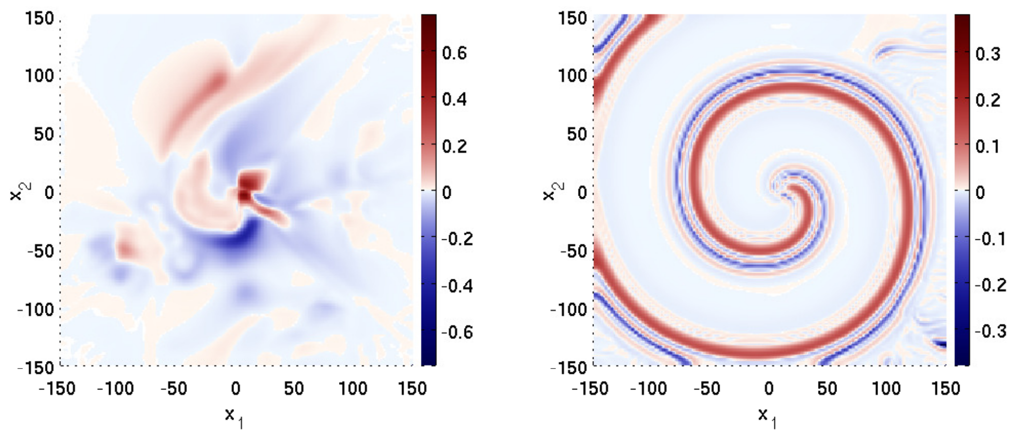

Figure 16. Example 7: Optimal control $\bar{u}$ at $t = 1$ (left) and at $t = 400$ (right) for $\mu = 0$.

in the case of $\mu > 0$. In this example, the sparsity of the optimal control is remarkable for sparse-parameters in the scale of $\mu = 10$. Comparing Fig. 16 with Fig. 17 presenting the optimal control for $\mu = 0$ and $\mu = 10$, both exemplarily at $t = 1$ and $t = 400$, we observe that it is sufficient to control only in small sub-domains of Ω to gain the required development.

All examples presented for the FitzHugh-Nagumo equations show that sparse controls are able to influence the trajectory in a satisfactory way. Even if the objective functional admits fairly large optimal values, the qualitative approximation of the desired trajectories is completely acceptable. Moreover, the use of sparse controls can accelerate the CG method of optimization. Finally, we should mention that the support of the computed sparse controls shows remarkable coincidence with the front part of the traveling waves or the spiral waves.

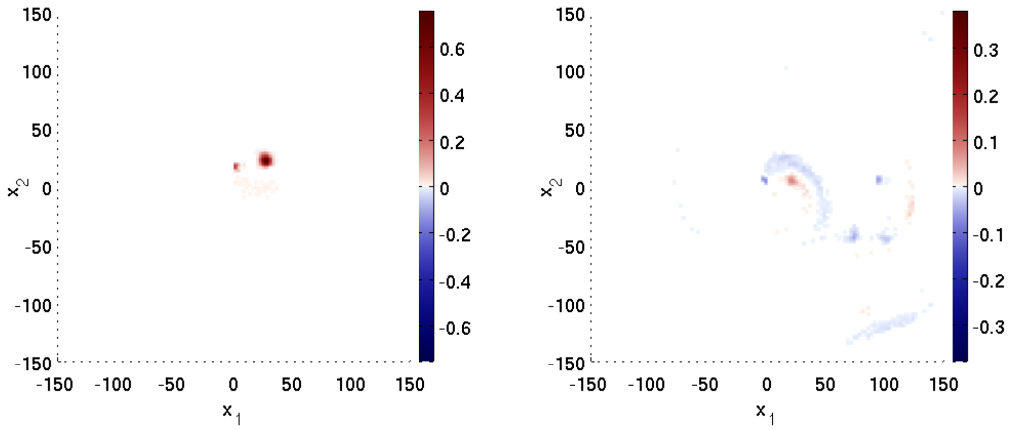

Figure 17. Example 7: Sparse optimal control $\bar{u}$ at $t = 1$ (left) and at $t = 400$ (right) for $\mu = 10$.

This somehow reflects the intuition of physicists.

References

[Alexandroff (2001)] P. S. Alexandroff: *Lehrbuch der Mengenlehre*, 7nd edn., Harri Deutsch, (2001).

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

[Brandão et al.(2008)] A. J. V. Brandão, E. Fernández-Cara, P. M. D. Paulo, M. A. Rojas-Medar: *Theoretical analysis and control results for the FitzHugh-Nagumo equation*, Electron. J. Differential Equations, No. 164, pp. 1-20 (2008).

[Breuer (2006)] J. Breuer: *Local control of complex patterns in excitable media*, Diploma thesis. TU Berlin, Institute of theoretical physics, (2006).

[Casas (1997)] E. Casas: *Pontryagin's principle for state-constrained boundary control problems of semilinear parabolic equations*, SIAM Journal on Control and Optimization, 35, pp. 1297-1327, (1997).

[Casas et al.(2012)] E. Casas, R. Herzog, G. Wachsmuth: *Optimality conditions and error analysis of semilinear elliptic control problems with L^1 cost functional*, SIAM Journal on Optimization, 22 (3), pp. 795-820, (2012).

[Casas et al.(2013)] E. Casas, R. Herzog, G. Wachsmuth: *Various ways to get sparsity in control problems of parabolic equations*, To appear in the Proceedings of the 1st IFAC Workshop on Control of Systems Modeled by Partial Differential Equations, CPDE, Paris 2013.

[Casas et al.(2008)] E. Casas, J. C. de los Reyes, F. Tröltzsch: *Sufficient second-order optimality conditions for semilinear control problems with pointwise state constraints*, SIAM J. Optim., 19(2), 616-643 (2008).

[Engel et al.(2013)] H. Engel, E. Kammann, F. Tröltzsch: *On the optimal control of the Schlögl-model*, To appear in Computational Optimization and Applications (2013), doi:10.1007/s10589-013-9550-y.

[FitzHugh (1961)] R. FitzHugh: *Impulses and physiological states in theoretical models of nerve membrane*, Biophys. Journal, 1, pp. 445-465 (1961).

[Jackson (1990)] D. E. Jackson: *Existence and regularity for the FitzHugh-Nagumo equations with inhomogeneous boundary conditions*, Nonlinear Analysis, Theory, Methods & Applications, vol. 14, No. 3, pp. 201-216 (1990).

[Kunisch and Wagner (I)] K. Kunisch, M. Wagner: *Optimal control of the bidomain system (I): the monodomain approximation with the Rogers-McCulloch model*, Nonlinear Anal., Real World Appl. 13(4), pp. 1525-1550 (2012).

[Kunisch and Wagner (II)] K. Kunisch, M. Wagner: *Optimal control of the bidomain system (II): Uniqueness and regularity theorems for weak solutions*, Ann. Mat. Pura Appl. doi:10.1007/s10231-012-0254-1.

[Kunisch and Wagner (III)] K. Kunisch, M. Wagner: *Optimal control of the bidomain system (III): Existence of minimizers and first-order optimality conditions*, ESAIM: Math. Model. Numer. Anal. doi:10.1051/m2an/2012058.

[Kunisch and Wang (2012)] K. Kunisch, L. Wang: *Time optimal controls of the linear FitzHugh-Nagumo equation with pointwise control constraints*, J. Math. Anal. Appl. 395, No. 1, pp. 114-130 (2012).

[Ladyzhenskaya et al.(1968)] O. A. Ladyzhenskaya and N. N. Ural'ceva: *Linear and Quasilinear Equations of Parabolic Type*, American Math. Society, Providence (1968).

[Mantel et al.(1996)] R. Mantel, D. Barkley, *Periodic forcing of spiral waves in excitable media*, Physical Review E, 54, pp. 4791-4802, (1996).

[Murray (1993)] J. D. Murray: *Mathematical Biology*, Biomathematics, vol. 19, 2nd edn. Springer, Berlin (1993).

[Nagumo (1962)] J. Nagumo: *An active pulse transmission line simulating nerve axon*, Proc. IRE, 50, pp. 2061 - 2071, (1962).

[Raymond and Zidani (1999)] J.-P. Raymond, H. Zidani, *Hamiltonian Pontryagin's Principles for control Problems governed by Semilinear Parabolic Equations*, Applied Math. and Optimization, 39, pp. 143-177, (1999).

[Schöll and Schuster (2007)] E. Schöll, H. G. Schuster: *Handbook of Chaos Control*, 2nd edn., Wiley-VCH, (2007).

[Schrader et al.(1995)] A. Schrader, M. Braune, H. Engel, *Dynamics of spiral waves in excitable media subjected to external periodic forcing*, Physical Review E, 52, pp. 98-108, (1995).

[Smoller (1994)] J. Smoller: *Shock Waves and Reaction-Diffusion Equations*, 2nd edn. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 258. Springer, NewYork, (1994).

[Stadler (2009)] G. Stadler: *Elliptic optimal control problems with L^1-control cost and applications for the placement of control devices*, Computational Optimization and Applications, 44 , pp. 159-181, (2009).

[Tröltzsch (2010)] F. Tröltzsch: *Optimal Control of Partial Differential Equations: Theory, Methods and Applications*, volume 112, American Math. Society, Providence (2010).

[Wachsmuth and Wachsmuth (2010)] G.. Wachsmuth and D. Wachsmuth: *Convergence and regularisation results for optimal control problems with sparsity function*, ESAIM Control Optimisation and Calculus of Variations, 17 (3), pp. 858-886, (2010).

[Yang and Zhang (2006)] J. Yang, M. Zhang: *Drift of Spiral Waves in Complex Ginzburg-Landau Equation*, Communications in Theor. Phys. 45, pp. 647-652 (2006).

[Zykov and Engel(2004)] V. Zykov, H. Engel: *Feedback-mediated control of spiral waves*, Communications in Physica D, vol. 199, pp. 243-263, (2004).